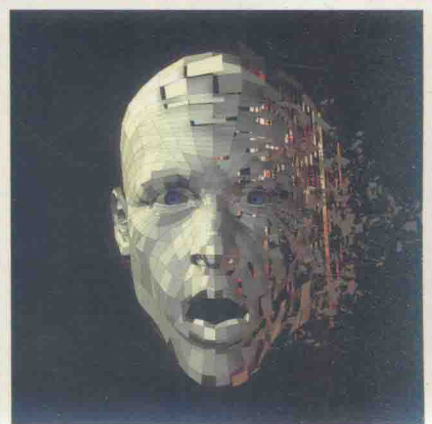


计 算 机



人脸表情自动识别

JISUANJI RENLIAN BIAOQING ZIDONG SHIBIE JISHU YU FANGFA

技术与方法

王文成◎著



中国水利水电出版社
www.waterpub.com.cn

计算机

人脸表情自动识别 技术与方法



文成◎著



中国水利水电出版社

www.waterpub.com.cn

·北京·

内 容 提 要

本书对计算机人脸表情自动识别技术与方法进行了分析与研究,其主要内容包括:人脸区域自动检测、人脸表情图像预处理、基于区域投影的人眼定位方法、基于 ARLBP 的人眼定位方法、基于特征块和 LBP 的表情识别方法、基于 Gabor 滤波器的人脸表情识别、基于证据理论的表情识别方法等。

本书结构合理,条理清晰,内容丰富新颖,可供相关工程技术人员参考使用。

图书在版编目(CIP)数据

计算机人脸表情自动识别技术与方法/王文成著.

—北京:中国水利水电出版社,2017.7

ISBN 978-7-5170-5723-9

I. ①计… II. ①王… III. ①图象识别 IV.
①TP391.41

中国版本图书馆 CIP 数据核字(2017)第 188053 号

书 名	计算机人脸表情自动识别技术与方法 JISUANJI RENLIAN BIAOQING ZIDONG SHIBIE JISHU YU FANGFA
作 者	王文成 著
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路 1 号 D 座 100038) 网址:www.waterpub.com.cn E-mail:sales@waterpub.com.cn 电话:(010)68367658(营销中心)
经 售	北京科水图书销售中心(零售) 电话:(010)88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	北京亚吉飞数码科技有限公司
印 刷	三河市天润建兴印务有限公司
规 格	170mm×240mm 16 开本 12.5 印张 162 千字
版 次	2018 年 8 月第 1 版 2018 年 8 月第 1 次印刷
印 数	0001—2000 册
定 价	64.00 元

凡购买我社图书,如有缺页、倒页、脱页的,本社营销中心负责调换

版权所有·侵权必究

自序

在人与人之间的交往中,面部表情作为一种非常重要的信息传递方式,能够传达很多语言所不能传达的信息。在人工智能和物联网技术快速发展的今天,如果人跟机器人之间也能够进行表情的沟通,实现人机之间的智能交互,那将会极大地促进人工智能技术的发展。因此,对计算机人脸表情进行分析和识别具有重要的研究意义,也是近十多年来模式识别领域研究的一个重要问题。本专著主要基于计算机人脸表情自动识别技术中人脸检测与预处理、人眼定位与特征提取、表情识别三个模块进行论述。全文根据内容的相关性共分为9章,其中第1章主要介绍了人脸表情识别研究的现状。第2章分析了基于肤色和 AdaBoost 方法相结合的人脸检测算法。第3章为表情图像预处理方法。第4章和第5章分别为基于区域投影的人眼自动定位方法和基于 ARLBP 的人眼定位方法。第6章、第7章和第8章为基于特征块和 LBP 特征的表情识别方法、基于 Gabor 特征的表情识别方法和基于证据理论的表情识别方法。第9章总结全文工作并对进一步的研究提出设想和建议。全文整体结构清晰,内容安排紧凑合理。

本专著的写作是根据作者多年的研究成果整理而成,主要工作是将计算机人脸表情识别相关研究内容系统化,有利于为初涉该领域的研究人员提供参考。文中所提出的算法能够通过实验验证,在此基础上进一步的研究并加以完善,将可以不断优化人脸表情自动识别系统。

