

非正面人脸表情识别方法综述

蒋斌¹ 甘勇¹ 张焕龙² 张秋闻¹

(郑州轻工业学院计算机与通信工程学院 郑州 450002)¹

(郑州轻工业学院电气信息工程学院 郑州 450002)²

摘要 人脸表情识别是生物特征识别的重要组成部分,也是人机交互领域的一项关键技术。然而多数方法只关注正面或接近正面的人脸图像及视频,并限制了正常的头部运动,因此不利于人脸表情识别的智能化发展。有鉴于此,首先介绍了人脸检测、头部姿态估计、特征提取和分类的方法,以探索非正面人脸表情识别系统的发展;然后重点介绍了非正面人脸表情特征提取和分类的方法,并对基于人脸关键点的非正面人脸表情识别方法、基于外貌特征的非正面人脸表情识别方法及基于姿态相关的非正面人脸表情识别方法进行了比较分析;最后总结了非正面人脸表情识别方法的研究现状,并对进一步的研究和发展方向进行了展望。

关键词 非正面人脸表情识别,人脸关键点,外貌特征,姿态相关

中图分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.03.007

Survey on Non-frontal Facial Expression Recognition Methods

JIANG Bin¹ GAN Yong¹ ZHANG Huan-long² ZHANG Qiu-wen¹

(College of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China)¹

(College of Electric and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China)²

Abstract Facial expression recognition is an important part in biometric feature recognition, and it is also a key technology of human-machine interaction. However, most methods only focus on the frontal or nearly frontal facial images and videos, and restrict the normal head movements, so it is bad for intelligent development of facial expression recognition. To handle this problem, firstly, the face detection, head pose estimation, facial expression feature extraction and classification methods were introduced for exploring the development of non-frontal facial expression recognition system. Secondly, the non-frontal facial expression feature extraction and classification methods were emphatically introduced, and the comparison and analysis of the facial key points-based non-frontal facial expression recognition algorithm, the appearance feature-based non-frontal facial expression recognition algorithm and the pose-depend-based non-frontal facial expression recognition algorithm were carried out. Finally, the current research on the non-frontal facial expression recognition was summarized, and the future research and development direction were projected.

Keywords Non-frontal facial expression recognition, Facial key points, Appearance feature, Pose-depend

1 引言

人脸表情识别是一项对人类面部表情进行表示与分析,进而使计算机能够识别甚至理解人类情绪的生物特征识别技术^[1-5]。表情识别最初被心理学家用来研究危机状态下的国民心理^[6]。与计算机科学融合后,该技术逐渐在疲劳驾驶、微笑检测^[7]、网络视频分析^[8]和移动手机服务^[9]等人机交互领域开展了一系列的尝试,并将触角伸向了情感机器人、车载辅

助系统、智能移动终端和互联网即时通信等领域。

尽管研究者渴望将人脸表情识别技术尽快投入实际应用中,但是在有效识别人脸表情的探索道路上仍然面临诸多技术层面的限制。2015年, Sariyanidi 等人在 PAMI 上撰文指出人脸表情识别研究面临着五大难题,它们分别是头部偏转、光照变化、配准误差、面部遮挡以及身份差异^[10]。而在这些难题中,头部偏转就是造成配准误差和面部遮挡的一个重要原因。心理学研究表明:头部偏转能够干扰人类的情感认

到稿日期:2018-05-21 返修日期:2018-09-16 本文受国家自然科学基金项目(61702464,61503173,61771432,61302118,61502435),河南省高等学校重点科研项目(16A520028,17B510011,14A520034),河南省科技攻关项目(182102210156),河南省高校科技创新人才项目(17HASTIT022),河南省青年骨干教师计划(2016GGJS-087),郑州轻工业学院博士科研基金项目(2014BSJJ077)资助。

蒋斌(1983—),男,博士,讲师,主要研究方向为图像处理与模式识别, E-mail: jiangbin@zzuli.edu.cn(通信作者);甘勇(1965—),男,博士,教授,硕士生导师,主要研究方向为图像处理与模式识别、计算机网络及应用;张焕龙(1981—),男,博士,讲师,硕士生导师,主要研究方向为图像处理与模式识别、智能监控系统设计与开发、公共区域异常行为检测与分析、多源传感器信息融合技术;张秋闻(1982—),男,博士,副教授,硕士生导师,主要研究方向为图像处理与模式识别。

知^[11-12],也在情感表达中具有重要的作用^[13],因此由头部偏转引申出的非正面人脸表情识别就成为了人脸表情识别实用化发展过程中不可避免的重点与难点。

近年来,非正面人脸表情识别受到了国家自然科学基金委员会的关注。2017年5月于美国华盛顿举办的IEEE国际自动人脸与手势识别会议(IEEE FG 2017),就将头部偏转状态下的表情识别作为第三届人脸表情识别与分析挑战赛(The Third Facial Expression Recognition and Analysis Challenge)的主题。虽然该问题的研究热度在不断提升,但是非正面表情识别领域还缺乏专门针对该问题进行分析的综述性文章,给本领域的研究工作造成了不少困难。本文按照时间顺序,分析了PAMI期刊发表的人脸表情识别综述,如表1所列。

表1 综述文献回顾

Table 1 Overview of survey papers

参考文献	关注要点	识别类型
[3]	从人脸检测、特征提取和表情分类的角度展开	正面表情识别
[14]	从声音、视觉以及自发表情出发	正面多模态情感识别
[10]	从人类情感研究的5个关键问题入手	正面和非正面表情识别
[15]	从RGB图像、三维图像和热图像进行考虑	正面和非正面表情识别

研究发现,文献[3,14]针对的都是正面(或接近正面)的人脸图像和视频的表情识别问题,没有考虑非正面表情样本的识别。文献[10]重点讨论了表情识别领域的5个疑难问题,将非正面表情识别列为子课题,但是并没有对此进行系统的论述。文献[15]着眼于人脸表情样本的变化,将RGB图像、三维图像和热成像等样本作为划分识别方法的重要依据。文献[15]中提到了大角度头部偏转对识别造成的影响,但是很显然,这并非其关注的主要内容。

针对上述问题与挑战,本文着重从以下3个方面来进行研究。

1)现有综述多以正面人脸表情识别系统为例展开介绍,忽略了非正面表情识别系统的特殊性;涉及到人脸检测和头部姿态估计的部分又单独割裂开来,无法构成一个完整的体系。本文分析了非正面人脸检测和头部姿态估计在非正面人脸表情识别系统中的重要作用,以为头部偏转状态下的表情识别方法提供完整的设计思路。

2)现有综述只是单纯罗列出现的非正面人脸表情识别方法,缺乏对现有方法原理及特点的分析。本文按照现有非正面表情识别方法的原理和特点进行归纳、分类,阐述了现有方法中隐含的规律,有利于研究者迅速理清非正面表情识别领域出现的研究思路和技术路线。

3)现有综述指出头部偏转是人脸表情识别面临的难题,却没有指出究竟难在何处,也没有告诉我们非正面表情识别的发展趋势。本文分析了非正面表情识别的技术难题及其产生的原因,基于研究进展和业界动态,分析了非正面人脸表情识别在未来的主要发展趋势。

2 非正面图像的获取

非正面人脸表情识别的步骤如图1所示。首先,从标准

数据库获取人脸样本,或通过自主采集的方式拍摄志愿者的人脸图像及视频,并对图像进行光照均衡化和尺度归一化的预处理;然后,依靠算法检测的手段判断出人脸区域的位置及头部偏转情况;最后,提取人脸数据的表情特征,进而判断特征从属于哪种表情,完成对非正面人脸样本的表情识别。

