

学科交叉视角下的情感识别研究进展

程 静¹ 刘光远²

(西南大学计算机与信息科学学院 重庆 400715)¹ (西南大学电子信息工程学院 重庆 400715)²

摘 要 人机交互环境下的情感识别是一个涉及心理学、计算机科学等的多学科交叉领域。试图从多学科交叉的角度介绍情感识别所需的心理学理论基础,从而为工程技术人员提供一定的理论背景。之后,根据待处理原始数据的不同,简介了构建此类系统现有的方法和技术;最后简述了当前研究工作的不足和进一步发展的方向,并对全文进行了小结。

关键词 情感识别,多学科交叉视角

中图分类号 TP399 **文献标识码** A

Research Progress of Affect Detection in Interdisciplinary Perspective

CHENG Jing¹ LIU Guang-yuan²

(School of Computer and Information Science, Southwest University, Chongqing 400715, China)¹

(School of Electronic and Information Engineering, Southwest University, Chongqing 400715, China)²

Abstract Affective detection in human-computer interaction environment is an interdisciplinary field which is related to psychology, computer science, etc. The basic affective theories were presented in the paper, which provide the engineer with the background of psychology knowledge. Then according to the differences of the original data, the existing methods and techniques were introduced. Finally a summary was showed and the insufficiency and development of recent research were presented.

Keywords Affect detection, Interdisciplinary perspective

1 引言

随着技术的发展,计算机已经成为现代生活不可或缺的一部分,成为人们工作、学习、娱乐的助手和伴侣。过去那种“以计算机为中心,人来适应计算机”的人机交互方式已经不能够适应社会发展的需要,人们越来越期待“以人为中心”的和谐的人机交互环境。同时,科学研究的有力证据表明,人机交互天生就是自然的、和谐的,遵从人际交互的基本原则。情感是人际交互的天然组成部分,具有重要作用,因而也将成为人机交互过程中不可或缺的关键要素。1997年,美国MIT实验室的Picard在其具有里程碑意义的专著《Affective Computing》^[1]中提出“情感计算”这一术语,其根本宗旨就是要建立能够主动识别和理解人类情感,并能对人类情感进行正确反馈的人机交互环境。之后,情感计算相关领域的研究得到前所未有的快速发展。众多学术团体和企业机构都投入至情感计算的研究领域。2005年及2009年,第一届和第二届“情感计算与智能交互学术会议”分别在北京和武汉召开。2010年,国际学术组织IEEE创办了情感计算专刊《Transactions of Affective Computing》,同时出版发行^[2,3]了大量的相关书籍。这些学术会议、期刊与书籍向人们介绍了人机交互

环境下情感计算的研究成果与进展。

Picard在《Affective Computing》一书中介绍了3种类型的情感计算应用系统:1)识别用户情感的应用系统;2)表达用户情感的应用系统;3)能够与用户进行自由情感交流的应用系统。其中,情感识别是这3类应用系统的基础与关键。试想计算机若不能识别用户的情感,如何能够对用户情感进行反馈和交流呢?同时,情感识别也是一项非常具有挑战性的任务,因为用户所处的情绪状态与用户个体的主观意识密切相关,较难直接测量和评价;另外,对待同一事物,不同个体所体验到的情绪会有很大差异。情感识别是一个涉及心理学、认知科学、计算机科学、生理学等多学科交叉的研究领域。

本文将从心理学和计算机科学的视角分别对人机交互过程中的情感识别系统所涉及的基础理论、数据、方法和技术进行简单介绍,以期让更多的工程技术人员不仅将视角停留在技术的层面,而是更多地关注技术背后的各种情绪心理学的理论基础,从而为更加科学地构建情感识别系统提供理论上的依据。本文第2节介绍情感识别所涉及的心理学理论基础;第3节介绍情绪识别系统常用的数据、方法和技术;第4节提出当前尚未解决的问题,从而展示未来的发展方向;最后小结全文。

到稿日期:2011-07-01 返修日期:2011-11-20 本文受教育部科学技术研究重大项目(311032),西南大学基本科研业务费专项资金项目(XDJK2009C027)资助。

程 静(1977—),女,博士生,讲师,主要研究方向为情感计算,E-mail:ejcat@swu.edu.cn;刘光远(1969—),男,教授,博士生导师,主要研究方向为情感计算。

2 情绪识别的相关心理学理论

早在 100 多年以前,心理学家、生理学家、哲学家们就展开了对情绪的相关研究,并形成了众多的情绪理论。在此,仅将其中与情感识别相关的情绪理论进行简单介绍。

2.1 关于表情的情绪理论

Darwin 于 1872 年在其论著《人和动物的表情》^[4]一书中就提及,人与动物的某些面部表情及身体动作表情有着许多相似之处。由此推知人类的某些外显情绪表达从动物类似的情绪表达中进化而来,人类情绪表达的原始方式本身具有某些生存价值的遗传模式。Ekman^[5]在对人类的跨文化面部表情的研究中发现,不同文化背景下的人类存在着一些共同的基本情绪。另外,从婴儿的外部表情中,人们也发现了一些天生固有的表情。

