

人机交互中的人脸表情识别研究进展

薛雨丽 毛 峡 郭 叶 吕善伟

(北京航空航天大学电子信息工程学院, 北京 100191)

摘 要 随着人机交互与情感计算技术的快速发展,人脸表情识别已成为人们研究的热点。为了阐明人机交互中人脸表情识别的研究方向及进展,该文从人脸表情数据库、表情特征提取、表情分类方法、鲁棒的表情识别、精细的表情识别、混合表情识别、非基本表情识别等方面对人脸表情识别的研究现状进行了分析。最后总结了人脸表情识别研究的热点及趋势,同时指出了人脸表情识别研究存在的局限性,并对人脸表情识别的发展进行了展望。

关键词 表情识别 情感表情 人脸表情数据库

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2009)05-764-09

The Research Advance of Facial Expression Recognition in Human Computer Interaction

XUE Yu-li, MAO Xia, GUO Ye, LV Shan-wei

(School of Electronic and Information Engineering, Beihang University, Beijing 100191)

Abstract Along with the rapid progress of the technologies in human computer interaction and affective computing, facial expression recognition has been actively researched. To clarify the research direction and development of facial expression recognition in human computer interaction, the research states of facial expression recognition are analyzed from the aspects of facial expression database, facial feature extraction, facial expression classification methods, robust facial expression recognition, fine-graded facial expression recognition, mixed facial expression recognition, and non-basic facial expression recognition. Finally, the study hotspots and trends of facial expression recognition are concluded, the limits in facial expression recognition are pointed out, and the expectation of the development of facial expression recognition is given.

Keywords facial expression recognition, emotional facial expression, facial expression database

1 引 言

人们对于人脸表情的研究可以追溯到19世纪,生物学家 Darwin 在《人类和动物的表情》一书中,就对人类的面部表情与动物的面部表情进行了研究和比较^[1]。心理学家 Mehrabian 提出,在人们的交流中,只有7%的信息是通过语言来传递,而通过面部表情传递的信息量却达到了55%^[2]。

人机交互模式已从语言命令交互阶段、图像用户界面交互阶段发展到自然和谐的人机交互阶段。同时,由麻省理工学院 Picard 教授提出的情感计算(affective computing)^[3]领域正蓬勃兴起,其目标是使计算机拥有情感,即能够像人类一样识别和表达情感,使人机交互更加人性化。为了使人机交互更加和谐与自然,新型的人机交互技术正逐渐成为研究热点。

人脸表情识别是人机交互与情感计算研究的重

基金项目: 国家自然科学基金项目(60572044); 国家高技术研究发展计划(863 计划)项目(2006AA01Z135); 高等学校博士学科点专项科研基金项目(20070006057)

收稿日期: 2008-10-11; **改回日期:** 2008-12-12

第一作者简介: 薛雨丽(1980 ~), 女。北京航空航天大学电子信息工程学院博士研究生。主要研究方向为情感计算、人机交互、图像处理与模式识别。E-mail: YuLi_Xue@ee.buaa.edu.cn

要组成部分。由于人脸表情包含丰富的行为信息,因此对人脸表情进行识别有利于了解人类的情感等心理状态,并可进行有效的人机交互。人脸表情识别涉及心理学、社会学、人类学、生命科学、认知科学、生物学、病理学、计算机科学等研究领域。可见,人脸表情识别的进展对提高人工情感智能水平和探索人类情感及认知能力极具科学意义,并将促进相关学科的发展。

2 人脸表情识别

随着计算机技术的发展,人脸表情识别技术也逐渐发展起来。20世纪90年代,对人脸表情识别的研究变得非常活跃。国外较为著名的研究机构有麻省理工学院、卡内基梅隆大学、匹兹堡大学、马里兰大学、加利福尼亚大学、代尔夫特理工大学等。国内的研究始于20世纪90年代末,近几年开展研究的单位较多,例如哈尔滨工业大学、中国科学院、清华大学、中国科学技术大学、浙江大学、南京理工大学、东南大学、北方交通大学、北京航空航天大学、北京科技大学、西南交通大学、大连理工大学、天津大学、湖南大学等都对手脸表情识别进行了研究。

本文将从人脸表情数据库、表情特征提取、表情分类方法、鲁棒的表情识别、精细的表情识别、混合表情识别、非基本表情识别等方面对手脸表情识别的研究现状进行分析。

2.1 人脸表情数据库

由于人脸表情数据库可为人脸表情识别提供必需的人脸表情数据,因此,建立丰富、有效的人脸表情数据库对手脸表情识别技术的发展具有重要意义。

目前的人脸表情数据库大致分为以下两类^[4]:一类人脸表情数据库是基于6种基本情感^[5](高兴、悲伤、害怕、厌恶、惊讶和生气),例如,马里兰大学人脸表情数据库^[6]包含40个不同种族和文化背景被试者的图像序列。每个图像序列持续大约9秒钟,包含1~3种基本表情。日本女性人脸表情数据库(JAFFE)^[7]包含10个日本女性的6种情感表情和中性表情共213个图像。

另一类人脸表情数据库则集中于提取人脸更加细微的表情。这些表情一般采用面部动作编码系统(FACS)进行编码,其可用44个不同的动作单元来描述表情特征的细微变化。例如卡内基梅隆大学机

器人研究所和心理学系共同建立的Cohn-Kanade AU-Coded人脸表情数据库^[8]就包含了单一运动单元和复合运动单元共23种表情。现有210个18岁到50岁成年人的脸部表情图像序列。

