

人脸表情识别综述

王大伟, 周 军, 梅红岩, 张素娥

WANG Dawei, ZHOU Jun, MEI Hongyan, ZHANG Su'e

辽宁工业大学 电子与信息工程学院, 辽宁 锦州 121001

School of Electronics & Information Engineering, Liaoning University of Technology, Jinzhou, Liaoning 121001, China

WANG Dawei, ZHOU Jun, MEI Hongyan, et al. Summary of facial expression recognition. *Computer Engineering and Applications*, 2014, 50(20): 149-157.

Abstract: As a research direction of the affective computing, facial expression recognition which constitutes the basis of emotion understanding, is the premise to complete human-computer interaction intelligent. Facial expression is so exquisite that it consumes a large amount of computation time and influences the timeliness and experience feeling from human-computer interaction intelligent. Consequently, facial feature extraction has become an important research topic in the area of facial expression recognition. The progress and stable framework for facial expression recognition in recent five years are generalized, a serial of algorithms applied in feature extraction and expression classification are summarized. Then, the main algorithms and their improvement are described in detail, and advantages and disadvantages among in different algorithms are analyzed and compared. In the same time, comparison with the other algorithms is also introduced. The challenges and shortcomings are pointed out by the research of practical problems in facial expression recognition application.

Key words: expression recognition; feature extraction; feature classification; algorithm improvement; algorithm comparison

摘 要: 人脸表情识别作为情感计算的一个研究方向, 构成了情感理解的基础, 是实现人机交互智能的前提。人脸表情的极度细腻化消耗了大量的计算时间, 影响了人机交互的时效性和体验感, 所以人脸表情特征提取成为人脸表情识别的重要研究课题。总结了国内外近五年的人脸表情识别的稳固框架和新进展, 主要针对人脸表情特征提取和表情分类方法进行了归纳, 详细介绍了这两方面的主要算法及改进, 并分析比较了各种算法的优势与不足。通过对国内外人脸表情识别应用中实际问题进行研究, 给出了人脸表情识别方面仍然存在的挑战及不足。

关键词: 表情识别; 特征提取; 表情分类; 算法改进; 算法对比

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.1312-0123

1 引言

人脸表情识别作为情感识别的一个方面, 在人际交流中起到非常重要的作用, 不仅是展示情绪, 更是传播情感信息与协调双方关系的重要方式。据心理学家 Mehrabian^[1]的研究表明, 在人类的日常交流中, 通过人脸表情传递的信息高达信息总量的 55%, 而通过声音、语言传递的信息分别占信息总量的 38% 和 7%。Schlosberg^[2]首先提出了描述表情的 3 个量化尺度: 注意-拒绝、高兴-不高兴和活跃程度, 到 20 世纪 70 年代美国心理学家

Ekman 和 Friesen^[3]通过充分的测试实验, 定义了人类的 6 种基本表情: 高兴(Happy)、生气(Angry)、吃惊(Surprise)、恐惧(Fear)、厌恶(Disgust)和悲伤(Sad), 随后 Ekman 和 Keltner^[4]更进一步地细化及完善了面部表情, 并提出了基于运动单元(AUs)来描述面部表情的面部动作编码系统(FACS)。

人脸表情识别为人工智能注入了新的血液, 对人工智能的研究开始了复兴, 掀起了又一次的研究高潮。Suwa 等人于 1978 年对表情识别做了一个最初的尝试^[5],

基金项目: 国家自然科学基金(No.61074014); 辽宁省教育厅重点实验室项目(No.LS2010079)。

作者简介: 王大伟(1989—), 男, 硕士生, 研究领域: 图像处理与模式识别; 周军(1966—), 女, 教授, 硕士生导师, 研究领域: 数据挖掘、图像处理; 梅红岩(1978—), 女, 博士研究生, 讲师, 研究领域: 粗糙集、数据挖掘; 张素娥(1987—), 女, 硕士生, 研究领域: 图像处理与模式识别。E-mail: wdw8027@gmail.com

收稿日期: 2013-12-10 **修回日期:** 2014-02-17 **文章编号:** 1002-8331(2014)20-0149-09

CNKI 网络优先出版: 2014-04-22, <http://www.cnki.net/kcms/doi/10.3778/j.issn.1002-8331.1312-0123.html>

Terzopoulos 和 Waters^[6]则在同一年运用简化的 Ekman-Friesen 模型,做了人脸视频序列的表情分析,到 1981 年有研究者用仿生学的方法从肌肉角度为面部表情建立模型,再到 1991 年 A.Pentland 和 K.Mase^[7]将人工的表情分类引入到计算机自动化处理中,伴随着图像处理、模式识别技术的稳固发展,使其引起了全世界的广泛关注,并形成了一个单独的研究方向——人脸表情自动识别。

随着数据库采集设备及技术的飞速发展,使得获取图片和视频的分辨率及像素越来越高,信息越来越细腻、越来越庞大,为人脸表情识别带来了更完整、更全面的信息。同时由于信息量过大,使得识别及分类算法的效率受到严重的影响。如何提高算法的效率就成为了分类识别算法中重要的研究方面,于是特征提取的方法被提了出来。其主要方法就是:不用原始图像进行表情分类识别,而在原始图像的基础上在确保不丢失重要信息的前提下来降低特征维数,提取出图像最为本质的信息。实验证明这种方法有效可行。实现特征提取后,进行人脸表情识别中最核心的一步——表情分类:将提取出来的特征彼此进行分析匹配,为所获取的图片标上相应的人脸表情类别。

本文对人脸表情识别的研究背景、发展过程进行了综述,着重介绍了人脸表情特征提取和人脸表情识别的主要算法,列举了近两年对这几种算法的改进以及其他算法的对比,最后对人脸表情识别发展趋势及所存在的问题和进一步的研究工作进行了分析。

