

# 情感可以计算——情感计算综述<sup>\*</sup>)

张迎辉 林学閻

(清华大学计算机科学与技术系 北京 100084)

**摘要** 情感计算是自然和谐人机交互技术的一个重要组成部分。本文综述了情感计算的背景、研究基础和核心技术,并对未来的发展进行了展望。

**关键词** 情感计算,人机交互

## Affect is Computable ——A Survey on Affective Computing

ZHANG Ying-hui LIN Xue-yin

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

**Abstract** Affective computing is crucial for the implementation of natural human computer interaction. A survey is given on the background, related research areas and key techniques of affective computing. Some future research directions are discussed.

**Keywords** Affective computing, Human computer interaction

### 1 引言

情感能力是人类智能的重要标志,情感在人与人的交流中必不可少。那么,计算机可以有情感吗?情感可以计算吗?情感计算的研究试图回答这些问题。

### 2 情感计算的背景

#### 2.1 情感计算问题的提出

早在 1985 年,人工智能的奠基人之一 Minsky 就提出了计算机与情感的问题,他在“*The Society of Mind*”<sup>[1]</sup>一书中写到:问题并不在于智能机器是否有情感,而是没有情感的机器怎么能是智能的?“情感计算”一词则是由 MIT 的 Picard 在 1997 年出版的“*Affective Computing*”<sup>[2]</sup>中提出来的,她把“情感计算”定义为“与情感有关、由情感引发或者能够影响情感的因素的计算。”情感计算的目标是赋予计算机感知、理解以及表达情感的能力。

#### 2.2 情感计算的背景

情感计算的提出,一方面是提高人机交互和谐性的要求,另一方面是实现真正意义上的人工智能的要求。

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。它从科学家的实验室走进了千千万万普通人的生活,成为人们工作、学习、娱乐的助手和伴侣。以往那种以计算机为中心、人来适应计算机的交互方式显然不能适应新的需求。人们期待的是以人为中心、计算机主动适应人、如同人与人的交流那样自然和谐的交互环境。今天,计算机有了一定的“听”(语音识别)、“说”(语音合成)、“看”(计算机视觉)、“读”(自然语言理解)的能力,唯独缺少一颗善解人意的“心”。在人与人的交流中,情感占据着重要的地位。少女对自己钟情的对象往往暗送秋波而不是直接表白;愤怒时人们圆睁双目、紧握拳头做出准备战斗的姿态,却很少直接说:“我生气了”。这种种情绪表现仿佛一套密码,掌握了解开密码的

钥匙,就可以通向人们的内心世界。基于人类表达意图和愿望的特点,计算机也需要掌握人类情感世界的钥匙,才能主动地、智能地为人类服务。

同时,由于神经生理学、脑科学的发展,尤其是脑成像技术的运用,人类对情绪中枢的运转和情绪活动的功能有了前所未有的清晰认识,“情绪是智能的一部分”已经成为越来越多人的共识。计算机模拟的人工生命能够在进化中获得情感能力,也从另一个侧面证实了情感对生命的价值。当计算机处理的问题越来越复杂、面临的环境越来越多变时,仅仅依靠经典的逻辑计算越来越显得捉襟见肘。正如“情商”与“智商”的有机融合构成了人类的智慧,情感计算与逻辑计算的有机融合才能实现真正的人工智能。



图1 情绪短路的示意图

LeDoux 和 Damasio 两位神经科学家的研究最能说明情感在智能决策中所起的作用<sup>[3]</sup>。神经科学家通常认为眼、耳等感官传送信息到丘脑,在此被转换成脑的信号语言;从丘脑再到新皮质的相应感觉皮质区,在此我们所感知对象的信息被加工整合;新皮质将信息传送到边缘系统,边缘系统做出适当的反应,再经脑反馈到机体各部分。但纽约大学的 Joseph