此外,还有一些人脸数据库包含了部分非基本表情,如Yale人脸数据库^[9]包含了困乏、眨眼等常见表情;PIE(pose, illumination, and expression)人脸数据库^[10]包含了眨眼等表情;AR人脸数据库^[11]包含了尖叫等表情;韩国人脸数据库(KFDB)^[12]包含有眨眼表情。

在国内,清华大学的清华人脸表情数据库^[13]包含常见的8类情感类表情和中文语音发音的说话类表情。北京航空航天大学毛峡教授等人建立了基于多种情感的北航人脸表情数据库^[14],包含有18种单一表情、3种混合表情和4种复杂表情,具有多种在其他人脸数据库中未曾出现的情感表情。中国科学院计算技术研究所的CAS-PEAL人脸数据库^[15]包含了7种类别(姿势、表情、饰品、光照、背景、距离和时间)的人脸图像。其中表情子数据集包含中性、微笑、皱眉、惊讶、闭眼和张嘴等6种表情。

2.2 表情特征提取

表情特征的提取方法分为基于运动与基于形变的两类^[16-17]。其中运动是指表情发生时面部特征随时间的变化,而形变则是指有表情的脸相对于中性表情脸的差异。

2.2.1 运动特征提取

面部表情图像的运动特征包括特征点运动距离的大小及运动方向,其可以反映人脸的表情及其变化。代表性的方法包括光流法^[18]、运动模型^[18-19]、特征点跟踪方法^[20]等。

佐治亚理工学院的Essa和麻省理工学院的Pentland使用光流法进行了运动估计,并使用面部肌肉运动模型描述了面部的运动^[18]。

马里兰大学计算机视觉研究室的Yacoob和Davis等人提出了面部的局部参数运动模型^[19],同时构建了面部运动的中级描述,并使用启发式规则对6种表情进行了分类。

萨诺夫公司的Zhu等人提出了一种检测28个面部特征点的方法^[21],其对各种表情和姿势的变化具有鲁棒性。该方法首先基于眼睛的位置用一个训练的人脸网格来评估面部特征的大致位置;然后,在预估位置的周围搜索精确的面部特征位置。

美国伦斯勒理工学院的Tong等人提出了多状

态层级方法用来跟踪面部特征点^[22]。其中统计形状模型的层级表达,用于表征人脸的全局形状约束和面部组件的局部结构细节;Gabor 小波和灰度级轮廓用于表征面部特征点;多状态局部形状模型用于处理面部组件的形状变化。这样通过一个多模跟踪方法即可动态评价面部组件的状态和面部特征点的位置。

2.2.2 形变特征提取

面部特征的形变可以由形状和纹理两方面来描述,又可以分为基于模型的方法和基于图像的方法两类。其中基于模型的方法有活动外观模型(AAM)、点分布模型(PDM)等;基于图像的方法有 Gabor 变换、主成分分析(PCA)等。

英国曼彻斯特大学的 Cristinacce 等人提出一种特征响应成对增强算法(PRFR)和 AAM 相结合的方法^[23],用以检测面部眼眶、鼻尖及嘴角等局部区域的特征点及边缘的一些特征。Tang 等人使用 AAM 来提取特征点,并结合局部形状特征识别表情^[24]。

法国格勒诺布尔国立理工学院 Hammal 等人先通过转移可信度模型(TBM)提取基本情感表达的轮廓信息,然后由基于规则的决策系统对距离系数进行分类^[25]。

美国加利福尼亚大学的 Yu 等人提出了一种基于规划的遗传方法^[26],该方法使用 Gabor 小波表征原始特征,并使用线性和非线性的算子合成新的特征,再由支持向量机对表情进行分类。

香港科技大学的 Liao 等人提出了基于两组特征的表情识别方法^[27]。其中一组特征采用了强度和梯度图的局部二元模式特征和 Gabor 滤波器的 Tsallis 能量;另一组特征通过对训练图像的线性判别分析(LDA)来获得。

希腊亚历士多德大学的 Kotsia 等人通过提取表情的纹理和形状信息,并采用神经网络进行融合来对基本表情和表情动作单元进行识别^[28]。

2.3 表情分类方法

目前用于表情分类的方法可以分为时空域方法和空域方法两类。其中时空域方法包括隐马尔可夫模型方法、回归神经网络方法、空时运动能量模板方法等;空域方法包括神经网络方法、支持向量机方法、AdaBoost 方法、基于规则推理的方法、主成分分析、独立分量分析、Fisher 线性判别分析等。近年来,隐马尔可夫模型、人工神经网络、支持向量机和

Adaboost 算法等成为主流的表情识别方法。

隐马尔可夫模型(HMM)是一种用于描述随机过程信号统计特性的概率模型。中科院自动化研究所的 Zhou 等人提出一种基于 Adaboost 的嵌入式隐马尔可夫模型^[29],实现了面部表情的实时识别。加拿大滑铁卢大学的 Miners 等人利用模糊隐马尔可夫模型进行了动态的面部表情识别^[30],并在人机交互中取得了很好的识别效果。