全文的写作得到了山东大学常发亮教授、陈振学博士的指导

作者

2017 年 5 月

目 录

自 序

第 1 章 绪 论	1
1.1 人脸表情识别研究的背景和意义	1
1.2 国内外研究的现状	8
1.3 存在的问题及发展趋势	23
1.4 人脸表情数据库	25
1.5 本书主要内容和结构安排	30
参考文献	33
第 2 章 人脸区域自动检测	57
2.1 基于肤色的人脸区域检测	58
2.2 基于 AdaBoost 的人脸检测	68
2.3 基于肤色+AdaBoost 方法的人脸检测	75
2.4 本章小结	77
参考文献	77
第 3 章 人脸表情图像预处理	80
3.1 表情图像的几何标准化	80
3.2 表情图像的去噪	88
3.3 图像的灰度变换	91
3.4 表情图像的灰度标准化	91
3.5 本章小结	94
参考文献	94
第 4 章 基于区域投影的人眼定位方法	96
4.1 人眼定位方法分析	96
4.2 基于区域投影的人眼定位方法	101

4.3 实 验	106
4.4 本章小结	110
参考文献	110
第 5 章 基于 ARLBP 的人眼定位方法	114
5.1 局部二值模式	114
5.2 自适应比率局部二值模式	118
5.3 实验结果及分析	127
5.4 本章小结	131
参考文献	132
第 6 章 基于特征块和 LBP 的表情识别方法	134
6.1 显著性特征区域的划分	134
6.2 LBP 特征的提取	137
6.3 LBP 特征降维	138
6.4 分类器设计	142
6.5 实验结果及分析	143
6.6 本章小结	148
参考文献	148
第 7 章 基于 Gabor 滤波器的表情识别方法	151
7.1 Gabor 特征提取	151
7.2 Gabor 特征分析	159
7.3 基于 Gabor 特征+Fisherfaces 的表情分析方法	162
7.4 实验设计	164
参考文献	169
第 8 章 基于证据理论的表情识别方法	172
8.1 信息融合与 D-S 证据理论	172
8.2 基于 D-S 证据理论的表情识别方法	178
8.3 实验结果及分析	183
8.4 本章小结	189
参考文献	189
第 9 章 总结与展望	192

第 1 章 绪 论

1.1 人脸表情识别研究的背景和意义

1.1.1 人脸表情识别研究的背景

在日常生活中,情绪能够表征人们对一个客观事物的态度,有助于了解他人的意图和看法。对于能够满足人们需要的事物会引起人肯定性质的情绪,如愉快、满意等;不能满足人们需要的事物会引起人否定性质的情绪,如愤怒、憎恨、哀怨等;而与需要无关的事物,会使人产生无所谓的情绪和情感。面部表情作为情绪表达的一种途径,包含了丰富的情绪信息和心理过程,它不但是人际交往的手段,也是人们进行非语言交流的一种有效方式。人们既可以通过面部表情准确、充分、微妙地表达自己的思想情感,也可以通过表情辨认对方的情感波动和内心世界。心理学家 Mehrabian 曾经指出^[1]:人与人之间的交流主要通过言辞、声音和面部表情完成,其中言辞占信息的 7%,声音占信息的 38%,而面部表情占信息的 55%。由此可见,面部表情在人与人之间的交流中包含了丰富的行为信息,对它的研究可以进一步了解人类对应的心理状态,而表情分析所要研究的就是如何自动、可靠、高效地获取人脸表情所传达信息的过程^[2]。

随着计算机科学技术和机器人技术的不断发展,人机交互

(Human to Computer Interface, HCI) 技术日益成为人工智能领域中研究的热点, 人与计算机之间交流感情问题开始受到广泛的关注^[3]。在机器系统中的多模式人机交互界面中, 表情与体态和声音的有机结合可以获得更加高效的人机信息传递。如果计算机和机器人系统能够像人类那样具有理解和表达情感的能力, 根据人所处的环境、心情、爱好和习惯等信息主动进行判断来帮助人类完成各项任务, 那将从根本上改变人与计算机之间的关系, 使计算机能够更好地为人类服务^[4-6]。因此, 基于计算机对表情识别技术进行研究, 让计算机更好地理解和表达情感, 成为社会发展的需要和历史发展的必然。