<sup>\*</sup>国家自然科学基金资助项目(60433030)。张迎辉 博士研究生,研究方向为计算机视觉、模式识别、情感计算等。

LeDoux 发现,在通往皮质的大神经通路之外,还有一束小的神经元直接从丘脑通向杏仁核。这条小而短的神经通路使杏仁核可直接获得感觉输入并赶在新皮质完成感觉登记前抢先做出反应。这种现象被称为“情绪短路”。图 1 是“情绪短路”的示意图<sup>[3]</sup>。这是一种急速而轻率的反应,可以帮助人类从危险中逃脱。LeDoux 这一发现揭示了某些情绪反应可以不需要意识、认知的参与,因为从丘脑到杏仁核的捷径绕过了新皮质。情绪本身即可触发行为,而无需理智的参与。同时,爱荷华大学的 Antonio Damasio 发现前额叶杏仁核通道受损的患者,其认知能力和智商一点也没降低,做决策的能力却严重受损,就连约会这样的平常小事也困惑不定,他们会在千百种选择中摇摆不定。这与仅依靠“理智”做决策的计算机不无相似之处。作为思维与情感的交汇点,前额叶杏仁核通道是通往情绪记忆之仓的要津。剥夺了杏仁核的情绪记忆,一切事物都变得索然无味。这些患者彻底“忘掉”了以往的情绪经验。基于以上研究,Damasio 提出了反传统的观点,认为在理性决策中,情绪的参与必不可少;情绪指引方向,逻辑再纵横捭阖。在复杂问题的决策中,我们往往面临太多的选择,生活交给我们的情绪经验在决策过程的一开始就指引我们筛选一些,突出另一些,以便做出最佳选择。

事实上,科学家在对情感起源的研究中发现,情感可以帮助人类的祖先趋利避害,对生命体适应严酷的生存环境起着举足轻重的作用,是人类在进化过程中为适应环境而获得的“心理工具”。通过计算机程序的模拟,虚拟的小狗 Sniffer 能够逐代进化出对猎物的喜爱、对天敌的恐惧等情感<sup>[4]</sup>。

人工智能发展到今天,人们意识到单纯的逻辑推理无法适应复杂多变的环境。对照人类的决策过程,可以发现人脑在记忆力、计算速度等方面远逊色于计算机,却能在短时间内做出最佳决策。如前文所述,这在很大程度上得益于人类情感的参与。这也印证了 Minsky 所说的:没有情感的机器怎么能是智能的? 计算机如具有情感,无疑将大大提高其决策速度和效率。有了情感的指引,决策过程将具有更加明确的目的性和方向性,而不是在庞大的求解空间中盲目地尝试。此外,情感计算机将具有更大的自主性,在遇到突发事件时,能自动地调整算法或行为,从而主动地、创造性地完成工作。通过情感记忆库,计算机还将能够及时总结经验教训,逐步具备自主学习的能力。情感计算的研究无疑将为人工智能的发展提供一条新的途径。

可以说,计算机现在已经具有了能力很强的“大脑”,能够存储大量信息,能够高速、准确地计算;具有了“耳朵”,能够听懂人类的语音;具有了“眼睛”,能够识别别人脸等外界事物;具有了嘴巴,能够发出合成语音。情感计算就是要为计算机安一颗善解人意的“心”,使它感知、理解和表达情感,从而建立一种更为和谐的人机交互环境,实现真正意义上的人工智能。

### 3 情感计算的研究基础

由于情感问题本身的复杂性,决定情感计算是一个多学科交叉的领域。情感计算的研究不仅要借鉴计算机科学其它领域的研究成果和研究手段,还要与心理学、生理学、神经科学、认知科学等学科的研究者展开广泛的合作。

在计算机科学领域中,情感计算与可穿戴技术、计算机视觉、语音技术、模式识别、机器学习、普适计算、计算机图形技术、虚拟现实等领域的关系尤为密切。

研究情感计算,首先要有真实自然的情感数据。借助可穿戴技术,可以满足获取情感数据的自然、舒适、可移动、保护隐私等要求。

其次,为识别情感,必须从情感数据中提取最能反映情感状态的特征。面部表情特征的提取,可以借助于计算机视觉的研究成果。语音特征的提取,可以借助于语音技术的研究成果。

提取了情感特征后,就要建立起外在的情感表现与内在的情感状态之间的映射关系,这是一个典型的模式识别问题。由于情感具有多种成分和多种表达方式,使得情感表现与情感状态之间的映射关系极为复杂,多模态融合的模式识别为情绪识别提供了理论框架。