人工神经网络(ANN)是对人脑神经系统的模拟,具有较强的自学习、自组织、联想及容错能力。美国卡内基梅隆大学的 Tian 等人提出一个基于 Gabor 小波特征和 FACS 系统的自动面部活动分析系统^[31],该系统采用神经网络作为分类器,识别率可达 90%。伊朗大不里士大学的 Seyedarabi 等人先从正面图像序列抽取面部特征点,然后通过光流自动跟踪这些特征点,最后使用 RBF 神经网络和模糊推理系统对特征向量进行分类^[32]。加拿大肯考迪亚大学的 Ma 等人先对人脸图像进行 2 维离散余弦变换(DCT),然后使用前馈神经网络作为分类器对表情进行分类^[33]。土耳其东地中海大学的 Soyel 等人将 3 维面部特征点的距离矢量通过神经网络识别 7 种基本情感,获得了 91.3% 的平均识别率^[34]。

支持向量机(SVM)作为一种模式识别方法,具有出色的学习性能和结构风险最小化的优点,其能够解决小样本情况下的学习问题。美国加利福尼亚大学的 Littlewort 等人提出了基于 Gabor 滤波的 SVM 分类器^[35],用于从连续的视频序列中识别表情。东南大学的 Xu 等人提出了一种层级结构的支持向量机树^[36],用于进行表情识别,获得了比其他方法更高的识别精确性和鲁棒性。

Adaboost 算法作为一种健壮、快速的检测算法,已成为对象检测领域的研究热点。清华大学的 Wang 等人使用基于 boosting Haar 特征的 look-up-table 类型的弱分类器来提取面部表情,自动的识别出 7 种表情^[37]。韩国科学技术院的 Jung 等人提出一种新型的矩形特征^[38],并使用 Adaboost 算法从所有可能矩形中选择矩形特征,取得了比使用 Haar 矩形特征更好的识别效果。但 AdaBoost 方法在小样本情况下的性能并不好^[39]。

2.4 鲁棒的表情识别

与受控环境下的人脸表情识别相比,鲁棒表情识别的研究相对较少。近些年,由于鲁棒人脸检测和鲁棒人脸识别的快速发展,逐渐有一些解决光照、

头部运动、噪声和局部遮挡情况的鲁棒表情识别的研究成果。

施乐帕罗奥多研究中心的 Black 等人、美国斯坦福大学的 Gokturk 等人研究了与视角无关的表情识别^[6, 40],使其对头部运动具有鲁棒性。例如, Gokturk 等人提出了一种鲁棒的与观察角度无关的表情识别方法^[40]。该系统使用一个3维的基于模型的跟踪器,同时鲁棒地提取每一帧中人脸的姿势和外形。

日本工学院大学的 Ebine 等人、日本佳能研究中心的 Matsugu 等人研究了与人无关的鲁棒表情识别^[41-42]。例如, Matsugu 等人描述了一个基于规则的算法^[42],用来识别人脸表情,解决了表情识别中与人无关、平移、旋转和尺度变换等问题。实验结果表明,对10余人的5600个静态图像的微笑表情的识别率达到了97.6%。

英国伦敦大学的 Shan 等人研究了对图像分辨率具有鲁棒性的表情识别^[43]。他们引入了一种低计算量的判别特征空间,在不同精度的图像上实现了鲁棒的表情识别。

更多的研究者则针对面部遮挡或噪声提出了一些表情识别方法。一些研究者使用局部特征进行鲁棒的表情识别。英国斯塔福德郡大学的 Bourel 等人提出了表情特征的局部表征^[44],并对分类器输出进行数据融合,用来对存在面部遮挡的视频序列进行表情识别。后来,他们又提出了基于局部空间几何面部模型和基于状态的面部运动模型的表情表示方法^[45],用于从视频序列鲁棒识别人脸表情。实验显示,其在面部有局部遮挡和特征有不同水平的噪声的情况下,表情的识别率有较少的下降。美国卡内基梅隆大学的 Gross 等人使用鲁棒主成分分析(RPCA)训练具有遮挡的人脸图像来得到活动外观模型(AAM)的人脸图像灰度变化模型,并提出了具有鲁棒性的活动外观模型方法^[46],其对于有遮挡情况下的人脸跟踪取得了很好的识别效果。

一些研究者也使用整体特征进行鲁棒表情识别。希腊亚历士多德大学的 Buciu 等人提出了基于 Gabor 小波的特征提取和两种分类器(最大相关分类器和余弦相似度度量)^[47],获得了对局部遮挡(如眼部遮挡和嘴部遮挡)鲁棒的表情识别。英国伦敦大学学院的 Towner 等人基于主成分分析提出了3种方法^[48],用来对不完整的人脸表情进行重构,并取得了较好的表情识别效果。清华大学的刘

晓旻使用模板匹配的方法检测块状区域内的遮挡情况^[49],对于遮挡部分的局部特征进行了丢弃。但其使用平均脸来进行模板匹配,在人脸结构差异较大的情况下,并不能获得好的识别效果。

基于对局部遮挡的鲁棒性观点,与考虑整个人脸的模型相比,采用局部面部信息的人脸模型具有较多优点。虽然基于局部特征点和局部几何表征的人脸模型对局部遮挡具有鲁棒性,但由于基于局部信息的方法会造成一些模型参数的丢失,因此分类器必须能够处理部分输入数据,并能进行重新训练。

以上研究虽提出了各种方法用来提取特征,以获得对面部遮挡的鲁棒性,但由于没有去除这些遮挡,因此难以避免引入错误。

2.5 精细的表情识别

对于面部表情的变化,存在以下两种表情分类的方法:一种是基于消息评价的方法,即把面部表情直接与情感状态相对应,如 Ekman 提出的6种基本表情^[5];另一种是基于符号表达的方法,即先把面部的各种运动和形变进行细致的编码,再根据表情字典将编码与某种表情对应,或用于描述各种表情。如 Ekman 和 Friesen 提出的面部动作编码系统^[5]。相对于基本表情的识别,对表情单元的识别实现了精细的表情识别。