基于计算机的人脸表情识别(Facial Expression Recognition, FER)是指利用计算机检测人脸, 对人脸的表情信息进行特征提取和分析, 最终实现计算机对人脸表情的理解与判断。通俗地说, 人脸表情识别就是计算机对人脸表情信息进行特征提取并归类的过程, 它使计算机能够获知人的表情信息, 进而推断人的心理, 从而实现人机之间的高级智能交互^[7-16]。它作为情感计算研究的重要分支, 能够有效地促进人机交互系统的发展, 同时也为表情合成、人体语言与自然语言的融合、人脸表情编码等领域的研究提供了理论基础。

最初与人脸表情相关的研究主要集中在生理学、心理学和认知科学部分。早在 19 世纪, 英国生物学家达尔文就对人类表情与动物表情之间的区别与联系进行了分析研究^[17], 证明了表情与性别、种族无关而具有一致性。1954 年, Schlosberg 提出了用来描述表情变化的 3 个量化尺度(注意/拒绝、高兴/不高兴、活跃程度)。尽管该方法无法实现对表情的分类, 但是为后续研究提供了理论支持^[18]。1971 年, 美国心理学家 Ekman 和 Friesen 对面部表情进行了重新划分, 他们定义了愤怒(Anger)、高兴(Happy)、悲伤(Sad)、惊讶(Surprise)、厌恶(Disgust)和恐惧(Fear) 6 种基本表情^[19], 并根据肌肉的运动建立了面部表情运动编码系统(Facial Action Coding System, FACS), 通过 44 个独立的脸部运

动单元(Active Units, AU)进行不同组合来描述人脸表情,其中一些运动单元如图 1.1 所示。该成果当时虽然主要应用于心理学领域,但是后来却为计算机视觉中人脸表情识别研究奠定了基础。

AU1  Inner Brow Raiser	AU2  Outer Brow Raiser	AU4  Brow Lowerer	AU5  Upper Lid Raiser	AU6  Cheek Raiser
AU7  Lid Tightener	AU9  Nose Wrinkler	AU10  Upper Lip Raiser	AU11  Nasolabial Deepener	AU12  Lip Corner Puller
AU13  Cheek Puffer	AU14  Dimpler	AU15  Lip Corner Depressor	AU16  Lower Lip Depressor	AU17  Chin Raiser
AU18  Lip Pucker	AU20  Lip Stretcher	AU22  Lip Funneler	AU23  Lip Tightener	AU28  Lip Suck

图 1.1 FACS 运动单元(AU)示例

1978 年, Suwa 等^[20]对表情识别进行了初次尝试,通过对连续图像序列中脸部固定的 20 个点的变化进行跟踪,并与原型模式进行对比识别,从而将表情识别引入机器视觉领域。20 世纪 90 年代,以 Mase 和 Pentland 为代表的研究人员利用光流法提取

表情特征,并构建了能够识别高兴、生气、厌恶、惊讶等 4 种表情的识别系统,精确率达到 80% 以上,利用计算机进行自动人脸表情分析的研究工作迅速展开^[20]。近几年,计算机技术的发展使得人脸表情识别成为多学科交叉研究的前沿热点^[21-26],研究方法也从二维图像识别扩展到三维图像领域,涉及的学科具体包括模式识别、图像处理、计算机视觉、人工智能、认知科学、生理学、心理学等。对人脸表情识别技术的研究可以促进其他学科的发展,而其他学科的发展也为人脸表情识别提供了理论指导。

