

人工智能语境下的情感交互设计

颜洪¹, 刘佳慧¹, 覃京燕²

(1.海南大学, 海口 570228; 2.北京科技大学, 北京 100083)

摘要: **目的** 研究情感交互系统在设计层面上的重要新兴趋势, 探讨以人为本的情感交互设计方案及其研究领域中的人工智能技术的实际应用。**方法** 在情感交互设计中, 机器需要实时捕捉用户的关键信息, 识别用户的情感状态, 利用多种线索最终决定合适的用户模型。首先通过分析人工智能技术在人类情感识别上的分类优势, 讨论如何利用面部表情、肢体动作、语言、心跳、呼吸等多模态用户特征, 使机器更加准确地识别人类情感; 其次通过分析经典的情绪状态的数学模型以及情绪表达与交互优化等关键要素, 针对用户的情感作出自然、合理、友好的反应; 最后探讨情感交互系统设计和模块化构建的可行性。**结论** 提出一种利用人工智能技术的情感交互系统的模块化设计方案。

关键词: 情感交互; 人工智能; 人机交互; 用户体验

中图分类号: TB472 **文献标识码:** A **文章编号:** 1001-3563(2020)06-0013-07

DOI: 10.19554/j.cnki.1001-3563.2020.06.003

Emotional Interaction Design in Artificial Intelligence Context

YAN Hong¹, LIU Jia-hui¹, QIN Jing-yan²

(1.Hainan University, Haikou 570228, China; 2.University of Science and Technology Beijing, Beijing 100083, China)

ABSTRACT: The work aims to study the important emerging trends in the design of emotional interaction systems, and explore the human-oriented emotional interaction design and the practical application of artificial intelligence technology in its field. In the emotional interaction design, the machine was required to capture the user's key information in real time, identify the user's emotional state, and use a variety of clues to finally determine the appropriate user model. Therefore, by analyzing the classification advantages of artificial intelligence technology in human emotion recognition, how to make the machine recognize human emotions more accurately with the multi-modal user features such as facial expressions, limb movements, language, heartbeat, and breathing was explored. Then, by analyzing classic mathematical models of emotional states, and key elements such as emotional expression and interaction optimization, a natural, reasonable, and friendly response was made to the user's emotions. Finally, the feasibility of the design and modular construction of emotional interaction system was discussed. A modular design of emotional interaction system based on artificial intelligence technology is proposed.

KEY WORDS: emotional interaction; artificial intelligence; human-computer interaction; user experience

人机共生 (Man Computer Symbiosis) 的设计理念提出以来, 关于交互系统的多维模型、范式理论、平台开发等研究取得了显著的成就^[1]。基于传感器的上下文情境感知 (Context Awareness)、意识与情绪感

知 (Consciousness Awareness), 基本实现了高效率、自然的人机交互方式^[2]。随着人工智能技术的广泛普及, 基于大数据的智能化决策与管理, 有效地解决了当代社会中的“数字鸿沟”等诸多问题^[3]。以用户为中

收稿日期: 2020-01-31

基金项目: 海南省自然科学基金项目 (619QN196); 海南省科协青年科技英才创新计划项目 (QCXM201911); 海南大学海南自贸区 (港) 研究专项项目 (19ZMQG17)

作者简介: 颜洪 (1984—), 男, 吉林人, 博士, 海南大学讲师, 主要从事个性化推荐、人机交互研究。

通信作者: 覃京燕 (1974—), 女, 四川人, 博士, 北京科技大学教授, 主要从事人工智能、交互设计等方面的研究。

心,亲切生动的、更加符合人类交流逻辑的新一代情感交互,已逐渐成为研究的主要目标^[4]。情感在人类行为中起着核心作用,对人的感知、学习、决策和注意力等机制产生了重大影响。1997年,MIT的Picard提出了情感计算(Affective Computing)的概念,通过赋予机器识别、理解和表达人类情绪的能力,使得计算机拥有更高的智慧。因此,情感交互需要给予机器洞察、理解和产生各类情绪的能力,最终使机器像人类一样进行真实、自然、赋有情感的交互^[5]。近年来,以人工智能技术为基础的情感交互的关注度不断增加,研究者们通过人工智能技术,在多媒体与多模态界面、信息可视化以及以计算机为基础的情感模型等领域,试图去处理日常生活中的现实问题。例如,在网络学习中,当学习者感到无聊或感兴趣、沮丧或满意时,情感交互系统适应调整计算机辅导方式^[6];当系统确定用户的情绪状态时,会提供可靠的心理健康咨询服务并从中收集有益信息^[7];当一个人在不确定或复杂的环境中工作时,能够处理情感信息的智能语音助手,便在此时表现出强大的灵活性^[8]。然而,人类的情感交流是非常复杂的交互过程,目前的情感交互研究仍存在诸多的挑战,例如用户信息的获取与情感识别、情绪的认知与情感建模、情感的表达与优化等交互问题。