以上这些研究说明了两个问题:(1)人类存在着不依赖于地域文化差异的基本情绪,这些基本情绪是人类在漫长适者生存道路上不断进化的结果,具有遗传的基础,即使在刚出生不久的婴儿身上也可以看到这些基本情绪的影子。至于究竟存在哪些基本情绪,不同的情绪心理学家给出了不同的定义。在情感识别领域,工程技术人员往往将 Ekman 关于 6 类基本情绪的定义作为待识别的目标情感,它们分别是快乐、悲伤、愤怒、惊奇、厌恶、恐惧。(2)人类的面部表情和身体姿势表情是人类心灵的窗口,通过分析个体的面部表情和身体姿势表情,可以识别隐藏于内心深处的真实情感。因此,在情感识别领域有大量研究工作集中在对面部表情和身体姿势表情的分析上。本文第 3.1 节和 3.3 节将分别介绍这两个领域相关的研究方法和技术。

2.2 詹姆斯-兰格的外周情绪理论

心理学中最早对情绪进行系统理论阐述的是詹姆斯-兰格的外周情绪理论,他们的观点是对外周生理反应(如心跳加速、身体发抖、血管收缩、气管痉挛等)的觉知就是情绪。由此可以推出,每一种情绪都对应着不同的生理唤醒模式。目前这一观点仍然存在着争议^[6,7],但是在情感识别领域已经得到广泛的应用,即人们可以通过采集人体的各种生理信号,如呼吸、脉搏、心率、皮肤电导等,来推测个体当前所处的情绪状态。本文第 3.4 节将对该领域所采用的方法和技术进行论述。

2.3 情绪的认知理论

Arnold 的“认知评价”理论是情绪认知理论中最具代表性的理论之一。她认为,情绪产生的基本过程是刺激-评价-情绪。即情绪体验是个体在对刺激事件觉知,并进行认知评价后,才得以产生的。她认为,刺激事件并不直接决定情绪性质,对同一刺激,个体认知对它的评价不同,会产生不同的情绪反应。比如,在森林里遇到一只熊和在动物园里看到一只熊所产生的情绪体验是不同的。Schachter 提出了“认知归因”理论来解释情绪的产生。他认为情绪既来自对自生理反应的反馈,也来自对引起这些生理反应的情境的认知归因^[8]。

情绪的认知理论尽管能够解释很多情绪现象,但是仍然存在以下不足:(1)用以区分不同情绪差异的最少数量的评价标准是什么;(2)如何解释那些根本无需进行认知评价的非常简单的情绪反应;(3)认知理论仅仅将情绪作为心理活动的副产品,而忽视了情绪的功能与作用。

尽管存在以上缺陷,但不妨碍由该理论衍生出相应的计算模型。Scherer 等人开发的专家系统 GENSE^[9]就是依据情绪的认知理论构造出来的。GENSE 的核心部件是一个将评价映射为不同情感的知识库,该知识库中存储了大量的启发式规则。这些规则表明了系统能够识别哪些情感,以及如何将对事件的评价映射至情感上。该系统成功地预测了 286 个情感片段的目标情感,其预测精度可以达到 77.9%。

2.4 情绪神经科学的相关理论

在过去的 20 年里,神经科学对情绪研究做出了巨大贡献。通过新技术、新方法的使用,如脑成像法(FMRI)、损伤法、电生理学方法等,人们不断加深了情绪与其对应的神经机制的认知和理解,从而为理解情绪的脑神经机制、特定精神疾病的病理学基础等提供了新视角^[10]。

大量情绪神经科学的证据表明,情绪对应的脑区是皮质下的组织和边缘系统区域,如扣带回、海马、杏仁核、下丘脑等。扣带回是有机体的警戒部位,能引起有机体的警觉和注意;海马则在情绪行为的背景调节中其中关键作用;杏仁核对愤怒和恐惧情绪起着关键作用;下丘脑、边缘系统和丘脑等部位存在着“奖励”中枢和“惩罚”中枢。同时还有大量证据显示,认知功能和情绪功能有许多重叠的脑区^[10]。这就表明包含了思考、目标评价、原因推理、记忆编码和检索的认知过程中必然会产生情绪体验。在人机交互系统的应用中,这一证据说明了仅仅拥有逻辑运算功能的系统存在着严重不足,人机交互环境中应该建立情感计算部件。

情感神经科学提供的证据表明,很多情感过程不一定能被意识所察觉^[11]。这个结论对情绪识别的研究者也是很有启发意义的。既然很多情绪体验不能被意识所察觉,那么仅仅依靠被试“自我报告”方法进行情绪研究是远远不够的,还需借助其他手段,如生理信号的监测、脑电的跟踪分析等,以辅助人类识别情感。