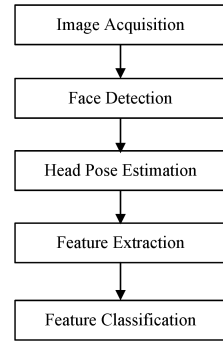


图1 非正面人脸表情识别的步骤

Fig. 1 Process of non-frontal facial expression recognition

对于图像获取而言,多数研究者倾向于依靠非正面人脸表情数据库来完成算法的验证工作。依据样本维度的不同,非正面人脸表情数据库可分为二维数据库(见表2)和三维数据库(见表3),这些数据库的标记分为表情和动作单元^[16]。

表2 二维非正面人脸表情数据库示例

Table 2 Examples of two-dimensional non-frontal facial expression databases

数据库	表情标记	头部偏转/(°)	样本类型
MMI ^[17]	动作单元	90	视频
BHU ^[18]	表情	-30	视频
ISL ^[19]	动作单元	±30	图像序列
CAS-PEAL ^[20]	表情	45	图像
RaFD ^[21]	表情	-90	图像
KDEF ^[22]	表情	±90	图像
AFFW ^[23]	表情	±45	视频
SFEW ^[23]	表情	±45	图像

表3 三维非正面人脸表情数据库示例

Table 3 Examples of three-dimensional non-frontal facial expression databases

数据库	表情标记	头部偏转/(°)	样本类型
BU-3DFE ^[24]	表情	±90	图像
BU-4DFE ^[25]	表情	±90	视频
Bosphorus ^[26]	动作单元	90	图像

从表2和表3可以看出,目前研究机构提供的非正面表情识别样本多以水平方向-90°到+90°角的头部偏转为主。其中,+45°和-45°的水平偏转是大家比较喜欢采用的拍摄角度,因为此时人脸面部的主要器官基本还在可视范围之内,降低了表情特征提取的难度;而极端的偏转角度,如+90°和-90°,使得面部器官被严重遮挡,呈现出“半脸”的形态,因此增加了识别难度。

国内的研究机构在二维非正面人脸表情数据库领域起步较早,如北京航空航天大学和中国科学院计算所分别构建了二维非正面动态表情数据库^[18]和二维非正面静态表情数据库^[20],但在三维非正面人脸表情数据库方面的研究还有待发展。除了表2和表3所列的数据库外,美国卡内基梅隆大学于2010年提供了一组收费的数据库The CMU Multi-PIE

Face Database,该数据库包含了不同姿态和表情的人脸数据。美国匹茨堡大学发布的数据库 Cohn-Kanade AU-Coded Expression Database 将在第三版增加头部偏移 30° 的表情视频。一些用于人脸检测、人脸验证和人脸识别的数据库也包括了非正面的人脸表情图像^[27-28]。其中的非正面人脸图片多以微笑表情为主,这与微笑识别领域的发展和具有笑脸识别功能的照相机、手机的普及有着密不可分的关系。此类数据库中家庭照片、新闻和电影图片占据了主要部分,在样本数量不足的情况下,可以考虑选择人脸数据库中具有明显表情和姿态角度的图片加以补充。也可以自行拍摄图片和视频,获取非正面表情的实验数据,而且这些图像更接近于真实的生活场景。但是,人脸数据库中的很多图片并没有具体的表情和姿态标记,这给非正面表情的识别研究带来了很大困扰。家庭和历史图片的背景以及过于复杂的人员情况,也给识别带来了不利因素。因此在实际应用中,人脸检测和头部姿态估计是非常重要的环节。

3 非正面人脸检测

在头部偏转角度不超过 45° 的情况下,经典的人脸检测方法(如 Viola-Jones 人脸检测器^[29])在处理非正面人脸检测问题时也可以实现部分非正面人脸图像的准确检测。然而,头部姿态的变换毕竟导致了面部器官位置的变化,使得面部器官分布的先验知识无效。在这种情况下,人脸检测器对非正面人脸图像的检测效果远低于对正面人脸图像的检测效果。而对于大于 45° 角的大角度头部偏转而言,面部器官的位置发生了剧烈变化,正面人脸检测方法更难确定人脸区域。因此,除了直接采用三维非正面人脸表情模型,或手工裁剪二维非正面人脸表情图像以外,必须要考虑针对非正面图像的人脸检测方法,以提升算法的自动化程度和实用性。

3.1 非正面人脸检测方法

文献[30]认为人在镜头前难免会改变头部姿态,获取多角度的人脸图像或许是解决该问题最有希望的方法。但是,多数实验并不具备获取测试对象的多角度人脸图像数据的条件,因此只能“舍近求远”,主要从如下3个方向来解决问题。

1)从局部特征出发,通过跟踪人脸局部区域或标记点的方式,实现非正面人脸检测。文献[31]从最优化问题入手,依靠多个局部分类器的反馈构建人脸检测的三维模型。文献[32]提出了一种采用“估计—校正—微调”步骤的人脸标记点检测方法,首先采用全卷积网络(Fully Convolutional Network)估算标记点的位置,其次通过点分布模型生成人脸标记点的起始形状,最后依靠加权的正则化标记点均值漂移(Regularized Landmark Mean-Shift)方法确定最终的人脸形状。文献[33]采用粒子滤波器方法,跟踪视频中侧面人脸的15个标记点。对每个标记点,采用100个粒子进行处理,根据后续帧的面部标记点位置的变化,测量出人脸表情的变化。但是该算法在2.5GHz的奔腾处理器上,针对每一帧图像的处理时间是15s,未达到实时跟踪的效果。综上可得,基于局部特征的非正面人脸检测方法可以在很大程度上排除背景因素的干扰,适用于非正面的人脸视频和图像序列;而且该方法从局部入手,弱化了姿态因素对检测的不利影响。但是,该

类方法多依靠人工的方式实现区域或标记点的初始定位,检测效果取决于跟踪方法的准确率。

2)从整体特征出发,对不同姿态的人脸图像进行训练,构建相应的人脸检测算子。文献[34]采用 Modified Census Transform 人脸检测器^[35],从不同的人脸偏转角度对人脸图像展开训练,并将检测到的人脸结果构成人脸跟踪器的低层部分,再依靠高层部分的相似度判决实现对人脸区域的跟踪。该方法对电视剧 *The Big Bang Theory* 第一季第一集中的人脸目标进行跟踪,取得了82%的准确率。因此,基于整体特征的非正面人脸检测方法通过不同角度的人脸检测器完成了对人脸的检测,从客观上回避了姿态因素的干扰,从而通过相似度判决获得了较高的检测准确率。然而,该类方法依靠人脸姿态的先验知识,一旦面临陌生样本,就很难做到有的放矢。

3)从深度特征出发,使目光并不停留在人脸图像的表层特征上。文献[36]提出了一种基于深度学习的动作单元检测方法,其核心是依靠不同姿态人脸图像调试(Fine-Tuning)得到的 VGG-Face 网络训练相应的专家系统,进而提取更多的判别特征用于最终的检测任务。文献[37]提出了基于区域化和多任务学习的卷积神经网络,实现多姿态条件下动作单元的检测和强度估计。文献[38]则深入到人脸的颜色空间中,依靠面部肤色与背景颜色的区别提取人脸颜色信息,进而判断出非正面人脸所在区域。因此,基于深度特征的非正面人脸检测方法利用了面部图像的深层信息,提高了信息的有效利用率,客观上有助于提高检测准确率。