2.5.1 对表情单元的识别

研究者们通常使用局部特征分析和光流法等运动分析技术来对表情单元进行识别。

卡内基梅隆大学的 Lien 通过区别表情单元(运动单元或混合运动单元)来理解面部肌肉的独立机制^[50]。它考虑了3个上面部表情单元和6个下面部表情单元,并对其进行自动识别和强度测量。这些表情是经常出现的表情,并存在细微差别。

匹兹堡大学的 Cohn 和卡内基梅隆大学的 Kanade 等人使用光流法来识别面部表情的细微变化^[51]。他们通过评价光流的分级算法来自动跟踪选取的面部特征,并对眉毛和嘴巴区域的动作单元和混合动作单元进行识别。

加利福尼亚大学的 Donato 等人采用光流法、主成分分析、独立成分分析、局部特征分析、线性判别分析、Gabor 小波和局部主成分等方法对12个面部单元进行了识别^[52],并对这些方法的识别效果进行了比较。

荷兰代尔夫特工业大学的 Pantic 等人建立了一个集成了各种方法的表情识别专家系统^[53],该系统

可以识别正面和侧面两个视角的 16 个动作单元和 6 种基本表情。

美国卡内基梅隆大学机器人研究所的 Tian 和 Kanade 等人开发了自动人脸分析系统^[54],先基于永久性和暂时性的面部特征分析面部表情,然后将面部表情的细微变化识别为面部动作单元,并分为中性表情、6 个上部动作单元和 10 个下部动作单元。

麻省理工学院媒体实验室的 Kapoor 等人提出了一种全自动的面部动作单元分析系统^[55],该系统先使用红外相机检测瞳孔,然后通过主成分分析提取参数,并用支持向量机来识别上部面部动作单元及其混合动作单元。

微表情是指持续时间不超过 1/4s 的简短表情。微表情经常不自觉地发生,可以表现非故意表达的情感,但大多数人不能够感知自己或他人的微表情。在 Diogenes 研究项目中,心理学家 Ekman 发现,微表情可以用于测谎。对于微表情的自动识别难度很大,目前还未有研究涉及。

2.5.2 对表情强度的识别

对人脸表情强度的识别有助于计算机理解人的情绪及其变化。由于不同个体表达同一种表情存在变化速度、幅度和持续时间上的差别,而且对面部表情强度的感知还依赖于如性别、文化背景和交流语境等诸多因素^[56-57],因此具有较大难度。

表情单元的强度可以从微量 (trace) 变化到最大 (maximum)^[58] (如图 1 所示)。微量指面部的运动刚刚能被注意到;轻微 (slight) 指外观上有一个或更多的变化可以被观察到;显著 (marked/pronounced) 可以由一系列与指定动作单元相关的准则确定,以区别于轻微;极度 (severe/extreme) 由类似于显著与轻微间的差别确定;最大 (maximum) 指具有最大的外观变化。



图 1 强度评价的等级
Fig. 1 Degrees of Intensity Evaluation

日本东京理科大学的 Kobayashi 等人通过神经网络对面部各器官的形变与表情强度之间的关系构建强度模型^[59],将“弱”和“强”的面部特征点位置输入神经网络,由一个输出神经单元给出 6 种基本

表情的强度值,取得了较高的识别率。

日本图像信息科技实验室的 Kimura 和大阪大学的 Yachida 先利用弹性网格模型提取面部表情形变信息,然后通过 K-L 变换将面部的多元运动向量投影到低维空间,再由投影向量的大小估计表情强度的大小,用于对每种表情构建一个强度模型^[60]。

美国匹兹堡大学电子工程系的 Lien 等人基于 FACS 系统进行面部特征提取与强度估计^[61],即先通过特征点跟踪、密度流跟踪、主成分分析、时空域高梯度成分分析等方法提取表情变化信息,然后将图像序列投影在特征空间用于估计表情的类型和强度。

日本大阪大学的 Wang 等人先通过 B 样条曲线构造表情变化模型来表征面部特征运动与表情变化的关系,然后用匹配图标记的方法跟踪面部特征点的运动,并依据面部特征点的位移轨迹来估计高兴、惊讶和生气表情序列的表情强度^[62]。

日本京都大学的 Yang 等人建立的实时表情识别系统^[63]是先将面部运动单元分为上部、中部和下部,然后计算不同特征点的距离变化,并由模糊规则推理人脸各部分的 6 种基本情感强度。

香港中文大学的 Lee 等人先通过提取表情变化特征,从面部运动轨迹中抽取表情强度值^[64];然后分别使用层叠神经网络和支持向量机通过建立表情强度模型来对表情强度值进行预测。其开发的视觉系统可以估计高兴、生气和悲伤等情感的表情强度。

湖南大学的李燕先采用等容特征映射进行表情特征降维,然后从高维的面部特征点运动轨迹中抽取出表示表情强度的 1 维流形^[65],并将支持向量机用于表情强度等级的度量。该论文对特定个体的高兴表情进行了强度度量,但没有对其他情感表情进行强度度量。