人脸表情识别系统主要由人脸检测与预处理模块、表情特征提取模块和表情分类模块等组成,其基本结构如图 1.2 所示。



图 1.2 人脸表情识别系统的主要组成

(1)人脸检测与预处理模块主要完成从输入的图像或视频中输出人脸表情图像。在自然场景下,一幅图像中所包含的信息不仅包括人脸还包含其他物体,因此需要首先将人脸从背景中检测并分离出来。由于获得的人脸区域受尺度变化、旋转和光照变化等情况的影响,还需要对人脸区域进行针对性预处理和归一化操作,其中人眼的检测及定位是进行人脸几何归一化的前提。

(2)表情特征提取模块是从人脸图像或图像序列中提取能够表征表情本质的特征信息,是人脸表情识别中的重要步骤,是提高表情分类性能的关键。通过特征提取和选择获得对分类更为

有利的特征数据,有时候由于特征维数较高存在冗余,需要在特征空间内进行特征筛选和降维。

(3)表情分类模块主要是将输入的人脸表情划分到相应的类别,判断人的情感并进行决策。根据选取的表情特征和不同的表情特征表示方式,选用相应的分类算法,训练出功能强大的分类器是表情分类的关键。

基于以上3个主要步骤,则一般的表情识别过程如图1.3所示。

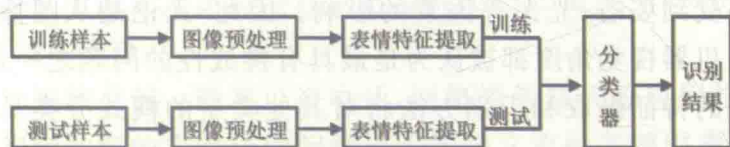


图 1.3 人脸表情识别的一般过程

对于一个模式识别问题,提取适合模式分类的模式特征非常重要,如果特征提取具有本征性,即使采用简单的分类器也能取得好的分类效果。通过对大量文献的总结,良好的表情特征提取结果应该具备以下几个条件:

- (1)能够完整地表示出人脸表情的本质特征。
- (2)去除噪声、光照及其他与表情无关的干扰信息。
- (3)数据表示形式紧凑,避免过高的维数。
- (4)不同表情类别的特征之间有较好的区分性。

1.1.2 人脸表情识别研究的意义

人脸面部表情自动识别技术之所以受到重视,是因为其有重要的研究意义,突出表现在对学科发展的促进作用和潜在的应用价值两个方面。

1. 对人脸面部表情识别的研究可以进一步促进多门相关学科的发展

人脸面部表情识别是一个典型的图像模式分析、理解和分类

问题,是多学科交叉研究的前沿热点,可为图像处理、计算机视觉、人工智能、模式识别、人机交互、认知科学、生理学、心理学等学科提供一系列良好的具体问题,有利于构建这些学科领域的基础试验平台用于尝试新方法、验证新理论、解释新现象。人脸表情识别问题的深入研究和最终解决可以有效地促进这些学科的发展,具有重要的科学意义和理论价值。

(1)作为一个模式识别问题,由于人脸为一个非刚性的柔体,不同人面相的差异导致所表现的同一种类型的表情差距较大,并且容易受到姿态、光照等因素的影响。因此,无论是从图像处理还是从机器视觉角度都被认为是最具有挑战性的问题之一,研究其良好的特征提取和识别方法将对其他类型的模式分类具有指导意义。

(2)从认知科学的发展角度看,关于人类的大脑为什么可以记住并识别成百上千张不同的面孔这个问题,在脑科学以及心理学领域中仍然是一个没有完全揭开的迷。如果人类通过一定的算法让计算机能够准确地识别人脸面部表情,这将必然对认知科学的发展、对了解人类认识世界的机制产生促进作用。

2. 人脸面部表情识别技术具有广阔的应用前景

人脸面部表情识别是情感理解的基础,是计算机理解人类情感的前提,也是人们探索智能、理解智能的有效途径。如何实现计算机的拟人化,使其能根据周围的环境以及对象的状态等内容,自适应地为交流对象提供最友好的操作环境,消除操作者和机器之间的障碍,已经成为下一代人机界面发展的目标,是智能机器人走进人们日常生活必须解决的问题,对建立多信息智能化人机交互系统有着重要意义。其研究成果可以用于以下领域:

(1)安全监控领域。随着机动车数量的大幅增加,道路安全已成为当今全球关注的热点问题之一,道路交通事故居高不下,酒后驾驶、疲劳驾驶和分心驾驶等违章驾驶现象普遍存在。据公安部网站消息,2011年全国共发生道路交通事故210 812起,造

成 62 387 人死亡,直接经济损失达 9.3 亿元,对人民的生命和财产造成了巨大的损失。如果能够通过监控司机面部表情及时识别出其疲劳、醉酒、分心等违章驾驶行为,并及时采取措施,则可及早排除由此引起的道路事故隐患,在一定程度上减少生命和财产损失。面部表情识别还可用于航空驾驶、核电站、钢厂、煤矿管理等着重强调安全和注意力的工作岗位,在岗者一旦出现疲劳、瞌睡的征兆,识别系统将及时发出警报避免险情发生。此外,将表情识别系统安装在公共场合,对表情可疑人员进行锁定和跟踪,将有利于治安人员将潜在的危险扼杀在摇篮之中,维护地方公共区域的安全。

(2) 视频交流。随着可视电话、网络视频聊天等应用技术的应用,视频数据的实时传输问题成为影响交流的关键因素之一。由于在人与人之间交谈时,面部器官本身的结构特征对于获取信息来说是次要的,而语言和面部表情才是传递信息的主要载体,因此在线视频对话时,可通过系统对视频的处理,提取表情信息并参数化,让网络仅仅传递对方的语言和表情类别,进而将参数用于驱动远程客户端上虚拟人头部模型产生相应的表情,从而达到低带宽在线视频表情对话的效果。

(3) 医疗领域。表情分析可以帮助医生分析病人的精神状态,使医生对病人的病情做出正确的诊断。人脸表情识别还可用于医疗电子护士系统,特别是在重症病房里,患者遇到紧急情况无法通过按钮进行紧急呼叫时,该系统可以全天候监测患者的面部表情,了解其身体状况的突然变化并报警,使医护人员对患者病情做出及时的治疗。此外,该技术还可在抑郁症、面瘫、精神病和中风病等疾病的辅助诊断和治疗方面发挥重要作用,记录并识别患者的精神状态,为临床医生提供数据参考。

(4) 远程网络教学。在当前的远程教育系统中,学生往往处于被动接收知识的状态,使得教学效果无法得到实时反馈,降低了学习效率。如果在系统中加入表情识别技术,老师就可以通过学生们的表情来得知他们对课程的掌握程度,以便随时调整教学

方案,提高教学质量。

(5)休闲娱乐。未来的网站将可以根据用户摄像头输入的画面信息,自动分析访问者的表情,从而根据分析结果决定在页面上放置什么样的广告图片,并播放相应的音乐;网络游戏中的角色,实时模拟网络游戏玩家的表情,增加娱乐的趣味性;基于表情识别技术可以开发宠物机器人,能根据用户的情绪状态做出相应的动作;另外,人脸表情识别技术在虚拟现实、人机交互、电脑游戏等许多领域都有着潜在的应用价值,如图 1.4 所示。

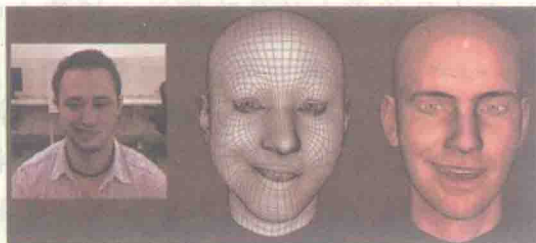


图 1.4 3D 人机交互

总之,人脸表情识别技术作为一种典型的高端生物特征识别技术具有重要的科学价值和广泛的应用前景。它是人机智能交互的重要基础,是当前国内外模式识别和人工智能领域的一个研究热点。它与人们的日常生活息息相关,对提高人们的生活质量具有重要作用。

1.2 国内外研究的现状

1.2.1 研究机构概况

计算机人脸表情识别技术在近几年的发展相当迅速,已经成为国际上研究的热点。近些年来,很多关于人脸表情识别方面的论文在计算机视觉领域的两大顶级国际期刊 PAMI (IEEE Transactions on Pattern Analysis and Machine Intelligence)、