2 人脸表情识别的一般步骤

经过几十年的发展,人脸表情自动识别已经形成了一个完整的体系,Pantic^[8]和 Fasel^[9]也分别从不同的角度系统地总结了整个完整体系。其主要的步骤有图片获取、图像预处理(主要有图片归一化、人脸检测与定位^[10])、特征提取及表情分类。其主要框架如图 1 所示。

2.1 图片采集

人脸表情识别是在人脸识别的基础上发展而来的,其图片采集的方式是大体相同的。首先,人脸识别与人脸表情识别的图像类型根据图像来源与色彩信息,又分别可以分为静态图像与动态图像、彩色与灰度。人脸识别的图像都可以是通过数码相机或监控设备来获取,但是这样获取的图像或者图像序列往往会存在背景复杂、人脸数目未知、人脸姿势等问题,而人脸表情识别更注重的是一个确定面部上的细微的表情变化,所以在人脸表情识别中一般采集正面的单人面部的图片,为了更方便地专注用算法的改进,人脸表情识别实验中常采用专

业人脸表情图像库,如:JAFPE、CK/CK+、MMI、Yale、AR、BU-3DFE、清华大学人脸表情视频数据库、BHU 人脸表情数据库、CAS-PEAL 人脸数据库。

2.2 图片预处理

本文主要介绍人脸表情识别,所以在一张原图上最希望的是能够完整地提取出整个脸部的特征信息。因此在进行特征提取之前往往需要对原始图片进行预处理。图片预处理主要分为两个方面:除去原图的复杂背景、图像归一化。

人脸检测技术的发展为人脸表情识别中如何除去原图的复杂背景提供了很好的解决方案。田巍等提出了一种基于人脸恒定特征的方法,通过阈值在 HSV 色彩空间中对人体肤色区域进行肤色分割,并采用相对重要性滤波及自适应区域归并^[11],最终将人脸从复杂背景中提取出来;姜军等提出了一种基于知识的快速人脸检测方法——广义几何投影方法,通过符合人脸生理结构特征的人脸镶嵌图模型^[12],针对人脸图像的灰度和边缘信息进行多级检测步骤,从而提取出人脸的区域;王智文等提出了一种基于模板匹配的人脸检测方法,通过人脸肤色混合高斯模型对人脸进行分割,并利用自适应模板匹配及二次匹配算法来减少类肤色背景的影响^[13],从而实现过滤掉复杂背景的效果;余龙华等提出了一种基于统计方式的马尔科夫模型的人脸识别,建立马尔科夫模型并通过采样窗对人脸图像进行采样及离散余弦变换,提取变换后的系数作为观察向量^[14],最后进行训练便得到一个人的 HMM。以上四种类型的人脸检测的方法可以实现除去原图复杂背景的目的。

由于图像采集的方式不同,所产生的图片格式及大小有所差异,为了图片在利用过程中的方便性,通常需要对原始图片进行归一化处理。几何归一化也称为位置校准,其思想是利用图像的不变矩寻找一组不变量参数使其能够避免其他的变换函数对图像变换的影响。它将有助于矫正由于成像距离或人脸姿势变化所造成的尺寸差异和角度倾斜,并以此解决人脸尺度变化和人脸旋转问题。具体包括人脸尺度归一化、平面人脸旋转矫正(歪头)、深度人脸旋转矫正(扭脸)三个环节。实现几何归一化后还可以通过拟合平面法^[15]、自商图法^[16]、Tan 提出的算法^[17]等方法来解决由于光照不均造成的灰度差异大的问题,即灰度归一化。经过多个步骤的配合最终完成原始图片的归一化处理。

2.3 人脸表情特征提取

表情特征提取是人脸表情识别中最重要的部分,有效的表情特征提取工作能够很大幅度地提升识别的性



图1 人脸表情自动识别主要步骤

能。经典的特征提取算法所该具备的几个条件^[18]:

- (1)将人脸表情的最本质特征完整地表示出来。
- (2)将与表情无关的干扰信息如光照、背景及噪声去除掉。
- (3)将图像维数降低,选取合适的数据结构,减小存储量。
- (4)将不同类别的表情特征进行较好的区分。

符合上述条件的特征提取算法,往往分为三个步骤进行实现。首先是原始特征的生成,即提取图片中所包含的所有信息,原始特征存在信息量过大,维数过高,信息冗余等问题。为了更好地展现出图片所包含的最本质信息,需对原始特征进行进一步加工,即以下两个步骤——特征降维和提取、特征分解,以达到降低维数,去除噪声干扰,得到对分类更为有利的特征数据。由 Calder Andrew J等在2001年在面部表情识别中引入的PCA^[19]与由T.Ahonen等在2004年在人脸识别中引入的LBP^[20]一直是特征提取中的两个经典的算法,下面对其进行简要的描述。

2.3.1 主成分分析法(PCA)

统计学上PCA的定义为,用几个较少的综合指标来代替原来较多的指标,而且这些较少的综合指标技能尽多地反映原来较多指标的有用信息,且相互之间又是无关的。

Dubussion也在文献[21]中提到了对特征用Principal Component Analysis(PCA,主成分分析法)进行分析,然后按照主元对某一识别任务的重要性将它们进行排列,采用前向逐步选择的方法选择和保留最有区分性的 k 个主元方向。

PCA的主要思想是通过协方差矩阵分析各个属性之间的相关性,选取出该图像的主成分来排除这些冗余信息,并形成变换矩阵;再通过该矩阵实现Karhunen-Loève变换(正交变换),将原有的高维图片所形成的向量进行降维。其过程如下:

步骤1 选取训练集。 N 幅人脸图片形成一个训练集 X ,表示为 $[x_1, x_2, \dots, x_N]$ 。其中每一张图片 $x_i(1 \leq i \leq N)$ 的矩阵组成都是 $m \times n$,设 $T=m \times n$,则每幅图片可以通过一个 T 维的向量表示,即

$$x_i = [x_{i1}, x_{i2}, \dots, x_{iT}]^T, 1 \leq i \leq N$$

步骤2 计算相关性。由于人脸表情特征之间的相关性,从而会引起信息的冗余。协方差矩阵是相关性的最好度量,通过计算协方差矩阵的特征值,并选取其中 M 个特征值来求出对应的特征向量,用选取出来的特征向量来组成变换矩阵 $P=[p_1, p_2, \dots, p_M]^T$,其中 $p_j=[p_{j1}, p_{j2}, \dots, p_{jT}]$, $1 \leq j \leq M$ 。在线性转换的过程中, p_j 起到权值的作用,根据属性对于人脸表情识别的贡献率来赋予相应的权值,甚至只保留贡献率最大的几个属性而

直接忽略掉其他属性,即:决定 M 的取值,该 M 对应的 p_j 便是人脸表情特征的主成分。

步骤3 通过正交变换实现降维。训练集 X 通过正交变换 P 得到的 Y ,其中 X 是 $T \times N$ 的矩阵, P 是 $M \times T$ 的矩阵, Y 是 $M \times N$ 的矩阵,即

$$Y=[y_1, y_2, \dots, y_N]=PX=\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \end{bmatrix} [x_1, x_2, \dots, x_N]$$

其中 $y_i=[y_{i1}, y_{i2}, \dots, y_{iM}]^T, 1 \leq i \leq N$ 。则

$$y_i = \begin{bmatrix} p_{11}x_i \\ p_{21}x_i \\ \vdots \\ p_{M1}x_i \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1T} \\ p_{21} & p_{22} & \dots & p_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \dots & p_{MT} \end{bmatrix} \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iT} \end{bmatrix}$$

即 x_i 与 y_i 通过线性变换形成了一对一的关系, x_i 可以由 y_i 来重新表示。由于 x_i 是 T 维向量, y_i 是 M 维向量,当 $M < T$ 时,属性值从 x_i 到 y_i 的映射便实现了降维。

由于PCA需要很大的存储空间以及计算复杂度,因此很多研究者都提出了相应的改进算法。张岩等在文献[22]中对经典的PCA算法进行了改进,提出对图像进行分块,然后对每一子块进行独立的PCA处理,求出测试样本子块与训练样本对应子块间的距离,最后将这些距离相加得到测试样本与训练样本的距离。Thai等在文献[23]中比较了2DPCA和PCA,由于PCA需要将图片矩阵转换成向量后进行转换且其维数特别高,而2DPCA直接对图像矩阵进行转换。最后运用大量实验验证了2DPCA在表现力及速度方面都远远高于PCA。Ying Wen等在文献[24]中引入了KPCA到人脸检测中,其是对PCA的一种非线性的扩展,采用非线性的方法抽取主成分并映射到高维空间中进行PCA分析,能够最大限度地提取指标的信息。楼中望等在文献[25]又对KPCA在提取特征时的不足提出了加权的改进算法W2KPCA-KNN,提高了识别的精度。

2.3.2 局部二值模式(LBP)

LBP通过计算图像中所包含的每个像素与其局部邻域的点在亮度上的序关系,然后对二值序关系进行编码形成局部二值模式,最后采用多区域直方图作为图像的特征描述^[20]。在给定一个局部区域内的像素位置 $g_c(x_c, y_c)$,设以其为中心点的窗口内邻近的8个点为 g_0, g_1, \dots, g_7 。以 g_c 灰度值为阈值对窗口内其他8个像素点做二值化处理^[26]:

$$d_i = \begin{cases} 1, & g_i > g_c \\ 0, & g_i < g_c \end{cases}, i=0, 1, \dots, 7$$

从而得到的结果是一个8位二进制数。其十进制的数值为:

$$LBP(x_c, y_c) = \sum_{i=0}^7 d(g_i - g_c) 2^i$$

二值化处理如图2所示。

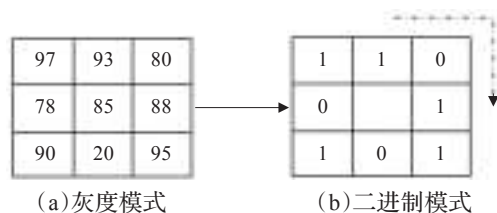


图2 像素点LBP处理

如图2所示,其二进制码为:11011010,对应的LBP值为: $10^7 + 10^6 + 10^4 + 10^3 + 10^1$,原始图像经过LBP处理后的效果如图3^[27]所示。

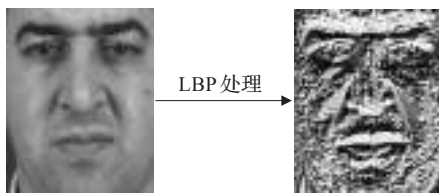


图3 LBP处理图像的效果

T.Ahonen 在文献[28]中进一步改进了LBP算法。他将一幅单独的图片划分为 R 份相同大小互不重叠的小块,分别进行LBP处理。转变后的LBP码,一共为256种。一幅图像中256个LBP码构成的直方图可看作纹理描述子。其计算公式为:

$$H_i = \sum_{x,y} I\{f_i(x,y)=i\}, i=0, 1, \dots, 255$$

其中 A 若为真,则 $I\{A\}=1$, 否则 $I\{A\}=0$ 。

$$I\{A\} = \begin{cases} 1, A \text{ 为真} \\ 0, \text{ 否则} \end{cases}$$

原始图像在改进的LBP处理时的划分效果如图4^[27]所示。

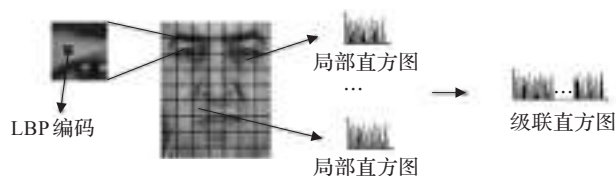


图4 改进LBP对图像的划分

由于各个区域包含的信息量不同,Ahonen 凭借经验主义,借用卡方检验为不同区域设定了不同的权重:

$$\chi^2(S, M) = \sum_{r,i} \frac{(S^r(i) - M^r(i))^2}{S^r(i) + M^r(i)}$$

其中 S, M 分别对应原始图像及模块图像的直方图。

由于LBP无法区分邻域像素点与中心像素点相等,并且高精度的特征提取会明显地增大内存占用量及影响运算速度,面对这些问题很多研究者提出了改进算法。Xu Jian 等在文献[29]中提出了DLBP(Double LBP),通过引入参数改善了像素点值轻微变化对LBP所产生的影响,能更好地表述图像的特征。Yuchun Fang 在文献[30]中提到了统一LBP(Uniform LBP)和旋转不变LBP

(Rotation Invariant Uniform LBP, Riu-LBP); Uniform LBP 将图像转变为二进制中01变换少于2次的归属为一类,在高维特征的支持下依赖采集密度来提高其精确性。LBP及Uniform LBP只能通过直方图解决平移问题,Riu-LBP解决了其弊端可以解决旋转问题。齐鸣鸣等在文献[31]中将多块MB-LBP(Multi-scale Block Local Binary Pattern)引入到人脸识别中,MB-LBP能表现图像块邻域内灰度级或者颜色的空间变化,反映图像的局部结构化特征,对图像局部纹理特征有很强的全局描绘能力。

2.3.3 其他提取算法

除了上述两种经典算法及改进外,许多研究者还纷纷提出了其他算法。根据提取对象的不同,特征提取主要分为基于运动和基于形变两类^[32]。前者用于动态图片,其表情特征提取有两类方法,一类是利用帧间信息来跟踪和定位特征,可称为基于帧内特征的方法,另一类不仅仅需要跟踪和定位,而且还要利用帧间信息来作为表情识别的特征,可称为基于序列特征的方法^[18]。后者由形状跟纹理的不同出发点又分为基于模型和基于图像两种。如表1所示。

2.3.4 部分算法实验分析

从表1中的各种特征提取算法中选取了光流法、ASM、Gabor算法,利用不同的表情数据库进行了实验分析,如下:

(1) 光流法

光流是用来测量人脸面部像素点运动的瞬时速度场,利用图像序列中像素在时间域上的变化以及相邻帧之间的相关性来找到上一帧跟当前帧之间存在的对应关系。在Cohn-Kanade人脸表情数据库中选取一个人的灰度图像运动序列,采用基于梯度约束的Lucas-Kanade光流法对关键区域嘴巴进行光流特征提取,如图5^[33]所示。



图5(a) 嘴巴区域的光流场



图5(b) 嘴巴特征点跟踪

如图5中所示的标记“+”的8个特征点,对应了32维的光流特征向量。基于梯度的光流估计方法通过微分的方式将运动场转化到了光流场,计算简单而又具有较好的效果。但在计算光流时涉及的参数选取、可靠性评价因子选择都具有一定的困难,将会严重影响结果的精度。

(2) ASM

主动外观模型ASM是一种基于统计的特征匹配方法。首先以目标物体的轮廓作为训练样本,提取其边界

表1 其他提取算法

	提取方法	参考文献	特点
动态图片 (运动)	光流法	[33-34]	在面部物理模型上利用运动图像序列中像素强度数据的时域变化以及相关性确定图像像素位置的运动情况
	特征点追踪法	[35-36]	首先在图像中提取Harris角点,针对特征点在图像平面运动运用卡尔曼滤波进行预测,最后利用预测结果进行区域内特征点匹配、外点剔除
模型	主动轮廓法(Snake)	[37-38]	引入作用方向可以自适应变化的外加强制力,使控制点能够不依赖于初始轮廓而快速地收敛到目标的真实轮廓
	ASM	[39-41]	基于点分布模型(PDM)用于提取目标轮廓的边界点集合,及改进后的基于点轮廓检测法(PCDM)的ASM
	AAM	[42-44]	在ASM基础上添加纹理信息,根据人脸图像特点提取特征
静态图片 (形变)	Gabor	[45]	将短时傅里叶变换作为窗函数,适合提取原始特征,但数据较大
	弹性图匹配法	[46-47]	对人脸图像中的部分关键特征点采取Gabor变换,以其关键点位置关系为边属性的属性图,从而将人脸识别问题转化为图匹配问题
	Fisher判别	[48]	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
	自动提取基准点	[49-50]	采用多检测器的方法,分别从人脸的正面跟侧面提取19和10个特征点,并为每个特征点赋予一个确定性因子,最后通过比较确定性因子来决定检测结果

点集合构造模型。在标准 IMM 人脸库中选取 n 幅图片,每幅图手工标记 k 个特征点以此来表示人脸表情显著的特征位置,单幅如图 6^[40]所示。



图6 ASM特征点集合模型

由图 6 中标记的特征点坐标转换成向量形式,即得到一个 $2k \times 1$ 的列向量。 n 幅人脸图像便形成一个 $2k \times n$ 的训练数据,迭代校准该训练数据形成一个平均轮廓,作为分类的先验模型。ASM 模型能够很好地覆盖人脸几何外形子空间,具有较高的点定位精度;但也存在不足,即其定位的准确性很大程度上取决于平均模型的初始化位置。