3 情感识别的数据、方法和技术

上一节简单介绍了情绪识别中最常用的情绪理论,它们能够为工程技术人员实际工作提供参考。本节将根据所处理的原始信号不同,介绍建立情感识别系统常用的方法和技术。这里将分别介绍基于面部表情、语音、身段表情、生理信号、文本和多模态的情感识别技术。

3.1 基于面部表情的情感识别

人脸面部表情是传播人类情感信息与协调人际关系的重要方式。据心理学家 J. A. Russe 的研究,在人类的日常交流中,通过语言传递的信息仅占信息总量的 7%,而通过面部表情传递的信息却达到信息总量的 55%。通过识别人脸面部表情可以较准确地识别对方的态度和内心世界。

20 世纪 70 年代,美国心理学家 Ekman 和 Friesen 在现代面部表情识别方面做了大量开创性工作,并为后来计算机自动进行表情识别奠定了坚实基础。1971 年,Ekman 定义了人类 6 种基本表情:高兴、愤怒、惊奇、恐惧、厌恶、悲伤;1978 年,他们开发了面部动作编码系统 FACS(Facial Action Coding System)^[12],该系统将人脸划分为若干运动单元 AU(Action Unit),AU 显示了人脸运动和表情之间的关系。

由于手工识别 AU 需要花费大量时间,尤其是在对动态视频的处理上,因此人们开始研究计算机自动进行表情识别。

1991年, K. Mase 和 A. Pentland^[13] 首先使用光流法来检测 FACS 中的运动单元 AU, 该系统可以识别高兴、生气、厌恶和惊奇 4 种表情, 识别率接近 80%。1997 年, Essa 等人^[14] 提出了基于视频的动态表情描述方法 FACS+。FACS+ 充分考虑了时间和空间变化特性, 可用于分析动态视频中的脸部数据, 克服了 FACS 的不足。2000 年, 我国的高文、金辉^[15] 提出转换 FACS 的方法, 它能够将运动单元转化为基于物理结构和肌肉模型的运动特征向量序列, 显著提高识别精度。2003 年, 北京科技大学的王志良教授^[16] 领导的团队将面部表情识别方法应用于机器人的情感控制研究中。东南大学的郑文明博士^[17] 在面部表情识别方面, 提出了基于核典型相关分析、偏最小二乘回归等多种识别方法, 并开发了自动面部表情识别系统。现在面部表情识别在情感识别领域已经得到了较快速发展。

3.2 基于语音信号的情感识别

语音是人类自然的交流手段, 除了自身所包含的语义信息外, 还携带有丰富的感情信息。随着当前计算机技术的发展, 从语音中提取情感信息的研究逐渐成为研究热点。早在 1972 年, Williams 发现人的情感变化对语言的基音轮廓有很大影响, 这是国外最早的关于语音情感的研究之一。之后, 1990 年, 美国 MIT 媒体实验室开发了一个“情感编辑器”, 它能够从人体的生理信号、语音信号、面部表情中识别出情感^[18]。1996 年, 日本东京的 Seikei 提出情感空间的概念, 并建立了语音情感模型。2000 年, Maribor 大学的 Vladimir Hozjan 等研究了基于多种语言的语音情感识别器^[19]。2009 年, 日本产业技术综合研究所开发了一个具有外显表情的机器人“HRP-4C”, 她能够识别主人语音信号中的情感, 并进行适当反馈^[20]。

语音信号的情感处理涉及不同语种间的差异, 发展也不尽相同, 我国的语音情感识别研究起步较晚, 发展相对滞后。2000 年, 东南大学的赵力等人进行一系列从语音信号中提取情感信息的研究^[21]; 2003 年, 北京科技大学信息工程学院也进行了基于 BDI Agent 的情感机器人语音识别研究; 另外, 微软亚洲研究院、台湾的一些大学和研究机构也在进行这方面的研究工作^[22]。

语音情感识别本质上是一个模式识别问题, 包含 3 个基本步骤: 预处理、特征提取和情感分类, 其中特征提取阶段是关键, 选择哪些特征作为情感识别的主要特征, 对后续情感分类效果影响很大。研究者根据心理学和语言学关于语音和韵律的研究成果, 采用语音的韵律特征作为语音情感识别的主要特征, 它们包括语音持续时间、基音频率(如平均基音频率、最大基音频率等)、语音信号的能量(如语音短时能量变化率等)、语音信号的振幅(如最大振幅、振幅平均变化率等)^[23, 24]。另有文献提出了语音信号的非线性特征 Teager 能量算子(TEO)的概念^[25], 而文献^[26]则将 TEO 应用于带噪声语音情感的识别。实验结果证明上述特征具有较高鲁棒性。语音的正弦模型也已在多个语音处理领域获得了应用。在这种模型中, 语音信号被假设可以由一组不同频率、幅度和相位的正弦波之和表示; 文献^[27]基于上述正弦模型分别研究了幅度特征、频率特征以及相位特征与情感的变化之间的关系。