目前,非正面人脸检测算法还存在着准确率和运算速度方面的局限性,尤其是在大角度人脸偏移的情况下,检测效果仍有待加强;而且姿态估计与目标检测存在很强的互补性^[39]。我们在调研中发现,头部姿态估计研究中也包含了一些非正面人脸检测方法,二者之间存在紧密联系。

3.2 头部姿态估计方法

非正面人脸检测方法的目的是确定图像中的人脸位置。在很多情况下,头部姿态估计也是检测过程中的重要环节。如果将二维人脸图像拓展到三维数据状态,那么头部姿态会更加复杂。除了水平转动(Yaw)之外,还需要考虑头部的垂直转动(Pitch)和左右转动(Roll)情况。因此,如何准确地判断头部的具体姿态,是研究者们面临的一个难题。

针对上述情况,文献[40]对头部姿态估计方法进行了梳理,将其分为外观模板法、检测阵列法、非线性回归法、流形嵌入法、弹性模型法、几何法、跟踪法,以及采用上述方案的混合法。其中,外观模板法、检测阵列法和弹性模型法强调对某一姿态人脸样本的建模;流形嵌入法和几何法侧重于人脸特征的提取。文献[41]通过使用一维 Gabor 特征,可以使得特征与姿态更加紧密相关。通过使用基于核函数的局部费舍尔判别分析,消除了姿态估计存在的非线性和多模态问题;非线性回归法侧重于建立姿态之间的映射关系;跟踪法则与目标跟踪领域的研究有交集,重点在于对视频中对象头部的跟踪。文献[42]提出了一种基于联合概率的目标跟踪方法,首先通过图结构方法定位目标区域,然后搭建目标跟踪的概率模型,采用粒子滤波跟踪单帧图像的目标,最后实现对行人肢体和头部运动方向的估计。

上述方法多依靠单一设备完成对头部姿态的跟踪与判断,与此相对,文献[43]则采用3个摄像头拍摄不同角度的头部图像,通过多视图粒子滤波跟踪器估计头部的位置,在多任务学习的框架下构建模型,实现对头部姿态的跟踪。文献[44]利用人脸影像变形(Face Morphing)技术,将结合深度参数的三维人脸模型变换到二维人脸图像上,并借助三维人脸模型的标定点确定二维图像的头部姿态。文献[45]采用深度学习中的卷积神经网络模型估计的头部姿态,在低分辨率、多模态的RGB-D数据上取得了较好的效果。从实验效果来看,该算法实现了实时处理。

至此,非正面表情识别的人脸检测工作已基本完成。在掌握了样本的人脸区域和头部姿态后,非正面人脸表情识别进入到最后的环节,即表情特征的提取和分类。

4 非正面表情识别

头部偏转动作的发生,给人脸表情图像带来了3方面的改变:1)头部姿态的变化使得部分面部器官被遮挡,造成了信息的缺失;2)头部姿态的变化导致人脸发生了不同程度的“形变”,导致人脸模型复杂化;3)头部姿态的变化打破了光照的均衡分布,改变了图像的灰度值大小,模糊了人脸图像的表情属性。上述改变造成了两个结果:1)表情特征的埋没;2)训练样本和测试样本不满足数据同分布的要求。针对这些情况,研究者分别从非正面表情特征提取和特征分类方面提出了3种主要的应对方案。

4.1 基于人脸关键点的方法

该类方法的主要思想是依靠二维几何模型或三维几何模型,跟踪人脸表情图像的关键点。例如,文献[46]列举了人脸侧面轮廓可选择的特征点,具体分布在额、眉、鼻、唇等与表情相关的人脸器官周围。在提取关键点特征之后,采用回归或分类方法实现对非正面人脸表情或动作单元的识别。

文献[47]提出了一种典型的基于人脸关键点的非正面表情识别方法。该方法手工标记了眉、眼、口周围的38个关键点,依靠几何特征和多分类器方法,对三维人脸表情数据库BU-3DFE中包含的相对于正面人脸图像偏转 0° 、 30° 、 45° 、 60° 和 90° 的非正面人脸图像的表情进行了识别。首先,计算相同人脸偏转角度下,表情图像和无表情图像之间关键点的位移量;然后,将位移量进行0均值和单位方差的标准化处理,处理结果将作为表情特征;最后,分别采用贝叶斯分类器、Parzen分类器、支持向量机(Support Vector Machine, SVM)和k-近邻分类器对表情特征展开识别。实验表明,各种识别算法对偏转 45° 的人脸表情图像具有最低的平均错误率;而对偏转 90° 的极端情况具有最高的平均错误率。其中,算法对正面图像(偏转 0°)的识别率反而低于偏转 45° 的人脸表情图像的识别率,文献[47]认为其原因在于非正面表情图像在某种方式上提供了人脸的深度信息,这一信息有助于改善识别效果。文献[33]采用粒子滤波器跟踪了人脸侧面序列的15个关键点,通过对比视频序列中有表情的第 t 帧和无表情的第1帧之间关键点的位置变化,判断出关键点的运动方向,并计算出二者之间的几何距离;然后结合表情发生的过程(表情开始—表情最大—表情结束)探索出不同表情的运动状态。该方法

对极端头部姿态条件下的人脸侧面样本展开识别时取得了87%的动作单元识别率。这项工作打破了传统研究只针对正面或接近正面的人脸样本展开识别的限制,与文献[47]相比,文献[33]着眼于动作单元,即局部表情和动作的识别。但是该方法限制了头部运动,因为一旦头部发生整体的刚性运动,对局部关键点的判断就会有失偏颇。文献[48]采取了基于视图的2D+3D主动外观模型(Active Appearance Model, AAM)的方法,通过模型拟合人脸图像的关键点位置,进而实现对头部姿态的估计;然后提出了一种双层广义判别分析方法用于人脸表情的识别,其对非正面人脸表情的识别率可达91.25%。该方法给出的示例图显示,实验在人脸向左/向右水平偏转 45° 的区域展开,且针对正面表情图像的识别率明显高于非正面表情图像的识别率。文献[49]与文献[47]都采用了多种方法来识别非正面人脸表情,区别在于后者在分类器的选择上进行了尝试,而前者在特征提取上展开了研究。首先,采用Pose-Dependent AAM获取不同姿态下的人脸关键点;其次,使用尺度不变特征转换(Scale-Invariant Feature Transform, SIFT)、局部二值模式(Local Binary Pattern, LBP)或离散余弦变换(Discrete Cosine Transform, DCT)提取关键点邻域的局部特征;最后,通过F-Score选择方法实现特征选择后,依靠SVM实现多视图表情的识别。其中,最高表情识别率出现在头部发生左右偏转 30° 的情况下。该算法的识别率也随着表情强度的增加而提高。文献[50]在自发人脸表情识别的研究中也涉及到了非正面人脸表情图像序列中的动作单元识别问题。其利用基于动态贝叶斯网络的联合概率人脸动作模型对面部运动进行建模,以此实现对头部运动状态下动作单元的识别。该文献在训练人脸形状模型时,针对左右偏转 30° 的人脸表情图像,采用了28个人工标注的关键点,在现有设备上取得了2.1帧/s的图像处理速度。文献[51]采用了BU-3DFE数据库提供的83个面部关键点,用稀疏SIFT方法实现对非正面人脸表情图像的特征提取,然后在联合贝叶斯理论框架下分别对 0° 、 30° 、 45° 、 60° 、 90° 的非正面表情图像进行识别。识别结果表明,在 30° 至 60° 之间的范围内,算法可以在BU-3DFE数据库上取得比正面人脸表情图像更高的识别率。