(3)Gabor

在 JAFFE 图像库中选取一幅图片,并选取 5 个尺度 8 个方向组成的 40 个不同的滤波器对图像进行滤波。滤波前后的对比图像如图 7 所示。

由图 7 可知,图像经过 Gabor 小波变换以后,从不同尺度、不同方向上提取了相关的特征,因此每幅图像变成多幅,每个像素点的特征用一组数据来描述,所以说可以获得丰富的特征信息,并且对于亮度和人脸姿态变化不敏感。然而,这种方法也存在不足之处,数据维数变成原图像矩阵的几十倍,给数据处理带来困难,很容

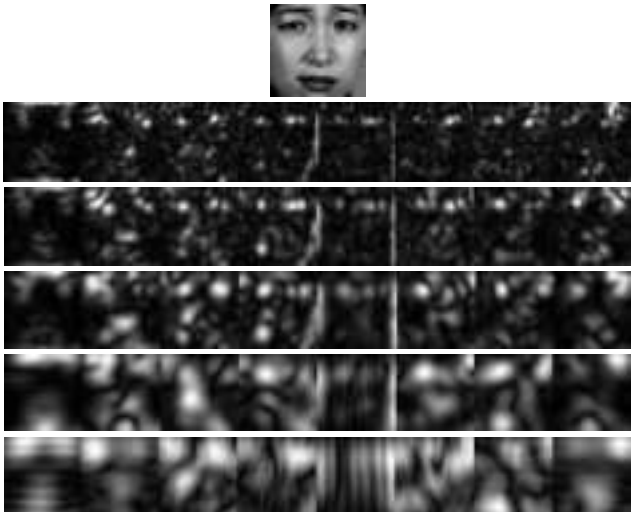


图7 Gabor滤波器组滤波后的图像

易造成维数灾难,需要对变换后的图像进行降维,这也增加了计算量。因此 Gabor 小波常常与基于形状的方法同时配合使用。

2.4 人脸表情分类

人脸表情分类包括表情识别与表情描述,表情识别是对面部信息的编码;表情描述分为情感映射与基本情感。其最终目的是判断提取的特征所对应的表情类别。BP 神经网络与 kNN 是分类中的两种经典算法,下面是其进行简要描述。

2.4.1 BP 神经网络

BP 神经网络作为一种非线性分类方法,实现了比线性方法更为复杂的分类性能。如图 8 所示。

作为仿生学的算法,神经网络充分模拟了神经元之间信息的接受、传递和处理。抽象出了以下三个关键点:节点本身的信息处理能力(数学模型),节点与节点

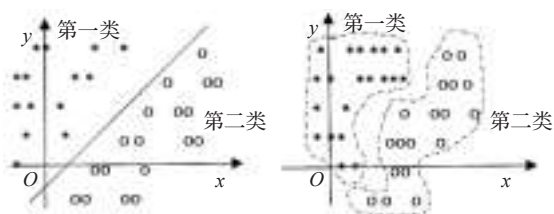


图8(a) 传统分类能力 图8(b) BP神经网络分类能力

之间连接(拓扑结构),相互连接的强度(通过学习来调整)。根据这三个关键点,神经网络的构成为:输入层、隐藏层、输出层、权值、阈值。基本模型^[51]如图9所示。

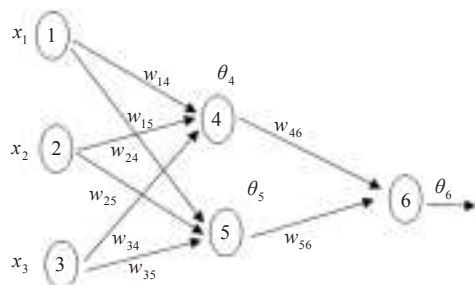


图9 神经网络基本模型

其中, $x_i (i=1, 2, 3)$ 是输入信号, $w_{jk} (j=1, 2, \dots, 5; k=4, 5, 6, \text{且} j < k)$ 为 j 节点到 k 节点的权值, $\theta_h (h=4, 5, 6)$ 是节点的阈值。

BP神经网络人脸表情分类主要步骤:

步骤1 网络初始化,给各连接权值分别赋一个区间 $(-1, 1)$ 内的随机数,设定误差函数 e , 给定计算精度值 ε 和最大学习次数 M 。

步骤2 随机选取第 k 个输入样本及对应期望输出。

步骤3 计算隐含层各神经元的输入和输出。

步骤4 利用网络期望输出和实际输出,计算误差函数对输出层的各神经元的偏导数 $\zeta_o(k)$ 。

步骤5 利用隐含层到输出层的连接权值、输出层的 $\zeta_o(k)$ 和隐含层的输出计算误差函数对隐含层各神经元的偏导数 $\zeta_h(k)$ 。

步骤6 利用输出层各神经元的 $\zeta_o(k)$ 和隐含层各神经元的输出来修正连接权值 $w_{ho}(k)$ 。

步骤7 利用隐含层各神经元的 $\zeta_h(k)$ 和输入层各神经元的输入修正连接权。

步骤8 计算全局误差。

步骤9 判断网络误差是否满足要求。当误差达到预设精度或学习次数大于设定的最大次数,则结束算法。否则,选取下一个学习样本及对应的期望输出,返回到步骤3,进入下一轮学习。

神经网络反馈学习过程中拓扑方式及误差的计算是研究的重点,曾涌现出许多改进算法,如:芬兰学者 Teuvo Kohonen 提出的自组织特征映射神经网络结构,学习向量量化 LVQ (Learning Vector Quantization) 神经网络;Muller 等在文献[52]中设计输出均方误差和这一

目标函数对权值和量子间隔进行统一训练,同时引入 Levenberg-Marquardt (LM) 算法降低目标函数陷入局部极小值的概率,实现了对量子神经网络的高效训练;Rumelhart、McClelland 提出的 BP 网络的误差反向后传 BP (Back Propagation) 学习算法;Moody、Darken 提出的 RBF 神经网络;Jeffrey 提出的 Elman 神经网络;Hopfield 提出的单层反馈神经网络——Hopfield 网络;Ackly、Hinton、Sejnowski 等以模拟退火思想为基础,对 Hopfield 模型引入随机机制,提出的 Boltzmann 机。