本文主要研究情感交互系统在设计方面的重要新兴趋势,并探讨以人为本的情感交互研究中的人工智能技术的实际应用,包括如何使机器正确地识别人的情感,减少分歧与不确定性;如何通过合适的情感模型进行合成与表达;最终根据现有技术条件,提出基于人工智能技术的情感交互系统的模块化设计方案。

1 用户的情感特征

交互过程中会产生各种形式的用户生理和行为数据。在检测情感特征中通常会使用多种传感器,在不解释输入数据的情况下,捕获有关用户的生理状态或行为数据。例如,摄像机可以捕获面部表情、身体姿态或手势,麦克风可以捕获语音信息,生理传感器通过直接测量生理数值(例如皮肤温度、心跳等)来检测用户的情绪状态。当下热门的研究主要集中在生理信号、面部表情、身体姿势和语音信息等方面。

1.1 生理信号

用户的生理信号是较为精确,并可被有效利用的特征之一^[9]。各国学者早已展开对生理信号情绪识别领域的研究,包括肌电信号EMG^[10]、血流量脉冲信号BVP^[11]、脑电信号EEG^[12]、呼吸信号RSP^[13]等。例如,牛洁^[14]利用脑电波信号传感器来测试用户在学习过程中的受挫状况。Nosu^[15]通过心率、呼吸频率、手指温度等生理信号识别和推测用户的心理状态。基于生理信号的情感识别具有较高的精准度,然而传统

生理传感器对人体有很强的侵入性,不利于大规模应用。

随着互联网和移动通讯技术的发展与可穿戴设备的普及,不拘泥于时间和地点,可充分利用生理特征方式的设备日益流行并趋于主流。人工智能在可穿戴设备上的情感交互研究也受到了高度关注,林文倩^[16]利用卷积神经网络,通过可穿戴设备采集多种人体生理参数,并尝试根据情绪的变化对机械臂发送不同的控制指令。Naima^[17]通过可穿戴设备检测人体的EMG、ECG、血氧饱和度、呼吸频率和运动状态等数据,利用人工智能技术实现了个人的身体和心理健康状况的实时监测。

1.2 面部表情

面部表情是反映人情绪状态最自然的途径之一^[18]。它是经由口部肌肉、眼部肌肉和脸部肌肉的变换而呈现出的各类情绪状态,而面部表情最丰富的区域是由口与眼部附近的肌肉群组成。著名心理学家Ekman在20世纪发现,人们在传达同一情感时,面部肌肉都具有相似的规律性,不受到性别、年龄、地域及受教育程度等多种因素的影响^[19]。之后Ekman将人类脸部表情划分为七大类,基于人脸部表情的分类见图1,惊讶(Surprise)、悲伤(Sad)、高兴(Happy)、恐惧(Fear)、厌恶(Disgust)、愤怒(Angry)、中性(Neutral),并基于运动单元(Action Unit, AU)提出面部动作编码系统(Facial Action Coding System, FACS)。

面部表情识别通常使用图像处理^[20]等手段,但具有占用计算资源较多,对一些物理条件如光照、遮挡等变化反应比较敏感的缺点,容易影响视觉传感器数据的实时处理和对情绪状态的推测精度。但随着人工智能的出现,识别的误差大幅度减少。刘小津^[21]面向人脸表情识别与人脸识别这两个关键技术应用,提出了对抗训练以及自适应深度度量学习两种方法,用以解析表情以及身份信息或身份与某些特征属性信息,研发了一种新的深度尺度的学习算法,解决了其所处领域常见的锚点选择问题并降低了运算量,其阈值参数可自适应学习。李卫^[22]基于小样本提出了一种改进的深度信念网络结构(Deep Belief Networks),其在辨识度上与卷积神经网络相持平,但训练方式的改善使得耗损的时间大幅度下降。

1.3 身体姿势

大多数研究者专注于面部表情(在情感检测方面进行的研究中,有95%以上使用面部作为研究对象),而忽略了通过其他渠道识别用户情感的信息^[23]。例如,即使人们不使用双手或大幅度的动作来主动地传达信息,但他们的身体仍在不断传递情感信息,如用脚轻敲、双臂交叉、头部倾斜等。实验心理学也已经证实了某些类型的运动与特定的情感存在一定关联^[24]。例如,为了避免恐惧,人会使自己的身体远离