表情强度的度量涉及到细微的面部运动,要求面部特征提取具有较高的准确性和鲁棒性,这使得人机交互中对表情强度进行识别的难度较大。

2.6 混合表情识别

混合表情是由多种情绪引起的复杂表情。某一混合表情并不等同于几种基本表情的简单叠加。与基本表情识别相比,混合表情识别对特征提取及识别方法的要求更高,识别难度更大。目前,对混合表情的研究较少,研究者主要采用模糊理论及较为成熟的模式识别方法对混合表情进行识别。

日本东京理科大学的 Kobayashi 等人将混合表

情图像的面部特征点位置输入到神经网络进行训练,对测试混合表情的识别率达到了70%^[66]。对于混合表情的选择标准是,当一种情感表情的强度变弱时,还能够感知到另一种情感表情。神经网络的输出有6个神经元,每个神经元对应一种情感,与获得的最大值神经元对应的情感即为识别结果。但该系统将表情识别结果归为单一情感,这明显不符合人们对混合表情的感觉。

日本国际模糊工程研究实验室 Ralescu 等人首先利用模糊逻辑对表情图像进行语言建模^[67],然后采用了基于模糊逻辑的定性建模对6种表情进行识别,并且对两种混合表情“悲伤且微笑表情”和“微笑且悲伤表情”进行识别。在表情建模中,由于同一表情图像可能被识别为几种表情,并有对应的强度,所以要采用几种表情来定义每一表情图像。

哈尔滨工业大学的金辉和中国科学院的高文提出了一种人脸面部混合表情识别系统^[68],首先把脸部分成各个表情特征区域,分别提取其运动特征,并按时序组成特征序列;然后通过分析不同特征区域所包含的不同表情信息的含义和表情的含量;最后通过概率融合来理解、识别任意时序长度的、复杂的混合表情图像序列。其各种表情的总体识别率达到96.9%。但该系统识别的仍旧是7种基本表情,不能满足人机交互中对混合表情识别的需要。

北京大学的武宇文考虑了基于脸部形状和结构特征的表情模糊性描述,并分析了基于脸部2维形状和结构特征描述表情模糊性的合理性和不足之处^[69]。

天津大学的左坤隆采用AAM提取表情特征,并用人工神经网络方法对提取的特征进行混合表情的分类,取得了稍优于Gabor小波方法的识别结果^[70]。

北京科技大学的Zhao等人提出了一种基于模糊核分类和支持向量机的面部复杂表情识别方法^[71]。

由于人类的情感和表情具有模糊性,因此对混合表情识别进行研究对人机交互中的模糊情感识别具有积极的促进意义。

2.7 非基本表情识别

虽然大多数人脸表情识别集中于情感表情和表情单元的识别,但在人机交互中,还需要从人脸表情中识别其他的一些状态,如疼痛、疲劳等。这些非基本表情的识别在许多人机交互中具有很好的应用

前景。

相对于基本情感表情和表情单元识别而言,非基本表情识别起步较晚。英国剑桥大学的El Kaliouby等人提出了一个能够推理复杂精神状态的系统^[72],该系统能够从表情和头部姿势视频中实时识别同意、不同意、精神集中、感兴趣、思考、不确定等精神状态。北京航空航天大学的薛雨丽、毛峡等人基于AdaBoost方法对包括打哈欠、好奇等9种表情进行了识别^[14]。美国加利福尼亚大学的Littlewort等人提出一个对自发的疼痛表情进行识别的系统^[73]。美国卡内基梅隆大学的Ashraf等人利用活动外观模型来识别疼痛表情^[74]。印度的Saradadevi等人通过对嘴巴的跟踪和对打哈欠的识别来检测驾驶员的疲劳状态^[75]。但是,大多数研究仍局限于对设定表情的识别,缺乏鲁棒性。例如,当对疲劳进行检测时,则可能会受到情感表情的干扰。

3 结 语

从表情识别的研究现状看,有以下几个热点及趋势:(1)提高表情识别的效率和速率,不断有新的表情识别方法和改进方案被提出;(2)解放以往需要人工辅助的工作,以实现表情识别的自动化;(3)从对静态图像转向动态视频表情序列的研究,以充分利用表情的时空信息;(4)随着3维人脸模型在人脸识别中的成功应用,表情识别研究也从2维向3维空间发展;(5)从对基本表情的识别逐渐转向更为复杂的精细表情、混合表情及非基本表情的识别;(6)与人脸跟踪、识别及合成等技术相结合来构建各种实时的表情识别系统。

然而目前开展的人脸表情识别研究仍存在以下几大局限:(1)多数研究仍停留在高兴、悲伤、惊讶、愤怒、嫌恶、害怕等基本表情识别的研究上,但基本表情并不能涵盖人类的主要表情,为了提高人工情感的表达能力,就需要识别更多的表情(如细微表情、混合表情、非基本表情);(2)缺乏对表情识别鲁棒性的研究,由于光照、姿势、遮挡等对面部表情识别有较大影响,并且表情模式也因人而异,因此这些因素都影响到表情识别的实用化。为了提高表情识别的鲁棒性,则需要借鉴人脸识别中成功的经验进行攻关;(3)人脸表情数据的不足。就目前的人脸表情数据库而言,一方面,多数数据库仍为静态灰度图像,不能满足彩色视频分析系统的需要;另一方

面,真实表情难以获得、表情的种类较少、对表情的标定还存在难度,这些都使得少数的表情数据库差强人意。表情数据的匮乏已成为表情识别的瓶颈问题,因此进一步的研究迫切需要更接近生活的基于自然情感的彩色视频表情数据库。