IJCV(International Journal of Computer Vision)和三大顶级国际会议 ICCV (IEEE International Conference on Computer Vision)、ECCV(European Conference on Computer Vision)、CVPR (IEEE Conference on Computer Vision and Pattern Recognition)以及其他重要国际期刊、会议中发表。国内外很多机构都在进行这方面的研究,吸引了大量的研究人员和基金支持。

1. 国外研究机构

近几年,美国、日本对人脸表情识别的研究变得非常活跃。此外,英国、德国、荷兰、瑞士等经济发达国家和印度、新加坡等发展中国家都有专门的研究组织进行这方面的研究,其中麻省理工学院的媒体实验室(Media Laboratory)^[20,27]、卡耐基梅隆大学(CMU)^[28,29]、剑桥大学^[30,31]、马里兰大学的计算机视觉实验室(Computer Vision Laboratory,CVL)^[32]、加利福尼亚大学的机器感知研究所(Machine Perception Laboratory,PRLab)^[33,34]、东京大学^[35]、韩国的庆熙大学^[36,37]、新加坡国立大学^[5,38,190]、瑞士的IDIAP研究院^[39]、日本的ATR研究所^[4]等机构的贡献尤为突出。例如,MIT媒体实验室的Essa和Pentland较早地提出了人脸表情自动分析与处理问题;Picard教授提出了情感计算(Effective Computing)的概念,将有关人类情感的研究推向更高的层次。

2. 国内研究机构

相对于国外的研究,国内对人脸表情识别的研究起步比较晚。较早开展这方面研究的是中科院计算所和哈尔滨工业大学的高文教授和金辉博士,他们早在20世纪90年代便开始了针对表情识别的相关研究^[40-43]。2003年,北京科技大学的王志良、陈峰军等人,将人脸表情识别应用于机器人的情感控制研究中^[44]。当前在国内开展表情识别研究的科研院所越来越多,如清华大学^[45,46]、东南大学^[47-50]、中国科技大学^[51]、浙江大学^[52-54]、北京

航空航天大学^[6,55]、西北工业大学^[56]、北京科技大学^[57,58]、中南大学^[59,60]、江苏大学^[61,62]、南京大学^[63]、华南理工大学^[64]等都有科研人员从事人脸表情识别的研究。2003年12月,在北京召开了首届中国情感计算与智能交互会议,代表着国内在情感计算方面的研究已经展开。2005年10月,我国召开的首届国际情感计算及智能交互会议受到了来自中国大陆、中国台湾、日本、韩国、美国及欧洲等国家和地区同行的积极响应,推动了我国在此国际前沿方向的发展。

1.2.2 理论研究现状

人脸表情识别的研究主要围绕人脸检测、脸部特征提取和表情分类3部分内容,它们共同决定了表情识别的效果。其研究成果主要体现在以下几个方面。

1. 人脸检测

人脸检测的基本思想是用某种方法对人脸进行建模,比较所有可能的待测区域与人脸模型的匹配度,从而得到可能存在人脸的区域。不同的角度有不同的分类方式,常见的分类方式为基于知识的人脸检测和基于统计学习的人脸检测^[65-67]。

(1) 基于知识的方法,该方法主要是提取人脸中不随个人差异而改变的特征,如灰度^[68]、颜色^[69-71]、纹理^[72]、形状^[73,74]等视觉信息。例如,Yun等^[75]在YCbCr色彩空间提取人脸候选区域,然后基于形态学处理方法,去除非人脸区域并定位人脸;Harasse等^[76]建立肤色模型,利用Bayes决策分类,这种方法的优点在于实现简单,然而在光照变化时会降低人脸检测结果的准确度;Huang等^[72]利用Gabor特征的局部空间描述能力设计了4个方向的滤波器,再利用PCA方法降维后基于多项式神经网络进行训练,取得了良好的检测率;Phimoltares等^[77]提出了一种与背景无关的人脸检测和特征定位方法,通过Canny算子检测边缘,再