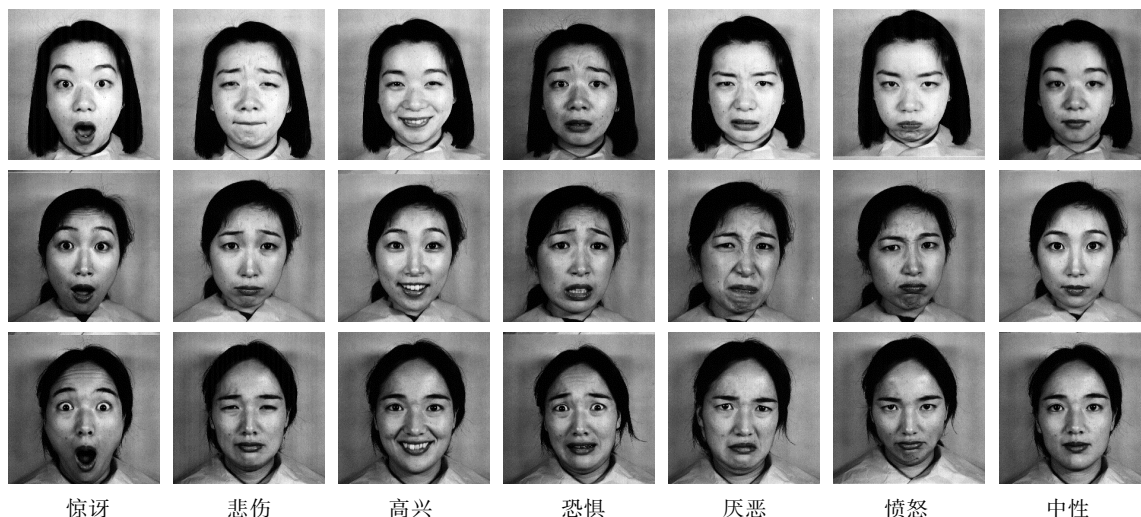


图 1 基于人脸部表情的情感分类

Fig.1 Emotion classification based on human facial expression

引起这种感觉的区域；在尝试幸福、惊喜或愤怒时，人们会将他们的身体转向引起这种感觉的方向。尽管该领域是新颖的，但仍有一些研究从手势中识别情绪状态^[25]，但学界对于这种检测情绪的方式并没有达成共识。通常，研究者们围绕身体的关节（手、膝盖、脖子、头部、肘部等）以及它们相互连接的身体部位之间的角度来构建情感特征^[26-27]。

1.4 语音信息

相较于动作与表情，人们擅长和最直接表达情感的方式就是语音。其中体现说话人情感的信息主要存在于语音信号的韵律特征中^[28]，包括基音频率、基音范围、声道的横截面面积、语音强度和语音速率等。另外，情感语音的信息也包含在音质特征中^[29]，如共振峰、梅尔频率倒谱系数（MFCC）、声道的横截面面积、线性预测系数（LPC）、Teager 能量算子等。

人工智能对于这些特征的识别具有较高的准确性和鲁棒性，这对语音情感识别的应用有着至关重要的作用。例如，Turk 等人创建了几种语音消息的神经网络识别器和识别器组合^[30]，识别器能够区分两种状态，分别为“激怒”（愤怒、幸福和恐惧）和“平静”（正常状态和悲伤），平均准确率为 77%。Bahreini^[31]提出了一种 FILTWAM 框架，该框架利用这项深度学习将用户语音在学习进程中的情感状况进行跟踪监控，进而及时规划授课重难点或者进度。Drigas^[32]提出了智能化的抑郁症患者治疗自适应框架，通过对患者的语音情感变化进行跟踪，用于辅导抑郁症患者重新获得情感理解和语言表达能力等。

2 基于人工智能技术的情感识别

人工智能概念始于 1956 年的达特茅斯会议，但当时计算机存储容量和网络还没有得到相应的发展，

因此无法存储人工智能需要的信息量，仅能通过基于规则的专家（Expert System）和统计算法解决简单问题，例如迷宫证明、数学定理等。很多研究者也尝试了不同的情感识别方法，包括使用随机森林^[33]、逻辑回归^[34]等，但在情感识别任务中，实现精准的预测是极其困难和复杂的，因为这需要将大量数据进行归纳和分类，然后计算机对这些特征进行编程以识别和理解情感。

进入 21 世纪以来，以神经网络^[35-37]为代表的机器学习，可以根据数据执行迭代学习，并调整学习结果（也可称为建模）。机器可以从这种迭代学习中掌握特定事件（事物）的“特征”。成功捕获事件的“特征”后，将其升华为一个“规则”，可以在下一次进行使用。神经网络在计算机视觉、语音识别和生物信息学等多个领域表现出了卓越的性能。深层结构的神经网络相较于传统的提取方式而言，可以有效地展示出数据之间的内在隐藏结构，提炼出对于分类或回归任务中有利的高层抽象特性。另一方面，随着移动技术和互联网的发展，可以随时随地收集数据，将其存储在云中，并平稳地使用存储的数据。因此，当前的人工智能是以大数据为基础，数据资源决定了人工智能的高度，收集的知识越多，人工智能就会越准确。深度学习与传统机器学习的情感推测对比见图 2，使用神经网络的方法明显优于往常的提取特征方式，特别是广泛存在于实际生活中的高维非结构化数据，神经网络在特征学习方面的优势日益增长。

在此基础上，不同模态间的特征融合，可以大幅提高推测精度。例如，在视频语义理解任务中，Ngiam^[38]将深度学习这一环节引入到多模态数据的特征学习工作中，整理了语音识别数据中视频和音频的不同信息模块。Srivastava^[39]将神经网络运用于图像检索过程中，利用图像数据中收集到的各类特