3.3 基于身体姿势的情感识别

面部表情和身体姿势表情都可以作为人类情感识别的处

理对象。达尔文在《人与动物表情》一书中用相当篇幅研究人类的身体姿势表情, 然而工程领域实现的相关识别系统却相对较少。Mota 和 Picard 首次进行了这方面的工作^[28], 他们通过一套身体压力评价系统(BPMS)推测用户在学习任务下的情绪状态。BPMS 是一个安装在学习者座椅上的薄膜压力坐垫, 它能够感知学习者身体姿势模式, 采用神经网络进行情感的分类识别, 识别精度达到了 87.6%。D'Mello 和 Graesser 进一步扩展了 Mota 和 Picard 的工作^[29], 他们从学习者的身体姿势中检测出了厌烦、困惑、沮丧、高兴等情绪, 并分别达到了 73%, 72%, 83%, 70% 的识别精度。文献^[30]讨论了钢琴演奏者的身体姿势表情, 演奏者要求用不同的情绪演奏同一个乐谱, 研究发现通过身体姿势能够识别不同情感, 其中头部运动的速度和方向是识别身体姿势情感的重要特征; 文献^[31]讨论了舞蹈者的身体姿势表情; 文献^[32]提出了一种通过身体姿势序列识别情绪的方法。

以身体姿势作为情感识别输入具备其他情感识别方式所不具备的优点^[33-35]: 1)能够在较远距离识别用户的情感; 2)能够同时监测多个用户的情感状态, 如群体情绪; 3)面部表情或者语音表情更容易掩饰和隐藏, 而身体姿势表情往往不易隐藏, 能够更加真实反应用户的情绪状态。目前这种情感识别方式已经开始受到研究者的关注。

3.4 基于生理信号的情感识别

詹姆斯-兰格的外周情绪理论为基于人体生理信号的情感识别提供了理论基础, 该理论认为每种基本情绪都对应着不同的生理唤醒模式。至今这一观点仍存在着广泛的争议^[6, 7], 但却得到了大量研究结论的支持^[36-38]。越来越多的工程技术人员以此为基础进行了基于生理信号的情感识别工作, 并获得可喜的成果。

美国 MIT 媒体实验室的 Picard 等人^[39]以图片诱发愤怒、厌恶、悲伤等 7 种情感, 用 5 种生理信号的 40 种特征对 8 类情感进行识别, 识别率高达 82.5%。法国的 Olivier Villon 等人^[40]采用 40 个被试, 讨论了被试情感激发的依赖和非依赖问题。韩国的 Kim 等人^[41]以音乐唤起被试悲伤、高兴、愤怒和愉悦 4 种情感, 用 4 种生理信号(心电、肌电、皮肤电导和呼吸)建立了 3 个被试的 90 组样本数据, 针对用户依赖模型识别率高达 95%。Lee MH 等^[42]从心率信号提取特征, 应用 Personal Digital Assistance(PDA)技术对人们的压力状态进行监视, 也取得了较好的成果。德国奥森堡大学的多媒体与信号处理实验室提取了诸如心电、皮肤电导等生理信号特征, 并比较了 SBS、SFS、ANOVA 等特征选择方法, 以及这些方法与 KNN、LDF、MLP 不同分类器结合进行情感识别的效果^[43]。文献^[44]以脑电为研究对象, 采用一种称为“混合自适应”(HAF)的特征提取算法和高阶交叉(HOCs)的特征选择方法进行情感识别, 识别率可以达到 85.17%。

国内基于生理信号的情感识别也获得了大量的研究成果。中国科学院心理研究所的李建平等^[45]采用电影片段作为情绪诱发素材, 以 92 名大学生作为被试, 测量他们在观看电影片段过程中多个反映自主神经生理指标(如呼吸频率、血压、心率等)的变化情感, 结果发现 5 种基本情感均有其自主神经反应的特异性模式。江苏大学^[46]从 55 组生理信号数据中提取出特征, 进行单因素方差分析(ANOVA), 发现了 7 种特征在不同情感上存在显著差异, 然后将这 7 种特征作为 Fisher、

Bayesian等分类器的输入,从而获得了较好的情感识别率。西南大学情感计算实验室在此领域进行大量的研究。温万惠等^[38]采用随机矩阵理论对8路生理信号进行相关性分析,发现皮肤电导、心率、心电和呼吸信号中的确存在可靠情感生理反应模式,从而为基于生理信号的情感识别奠定了可靠性基础。刘光远等针对多生理信号情感识别过程中的特征冗余以及在大样本数据下传统特征降维算法效率普遍不高的现状,采用基于智能搜索的组合优化思想进行生理信号情感特征选择,分别尝试了蚁群算法^[47]、免疫机制^[48]、遗传算法^[49]、粒子群算法^[50]等方法,以筛选出优化的特征组合。

3.5 基于文本的情感识别

文字能够表现人类的情感。古往今来,文人骚客们正是通过文字来抒发内心的情感,通过计算机技术自动识别出文本中所隐含的情感,却是近几年的事情。随着互联网的发展,Web文本成为人们获取信息、发表观点和交流情感的重要平台。如何从这些Web文本中挖掘出各种意见、观点、态度,获取人们的情感倾向,即所谓的“观点挖掘”、“意见挖掘”或“文本的情感识别”,已经成为网络管理者、政府机构、各商务团体较为关注的问题之一,能够广泛使用在网络评论分类、产品声誉分析、舆情跟踪与控制等领域。