2.4.2 k-最近邻学习法

k-最近邻学习法 (kNN) 作为一种惰性学习法,在进行分类时并不像急切学习法一样,在接收到训练元组时只是简单地存储它(或只是稍加处理),仅有接收到测试元组时才进行泛化^[51]。以距离为度量,计算并得出离测试最近的 k 个元素,根据这 k 个元素所属类别来标记测试元组。

kNN 算法中的几个关键点:

(1) 一个所要估算对象所属类别的类别集,即用来标记训练集里每个对象所属的类。

(2) 用来计算对象之间距离的算法公式:

$$\textcircled{1} \|x - x\| = 0$$

$$\textcircled{2} \|x - y\| > 0$$

$$\textcircled{3} \|x - y\| + \|y - z\| > \|x - z\|$$

其中 x, y, z 分别表示三角形的三个点。距离的定义方法只要符合上面的三个条件即可自定义一种有效的距离公式。最常用的两个距离公式^[53-54]为:

① 欧式距离 (L2 测度)

$$d(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$$

② 明氏距离 (Lp 测度)

$$d(x, y) = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p}$$

③ 余弦距离 (夹角距离)

$$d(x, y) = -\cos(x, y) = -\frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$$

由于二元属性与如收入之类属性的巨大差异,在计算距离之前将每个属性的值规范化。

(3) 所选取进行比较的最近邻的个数 k 。

所比较个数 k 的选取在很大长度上决定了该算法的准确性。如图 10 所示:图 10(a) 中展示了若 k 值取值太小时,所鉴别的对象有可能会跟孤立点的属性相同。图 10(b) 中展示了适当的 k 值能够确定待测对象的类别。图 10(c) 中展示了若 k 值取值太大时,所鉴别对象的类别可能会发生巨大的误差。

表2 其他分类方式

	分类方法	参考文献	特点
时空域	隐马尔科夫模型 算法(HMM)	[58-59]	一个双重随机过程,其分为两部分:一个是Markov链,输出状态序列;另一个随机过程,输出观察值序列
	支持向量机方法 (SVM)	[60-61]	能够出色地解决小样本、非线性及高维模式识别问题,并应用到函数拟合问题中
	相关向量机 (RVM)	[62]	一种基于核函数的模式识别方法,是一个概率的稀疏核模型,采用贝叶斯方法,并引入权重的先验,而权重由超参数一对一支配,其值通过数据反复迭代进行估计
空域	AdaBoost算法	[63-64]	一种迭代算法,其核心思想是针对同一个训练集训练不同的分类器(弱分类器),然后把这些弱分类器集合起来,构成一个更强的最终分类器(强分类器)
	Fisher线性判别 分析(FDA)	[65]	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

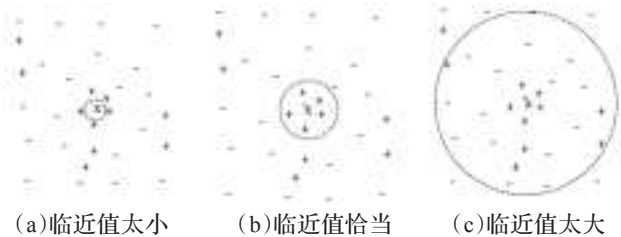


图10 k 不同取值的效果

由于kNN算法存在分类效率低以及同等对待各个特征的原因降低了分类的准确程度,对此随后有部分研究者对其进行了改进。Zhang Bin等在文献[55]中提出了超球体搜索方法,将搜索空间局限在一个超球体内,缩短了寻找 k 个近邻的时间,提高了算法效率。Gongde Guo等在文献[56]中提出了一种kNN模型的改进方法,其在训练数据集上构建模型簇,减少了对 k 值的依赖,从而提高了效率及准确性。Xi Bin Jia等在文献[57]中提到了FkNN(Fuzzy kNN)的思想,将模糊集的概念引用到kNN中,提升了KNN的效率及执行能力。

2.4.3 其他分类算法

除了上述的经典算法及改进外,中外学者还提出了许多其他算法。所提及到的算法根据分类器选择和设计是否利用时间条件分为时空域与空域两类。如表2所示。

2.4.4 部分算法实验分析

从表2中的各种分类算法中选取了HMM、SVM算法,利用不同的表情数据库进行了实验分析,得出其分类识别率,如下:

(1)HMM

通过提取的特征构成观察向量,人脸的表情特征可以抽象为HMM的状态序列,每种表情对应一个HMM,并使用概率矩阵描述状态的出现及其转移用。在整个过程中主要涉及到求值评估、识别解码、训练学习三个问题,可通过前向-后向、Viterbi、Baum-Weleh三个经典

算法依次解决。识别的主要在于分析和建立HMM,训练后的识别相对简单。选取CMU人脸表情数据库进行训练和测试,其相应的识别率如表3^[58]所示。

表3 HMM 分类识别率

表情	正确	错误	正确率/(%)
中性	100	0	100
高兴	93	7	93
惊讶	90	10	90
愤怒	87	13	87
恐惧	82	18	82
厌恶	80	20	80
悲伤	78	22	78

如表3可知,HMM算法在CMU中的表情识别达到了87.14%的较高平均识别率。正是由于隐Markov模型可以追踪表情图片中隐含的状态转移特征,能够良好地预测过程中的状态。但是在参数估计问题上,Baum-Weleh算法只能使模型参数收敛到局部最优值,很难实现全局最优值。

(2)SVM

支持向量机是一个二元分类器,通过非线性函数将样本映射到一个特征空间内,使两个类在特征空间中线性可分,并寻找样本在特征空间中的最优线性分类超平面,使得训练集中的样本点距离分类面尽可能远,即使分类间隔最大。在日本标准表情库JAFPE中进行训练与测试,其识别率如表4^[61]所示。