上述方法都需要大量已标记关键点的样本参与训练,但是非正面人脸关键点的跟踪与检测却是一个难题。文献[52]使用深度学习中的受限玻尔兹曼机(Restricted Boltzmann Machines, RBM)构建正面和侧面表情图像的模型,采用3-way RBM捕获从正面图像过渡到非正面图像的人脸特征点。该方法的局限性在于:1)采用的侧面人脸图像偏转幅度不大;2)算法需要标定初始点位置。文献[53]提出了一种级联的姿态估计方法,用于估计头部偏转、人脸部件和人脸关键点。其依靠梯度增强随机蕨(Gradient Boosted Ferns, GBF)获得了比随机蕨(Fern)方法更好的效果。该算法在英特尔奔腾3.2GHz CPU上,采用C++编程,处理一幅测试图像仅需0.55ms。文献[54]提出了一种基于数据驱动的人脸关键点定位方法。该模型建立了关键点和周围外貌特征之间的联系,每次运行过程中,每一个特征都会采用加权投票的方式来预测关键点的位置,其中权重的大小会依据特征的判别能力

进行预先计算。该算法在具有较大头部偏转的表情图像上取得了较高的定位准确率,但仍需大量已标记关键点的样本来完成模型的初始化。该文献也测试了算法在设备上的运行时间,即在英特尔 Xeon E5-2670 工作站上,采用 Matlab 编程,对一幅大小为 900×600 的人脸图像进行人脸检测需要 42.8 s,完成关键点定位还需要 25.5 s。不过该文献也指出,与其他类似方法相比,基于数据驱动的人脸关键点定位方法已经减少了运行时间。文献[55]在低层模型中使用混合模型提取人脸部件的形变。因为不同人脸部件的状态高度相关,同时表情和姿态又是引发人脸形变的主要原因,所以在高层模型中通过自动结构学习和概率模型的参数估计,就可以学习人脸部件、表情和姿态之间的联系,实现对多姿态人脸表情关键点的跟踪。文献[56]在粒子滤波器的框架下,使用可变强度模板(Variable-Intensity Templates)同时估计一组表情视频的头部姿态与面部表情。在相对于正面人脸水平偏转 $\pm 40^\circ$ 的方向上,该算法取得了 90.00% 以上的表情识别率。该方法在自采集的视频上取得了成功,但是却无法在偏转幅度较大及光照条件变化的情况下保持较高的识别率。

总之,基于人脸关键点的方法对非正面人脸关键点的跟踪与检测问题进行了多方面的探索,但非正面人脸关键点自动标记的难度较大。针对这一问题,研究者们又提出了基于外貌特征的非正面表情识别方法。

4.2 基于外貌特征的方法

该类方法的主要思想是获取不同姿态人脸图像的全局或局部表情特征,降低图像中与表情无关的因素的不利影响。

文献[57]是典型的依靠整体外貌特征实现非正面人脸表情识别的方法。该文献在 BU-3DFE 数据库上比较了多种特征提取方法和分类器,结果表明:1) SIFT 方法的效果优于方向梯度直方图(Histogram of Oriented Gradient, HoG)方法和 LBP 方法的效果;2) 流形学习中的局部保留投影(Local Preserving Projection, LPP)方法的效果优于主分量分析(Principal Component Analysis, PCA)方法和线性分量分析(Linear Discriminant Analysis, LDA)方法的效果;3) 分类器融合方法的效果并不比单独使用 SIFT 方法的效果明显。上述结论启发了研究者们对非正面人脸表情识别的探索^[58-59],文献[49]也采用了通过多种方法来获取非正面人脸表情特征,进而对比分析的研究思路。所不同的是,相对于基于关键点的方法, LPP, PCA 和 LDA 3 种方法更适用于对高维人脸表情图像矩阵进行处理。因此,研究者们更倾向于将这些典型的子空间或流形学习方法应用在基于外貌特征的非正面人脸表情识别领域。文献[60]提出了一种判别子空间的方法,称之为判别共享高斯过程潜变量模型(Discriminative Shared Gaussian Process Latent Variable Model)。该方法首先学习同一表情的多视图判别流形,进而在表情流形中对人脸表情展开识别。但是,该方法的实验对象局限于偏转 30° 范围内的表情图像。

文献[61]是典型的依靠局部外貌特征实现非正面人脸表情识别的方法。该文献分别采用了 LBP、局部 Gabor 二值模式(Local Gabor Binary Patterns, LGBP)以及多尺度 LBP 等方法获得非正面人脸图像的局部表情特征。以 15° 为间隔,分别针对 0° 至 90° 的人脸偏转头像进行识别。最佳识别效果出

现在接近正面的 15° ,不过 90° 偏转的表情识别率并不低于小角度偏转情况下的识别率。文献[62]通过区域协方差矩阵实现了对多姿态表情图像的局部特征表示,然后在高斯混合模型(Gaussian Mixture Model, GMM)的基础上提出了一种新的贝叶斯判别分析(Bayes Discriminant Analysis, BDA)算法,即 BDA/GMM 方法。实验结果表明,该方法在 BU-3DFE 数据库上取得了 31.77% 的平均错误率。文献[63]提出了基于局部约束线性编码(Locality-constrained Linear Coding)方法的双层模型(Bi-layer Model),依靠局部外貌特征,针对 Multi-PIE 数据库和 BU-3DFE 数据库中水平偏转 0° 至 90° 的非正面表情图像,分别取得了 86.30% 和 74.58% 的平均识别率。

除上述方法外,基于外貌特征的方法还可以不单纯依靠全局或局部外貌特征来识别非正面表情。具体方式如下:

1) 特征融合。文献[64]提出了一种基于类特征的参数核特征空间方法(Parametric Kernel Eigenspace Method Based on Class Features, pKEMC)。该方法分为两步:①利用核方法等特征空间技术获取非线性判别的映射子空间;②通过映射子空间,构成包含不同姿态的各类表情样本的流形,进而依靠参数特征空间技术实现表情的分类。基于 Multi-PIE 数据库的实验结果表明,该方法在人脸局部区域取得的表情识别率高于在人脸全局范围内的识别率。文献[65]采用局部外貌特征和全局几何特征相结合的方法构建多姿态人脸表示,进而依靠分层贝叶斯模型判断人脸表情。文献[66]提出了基于随机森林的非正面表情识别算法。该方法在随机森林的框架下整合了人脸表情视频的静态(空间)特征和动态(时空)特征,提高了随机森林方法面对不同头部姿态表情样本时的识别率。

2) 空间转换。文献[67]的思路是将每种姿态的三维表情图像映射到二维空间,再依据映射后数据的缺失程度来区分不同姿态之间的差异。

3) 深度学习。文献[68]提出了加权混合深度神经网络方法。该方法将人脸灰度图像和相应的 LBP 图像作为特征提取对象,前者依靠 VGG16 模型,后者采用卷积神经网络,分别提取人脸图像的表情特征。两种特征加权融合后作为最终的表情特征,被 softmax 分类器识别。文献[69]提出一种加权中心回归自适应特征映射(Weighted Center Regression Adaptive Feature Mapping)方法,对存在个体和姿态变化的 CK+, RaFD 和 ADFES 数据库中的人脸图像进行了表情识别测试,并与具有深度学习结构的不同算法进行了比较。文献[70]用训练好的多通道卷积神经网络对人脸图像、眼部图像和口部图像分别进行特征提取;然后将特征输入到联合多尺度融合网络,获得表情特征,并判断头部偏转角度;最后根据头部偏转角度将特征输入到对应的 Pose-aware Expression Network 中,得到对 BU-3DFE, KDFE 和 SFEW 数据库的表情识别结果。该方法对单一数据库的识别效果较好,但是当训练集和测试集分别采用不同的数据库时,表情识别率会大幅度下降。从运行时间来看,该方法在两个 GTX 1080 GPU 上对每幅图像的训练时间与测试时间分别是 24 ms 和 11 ms。