文本情感识别研究始于20世纪90年代末期,最初的工作主要集中在对词语粒度级的文本识别其情感(褒/贬)。Hatzivassiloglou和Mckcown^[51]从大语料库华尔街日报中挖掘出大量的形容词性的评价词语;Wiebe等人^[52]采用相似度分布的词聚类方法进行了类似工作;Riloff等人^[53]手工选取一些种子评价词语,之后使用迭代的方法获取了名词词性的评价词语;Turney等^[54]提出了点互信息方法判别某词是否是评价词语,该方法适用于各种词性的评价词语的识别,但是较依赖种子褒/贬词语集合;Kim等学者利用词典(如WordNet)将手工采集的种子评价词语进行扩展来获取大量的评价词语^[55,56],这种方法简单易行,但是较依赖于种子评价词语的个数和质量。之后,随着研究工作的推进,文本情感识别的粒度由词语级扩展至篇章级。Das和Chen^[57]开发了一种从股票留言板中挖掘投资者情绪的方法;Denecke^[58]采用基于SentiWordNet情感词汇抽取,采用机器学习方法,在电影评论上进行试验,以确定文本的褒贬极性;Polpinij和Ghose^[59]通过词典,以名词、动词为基础建立本体,对在线顾客评论进行分析。

国内对文本情感识别的研究起步较晚,近几年它也逐渐成为大量学者关注的焦点。王素格和魏英杰^[60]研究了停用词表对中文文本情感识别的影响;徐琳宏、林鸿飞等^[61]采用手工分类和自动分类相结合办法构造情感词汇本体库,以此为基础来实现对中文文本的情感识别。文献^[62]则对当前文本情感挖掘中常用的方法和技术进行了综述。

3.6 多模态的情感识别

“多模态情感识别”就是指利用人脸表情、语音、生理信号等多个通道信息进行情感识别,从而达到提高情感识别效果的目的。尽管“多模态”的方法被认为是克服“单模态”方法识别率低的重要手段之一,但是目前真正实现的多模态情感识别系统还较少。Scherer和Ellgring^[63]采集了面部表情、语音和身体姿势3个通道的特征,用以识别14种情感,识别率达到了79%。如果单独采用面部表情特征和语音特征进行单模态情感识别,识别率分别为52.2%,52.5%。可见,多模态

方法能显著改善情感识别系统的识别率。Castellano等^[64]通过检测面部表情、语音和手势3个通道信息,来识别8种情感,多模态的情感识别率达到了78.3%,而单独使用面部表情、语音和手势信息其识别率分别为48.3%,67.1%,57.1%。Kapoor和Picard^[65]在儿童游戏的环境下,根据面部表情特征、身体姿势以及上下文信息(如游戏当前的难度)来推导儿童的兴趣水平,多模态的情感识别精度是86%。若分别采用面部表情、身体姿势和上下文信息进行单模态情感识别,识别精度分别为67%,82%,57%。黄程韦和赵力等人^[66]通过噪音刺激和观看电影片段诱发被试的烦躁情绪和喜悦情绪,采集了相应情感状态下语音信号和心电信号,基于心电信号和语音信号的单模态情感识别率分别为71%和80%,而二者的多模态识别率则达到90%以上。李爱军等人^[67]研究了音频和视频2个模态信息对跨文化情绪感知的作用,结果发现音频和视频信息在跨文化的情感识别中起着不同作用,面部视频信息的加入有助于情感识别。另外,跨文化的研究表明,情感识别存在普遍的心理基础。

4 研究展望

当前情感识别得到了较为快速的发展,展望未来,情感识别中尚有许多问题需要探索 and 解决。未来的研究方向可能会集中在以下几个方面。

1)在数据采集方面,可以考虑进一步扩大当前数据采集的范围。其中用户对自身情绪的“自我报告”将是一个特别值得关注的数据库,可以通过设计详细的问卷来获取用户情绪主观体验的自我报告^[68]。另外,情感事件总是发生在一定的上下文环境中,现有的情感识别技术忽视了对上下文的监测,通过获取用户情绪的上下文环境来辅助情感识别将成为情感识别进一步的研究方向^[69,70]。

2)在情绪识别的基础理论方面,还应该关注除情绪心理学之外更为广泛的心理学理论,如社会心理学^[71,72]、发展心理学^[73]等。用户的情感状态不是一个独立的系统,必然会受到用户的动机、目标、愿望的影响,因此文献^[68]建议进一步拓展情感识别的任务范围,将情感识别拓展为用户精神状态的识别。另外,在情绪识别领域,群体情绪对个体情绪影响的研究完全被忽视,而事实上用户的情绪往往不是孤立产生的。研究在社会环境下群体情绪如何对个体情绪施加影响,从而影响个体与计算机的交互,是情感识别另一个亟待解决问题。以上两方面问题都需要涉及大量社会心理学的理论知识。