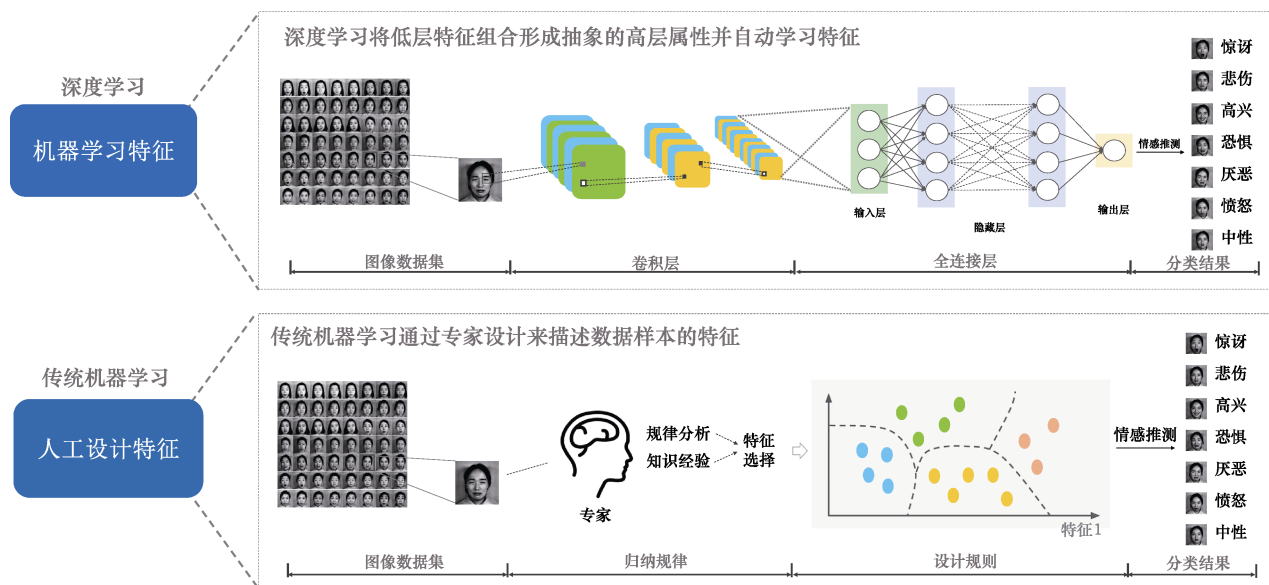


图 2 深度学习与传统机器学习的情感推测对比

Fig.2 Comparison of emotion estimation between deep learning and traditional machine learning

征点作为图像模态,关于图像的文本备注则可以作为文本模态,搭建同一深度数据的神经网络模型,且将练习中获得的抽象语义特征用以图像分类。这些研究均表明,多模态数据的深度学习可以有效地提升推测精准度,因此,深入分析模态间的本质特征,并有效地进行特征融合,有助于提高情绪预测的精度,实现更好的用户体验。然而,松散的数据结构以及开放性的检测环境导致情绪状态的推测难度加大,如何构建精准的分类模型,是实现个性化情感交互系统的一个重要科学问题。既需要考虑到单个行为,又需要兼顾行为之间的互补性,还需要通过反复迭代确定各种权重的因果关系,基于用户特征的预测评价与潜在变量,即情绪状态客观评价之间的假设关系,最终实现最佳的分类效果。此外,在多模态数据融合上,研究多采用决策级融合的方式。虽然这种融合方式比较快速、有效,但是仍需要分析对比其他的融合方式,多模态情感数据融合框架见图 3,决策级融合、特征级融合和双向融合在情绪状态推测应用中的优劣,是一个值得深入研究的问题。

3 情感模型

情感建模作为情感表达和情感识别的关键因素,重点是利用搭建的情感状态的数学模型,可以增加现实的反应和感受情绪的意义。在情感建模中,目前维度情感模型在实际应用中占据多数,它把人类多方位的情感体验转化为欧氏距离空间进行描述,通过纬度、取值两种不同的因素可以组合出一种特定的情感状态。Russell^[40]的 VA (Valence-Arousal) 二维情感模型是如今使用量较高的数学模型, V 表示价效维度, A 表示唤醒维度。目前认可度相对较高的 PAD 三维情感模型,便是由 Mehrabian^[41]在此基础上提出