总之,基于外貌特征的方法回避了关键点标注和跟踪的工作,提高了算法的自动化程度。但是,算法的识别对象多集中在水平偏转 45° 范围内的人脸表情图像,面对大角度偏转的

人脸表情图像时,算法的识别率还有待加强。

4.3 基于姿态相关的方法

该类方法的主要思想是建立不同姿态样本之间的联系,借助一种姿态的表情信息识别出另一种姿态的表情类别。首先,提取多姿态图像的表情特征;然后,利用这些特征构建分类或回归模型;最后,建立起不同姿态样本之间的联系。该类方法依据姿态关系的不同,又可细分为两类。

1) 多数人认为人脸的偏转造成了部分表情区域的遮挡和人脸形变,暗示了正面人脸图像蕴含的表情信息比非正面人脸图像多。因此,研究者希望建立正面人脸表情图像与非正面人脸表情图像之间的联系,依靠正面表情图像相对丰富的表情特征来提高算法对非正面表情图像的识别率。文献[71]提出了多种线性回归方法来构建非正面人脸图像与正面人脸图像之间的联系,目的是依靠正面人脸图像“弥补”非正面人脸图像缺失的信息。首先依靠 LBP 和 HoG 提取表情特征,然后通过不同的映射方法,如 PSR (Pose Specific Regression),构建不同姿态表情特征之间的关联,最后利用 SVM 识别样本表情。该文献还给出了在 BU-3DFE 和 Multi-PIE 数据库上的表情识别率和算法运行时间,实验结果显示,针对 BU-3DFE 数据库的 4200 个测试样本,算法中运行时间最短的是 49 s,最长达到 85 s;而对于 Multi-PIE 数据库的 1885 个测试样本而言,算法中运行时间最短的是 162 s,最长达到 282 s。文献[72]在文献[66]的基础上又提出了非线性映射的方法,其表情识别率提高了近 2 个百分点。文献[73]分别采用 4 种映射函数,即线性回归、支持向量回归、相关向量回归和高斯过程回归,来实现非正面人脸模型到正面人脸模型的映射,并比较它们在偏转 15°,30°和 45°时对非正面人脸表情图像的识别效果。文献[74]首先采用在线 AAM 标记 39 个人脸标记点,利用线性判别分析将向量映射到低维流形空间;然后使用 GMM 训练流形数据,估计样本处于某种姿态的可能性;再利用耦合尺度高斯过程模型(Coupled Scaled Gaussian Process Regression)来构建非正面人脸图像与正面人脸图像之间标记点的映射关系;最后通过 SVM 实现表情识别。文献[75]借鉴了人脸识别中的人脸摆正(Face Frontalization)方法,将非正面人脸图像转换成正面人脸图像,同时采用生成

式对抗网络(Generative Adversarial Networks, GAN)保留图像的表情特征。该方法针对 Multi-PIE 数据库和 BU-3DFE 数据库中水平偏转 0°至 90°的非正面表情图像,分别取得了 86.76%和 73.13%的平均识别率。

2) 另一些研究者认为人脸表情识别率与所采用的实验数据有关,对正面图像的表情识别率不一定高于对非正面图像的表情识别率。基于上述考虑,研究者们并不局限于建立正面人脸图像与非正面人脸图像之间的联系。文献[58]首先分别采用 LBP 和 SIFT 提取表情特征,依靠 LDA 完成特征降维;然后使用一种非正面姿态的特征估计出另一种非正面姿态的特征,并将这一过程转化为最小均方回归问题(其中姿态之间的变化矩阵就是最小均方回归问题中的回归系数矩阵);最后,通过群稀疏降阶回归(Group Sparse Reduced-Rank Regression, GSRRR)模型实现了对非正面样本的最优特征选择。该算法在 Multi-PIE 数据库和 BU-3DFE 数据库上都取得了比基线算法更高的识别率。文献[76]依靠 Gabor 小波提取人脸表情图像的局部特征,采用偏最小二乘回归法(Partial Least Squares, PLS)将人脸表情样本按照不同姿态两两组合。其在 BU-3DFE 数据库上对 6 种表情取得了 86.60%的平均识别率。文献[77]借鉴了说话人识别中的高斯混合模型-通用背景模型(Gaussian Mixture Model-Universal Background Model, GMM-UBM)的思想,提出了有监督超向量编码(Supervised Super-vector Encoding, SSE)方法。首先构建数据集的 GMM(被称为 UBM),再以 UBM 为参照物,建立多姿态人脸表情图像的 GMM,最后实现对不同姿态人脸表情的识别。

总之,基于姿态相关的方法主要依靠线性回归和建模的方式构建不同姿态人脸图像之间的联系,进而识别出表情。该类方法并不局限于小角度偏转条件下的表情识别,但是其识别率仍有待提高。

综上所述,表 4 对 3 类方法的性能进行了比较,表 5 列举了其中部分算法的运行时间。需要说明的是:1)表 4 和表 5 中的横线“—”表示没有在文献中查到相关内容;2)为保证计算单位一致,表 5 中 PSR 的运行时间是按照文献提供的数据计算得到的结果。

表 4 典型识别方法的性能对比

Table 4 Performance comparison of typical recognition algorithms

识别方法	表情特征	数据库	偏转角度/(°)	训练-测试方法	表情数量	识别率/%
SVM ^[46]	人脸关键点的几何形变特征	BU-3DFE	0~90	5 组交叉验证	6	66.50
DCT ^[49]	人脸关键点的几何形变特征	BU-3DFE	-90~90	—	6	75.10
贝叶斯方法 ^[51]	人脸关键点的几何形变特征	BU-3DFE	0~90	Holdout 验证	6	78.35
可变强度模板 ^[56]	人脸关键点的几何形变特征	自制数据库	-40~40	—	5	90.10
SIFT+LPP ^[57]	整体外貌特征	BU-3DFE	0~90	2 组交叉验证	6	73.06
LGBP ^[61]	局部外貌特征	BU-3DFE	0~90	10 组交叉验证	6	67.96
	局部外貌特征	Multi-PIE	0~90	10 组交叉验证	6	80.60
BDA+GMM ^[62]	局部外貌特征	BU-3DFE	-45~45	5 组交叉验证	6	68.23
pKEMC ^[64]	整体/局部外貌特征	Multi-PIE	0~45	10 组交叉验证	4	77.25
	正面/非正面表情的相关特征	Multi-PIE	-90~90	5 组交叉验证	6	81.96
PSR ^[71]	正面/非正面表情的相关特征	BU-3DFE	0~90	5 组交叉验证	6	77.87
	多姿态表情的相关特征	Multi-PIE	0~90	Holdout 验证	6	81.70
GSRRR ^[58]	多姿态表情的相关特征	BU-3DFE	0~90	Holdout 验证	6	78.90
	多姿态表情的相关特征	BU-3DFE	-90~90	—	6	86.60
PLS ^[76]	多姿态表情的相关特征	Multi-PIE	—	Holdout 验证	4	81.80
SSE ^[77]	多姿态表情的相关特征	BU-3DFE	-45~45	5 组交叉验证	6	76.60
	多姿态表情的相关特征	BU-3DFE	-45~45	5 组交叉验证	6	76.60