3)在情绪识别的技术层面,当前大多数的情感识别系统尽可能获取用户的各种特征,带来的最直接问题就是系统中存在高维特征,导致维数灾难,从而显著降低了分类器的效率。如何针对不同的数据库,采用合适的特征降维方法,选择数量适当且分类效果好的特征,至今仍在广泛的研究中。另外,高效、稳定的情感识别算法仍将是未来研究的热点。将现有的几种主要算法各取所长、集成使用,是解决该问题的有效途径。最后在分类器的选择上,如何根据特征选择算法选取合适的分类器,也是目前研究较热的课题。

4)对系统识别结果的处理方面,当前大部分系统是将情感识别结果标志为几种离散的情感,如愤怒、悲伤、高兴等。但也有系统将情感识别结果标志为“效价-激活”维度上的一个点^[74],究竟选择哪种方式更为适合,至今仍是一个开放的问题。

结束语 情感识别是一个涉及心理学、生理学、认知科学、计算机科学的多学科交叉的研究领域,本文试图从学科交叉的角度对情感识别系统中所涉及的基本理论、数据、方法和技术进行较为全面的介绍。情绪心理学为人们构造情感识别系统提供了理论基础,而计算机科学则辅助人们搜集情感数据,并为系统实现提供了方法和技术。由于情感是一个包含内在主观体验、外显表情(如面部表情、身段表情、语音表情等)和生理激活的多成分、多维量的心理活动,因此本文分别从面部表情、语音、身体姿势、生理信号、文本的角度介绍了实现情感识别系统常用的基本理论、实现技术和方法。最后本文对亟待解决的研究问题进行了简要分析和介绍。

参 考 文 献

[1] Picard R W. *Affective Computing* [M]. London: The MIT Press, 1997

[2] Gkay D, Yildirim G. *Affective Computing and Interaction; Psychological, Cognitive and Neuroscientific Perspectives* [M]. Pennsylvania: Information Science, 2010

[3] Calvo R A, D'Mello S. *New Perspectives on Affect and Learning Technologies* [M]. German: Springer, 2010

[4] Darwin C. *The Expression of the Emotions in Man and Animals* [M]. England: John Murray, 1872

[5] Ekman P, Friesen V. Constants across cultures in the face and emotion [J]. *Journal of Personality and Social Psychology*, 1971, 17(2):124-129

[6] 李建平, 张平, 王丽芳, 等. 5种基本情绪自主神经反应模式特异性的实验研究[J]. *中国行为医学科学*, 2005, 14(3):257-259

[7] Christie C, Friedman H. Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach [J]. *International Journal of Psychophysiology*, 2004, 51(2):143-153

[8] 黄希庭. *心理学导论(第二版)* [M]. 北京: 人民教育出版社, 2007:487-495

[9] Scherer K R. Studying the Emotion-Antecedent Appraisal Process: An Expert System Approach [J]. *Cognition and Emotion*, 1993, 7(4):325-355

[10] Dalgleish T, Dunn B, Mobbs D. *Affective Neuroscience: Past, Present, and Future* [J]. *Emotion Review*, 2009, 1(4):355-368

[11] Ohman A, Soares J. Emotional Conditioning to Masked Stimuli: Expectancies for Aversive Outcomes Following Nonrecognized Fear-Relevant Stimuli [J]. *Journal of Experimental Psychology: General*, 1998, 127(1):69-82

[12] Ekman P, Friesen W. *Facial Action Coding System: A Technique for the Measurement of Facial Movement* [M]. California: Consulting Psychologists Press, 1978

[13] Mase K, Pentland A. Recognition of Facial Expression from Optical Flow [J]. *IEICE Transactions*, 1991, E74(10):3474-3483

[14] Essa I A, Pentland A P. Coding, analysis, interpretation, and recognition of facial expressions [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7):757-763

[15] 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

[16] 王志良, 陈峰军, 薛为民. 人脸表情识别方法综述[J]. *计算机应用与软件*, 2003, 12(1):63-66

[17] 郑文明. 基于核函数的判别分析研究[D]. 南京: 东南大学, 2004

[18] Chennoukh S, Gerrits A, Miet G, et al. Speech enhancement via frequency bandwidth extension using line spectral frequencies

[C]// *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. Salt Lake City: IEEE Computer Society, 2001:666-668

[19] 赵腊生, 张强, 魏小鹏. 语音情感识别研究进展[J]. *计算机应用研究*, 2009, 26(2):428-432

[20] 方恨少. 日本新型女机器人 HRP-4C 会说话表情丰富 [EB/OL]. <http://info.china.alibaba.com/news/detail/v5000441-d1004571420.html>