基于以上的分析,本文认为在人脸表情识别领域,尽管基础研究较为丰富,在某些方面已取得可喜的成绩,但与人类的表情识别及情感能力相比,仍有很大差距,主要表现在对精细表情的识别及情感理解能力不足、表情识别的鲁棒性差等。现今,人脸识别系统已经推向商业应用,而表情识别系统仍难以满足实际需要。未来的人机交互中,情感将是不可忽视的因素,而表情又是情感交流的门户,表情识别必将在情感化人机交互中占有重要地位。所以,在当今科学高度发展、智能识别技术日益进步的条件下,需要以勇气与智慧来挑战表情识别的难点问题,并进行创新性研究,以便早日使机器接近人类的情感智能水平。

参考文献 (References)

- 1 Darwin C. The Expression of the Emotions in Animals and Man [M]. London, UK: John Murray, 1872.
- 2 Mehrabian A. Communication without words [J]. Psychology Today, 1968, 2(4): 53-56.
- 3 Picard R W. Affective Computing [M]. Cambridge, MA, USA: MIT Press, 1997.
- 4 Gross R. Face databases [A]. In: Li S, Jain A, eds. Handbook of Face Recognition [M]. New York: Springer, 2005.
- 5 Ekman P, Friesen W. Constants across cultures in the face and emotion [J]. Journal of Personality and Social Psychology, 1971, 17(2): 124-129.
- 6 Black M J, Yacoob Y. Recognizing facial expressions in image sequences using local parameterized models of image motion [J]. International Journal of Computer Vision, 1997, 25(1): 23-48.
- 7 Lyons M, Akamatsu S, Kamachi M, et al. Coding facial expressions with Gabor wavelets [A]. In: Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition [C], Nara, Japan, 1998: 200-205.
- 8 Kanade T, Cohn J, Tian Y L. Comprehensive database for facial expression analysis [A]. In: Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition [C], Grenoble, France, 2000: 46-53.
- 9 Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces: vs. fisherfaces: recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720.
- 10 Sim T, Baker S, Bsat M. The CMU pose, illumination, and expression (PIE) database [A]. In: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition [C], Washington, DC, USA, 2002: 46-51.
- 11 Martinez A M, Benavente R. The AR face database [R]. Technical Report 24, The Computer Vision Center (CVC), Barcelona, Spain, 1998.
- 12 Hwang B W, Roh M C, Lee S W. Performance evaluation of face recognition algorithms on Asian face database [A]. In: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition [C], Seoul, South Korea, 2004: 278-283.
- 13 Wu Dan, Lin Xue-yin. The design and realization of a video database for facial expression analysis [J]. Computer Engineering and Applications, 2004, 40(5): 177-180. [吴丹, 林学茵. 人脸表情视频数据库的设计与实现 [J]. 计算机工程与应用, 2004, 40(5): 177-180.]
- 14 Xue Yu-li, Mao Xia, Zhang Fan. Design and realization of BHU facial expression database [J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(2): 224-228. [薛雨丽, 毛峡, 张帆. BHU 人脸表情数据库的设计与实现 [J]. 北京航空航天大学学报, 2007, 33(2): 224-228.]
- 15 Gao W, Cao B, Shan S, 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.
- 16 Pantic M, Rothkrantz L J M. Automatic analysis of facial expressions: the state of the art [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(12): 1424-1446.
- 17 Fasel B, Luetttin J. Automatic facial expression analysis: a survey [J]. Pattern Recognition, 2003, 36(1): 259-275.
- 18 Essa I, Pentland A. Coding, analysis, interpretation, and recognition of facial expressions [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 757-763.
- 19 Yacoob Y, Davis L. Recognizing human facial expressions from long image sequences using optic flow [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(6): 636-642.
- 20 Cohn J, Zlochower A, Lien J J, et al. Automated face coding: a computer-vision based method of facial expression analysis [A]. In: Proceedings of the 7th European Conference on Facial Expression Measurement and Meaning [C], Salzburg, Austria, 1997: 329-333.
- 21 Zhu Z, Ji Q. Robust pose invariant facial feature detection and tracking in real-time [A]. In: Proceedings of the 18th International Conference on Pattern Recognition [C], Hong Kong, China, 2006: 1092-1095.
- 22 Tong Y, Wang Y, Zhu Z W, et al. Facial feature tracking using a multi-state hierarchical shape model under varying face pose and facial expression [A]. In: Proceedings of the 18th International Conference on Pattern Recognition [C], Hong Kong, China, 2006: 283-286.
- 23 Cristinacce D, Cootes T, Scott I. A multi-stage approach to facial feature detection [A]. In: Proceedings of the 15th British Machine Vision Conference [C], London, UK, 2004: 277-286.
- 24 Tang F Q, Deng B. Facial expression recognition using AAM and