表5 部分识别方法的运行时间

Table 5 Running time of some recognition algorithms

识别方法	使用设备	编程语言	样本大小	训练/测试样本数量	每幅图像运行时间
GBF ^[53]	英特尔奔腾 3.2GHz CPU	C++	150×150	2000/1800	0.55 ms
数据驱动法 ^[54]	英特尔 Xeon E5-2670 工作站	Matlab	900×600	—	25.5 s
多通道卷积神经网络 ^[70]	两个 GTX 1080 GPU	—	224×224	16 515/—	24 ms
PSR ^[71]	英特尔 2.3GHz 赫兹双核	—	—	—/4200	20 ms

通过分析现有研究成果,我们发现:1)多数算法的识别率尚不及90%,其中在BU-3DFE数据库和Muti-PIE数据库上,基于姿态相关的方法取得了最高的识别率,基于人脸关键点的方法和基于外貌特征的方法在识别率上不相伯仲;2)对基于人脸关键点的方法而言,人脸关键点的数量增加时,识别率提高;3)对基于外貌特征的方法而言,整体与局部外貌特征相结合的方法一般能取得比单一特征更高的识别率;4)基于姿态相关的方法主要是将不同面部姿态之间的表情图像识别问题转化为回归问题,依靠线性回归与多特征融合的方式实现对非正面人脸表情的识别,相较于前两种方法,更强调“协作”;5)从上文对各个算法的描述中不难看出非正面人脸图像的表情识别率并不一定低于正面人脸图像的表情识别率,这是侧面人脸图像具有一定深度信息的缘故造成的,而且大角度偏转的表情识别率一般低于小角度偏转的表情识别率;6)如表5所列,实验环境的不同和细节的缺失使得我们很难从运行时间上对不同算法进行精确的比较,但就算法的自动化程度而言,手工标注人脸关键点的方式仍需改进。算法的实时性将是研究者们需要解决的一个难题。

4.4 挑战与展望

非正面表情识别主要研究头部偏移状态下的人类面部表情的类别判断,从而为人机交互等研究提供技术支持。完整的非正面表情识别系统可分为非正面人脸检测、头部姿态估计和非正面表情特征分类3部分。为了解决头部偏转造成的表情信息缺损的问题,研究者提出了3种方案。1)依靠二维或三维几何模型跟踪人脸表情图像的关键点,通过关键点获取面部表情特征,从局部上回避头部偏转问题的干扰;但是该方法在初始化阶段多需要人工标记关键点,自动化程度偏低。2)获取不同姿态人脸图像的全局或局部表情特征,以外貌信息替代关键点的信息,避免了人工标记的压力;但是在头部偏转角度过大时,该算法的表情识别率仍有待提高。3)建立不同姿态样本之间的联系,借助一种姿态的表情信息识别出另一种姿态的表情类别。该方法借鉴了以上两类方法的手段;但是其识别率仍较低。

通过分析现有研究成果,我们认为非正面表情识别还有以下问题值得研究。

1)非正面表情特征的提取

在头部偏转条件下,同一幅人脸表情图像就具备身份、表情和姿态等多种不同的描述特征;而且同以人脸为载体,这些特征你中有我,我中有你。因此,如何有效地对非正面人脸表情图像进行数据组织,在图像元素的河沙中淘出表情特征这

一金矿,是非正面表情识别研究所必须考虑的问题。这需要借鉴最新的机器学习成果,探寻数据中蕴含的表情特征,迁移学习便是一种有益的尝试。文献[78]研究了在头部姿态变化时,表情特征在视频中的迁移过程。文献[79]依靠迁移学习模型,测量了9种姿态下7个动作单元的强度。文献[80]则将迁移学习直接用于识别领域,提出了一种转导迁移正则化最小二乘回归模型,首先构建不同姿态人脸图像的判别子空间,然后再通过SVM实现对样本的表情分类。文献[81]在表情识别算法竞赛(Emotion Recognition in the Wild, EmotiW)中依靠迁移学习和深度卷积神经网络对非正面表情视频展开识别。

2)大数据条件下的非正面表情识别

头部偏转造成了人脸外观的扭曲与形变,即使是同一个人的同一种表情,在不同的姿态下,其数据分布也会不同。这将极大地降低算法的推广性和识别率,使得姿态成为横亘在表情样本之间的鸿沟。在有限的表情个体之间建立起跨越姿态的“表情桥梁”,是小样本条件下非正面表情识别很难达到的要求。然而百度大脑计划的阶段性成果已经展示了“互联网+”时代大数据处理的优势,因此搭乘大数据与物联网发展的列车,通过海量样本构建非正面表情模型,克服姿态、个体之间的数据分布差异对表情识别的干扰,提升表情识别算法的性能,将是推进该领域商业化发展的一条可行之路。

3)非正面的微表情或自发表情识别

微表情或自发表情识别研究在美剧*Lie To Me*的影响下成为了表情识别在大众心目中一个重要的代名词。该领域的发展也从心理学的理论领域向着更加实用化的方向迈进。文献[82]针对特定个体构建表情模型,对小角度头部运动条件下的自发表情具有一定的跟踪和表示效果,并计划在表情识别系统上继续探索。总而言之,非正面状态下的相关表情识别研究正结合各方面的成果,向着更加实用化的目标迈进。

结束语 非正面表情识别是表情识别实用化发展的关键步骤,随着人工智能、机器学习等信息类学科的发展,该方向的研究有望克服诸多难题的限制,并投入到信息化发展的浪潮中,进一步提高人们的生活品质。

参 考 文 献

- [1] EKMAN P, SORENSON E, FRIESEN W. Pan-Cultural Elements in Facial Displays of Emotion [J]. *Science*, 1969, 164 (3875): 86-88.
- [2] AVIEZER H, TROPE Y, TODOROV A. Body Cues, Not Facial Expressions, Discriminate between Intense Positive and Negative Emotions [J]. *Science*, 2012, 338 (6111): 1225-1229.
- [3] 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.
- [4] TIAN Y L, KANADE T, COHN J. Recognizing Action Units for Facial Expression Analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23 (2): 97-115.
- [5] GAO W, JIN H. Analysis and research of facial expression image [J]. *Chinese Journal of Computers*, 1997, 20 (9): 782-789. (in Chinese)