[21] 赵力, 钱向民, 邹采荣, 等. 语音信号中的情感识别研究[J]. *软件学报*, 2001, 12(7):1050-1055

[22] 林奕琳, 韦岗, 杨康才. 语音情感识别的研究进展[J]. *电路与系统学报*, 2007, 12(1):90-98

[23] Hyunk H, Kme H, Kwak Y K. Emotional feature extraction based on phoneme information for speech emotion recognition [C]// *Proceedings of the 16th IEEE International Symposium on Robot & Human Interactive Communication*. ROMAN: IEEE Computer Society, 2007:802-806

[24] Morrison D, Wang Rui-li, Silval D. Ensemble methods for spoken emotion recognition in call-centres [J]. *Speech Communication*, 2007, 49(2):98-112

[25] Kaiser F. On a simple algorithm to calculate the energy of a signal [C]// *Proceedings of IEEE, International Conference on Acoustics, Speech, and Signal Processing*. Xi'an: IEEE Computer Society, 1990:381-384

[26] 林奕琳. 基于语音信号的情感识别研究[D]. 广州: 华南理工大学, 2006

[27] Ramamohan S, Dandapat S. Sinusoidal model-based analysis and classification of stressed speech [J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2006, 14(3):737-746

[28] Mota S, Picard R. Automated Posture Analysis for Detecting Learner's Interest Level [C]// *Proceedings of Conference on Computer Vision and Pattern Recognition Workshop*. Wisconsin: IEEE Computer Society, 2003:49

[29] D'Mello S, Graesser A. Automatic Detection of Learner's Affect from Gross Body Language [J]. *Applied Artificial Intelligence*, 2009, 23(2):123-150

[30] Castellano G, Mortillaro M, Camurri A, et al. Automated analysis of body movement in emotionally expressive piano performances [J]. *Music Perception: An Interdisciplinary Journal*, 2008, 26(2):103-119

[31] Camurri A, Lagerlöf I, Volpe G. Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques [J]. *International Journal of Human-Computer Studies*, 2003, 59(1):213-225

[32] Castellano G, Santiago D V, Camurri A. Recognizing Human Emotions from Body Movement and Gesture Dynamics [C]// *Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction*. Berlin, Heidelberg: Springer, 2007:73-81

[33] Demeijer M. The Contribution of General Features of Body Movement to the Attribution of Emotions [J]. *Journal of Nonverbal Behavior*, 1989, 13(4):247-268

[34] Coulson M. Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence [J]. *Journal of Nonverbal Behavior*, 2004, 28(2):117-139

[35] Montepare J, Koff E, Zaitchik D, et al. The Use of Body Movements and Gestures as Cues to Emotions in Younger and Older Adults [J]. *Journal of Nonverbal Behavior*, 1999, 23(2):133-152