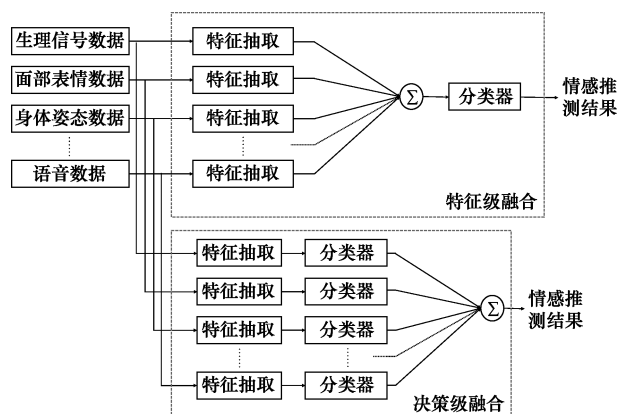


图 3 多模态情感数据融合框架

Fig.3 Framework of multimodal emotional data fusion

了愉悦度、唤醒度和优势度三个维度。

随着自然选择理论的发展,情感具有进化的历史,并在不同文化之间共享。Plutchik^[42]提出了一种分类法,以车轮模型的形式对情绪进行分类,Plutchik 情感模型,其中包含了八种基本情绪:悲伤、愤怒、信任、厌恶、惊奇、喜悦、期待和恐惧。在他的分类法中,情感可以再次混合形成复杂的形式、人格特质和心理病理学。在此基础上,Izard^[43]提出了十种基本情绪:喜悦、兴趣、鄙视、悲伤、愤怒、恐惧、害羞、惊奇、内疚和厌恶。Izard 认为情绪是人类进化的结果,每一种情绪都与一个简单的大脑回路相关,其中不涉及复杂的认知成分。

除此之外,Ortony^[44]提出的 OCC 情感模型将二十二种基本的情感依据成因分为三类,分别为仿生代理的动作、整件事情的结果和对象的观感,且进一步划分了这三类情感的主次关系,能够描述特定情感的起因和过程。Picard 的隐马尔可夫模型^[45],有三种情感状态——高兴、悲伤、感兴趣,而且可以拓展到多

种情形下的情感状态,其通过观测模型中情绪变动的上升时间、峰值间隔的频率高低等情感状况获得,进而演算得到一种最接近的情感状态。Kesteren^[46]提出的分布式情感模型,主要是应对外界干扰开发的一种分布式情感模型。

4 情感表达与优化

在情感交互系统设计中,及时、准确而有效地识别用户情感与自然、生动而合理地表达情感同等重要。基于人工智能的情感表达已应用于多种生活场景当中。例如, Liu^[47]从自主神经系统 (Autonomic Nervous System, ANS) 的反应中识别出玩家的焦虑程度,并用它来适应乒乓球比赛中遇到的困难;他们使用类似的方法来识别儿童在玩机器人时的娱乐性,并让机器人改变其行为以获得最大化娱乐性^[48]; ANS 响应已被用于识别运动康复中的工作量水平,并调整运动以最佳地挑战患者^[49];在意识到睡意的情况下警告汽车驾驶员的问题^[50-51]等。

在此基础上, Mello^[52]等人研究了机器的多模态情感表达,采用一组参数来衡量人在接受单模和多模刺激时可能的状态,分别为:中性、无聊、困惑、吸引和无奈。Mellod 研究了单模态(视觉)、双模态(视觉、听觉)、多模态(嗅觉、视觉、听觉)的刺激分别对人类情感的不同影响。如果界面可以设计成用多种模态来刺激用户的方式,那么情感表达在情感交互设计中就变得非常关键。例如,日本软银的 Pepper 可以识别并应对各种情绪状态,根据对方的情绪以适当的手势和言语作出回应。当 Pepper 认识到“对方对互动感到失望”时,它会尝试道歉。

同时,在情感表达的模型设计和构建中,情感交互系统需要具有自适应性,即情感模型是针对于个性化服务并可在特定状态下进行自动编辑,当然这需要通过理解用户的实际情况反馈并调节其适合的表达模式。情感交互系统的自适应设计见图 4,由于交互系统都具备实时获取用户生理、行为特征的能力,并通过上下文感知、语音交互及多模态界面等技术手段,构筑更加完善的用户画像 (User Profile),进而提供更加个性化的需求信息。随着数据的不断累积,交互效果也会变得越好,在不断迭代中实现更加优质的用户体验。

5 情感交互系统的模块化设计方案

情感交互的最终目标是使计算机在理解人类情感状态的前提下,进行合理调整,顺应用户情感的迁移。基于以上论述,在人工智能的语境下,情感交互系统的重点是使用各种感知方式来识别,解释和响应人类的情感。因此提出了具有更好的灵活性和可拓展性的基于人工智能技术的模块设计方案,见图 5,其