识别情感的目的,是为了进行情感交流。情感交流是人与人交流的一个重要方面,人类是计算机最好的老师。借助机器学习的手段,让计算机从人与人交流的行为中学习情感交流的技能,可以收到事半功倍的效果。

情境是推动情感产生、变化的重要外因,能够为情感识别、情感交流提供重要信息。离开了情境,情感识别和情感交流都将面临巨大的困难。这是因为,识别情感并不是情感计算的最终目的,情感计算的最终目的是使计算机理解人的需求,更好地为人类服务。需求和意图是与情境密切关联的。普适计算为情境的研究提供了很好的平台。在普适计算环境中,通过传感器、摄像头、无线网络等设备,能够随时感知人的位置、动作等情境信息。这些信息对情感计算的研究具有极大的价值。同时,情绪与需求之间有着密不可分的联系,需求的满足激发正性的情绪,需求无法满足引发负性的情绪。情感计算将识别环境中个体的情感状态,并结合情境信息推测其需求。从这个意义上,情感计算也可以推动普适计算的发展,填补现有的普适计算环境中对人类情感的关注和体贴的空白。例如,孩子的哭声往往意味着孩子的需求没有得到满足,孩子希望得到注意和照顾。

情感表达需要借助计算机图形技术生成具有真实感的人脸表情动画,借助语音合成技术生成具有真实感的语音,增加虚拟人物的情感表现力。此外,借助虚拟现实技术和设备,虚拟人在三维世界中还可以展示更丰富的神态、表情和动作。另一方面,目前虚拟世界中更多的是人与物的交互,而缺少人与人(或虚拟人)的交互。虽然出现了虚拟主持人,但他们与人工交流情感的能力非常有限。事实上,即便是在虚拟的世界中,人们仍然有交流情感的需要,虚拟人物之间也会出现喜爱、害怕、厌恶、敌对等情绪。情感计算技术如何与虚拟现实技术相结合,改变虚拟世界“冷冰冰”的状况,是一个值得研究和探讨的课题。

情感计算在计算机科学领域还是一片待开垦的沃土,但情绪心理学、生理学、神经科学、认知科学等领域早已把情绪作为研究对象。这些领域的研究成果为情感计算提供了坚实的基础,为认识人类情绪提供了依据和手段。以下简介与情感计算联系最密切的情绪心理学的部分研究成果,主要内容包括情绪心理学对情绪的定义以及情绪测量的方法。

心理学把心理现象分为“知、情、意”三个过程,即认知过程、情感过程和意志过程。认知过程是对外界刺激事件本身特性的反映。意志过程是认知活动的能动方面和自觉调解方面。情感则同人的切身需要和主观态度联系在一起。情绪可以发动、干涉、组织或破坏认知过程或意识行动;反之,对事物的评价也可以发动、转移、或改变情绪反应和体验<sup>[5]</sup>。心理学

家对情绪的定义还有不同的见解,但根据不同的定义可以概括出情绪具有如下特点<sup>[5]</sup>:

1)情绪是多成分的复合过程。它包括内在体验、外显表情和生理激活这三种成分。

2)情绪具有多维量结构。

3)情绪是生理和心理多水平整合的产物。

与情绪的成分相对应,情绪心理学对情绪的测量方法主要有主观体验测量方法、表情测量方法和生理测量方法。

采用主观体验测量方法时要求被试主动报告自己的情绪体验。报告时可以采用形容词检表或维量等级量表的形式。形容词检表包含一系列描述情绪的形容词,被试通过内省从检表中选出符合自身当时情绪状态的词汇用来确认自身的情绪体验。维量等级量表的理论基础是情绪具有多维量结构,可以视为一个在多维空间中连续变化的量。测试时要求被试标出每个维度上情绪分量的值或等级。