- local facial features [A]. In: Proceedings of the Third International Conference on Natural Computation [C], Haikou, China, 2007: 632-635.
- 25 Hammal Z, Couvreur L, Caplier A, *et al.* Facial expression classification: an approach based on the fusion of facial deformations using the transferable belief model [J]. International Journal of Approximate Reasoning, 2007, **46**(3): 542-567.
- 26 Yu J G, Bhanu B. Evolutionary feature synthesis for facial expression recognition [J]. Pattern Recognition Letters, 2006, **27**(11): 1289-1298.
- 27 Liao S, Fan W, Chunga C S, *et al.* Facial expression recognition using advanced local binary patterns, Tsallis entropies and global appearance features [A]. In: Proceedings of the 2006 IEEE International Conference on Image Processing [C], Atlanta, GA, USA, 2006: 665-668.
- 28 Kotsia I, Zafeiriou S, Pitas L. Texture and shape information fusion for facial expression and facial action unit recognition [J]. Pattern Recognition, 2008, **41**(3): 833-851.
- 29 Zhou X, Huang X, Xu B, *et al.* Real-time facial expression recognition based on boosted embedded hidden markov model [A]. In: Proceedings of the Third International Conference on Image and Graphics [C], Hong Kong, China, 2004: 290-293.
- 30 Miners B W, Basir O A. Dynamic facial expression recognition using fuzzy hidden Markov models [A]. In: Proceedings of the International Conference on Systems, Man and Cybernetics [C], Hawaii, USA, 2005: 1417-1422.
- 31 Tian Y, Takeo K, Cohn J F. Recognizing Facial Actions by Combining Geometric Features and Regional Appearance Patterns [R]. CMU-RI-TR-01-01, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA, 2001.
- 32 Seyedarabi H, Aghagolzadeh A, Khanmohammadi S. Recognition of six basic facial expressions by feature-points tracking using RBF neural network and fuzzy inference system [A]. In: Proceedings of the 2004 IEEE International Conference on Multimedia and Expo [C], Taipei, Taiwan, China, 2004: 1219-1222.
- 33 Ma L, Khorasani K. Facial expression recognition using constructive feedforward neural networks [J]. IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics, 2004, **34**(3): 1588-1595.
- 34 Soyel H, Demirel H. Facial expression recognition using 3D facial feature distances [A]. In: Proceedings of the 4th International Conference on Image Analysis and Recognition [C], Montreal, Canada, 2007: 831-838.
- 35 Littlewort G, Bartlett M S, Fasel I, *et al.* Analysis of machine learning methods for real-time recognition of facial expressions from video [R]. MPLab TR 2003.05, Machine Perception Laboratory Institute for Neural Computation, University of California, San Diego, CA, USA, 2003.
- 36 Xu Q Z, Zhang P Z, Yang L X, *et al.* A facial expression recognition approach based on novel support vector machine tree [A]. In: Proceedings of the 4th International Symposium on Neural Networks [C], Nanjing, China, 2007: 374-381.
- 37 Wang Y B, Ai H Z, Wu B, *et al.* Real time facial expression recognition with adaboost [A]. In: Proceedings of the 17th International Conference on Pattern Recognition [C], Cambridge, UK, 2004: 926-929.
- 38 Jung S U, Kim D H, An K H, *et al.* Efficient rectangle feature extraction for real-time facial expression recognition based on AdaBoost [A]. In: Proceedings of the 2005 Ieee/Rsj International Conference on Intelligent Robots and Systems [C], Edmonton, Canada, 2005: 3634-3639.
- 39 Guo G, Dyer C R. Simultaneous feature selection and classifier training via linear programming: a case study for face expression recognition [A]. In: Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C], Madison, Wisconsin, USA, 2003, **1**: 346-352.
- 40 Gokturk S B, Bouguet J Y, Tomasi C, *et al.* Model-based face tracking for view-independent facial expression recognition [A]. In: Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition [C], Washington, DC, USA, 2002: 287-293.
- 41 Ebine H, Nakamura O. The recognition of facial expressions considering the difference between individuality [J]. Transactions of the Institute of Electrical Engineers of Japan, 1999, **119-C**(4): 474-481.
- 42 Matsugu M, Mori K, Mitari Y, *et al.* Facial expression recognition combined with robust face detection in a convolutional neural network [A]. In: Proceedings of the International Joint Conference on Neural Networks [C], Portland, OR, USA, 2003: 2243-2246.
- 43 Shan C F, Gong S G, Mcowan P W. Robust facial expression recognition using local binary patterns [A]. In: Proceedings of the 2005 International Conference on Image Processing [C], Genoa, Italy, 2005: 2225-2228.
- 44 Bourel F, Chibelushi C C, Low A A. Recognition of facial expressions in the presence of occlusion [A]. In: Proceedings of the 12th British Machine Vision Conference [C], Manchester, UK, 2001: 213-222.
- 45 Bourel F, Chibelushi C C, Low A A. Robust facial expression recognition using a state-based model of spatially-localised facial dynamics [A]. In: Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition [C], Washington, DC, USA, 2002: 113-118.
- 46 Gross R, Matthews I, Baker S. Active appearance models with occlusion [J]. Image and Vision Computing, 2006, **24**(6): 593-604.
- 47 Buciu I, Kotsia I, Pitas I. Facial expression analysis under partial occlusion [A]. In: Proceedings of the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing [C], Philadelphia, PA, USA, 2005, **V**: 453-456.
- 48 Towner H, Slater M. Reconstruction and recognition of occluded facial expressions using PCA [A]. In: Proceedings of the Second International Conference on Affective Computing and Intelligent Interaction [C], Lisbon, Portugal, 2007: 36-47.
- 49 Liu Xiao-min. Facial Expression Recognition in Human Computer Interaction [D]. Beijing: Tsinghua University, 2006. [刘晓旻. 人机交互界面中的人脸表情识别 [D]. 北京: 清华大学, 2006.]