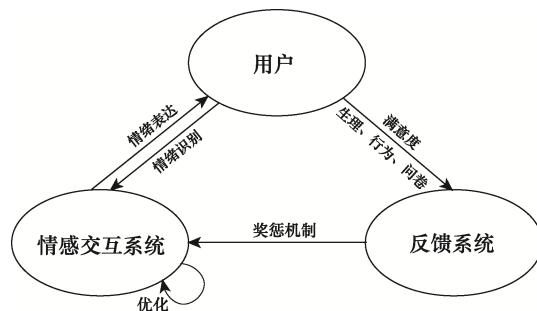


图 4 情感交互系统的自适应设计

Fig.4 Adaptive design of emotional interaction system

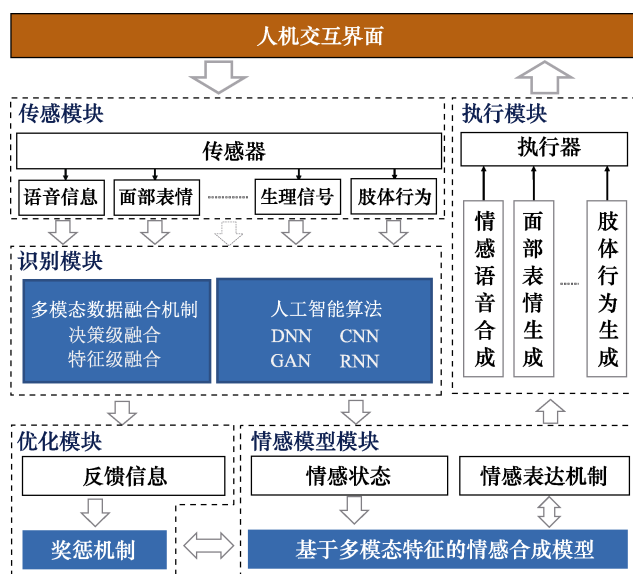


图 5 基于人工智能的情感交互系统的模块化设计方案

Fig.5 Modular design of emotional interaction system based on artificial intelligence

主要分成传感模块、执行模块、识别模块、情感计算模块及优化模块五个方面。

传感模块是情感交互系统的基础,它利用麦克风、摄像头、可穿戴设备、眼动仪、kinect 等多种传感设备收集生物识别数据,包括大脑中的电信号、心电信号、电热信号、情感语音、面部表情和肢体行为等。传感模块就相当于人体上的神经元,无时无刻地感受外界物质的刺激。

识别模块是情感交互的基础,需要对收集到的数据进行预处理,然后将数据样本的特征进行提取、分析等。相关研究表明^[53]利用深度学习的方法,如循环神经网络 (RNN)、卷积神经网络 (CNN)、对抗神经网络 (GAN) 显示出更高的推测潜力。它们既可以自动生成检测用户情绪的主观体验的最优方法,又可以通过存储设备对用户的长期行为进行建模。此外,基于多模态用户信息的情感识别可以选择不同模态组合,并且模态间的数据融合方式也是可以根据实际情况有不同的选择。

情感模型模块是情感识别和情感表达的关键部分,通过建立多模态情感状态的数学模型,实现更加

合理的情感理解和反馈。情感模型模块需要在认知科学基础下研究用户的各种心理现象,通过判断情感信息的构成,对用户的情绪作出合理的反应。情感的表达作为经验和经典理论研究的分支,需要构建具有内在效用的人类情感模型,从现实中归纳认知或心理现象及理论,通过实验以验证假设推论。

执行模块是负责情感交互系统的信息输出任务,根据情感模型模块作出的决策执行相应的反馈,最终通过如面部表情、肢体动作、合成语音等情感表达方式完成情感交互。

在情感交互系统中,优化模块是提供个性化服务的关键环节,由于无法找到适用于所有系统的情感模型方案,情感建模仍然高度依赖用户的反馈信息,所以具有自适应的优化模块显得尤为重要。优化方案通常在受控的交互环境中完成,并且用户对特定产品的反应需要得到监控,为系统提供有价值的反馈信息,从而建立良好的、个性化的奖惩优化机制,实现更好的用户体验。

通过人工智能技术,可以解决以往情感交互系统设计面临的技术瓶颈,如判断用户情感的多种生理或行为特征信息及上下文环境,提高多模态用户情感状态推测的准确性,以及模块化设计的灵活性及可拓展性。当前,深度学习+大数据已经成为人工智能发展的主要技术路线,这些都使人工智能在情感交互的性能表现日渐凸显。

6 结 语

在过去的二十年中,情感科学、计算机科学和电子学领域的进步,高效的人工智能算法的开发以及创建无所不在的可穿戴技术,使得情感交互变得可能。这些新技术已成为人们现实生活的一部分,有助于不断改善生活质量,并允许获取大量数据,这些数据可用于开发复杂的人工智能算法以实现可靠的情感算法计算模型。情感计算已发展出许多方法,最终出现了新的研究问题、挑战和机遇,使人们对情感的理解和认知又向前迈进了一步。

参考文献:

- [1] DIX A, FINLAY J, ABOWD G, et al. Human-Computer Interaction[M]. Upper Saddle River: Prentice Hall, 2003.
- [2] 吴敏, 刘振焄, 陈略峰. 情感计算与情感机器人系统[M]. 北京: 科学出版社, 2018.
WU Min, LIU Zhen-cao, CHEN Lue-feng. Affective Computing and Affective Robot System[M]. Beijing: Science Press, 2018.
- [3] NEAPOLITAN R E, JIANG X. Artificial Intelligence: with an Introduction to Machine Learning[M]. New York: Chapman and Hall/CRC, 2018.
- [4] 毛峡, 薛雨丽. 人机情感交互[M]. 北京: 科学出版社, 2011.
- MAO Xia, XUE Yu-li. Human Computer Emotional Interaction[M]. Beijing: Science Press, 2011.
- [5] PICARD R W. Affective Computing[M]. Cambridge: MIT Press, 2000.
- [6] GU X, SHI Z, WANG Z. Research on Affective Robot Speech Recognition Technology Based on BDI Agent Technology[J]. Application Research of Computers, 2003(4): 24-26.
- [7] 沈林翔. 基于云计算的心理健康服务平台的研究与实现[D]. 南京: 南京邮电大学, 2016.
SHEN Lin-xiang. Research and Implementation of Mental Health Service Platform Based on Cloud Computing[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2016.
- [8] POLYAKOV E V, MAZHANOV M S, ROLICH A Y. Investigation and Development of the Intelligent Voice Assistant for the Internet of Things Using Machine Learning[C]. Piscataway: IEEE, 2018.
- [9] GOODMAN J B. Physiological Signal Monitoring System: U.S. Patent 6,616,613[P]. 2003.
- [10] PARKER P, ENGLEHART K, HUDGINS B. Myoelectric Signal Processing for Control of Powered Limb Prostheses[J]. Journal of Electromyography and Kinesiology, 2006, 16(6): 541-548.
- [11] MARCOVECCHIO A F, GEHEB F, BOUCHER D R. Processing Pulse Signal in Conjunction with Accelerometer Signal in Cardiac Resuscitation: U.S. Patent 7,488,293[P]. 2009.
- [12] SUBASI A, ERCELEBI E. Classification of EEG Signals Using Neural Network and Logistic Regression[J]. Computer Methods and Programs in Biomedicine, 2005, 78(2): 87-99.
- [13] VEDAM S S, KEALL P J, KINI V R. Acquiring a Four-Dimensional Computed Tomography Dataset Using an External Respiratory Signal[J]. Physics in Medicine & Biology, 2002, 48(1): 45.
- [14] 牛洁. 注意力相关脑电的特征提取及分类方法研究[D]. 西安: 西安电子科技大学, 2013.
NIU Jie. Feature Extraction and Classification of Attention-Related EEG[D]. Xi'an: Xi'an University of Electronic Technology, 2013.
- [15] NOSU K, KUROKAWA T. A Multi-Modal Emotion-Diagnosis System to Support e-Learning[C]. Piscataway: IEEE, 2006.
- [16] 林文倩. 生理信号驱动的情绪识别及交互应用研究[D]. 杭州: 浙江大学, 2019.
LIN Wen-qian. Emotion Recognition and Interaction Application Driven by Physiological Signal[D]. Hangzhou: Zhejiang University, 2019.
- [17] NAIMA R, CANNY J. The Berkeley Tricorder: Ambulatory Health Monitoring[C]. Piscataway: IEEE, 2009.
- [18] CHEN J, LUO N, LIU Y. A Hybrid Intelligence-Aided Approach to Affect-Sensitive E-Learning[J]. Computing, 2016(1-2): 215-233.
- [19] EKMAN P. Facial Expressions of Emotion: an Old Controversy and New Findings[J]. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 1992, 335(1273): 63-69.
- [20] 黄雪琳. 人脸的几何特征提取与查询[D]. 泉州: 华侨大学, 2004.
HUANG Xue-lin. Extraction and Query of Geometric Features of Human Face[D]. Quanzhou: Huaqiao University, 2004.
- [21] 刘小泮. 基于深度学习的人脸图像识别技术的研究[D]. 长春: 中国科学院大学 (中国科学院长春光学精