面部表情测量方法中最主要的是 Ekman 的“面部肌肉活动编码系统”(FACS)<sup>[6]</sup>。FACS 采用的测量单位是面容活动单元(AU)。FACS 共列出 24 种单一活动单元和 19 种复合的活动单元。FACS 制定了一个手册,它包括单一和复合活动单元引起面容变化的详细描述,并提供照片和影片相对照。FACS 是迄今为止最为详尽、精细的面部运动测量技术。只要使用单一活动单元和复合活动单元的规则,任何面部运动都能用它的元素成分做出分析。但是 FACS 只是对面部活动的测量和记录,而不是对情绪的测量。电刺激等非情绪原因也可引起面部活动的变化。

生理测量方法是以生理变化作为情绪的客观指标,其依据在于情绪可以激活自主神经系统。因此人处于情绪状态下,可以表现出许多生理反应。最主要的生理指标包括心率、皮肤电反应、呼吸的频率和深度、体内神经化学物质的分泌量或排出量<sup>[5]</sup>。

这些测量方法为情感计算提供了研究和测量情绪的手段。情绪的多成分多维特征决定了没有哪一种单一的度量能够完全揭示它的规律。研究者普遍认为,多种方法综合运用,多种指标综合分析,将是确立情绪研究方法论的有效途径。在情感计算领域,生理指标经常被采用,人们希望从中得到情绪的客观指标。有机体各机能系统的变化常常标示着各个情绪维量。例如心血管系统的某些活动是情绪唤醒维量的标志;消化腺反应可能是情绪紧张维量的参数。表情,特别是面部表情也引起了研究者极大的兴趣。但是,由于人们可能会修饰自身的表情,相同情绪体验下不同个体的表情可能有差异,单独运用表情很难测量真实的情绪。对于自然场景下的情绪,可以采用主观报告的方式测量情绪,即设计各种问卷和量表,让被试报告自己的情绪。主观报告可以反映被试真实的情绪体验,在情绪模式识别的训练阶段可以搜集到可靠的数据。

#### 4 情感计算的核心技术

要实现情感计算,形成一个完整的情感交流过程,涉及到一系列理论和技术,其核心技术主要包括:情感信号的获取、情感状态的识别、情感理解和反馈以及情感表达。

情感信号的获取是指通过各种设备记录情感信号。情感信号包括语音、面部表情、手势、站姿等体态语以及脉搏、皮肤电等生理指标。在生理指标的测量方面,传统的多导生理仪不仅形体笨重不易移动,而且测试时要在被试身上固定导线

和电极,容易干扰被试的情感。随着可穿戴技术的发展,出现了一些能够测量生理信号的可穿戴设备。例如,能够测量脉搏的手镯和测量心率的胸针、项链等。又如,集成在鼠标上的生理信号测量设备。其中,IBM 研制的情感鼠标可以测量脉搏、体温、皮肤电反应等生理指标。这些设备要么与用户形影不离,要么直接融入人机交互界面中,大大方便了情感信号的获取。值得注意的是,由于情感是一个多成分的复合过程,包括内在体验、外显表情和生理激活这三种成分,采集情感信号时往往要动用多种设备,同步记录各通道上情感信号的变化。例如,在清华大学计算机系与中科院心理所合作采集的大规模动态情感数据库中,使用多导生理仪记录了心率、脉搏、呼吸、皮肤电等生理指标,同时拍摄了被试的正面和侧面的面部表情视频,并记录了被试的语音。此外,在实验进行中和实验结束后,分别请被试填写情感的主观体验问卷。这样,就可以从生理激活、外显表情和内在体验三方面反映情感的全貌,为进一步的研究提供丰富、全面的数据。