- 50 Lien J J J. Automatic Recognition of Facial Expression Using Hidden Markov Models and Estimation of Expression Intensity [D]. Pittsburgh, PA, USA: Carnegie Mellon University, 1998.
- 51 Cohn J F, Zlochower A J, Lien J J, *et al.* Feature-point tracking by optical flow discriminates subtle differences in facial expression [A]. In: Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition [C], Nara, Japan, 1998: 396-401.
- 52 Donato G, Bartlett M S, C. Hager J, *et al.* Classifying facial actions [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, **21**(10): 974-989.
- 53 Pantic M, Rothkrantz L J M. Expert system for automatic analysis of facial expressions [J]. Image and Vision Computing, 2000, **18**(11): 881-905.
- 54 Tian Y I, Kanade T, Cohn J F. Recognizing action units for facial expression analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, **23**(2): 97-115.
- 55 Kapoor A, Qi Y, Picard R W. Fully automatic upper facial action recognition [A]. In: Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures [C], Nice, France, 2003: 195-202.
- 56 Matsumoto D, Kasri F, Kookan K. American-Japanese cultural differences in judgements of expression intensity and subjective experience [J]. Cognition and Emotion, 1999, **13**(2): 201-218.
- 57 Niedenthal P M, Halberstadt J B, Margolin J, *et al.* Emotional state and the detection of change in facial expression of emotion [J]. European Journal of Social Psychology, 2000, **30**(2): 211-222.
- 58 Ghent J, Mcdonald J. Photo-realistic facial expression synthesis [J]. Image and Vision Computing, 2005, **23**(12): 1041-1050.
- 59 Kobayashi H, Hara F. Recognition of six basic facial expressions and their strength by neural network [A]. In: Proceedings of the IEEE International Workshop on Robot and Human Communication [C], Tokyo, Japan, 1992: 381-386.
- 60 Kimura S, Yachida M. Facial expression recognition and its degree estimation [A]. In: Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C], San Juan, Puerto Rico, 1997: 295-300.
- 61 Lien J J J, Kanade T, Cohn J F, *et al.* Subtly different facial expression recognition and expression intensity estimation [A]. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C], Santa Barbara, CA, USA, 1998: 853-859.
- 62 Wang M, Iwai Y, Yachida M. Recognizing degree of continuous facial expression change [A]. In: Proceedings of the 14th International Conference on Pattern Recognition [C], Brisbane, Australia, 1998: 1188-1190.
- 63 Yang D, Kunihiro T, Shimoda H, *et al.* A study of real-time image processing method for treating human emotion by facial expression [A]. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics [C], Tokyo, Japan, 1999: 360-364.
- 64 Lee K K, Xu Y. Real-time estimation of facial expression intensity [A]. In: Proceedings of the IEEE International Conference on Robotics & Automations [C], Taipei, Taiwan, China, 2003: 2567-2572.
- 65 Li Yan. Research on Methods for Facial Expression Intensity Measurement Based on Video [D]. Changsha: Hunan University, 2006. [李燕. 基于视频的面部表情强度度量方法研究 [D]. 长沙: 湖南大学, 2006.]
- 66 Kobayashi H, Hara F. Recognition of mixed facial expressions by neural network [A]. In: Proceedings of the IEEE International Workshop on Robot and Human Communication [C], Tokyo, Japan, 1992: 387-391.
- 67 Ralescu A, Iwamoto H. Recognition of and reasoning about facial expressions using fuzzy logic [A]. In: Proceedings of the 2nd IEEE International Workshop on Robot and Human Communication [C], Tokyo, Japan, 1993: 259-264.
- 68 Jin Hui, Gao Wen. The human facial combined expression recognition system [J]. Chinese Journal of Computers, 2000, **23**(6): 602-608. [金辉, 高文. 人脸面部混合表情识别系统 [J]. 计算机学报, 2000, **23**(6): 602-608.]
- 69 Wu Yu-wen. Studies on Facial Expression Recognition Based on 2D Shape and Structural Features [D]. Beijing: Peking University, 2005. [武宇文. 基于脸部二维形状与结构特征的表情识别研究 [D]. 北京: 北京大学, 2005.]
- 70 Zuo Kun-long. A Study of Automatic Facial Expression Analysis and Recognition System [D]. Tianjin: Tianjin University, 2004. [左坤隆. 人脸表情自动分析与识别系统的研究 [D]. 天津: 天津大学, 2004.]
- 71 Zhao H, Wang Z L, Men J H. Facial complex expression recognition based on fuzzy kernel clustering and support vector machines [A]. In: Proceedings of the Third International Conference on Natural Computation [C], Haikou, China, 2007: 562-566.
- 72 El Kaliouby R, Robinson P. Real-time inference of complex mental states from facial expressions and head gestures [A]. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop [C], Washington, DC, USA, 2004: 154-159.
- 73 Littlewort G C, Bartlett M S, Lee K. Faces of pain: automated measurement of spontaneous facial expressions of genuine and posed pain [A]. In: Proceedings of the 9th International Conference on Multimodal Interfaces [C], Nagoya, Aichi, Japan, 2007: 15-21.
- 74 Ashraf A B, Lucey S, Cohn J F, *et al.* The painful face: pain expression recognition using active appearance models [A]. In: Proceedings of the 9th International Conference on Multimodal Interfaces [C], Nagoya, Aichi, Japan, 2007: 9-14.
- 75 Saradadevi M, Bajaj P. Driver fatigue detection using mouth and yawning analysis [J]. International Journal of Computer Science and Network Security, 2008, **8**(6): 183-188.