- 高文,金辉. 面部表情图像的分析与识别[J]. 计算机学报,1997,20(9):782-789.
- [6] EKMAN P, COHEN L, MOOS R, et al. Divergent Reactions to The Threat of War[J]. *Science*,1963,139(3550):88-94.
- [7] YANG S F, KAFAI M, AN L, et al. Zapping Index: Using Smile to Measure Advertisement Zapping Likelihood[J]. *IEEE Transactions on Affective Computing*,2014,5(4):432-444.
- [8] TEIJEIRO-MOSQUERA L, BIEL J, ALBA-CASTRO J, et al. What Your Face Vlogs About: Expressions of Emotion and Big-Five Traits Impressions in YouTube[J]. *IEEE Transactions on Affective Computing*,2015,6(2):193-205.
- [9] RADHAMANI V, DALIN G. A Supporting Survey to Step into a Novel Approach for Providing Automated Emotion Recognition Service in Mobile Phones[C]// *Proceedings of the Second International Conference on Inventive Systems and Control*. New York: IEEE Press,2018:35-39.
- [10] SARIYANIDI E, GUNES H, CAVALLARO A. Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and Recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2015,37(6):1113-1133.
- [11] LYONS M J, CAMPBELL R, PLANTE A, et al. The Noh Mask Effect Vertical Viewpoint Dependence of Facial Expression Perception[C]// *Proceedings of the Royal Society B, Biological Sciences*. London: Royal Society of London,2000:2239-2245.
- [12] NICHOLLS R, WOLFGANG J, CLODE D, et al. The Effect of Left and Right Poses on the Expression of Facial Emotion[J]. *Neuropsychologia*,2002,40(10):1662-1665.
- [13] SAMANTA A, GUHA T. On The Role of Head Motion in Affective Expression [C] // *IEEE International Conference on Acoustics, Speech and Signal Processing*. New York: IEEE Press,2017:2886-2890.
- [14] ZENG Z H, PANTIC M, ROISMAN I, et al. A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2009,31(1):39-58.
- [15] CORNEANU A, OLIU M, COHN F, et al. Survey on RGB,3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-Related Applications [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2016,38(8):1548-1568.
- [16] EKMAN P, FRIESEN W. Facial Action Coding System: A Technique for The Measurement of Facial Movement[M]. Palo Alto: Consulting Psychologists Press,1978.
- [17] PANTIC M, VALSTAR M, RADEMAKER R, et al. Web-Based Database for Facial Expression Analysis[C]// *IEEE International Conference on Multimedia and Expo*. Amsterdam, Netherlands, New York: IEEE Press,2005:317-321.
- [18] XUE Y L, MAO X, ZHANG F. Beihang University Facial Expression Database and Multiple Facial Expression Recognition [C]// *IEEE International Conference on Machine Learning and Cybernetics*. New York: IEEE Press,2006:3282-3287.
- [19] TONG Y, LIAO W H, JI Q. Facial Action Unit Recognition by Exploiting Their Dynamic and Semantic Relationships[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2007,29(10):1683-1699.
- [20] GAO W, CAO B, SHAN S G, et al. The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations [J]. *IEEE Transactions on Systems, Man and Cybernetics, Part A*,2008,38(1):149-161.
- [21] LANGNER O, DOTSCHE R, BIJLSTRA G, et al. Presentation and Validation of The Radboud Faces Database[J]. *Cognition & Emotion*,2010,24(8):1377-1388.
- [22] LUNDQVIST D, FLYKT A, ÖHMAN A. The Karolinska Directed Emotional Faces-KDEF[R]. Stockholm: Department of Clinical Neuroscience, Psychology section, Karolinska Institutet,1998.
- [23] DHALL A, MURTHY O V, GOECKE R, et al. Video and Image based Emotion Recognition Challenges in the Wild: EmotiW 2015[C]// *ACM International Conference on Multimodal Interaction*. New York: ACM Press,2015:423-426.
- [24] YIN L J, WEI X Z, SUN Y, et al. A 3D Facial Expression Database for Facial Behavior Research[C]// *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press,2006:211-216.
- [25] ZHANG X, YIN L J, COHN J, et al. A High-Resolution 3D Dynamic Facial Expression Database[C]// *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press,2008:1-6.
- [26] SAVRAN A, ALYÜZ N, DİBEKLİOĞLU H. Bosphorus Database for 3D Face Analysis[C]// *Proceedings of the First European Workshop on Biometrics & Identity Management*. New York: IEEE Press,2008:47-56.
- [27] ZHU X X, RAMANAN D. Face Detection, Pose Estimation and Landmark Localization in the Wild[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE Press,2012:2879-2886.
- [28] SETTY S, HUSAIN M, BEHAM P, et al. Indian Movie Face Database: A Benchmark for Face Recognition Under Wide Variations[C]// *IEEE International Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*. New York: IEEE Press,2013:1-5.
- [29] VIOLA P, JONES M. Robust Real-Time Face Detection[J]. *International Journal of Computer Vision*,2004,57(2):137-154.
- [30] YACOOB Y, DAVIS L, BLACK M, et al. Looking at People in Action[M]// Cipolla R, Pentland A, Eds. *Computer Vision for Human-Machine Interaction*. Cambridge: Cambridge University Press,1998:171-187.
- [31] CECH J, FRANC V, MATAS J. A 3D Approach to Facial Landmarks: Detection, Refinement, and Tracking[C]// *IEEE International Conference on Pattern Recognition*. New York: IEEE Press,2014:2173-2178.
- [32] ZHANG H W, LI Q, SUN Z, et al. Combining Data-Driven and Model-Driven Methods for Robust Facial Landmark Detection [J]. *IEEE Transactions on Information Forensics and Security*, arXiv:1611.10152.
- [33] PANTIC M, PATRAS I. Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments from