表4 SVM 分类识别率

表情	正确	错误	正确率/(%)
中性	10	0	100
高兴	8	2	80
惊讶	10	0	100
愤怒	10	0	100
恐惧	6	4	60
厌恶	10	0	100
悲伤	7	3	70

如表4可知,SVM算法在JAFPE中的表情识别达到87.1%的高平均识别率。正是由于SVM避开了从归纳到演绎的传统过程,高效地实现了从训练样本到预报样本的推理转化,降低了算法的复杂度并具有良好的鲁棒性。但是支持向量机只适合应用于小样本数据的分类情况,在大样本情况下,由于矩阵的计算需要耗费大量的内存资源,运行效率会直接受到影响。

3 人脸表情识别的进一步挑战

人脸表情识别相应的理论及技术已经取得了很大的进步,但是依然存在着亟须解决及优化的问题:

(1)由于每个人的差异性,人脸外貌、表情、肤色等都有可能不同,具有模式的可变性。

(2)各种算法往往需要将人脸上存在的人为表情作为前景提出之后才能发挥作用,实用性不是很大。

(3)即使手工提取出人脸作为前景,但是人脸上依然可能存在的眼睛、胡须等附属物的影响无法排除,鲁棒性差。

(4)存在于三维现实生活空间中的人脸往往会受到光照而产生灰度变化。

(5)表情不够精致,人类表情并非只局限于6种基本表情,以后应该加强人脸表情的视觉认知加工等认知心理学的基础研究。

(6)大多数的论文研究都是将已有的算法进行叠加组合来提高研究水平,缺少新算法的提出。

人脸表情识别难度较大,主要难在人脸都是塑性变形体,无论对于检测定位还是识别,如何将各种信息最大限度、有机地结合起来加以利用,是有效提高人脸表情识别系统效率的手段^[66]。

4 结束语

随着类似于Google眼镜等智能佩戴设备的推广,大数据时代已经到来,使得实时设备每时每刻需要处理的面部表情的数量急剧增加,所以为了使人脸表情识别的应用更具有前瞻性及利用价值,在算法的改进中不能只局限于现有的已知模式下的数据库,而应该转变数据思维^[67],在结构化、半结构化、非结构化的复杂数据源上以更具有实用价值的指导下设计新的算法或改进。

参考文献:

- [1] Mehrabian A, Russell J A. An approach to environmental psychology[M]. Cambridge, MA, US: The MIT Press, 1974.
- [2] Schlosberg H. Three dimensions of emotion[J]. Psychological Review, 1954, 61(2): 81-88.
- [3] Ekman P, Friesen W V. Facial action coding system: a technique for the measurement of facial movement[M]. Palo Alto: Consulting Psychologists Press, 1978: 271-302.

- [4] Keltner D, Ekman P. Facial expression of emotion[J]. American Psychologist, 1993, 48(4): 384-392.
- [5] Suwa M, Sugie N, Fujimora K. A preliminary note on pattern recognition of human emotional expression[C]//International Joint Conference on Pattern Recognition, 1978: 408-410.
- [6] Terzopoulos D, Metaxas D. Dynamic 3D models with local and global deformations: deformable superquadrics[C]//IEEE 1990 3rd International Conference on Computer Vision, Osaka, Japan, 1990: 606-615.
- [7] Mase K, Pentlad A. Recognition of facial expression from optical flow[J]. IEICE Transactions on Information and Systems, 1991, E74-D(10): 3474-3483.
- [8] Pantic M, Rothkrantz L. Automatic analysis of facial expressions: the state of the art[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(12): 1424-1445.
- [9] Fasel B, Luetttin J. Automatic facial expression analysis: a survey[J]. Pattern Recognition, 2003, 36(1): 259-275.
- [10] Yang M, Kriegman D J, Ahuja N. Detecting faces in images: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(1): 34-58.
- [11] 田巍, 庄镇泉. 基于HSV色彩空间的自适应肤色检测[J]. 计算机工程与应用, 2004, 40(14): 81-85.
- [12] 姜军, 张桂林. 一种基于知识的快速人脸检测方法[J]. 中国图象图形学报, 2002, 7(1): 8-12.
- [13] 王智文, 蔡启先, 陈劲枫, 等. 利用肤色分割和自适应模版匹配的人脸检测[J]. 广西工学院学报, 2013, 24(1): 1-8.
- [14] 余龙华, 王宏, 钟洪声. 基于隐马尔科夫模型的人脸识别[J]. 计算机技术与发展, 2012, 22(2): 25-28.
- [15] 白高峰. 一种基于平面拟合的图像恢复方法[J]. 计算机应用, 2004, 24(11): 126-127.
- [16] 王海涛, 刘俊, 王阳生. 自商图像[J]. 计算机工程, 2005, 31(18): 178-179.
- [17] Tan Xiaoyang, Triggs B. Enhanced local texture feature sets for face recognition under difficult lighting conditions[J]. Analysis and Modeling of Faces and Gestures, 2007, 4778(10): 168-182.
- [18] 刘晓旻, 谭华春, 章毓晋. 人脸表情识别研究的新进展[J]. 中国图象图形学报, 2006, 11(10): 1359-1368.
- [19] Calder J A. A principal component analysis of facial expressions[J]. Vision Research, 2001, 41(9): 1179-1208.
- [20] Ahonen T, Hadid A, Pietikainen M. Face recognition with local binary patterns[C]//Computer Vision-ECCV 2004, 2004, 3021: 469-481.
- [21] Duhuisson S, Devoine F, Masson M A. A solution for facial expression representation and recognition[J]. Signal Processing: Image Communication, 2002, 17(9): 657-673.
- [22] 张岩, 武玉强. 一种改进的模块PCA人脸识别新方法[J]. 计算机工程与应用, 2011, 47(26): 216-218.