- 密机械与物理研究所), 2019.
- LIU Xiao-feng. Face Image Recognition Based on Deep Learning[D]. Changchun: University of Chinese Academy of Sciences (Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences), 2019.
- [22] 李卫. 深度学习在图像识别中的研究及应用[D]. 武汉: 武汉理工大学, 2014.
- LI Wei. Research and Application of Deep Learning in Image Recognition[D]. Wuhan: Wuhan University of Technology, 2014.
- [23] YANG L, JIANG D, XIA X. Multimodal Measurement of Depression Using Deep Learning Models[C]. New York: ACM, 2017.
- [24] VINES B W, WANDERLEY M M, KRUMHANS L C L. Performance Gestures of Musicians: What Structural and Emotional Information Do They Convey?[C]. Berlin: Springer, 2003.
- [25] YORUK E, KONUKOGLU E, SANKUR B. Shape-Based Hand Recognition[J]. IEEE Transactions on Image Processing, 2006, 15(7): 1803-1815.
- [26] GLOWINSKI D, CAMURRI A, VOLPE G. Technique for Automatic Emotion Recognition by Body Gesture Analysis[C]. Piscataway: IEEE, 2008.
- [27] KAPUR A, VIRJI-BABUL N. Gesture-based Affective Computing on Motion Capture Data[C]. Berlin: Springer, 2005.
- [28] KWON O W, CHAN K, HAO J. Emotion Recognition by Speech Signals[C]. Australia: ISCA Archive, 2003.
- [29] 余伶俐, 蔡自兴, 陈明义. 语音信号的情感特征分析与识别研究综述[J]. 电路与系统学报, 2007, 12(4): 76-84.
- YU Ling-li, CAI Zi-xing, CHEN Ming-yi. Review of Affective Feature Analysis and Recognition of Speech Signals[J]. Journal of Circuits and Systems, 2007, 12(4): 76-84.
- [30] TURK M, PENTLAND A. Face Recognition Using Eigenfaces[C]. Piscataway: IEEE, 1991.
- [31] BAHREINI K, NADOLSKI R, WESTER W. Towards Multimodal Emotion Recognition in E-Learning Environments[J]. Interactive Learning Environments, 2016, 24(3): 590-605.
- [32] DRIGAS A S, IOANNIDOU R E. Artificial Intelligence in Special Education: a Decade Review[J]. International Journal of Engineering Education, 2012, 28(6): 1366-1372.
- [33] 陆兵, 顾苏杭. 基于级联特征的随机森林运动目标跟踪算法[J]. 计算机技术与发展, 2019(5): 1-6.
- LU Bing, GU Su-hang. Random Forest Moving Target Tracking Algorithm Based on Cascade Feature[J]. Computer Technology and Development, 2019(5): 1-6.
- [34] 胡祖辉, 施佳. 高校学生上网行为分析与数据挖掘研究[J]. 中国远程教育, 2017(02): 26-32.
- HU Zu-hui, SHI Quan. Analysis of University Students' Internet Access Behavior and Data Mining Research[J]. China Distance Education, 2017(2): 26-32.
- [35] 徐振国, 张冠文, 孟祥增, 等. 基于深度学习的学习者情感识别与应用[J]. 电化教育研究, 2019, 40(2): 87-94.
- XU Zhen-guo, ZHANG Guan-wen, MENG Xiang-zeng, et al. Emotion Recognition and Application of Learners Based on Deep Learning[J]. Research on Educational Technology, 2019, 40(2): 87-94.
- [36] BENGIO Y, LAMBLIN P, POPOVICI D. Greedy Layer-Wise Training of Deep Networks[J]. In Advances in Neural Information Processing Systems, 2007(4): 153-161.
- [37] HINTON G E, SALAKHUTDINOV R. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(6): 504-507.
- [38] NGIAM J, KHOSLA A, KIM M. Multimodal Deep Learning[C]. Washington: Proceedings of International Conference on Machine Learning, 2011.
- [39] SRIVASTAVA N, HINTON G, KRIZHEVSKY A. Dropout: a Simple Way to Prevent Neural Networks from Overfitting[J]. Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [40] RUSSELL S J, SUBRANABUAB D. Provable Bounded-Optimal Agents[J]. Journal of Artificial Intelligence Research, 1995, 2(1): 575-609.
- [41] MEHRABIAN A. Basic Dimensions for a General Psychological Theory[M]. Cambridge: Oelgeschelager, 1980.
- [42] PLUTCHIK R. Emotion: a Psyche-Volutionary Synthesis[M]. New York: Harper & Row Publishers, 1980.
- [43] IZARD C E. Human Emotions[M]. New York: Plenum Press, 1977.
- [44] ORTONY A, CLORE G L, COLLINS A. The Cognitive Structure of Emotion[J]. Contemporary Sociology, 1988, 18(6): 2147-2153.
- [45] 王志良. 人工心理[M]. 北京: 机械工业出版社, 2006.
- WANG Zhi-liang. Artificial Psychology[M]. Beijing: Machinery Industry Press, 2006.
- [46] KESTEREN A, AKKER R, POEL M, et al. Simulation of Emotions of Agents in Virtual Environments Using Neural Networks[C]. Piscataway: IEEE, 2000.
- [47] LIU C, AGRAWAL P, SARKAR N. Dynamic Difficulty Adjustment in Computer Games Through Real-Time Anxiety-Based Affective Feedback[J]. International Journal of Human-Computer Interaction, 2009, 25(6): 506-529.
- [48] LIU C, CONN K, SARKAR N. Online Affect Detection and Robot Behavior Adaptation for Intervention of Children With Autism[J]. IEEE Transactions on Robotics, 2008, 24(4): 883-896.
- [49] NOVAK D, MIHELJ M, ZIHERL J. Psychophysiological Measurements in a Biocooperative Feedback Loop for Upper Extremity Rehabilitation[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2011, 19(4): 400-410.
- [50] HEALEY J A, PICARD R W. Detecting Stress during Real-World Driving Tasks Using Physiological Sensors[J]. IEEE Transactions on Intelligent Transportation Systems, 2005, 6(2): 156-166.
- [51] PICOT A, CHARBONNIER S, CAPLIER A. On-line Detection of Drowsiness Using Brain and Visual Information[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2011, 42(3): 764-775.
- [52] D'MELLO S K, DOWELL N, GRAESSER A. Unimodal and Multimodal Human Perception of Naturalistic Non-Basic Affective States during Human-Computer Interactions[J]. IEEE Transactions on Affective Computing, 2013, 4(4): 452-465.
- [53] CALVO R A, D'Mello S. Affect Detection: an Interdisciplinary Review of Models, Methods, and Their Applications[J]. IEEE Transactions on Affective Computing, 2010, 1(1): 18-37.