- [36] Cacioppo J T, Berntson G G, Larsen J T, et al. The psychophysiology of emotion [J]. *The Handbook of Emotion*, 1993, 2(4): 173-191
- [37] Rainville P, Bechara A, Naqvi N, et al. Basic emotions are associated with distinct patterns of cardio respiratory activity[J]. *International Journal of Psychophysiology*, 2006, 61(1): 5-18
- [38] 温万惠, 邱玉辉, 刘光远, 等. 情感生理反应样本库的建立与数据相关性分析 [J]. *中国科学: 信息科学*, 2011, 41(1): 77-89
- [39] Picard R W, Healey J. *Affective wearable*[C]//*Proceedings of the First International Symposium on Wearable Computers*. Washington, DC, USA; IEEE Computer Society Press, 1997: 231-240
- [40] Villon O, Lisetti C. Toward Recognizing Individual's Subjective Emotion from Physiological Signals in Practical Application[C]//*Proceedings of 20th IEEE International Symposium on Computer Based Medical Systems*. Maribor, Slovenia; IEEE Computer Society, 2007: 357-362
- [41] Kim J, Andre E. Emotion recognition based on physiological changes in music listening[J]. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 30(12): 2067-2083
- [42] Lee M, Yang G, Lee H K, et al. Development Stress monitoring System based on Personal Digital Assistant (PDA)[C]//*Proceedings of the 26th Annual International Conference of the IEEE EMBS San Francisco*. CA, USA; IEEE Computer Society, 2004: 312-320
- [43] <http://emotion-research.net/>
- [44] Panagiotis P C, Leontios J H. Emotion Recognition from Brain Signals Using Hybrid Adaptive Filtering and Higher Order Crossings Analysis[J]. *IEEE Transactions on affective Computing*, 2010, 1(2): 81-97
- [45] 李建平, 郭念锋, 等. 情绪自主神经特异性研究及进展[J]. *心理科学*, 2005, 28(3): 744-746
- [46] Li Lan, Chen Ji-hua. Emotion recognition Using Physiological Signals[C]//*Proceedings of 16th International conference on artificial reality and telexistence*. Hangzhou, China; IEEE Computer Society, 2006: 437-446
- [47] 鲁舜, 刘光远. 用自适应蚁群算法的生理信号情感状态识别[J]. *计算机应用*, 2009, 29(6): 146-148
- [48] 吴光华, 刘光远, 龙正吉. 免疫机制对皮肤电信号情感特征选择的影响[J]. *计算机应用研究*, 2010, 27(12): 4558-4560
- [49] 牛晓伟, 刘光远. 用于生理信号情感识别的自适应遗传算法[J]. *计算机工程与设计*, 2008, 29(14): 3726-3728
- [50] 程德福, 刘光远, 邱玉辉. 双重结构粒子群和 KNN 在生理信号情感识别中的应用[J]. *计算机应用*, 2009, 29(5): 1423-1425
- [51] Hatzivassiloglou V, McKeown K R. Predicting the semantic orientation of adjectives[C]//*the Proceedings of the EACL'97*. Morristown; ACL, 1997: 174-181
- [52] Wiebe J. Learning subjective adjectives from corpora[C]//*the Proceedings of the 17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence*. Menlo Park; AAAI, 2000: 735-740
- [53] Riloff E, Wiebe J. Learning extraction patterns for subjective expressions[C]//*the Proceedings of the 2003 conference on Empirical methods in natural language processing*. Morristown; ACL, 2003: 105-112
- [54] Turney P, Littman M L. Measuring praise and criticism: Influence of semantic orientation from association[J]. *ACM Transactions on Information Systems*, 21(4): 315-346
- [55] Kim S M, Hovy E. Automatic detection of opinion bearing words and sentences[C]//*the Proceedings of the International Joint Conference on Natural Language Processing*. Morristown; ACL, 2005: 61-66
- [56] Kim S M, Hovy E. Identifying and analyzing judgment opinions [C]//*the Proceedings of the Joint Human Language Technology/North American Chapter of the ACL Conference*. Morristown; ACL, 2006: 200-207
- [57] Sanjiv R D, Chen M Y. Yahoo! For Amazon: Sentiment Extraction from small Talk on the Web [J]. *Management Science*, 2007, 53(9): 1375-1388
- [58] Denecke K. Using SentiWordNet for Multilingual Sentiment Analysis[C]//*the Proceedings of the International Conference on Data Engineering*. Cancun; IEEE Computer Society, 2008: 507-512
- [59] Polpinij J, Ghose A K. An Ontology-based Sentiment Classification Methodology for Online Consumer Reviews[C]//*IEEE/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. Sydney; IEEE Computer Society, 2008: 518-524
- [60] 王素格, 魏英杰. 停用词表对中文文本情感分类的影响[J]. *情报学报*, 2008, 27(2): 175-179
- [61] 徐琳宏, 林鸿飞, 潘宇, 等. 情感词汇本体的构造[J]. *情报学报*, 2008, 27(2): 180-185
- [62] 赵妍妍, 秦兵, 刘挺. 文本情感分析[J]. *软件学报*, 2010, 21(8): 1834-1848
- [63] Scherer K, Ellgring H. Multimodal Expression of Emotion: Affect Programs or Componential Appraisal Patterns [J]. *Emotion*, 2007, 7(1): 158-171
- [64] Castellano G, Kessous L, Caridakis G. Emotion Recognition through Multiple Modalities: Face, Body Gesture, Speech [J]. *Emotion*, 2008, 8(6): 92-103
- [65] Kapoor A, Picard R W. Multimodal Affect Recognition in Learning Environments[C]//*the Proceedings of 13th annual ACM international conference on multimedia*. New York, USA; ACM, 2005: 677-682
- [66] 黄程韦, 赵力, 等. 基于语音信号与心电信号的多模态情感识别 [J]. *东南大学学报: 自然科学版*, 2010, 40(5): 895-900
- [67] 李爱军, 邵鹏飞, 党建武. 情感表达的跨文化多模态感知研究 [J]. *清华大学学报: 自然科学版*, 2009, 49(S1): 1393-1401
- [68] Reisenzein R. Broadening the Scope of Affect Detection Research [J]. *IEEE Transactions on Affective Computing*, 2010, 1(1): 42-45
- [69] Keltner D, Haidt J. Social Functions of Emotions [J]. *Emotions: Current Issues and Future Directions*, 2001, 1(1): 192-213
- [70] Rime B, Elicit E. the Social Sharing of Emotion: Theory and Empirical Review [J]. *Emotion Review*, 2009, 1(1): 60-85
- [71] Salovey P. Introduction: Emotion and Social Processes, *Handbook of Affective Sciences* [M] // Davidson R J, Scherer K R, Goldsmith H, eds. , Oxford Univ. Press, 2003
- [72] Chiao J Y, Iidaka T, Gordon H L, et al. Cultural Specificity in Amygdala Response to Fear Faces [J]. *Journal of Cognitive Neuroscience*, 2008, 20(1): 2167-2174
- [73] Peterson G. Cultural Theory and Emotions, . *Handbook of the Sociology of Emotions* [M] // Stets J, Turner J, eds. Springer, 2006
- [74] Russell J A. Core Affect and the Psychological Construction of Emotion [J]. *Psychological Review*, 2003, 110(1): 145-172