- [23] Thai L H, Ha D T T. Facial expression representation and classification using 2DPCA[C]//IEEE 2012 International Conference on Control, Automation and Information Sciences (ICCAIS), 2012: 42-47.
- [24] Wen Ying, He Lianghua, Shi Pengfei. Face recognition using difference vector plus KPCA[J]. Digital Signal Processing, 2012, 22(1): 140-146.
- [25] 楼中望, 姚明海, 瞿心昱, 等. 基于 W2KPCA-KNN 算法的人体异常行为识别[J]. 计算机系统应用, 2011, 20(2): 157-160.
- [26] 刘伟锋, 李树娟, 王延江. 人脸表情的 LBP 特征分析[J]. 计算机工程与应用, 2011, 47(2): 149-152.
- [27] Marcel S, Rodriguez Y, Heusch G. On the recent use of local binary patterns for face authentication[J]. International Journal on Image and Video Processing-Special Issue on Facial Image Processing, 2007: 6-34.
- [28] Ahonen T, Hadid A, Pietikainen M. Face description with local binary patterns: application to face recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(12): 2037-2041.
- [29] Xu Jian, Ding Xiaoqing, Wang Shengjin, et al. Background subtraction based on a combination of texture, color and intensity[C]//ICSP 2008 9th International Conference on Signal Processing, Beijing, China, 2008: 1400-1405.
- [30] Fang Yuchun, Luo Jie, Lou Chengsheng. Fusion of multi-directional rotation invariant uniform LBP features for face recognition[C]//IITA 2009 3rd International Symposium on Intelligent Information Technology Application, Nanchang, China, 2009: 332-335.
- [31] 齐鸣鸣, 向阳. 基于 MB-LBP 和改进的 LFDA 的人脸识别[J]. 计算机科学, 2012, 39(6): 266-269.
- [32] 蒋斌, 贾克斌, 杨国胜, 等. 人脸表情识别的研究进展[J]. 计算机科学, 2011, 38(4): 25-31.
- [33] 孔健. 基于 MLBP-TOP 与光流多特征的人脸表情融合识别研究[D]. 江苏镇江: 江苏大学, 2009.
- [34] 金辉, 高文. 基于特征流的面部表情运动分析及应用[J]. 软件学报, 2003, 14(12): 2098-2105.
- [35] Bourel F. Robust facial expression recognition using a state-based model of spatially-localized facial dynamics[C]//IEEE 5th International Conference on Automatic Face and Gesture Recognition, Washington, DC, USA, 2002: 106-111.
- [36] 许竞, 姜波. 摄像机运动下特征点追踪方法研究[C]//2011 年振动与噪声测试峰会, 北京, 中国, 2011: 95-97.
- [37] Kass M, Witkin A, Terzopoulos D. Snakes: active contour models[C]//International Conference on Computer Vision, London, United Kingdom, 1987: 259-268.
- [38] 李熙莹, 倪国强. 一种自动提取目标的主动轮廓法[J]. 光子学报, 2002, 31(5): 606-610.
- [39] Cootes T F, Taylor C J. Active shape models—their training and application[J]. Computer Vision and Image Understanding, 1995, 61(1): 38-59.
- [40] 范小九, 彭强, Chen J X, 等. 一种改进的 AAM 人脸特征点快速定位方法[J]. 电子与信息学报, 2009, 31(6): 1354-1358.
- [41] Cootes T F, Edwards G J, Taylor C J. Active appearance models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(6): 681-685.
- [42] Chang J Y, Chen J L. Automated facial expression recognition system using neural networks[J]. Journal of the Chinese Institute of Engineer, 2000, 24(3): 345-356.
- [43] Edwards G J, Cootes T F, Taylor C J. Face recognition using active appearance models[C]//Computer Vision-ECCV'98, 1998, 1407: 581-695.
- [44] Edwards G J, Cootes T F, Taylor C J. Advances in active appearance models[C]//IEEE 1999 7th International Conference on Computer Vision, Kerkyra, 1999: 137-142.
- [45] Liu Chengjun, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition[J]. IEEE Transactions on Image Processing, 2002, 11(2): 467-476.
- [46] Wiskott L, Fellous J M, Kruger N, et al. Face recognition by elastic bunch graph matching[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 775-779.
- [47] 张文超, 山世光, 张洪明, 等. 基于局部 Gabor 变化直方图序列的人脸描述与识别[J]. 软件学报, 2006, 17(12): 2508-2517.
- [48] Zou Bin, Li Luoqing, Xu Zongbin, et al. Generalization performance of Fisher linear discriminant based on Markov sampling[J]. IEEE Transactions on Neural Networks and Learning Systems, 2013, 24(2): 288-300.
- [49] Pantic M, Rothkrantz L. Facial action recognition for facial expression analysis from static face images[J]. IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics, 2004, 34(3): 1449-1461.
- [50] Pantic M, Rothkrantz L. Expert system for automatic analysis of facial expression[J]. Image Vision Computing, 2000, 18(11): 881-905.
- [51] Han Jiawei, Kamber M. 数据挖掘概念与技术[M]. 范明, 孟小峰, 译. 第2版. 北京: 机械工业出版社, 2011.
- [52] Muller S, Wallhoff F, Hulsken F, et al. Facial expression recognition using pseudo 3-D hidden Markov models[C]//IEEE 2002 16th International Conference on Pattern Recognition, Quebec City, Canada, 2002: 32-35.
- [53] Thrk M, Pentland A. Eigen faces for recognition[J]. Journal of Cognitive Neuroscience, 1995, 3(1): 71-86.
- [54] 孙即祥. 现代模式识别[M]. 北京: 国防科技大学出版社, 2002.