- Face Profile Image Sequences[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 2006, 36(2): 433-449.
- [34] ROTH M, BAUML M, NEVATIA R, et al. Robust Multi-Pose Face Tracking by Multi-Stage Tracklet Association[C]// *IEEE International Conference on Pattern Recognition*. New York: IEEE Press, 2012: 1012-1016.
- [35] KÜBLBECK C, ERNST A. Face Detection and Tracking in Video Sequences Using The Modified Census Transformation[J]. *Image & Vision Computing*, 2006, 24(6): 564-572.
- [36] TANG C G, ZHENG W M, YAN J W, et al. View-Independent Facial Action Unit Detection[C]// *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2017: 878-882.
- [37] BATISTA C, ALBIERO V, BELLON P, et al. AUMPNet: Simultaneous Action Units Detection and Intensity Estimation on Multipose Facial Images Using A Single Convolutional Neural Network[C]// *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2017: 866-871.
- [38] BAN Y, KIM S K, KIM S, et al. Face Detection Based on Skin Color Likelihood[J]. *Pattern Recognition*, 2014, 47(4): 1573-1585.
- [39] CHEN Y D, LI R F. A hierarchical model for joint object detection and pose estimation[J]. *Journal of Computer Research and Development*, 2015, 52(11): 2431-2440. (in Chinese)
陈耀东, 李仁发. 一种层次化的联合识别模型[J]. *计算机研究与发展*, 2015, 52(11): 2431-2440.
- [40] MURPHY-CHUTORIAN E, TRIVEDI M. Head Pose Estimation in Computer Vision: A Survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(4): 607-626.
- [41] MA B P, SHAN S G, CHEN X L, et al. Robust appearance-based method for head pose estimation[J]. *Journal of Software*, 2009, 20(6): 1651-1663. (in Chinese)
马丙鹏, 山世光, 陈熙霖, 等. 一种基于图像观的鲁棒姿态估计方法[J]. *软件学报*, 2009, 20(6): 1651-1663.
- [42] FLOHR F, DUMITRU-GUZU M, KOOIJ J, et al. A Probabilistic Framework for Joint Pedestrian Head and Body Orientation Estimation[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(4): 1872-1882.
- [43] YAN Y, RICC E, SUBRAMANIAN R, et al. A Multi-Task Learning Framework for Head Pose Estimation Under Target Motion[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 38(6): 1070-1083.
- [44] KONG G S, MBOUNA R. Head Pose Estimation from A 2D Face Image Using 3D Face Morphing with Depth Parameters [J]. *IEEE Transactions on Image Processing*, 2015, 24(6): 1801-1808.
- [45] MUKHERJEE S, ROBERTSON N. Deep Head Pose: Gaze-Direction Estimation in Multimodal Video[J]. *IEEE Transactions on Multimedia*, 2015, 17(11): 2094-2107.
- [46] PANTIC M, PATRAS I, ROTHKRANTZ L. Facial Mimics Recognition from Face Profile Image Sequences[R]. Delft: Delft University of Technology, 2002.
- [47] HU Y X, ZENG Z H, YIN L J, et al. A Study of Non-Frontal-View Facial Expressions Recognition[C]// *IEEE International Conference on Pattern Recognition*. New York: IEEE Press, 2008: 1-4.
- [48] SUNG J, KIM D. Pose-Robust Facial Expression Recognition Using View-Based 2D + 3D AAM[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 2008, 38(4): 852-866.
- [49] HESSE N, GEHRIG T, GAO H, et al. Multi-View Facial Expression Recognition Using Local Appearance Features[C]// *IEEE International Conference on Pattern Recognition*. New York: IEEE Press, 2012: 3533-3536.
- [50] TONG Y, CHEN J X, JI Q. A Unified Probabilistic Framework for Spontaneous Facial Action Modeling and Understanding[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(2): 258-273.
- [51] ZHENG W M, TANG H, LIN Z C, et al. A Novel Approach to Expression Recognition from Non-Frontal Face Images[C]// *IEEE International Conference on Computer Vision*. New York: IEEE Press, 2009: 1901-1908.
- [52] WU Y, WANG Z G, JI Q. Facial Feature Tracking Under Varying Facial Expressions and Face Poses Based on Restricted Boltzmann Machines[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE Press, 2013: 3452-3459.
- [53] ZHANG Z P, ZHANG W, DING H J, et al. Hierarchical Facial Landmark Localization via Cascaded Random Binary Patterns [J]. *Pattern Recognition*, 2014, 48(4): 1277-1288.
- [54] SMITH B M, BRANDT J, LIN Z, et al. Nonparametric Context Modeling of Local Appearance for Pose- and Expression-Robust Facial Landmark Localization[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE Press, 2014: 1741-1748.
- [55] WU Y, WANG Z H, JI Q. A Hierarchical Probabilistic Model for Facial Feature Detection[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE Press, 2014: 1781-1788.
- [56] KUMANO S, OTSUKA K, YAMATO J, et al. Pose-Invariant Facial Expression Recognition Using Variable-Intensity Templates[C]// *Proceedings of the 8th Asian Conference on Computer Vision*. New York: IEEE Press, 2007: 324-334.
- [57] HU Y X, ZENG Z H, YIN L J, et al. Multi-View Facial Expression Recognition[C]// *Proceedings of the 8th International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2008: 1-6.
- [58] ZHENG W M. Multi-View Facial Expression Recognition Based on Group Sparse Reduced-Rank Regression[J]. *IEEE Transactions on Affective Computing*, 2014, 5(1): 71-85.
- [59] RUDOVIC O, PATRAS I, PANTIC M. Coupled Gaussian Process Regression for Pose Invariant Facial Expression Recognition[C]// *Proceedings of the 11th European Conference on Computer Vision*. New York: IEEE Press, 2010: 350-363.
- [60] 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.
- [61] MOORE S, BOWDEN R. Local Binary Patterns for Multi-View Facial Expression Recognition[J]. *Computer Vision Image Understanding*, 2011, 115(4): 541-558.
- [62] ZHENG W M, TANG H, LIN Z C, et al. Emotion Recognition from Arbitrary View Facial Images[C] // *Proceedings of the 11th European Conference on Computer Vision*. New York: IEEE Press, 2010: 490-503.
- [63] WU J, LIN Z, ZHENG W, et al. Locality-Constrained Linear Coding Based Bi-layer Model for Multi-view Facial Expression Recognition[J]. *Neurocomputing*, 2017, 239(C): 143-152.
- [64] YUN W, KIM D, PARK C, et al. Multi-View Facial Expression Recognition Using Parametric Kernel Eigenspace Method Based on Class Features[C] // *IEEE International Conference on Systems, Man, and Cybernetics*. New York: IEEE Press, 2013: 2689-2693.
- [65] MAO Q R, RAO Q Y, YU Y B, et al. Hierarchical Bayesian Theme Models for Multipose Facial Expression Recognition[J]. *IEEE Transactions on Multimedia*, 2017, 19(4): 861-873.
- [66] DAPOGNY A, BAILLY K, DUBUISSON S. Dynamic Pose-Robust Facial Expression Recognition by Multi-View Pairwise Conditional Random Forests [OL]. <https://ieeexplore.ieee.org/document/7934067>.
- [67] VIERIU L, TULYAKOV S, SEMENIUTA S, et al. Facial Expression Recognition under a Wide Range of Head Poses[C] // *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2015: 1-7.
- [68] YANG B, CAO J M, NI R R, et al. Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images[J]. *IEEE Access*, 2018, 6: 4630-4640.
- [69] WU B F, LIN C. Adaptive Feature Mapping for Customizing Deep Learning Based Facial Expression Recognition Model[J]. *IEEE Access*, 2018, 6: 12451-12461.
- [70] LIU Y Y, ZENG J B, SHAN S G, et al. Multi-channel Pose-aware Convolution Neural Networks for Multi-view Facial Expression Recognition[C] // *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2018: 458-465.
- [71] JAMPOUR M, MAUTHNER T, BISCHOF H. Pairwise Linear Regression: An Efficient and Fast Multi-View Facial Expression Recognition[C] // *Proceedings of the 11th IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2015.
- [72] JAMPOUR M, LEPETIT V, MAUTHNER T, et al. Pose-specific Non-linear Mappings in Feature Space towards Multiview Facial Expression Recognition[J]. *Image & Vision Computing*, 2017, 58: 38-46.
- [73] RUDOVIC O, PATRAS I, PANTIC M. Regression-Based Multi-View Facial Expression Recognition[C] // *Proceedings of the 20th International Conference on Pattern Recognition*. New York: IEEE Press, 2010: 4121-4124.
- [74] RUDOVIC O, PANTIC M, PATRAS I. Coupled Gaussian Processes for Pose-Invariant Facial Expression Recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(6): 1357-1369.
- [75] LAI Y H, LAI S H. Emotion-Preserving Representation Learning via Generative Adversarial Network for Multi-view Facial Expression Recognition[C] // *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2018: 263-270.
- [76] GÜNEY F, ARAR N M, FISCHER M, et al. Cross-Pose Facial Expression Recognition[C] // *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2013.
- [77] TARIQ U, YANG J C, HUANG T S. Supervised Super-Vector Encoding for Facial Expression Recognition[J]. *Pattern Recognition Letters*, 2014, 46(3): 89-95.
- [78] HOSOI T. Head Pose and Expression Transfer Using Facial Status Score[C] // *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2017: 573-580.
- [79] ZHOU Y, PI J, SHI B E. Pose-Independent Facial Action Unit Intensity Regression Based on Multi-Task Deep Transfer Learning[C] // *IEEE International Conference on Automatic Face & Gesture Recognition*. New York: IEEE Press, 2017: 872-877.
- [80] ZHENG W M, ZONG Y, ZHOU X Y, et al. Cross-Domain Color Facial Expression Recognition Using Transductive Transfer Subspace Learning[J]. *IEEE Transactions on Affective Computing*, 2018, 9(1): 21-37.
- [81] KAYA H, GÜRPINAR F, SALAH A A. Video-Based Emotion Recognition in the Wild using Deep Transfer Learning and Score Fusion[J]. *Image & Vision Computing*, 2017, 65: 66-75.
- [82] WEBER R, BARRIELLE V, SOLADIÉ C, et al. Unsupervised Adaptation of A Person-Specific Manifold of Facial Expressions [J]. *IEEE Transactions on Affective Computing*, 2018. doi: 10.1109/TAFFC.2018.2807430.