情感状态的识别是指对采集的情感信号进行分析,识别出内在的情绪。由于情感状态是一个反映内在体验的、隐藏在多种类别的情感信号中的量,因此情感状态与情感信号之间是一个复杂的多对多的映射。在早期的研究中,往往采用演员表演的情感片段作为实验数据,而非真实的情感体验。随着研究的深入,目前研究者的重点已经转移到识别真实的情感。无论是在生活中捕捉人们的真情流露,还是在精心设计的实验中激发各种情绪体验,研究者都在努力提高情感数据的自然度。例如,MIT 情感计算研究小组捕捉并识别了多种生活情境中的情感状态,包括司机在驾驶过程中的困倦,学生在学习过程中的感兴趣、困惑、厌倦等情感状态。采集的情感数据经过情绪标注,就可以用来训练识别情感的模型。通过选择合适的模型和优化模型的参数,可以提高情感的识别率。目前,计算机不仅可以识别情感的类别,还可以计算情感的强度。不仅可以识别静态图像中的表情,还可以分析连续视频中的情感变化。多模态信息融合的技术,使计算机可以同步分析视频、音频、生理指标等多种信息,大大提高了识别的准确率。例如,MIT 情感计算研究小组综合利用面部表情、坐姿、游戏难度等三个模态的信息,识别游戏者的感兴趣、厌倦、困惑等情感,与只用面部表情识别情感相比,识别率有相当大的提高<sup>[7]</sup>。当然,情绪识别也面临一些挑战。首先,情感的变化受到情境的影响,当情境发生有利的变化时,往往会激发正性的情感,反之亦然。如何在情感识别中加入情境的因素,如何对情境进行建模,都是值得研究的问题。其次,情感表现因人而异,在情感识别中如何处理个性和共性的矛盾,也是研究者非常关注的问题。

情感理解和情感反馈是指计算机能够分析人的情感产生的原因,并对此做出合理、恰当的情感反应。例如,当用户表现出厌倦的情绪时,计算机主动提供更加新鲜有趣的内容。当用户表现出困惑的情绪时,计算机及时地鼓励用户并提供有价值的建议。情感理解和情感反馈是情感计算的核心,具有了这样的能力,计算机才能真正给人以情感的支持。目前,在有限的几种情境中,计算机被赋予了情感理解和情感反馈的能力。这些系统中,计算机的情感反馈是由设计者预先设定的,计算机根据用户的反应搜索最合适的反馈方式。这些系统既无法产生新的情感,也无法处理系统中未定义的突发状况。随着人类对自身情感产生和反馈机制的研究不断深入,计算机的情感理解和情感反馈机制将不断完善。

情感表达是指通过计算机交互设备向用户表达情感。拟人化的情感表达是目前的主流,例如通过语音、动画等方式表达情感。此外,借助音乐、色彩等,计算机也可以传递情感信息。随着情感计算研究的深入,情感表达的广度和深度都有了巨大的进步。现在,计算机可以综合运用面部表情、肢体语言、语音等模态,表达快乐、愤怒、悲伤等情感,甚至能够在一定的情境下表达出同情、困惑等细微情感。在面部图像合成领域,计算机已经能够合成出混合表情的面部图像。计算机合成的语音也不再机械单调,能够体现出快乐、愤怒、悲伤等不同情感基调。为表达情感而设计的机器人动作也出现了。MIT的类人机器人小组(humanoid robotics group)研制的机器人 Kismet 能够以眼睛、嘴唇、耳朵的动作表现快乐、悲伤、厌倦等情绪,综合了表情合成、语音合成、体态语等领域的研究成果,集中展示了情感机器人的魅力。表情、语音、体态语等外在表现在情感这一新的层次上达到了和谐统一。图 2 展示了 Kismet 的不同表情。

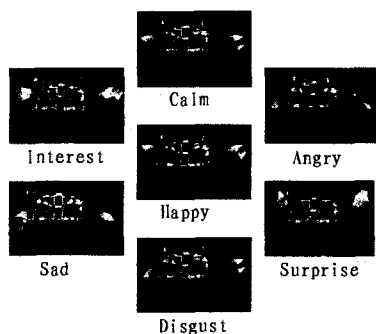


图 2 拟人机器人 Kismet 的不同表情

**结束语** 在情感计算中,计算机领域的科学家第一次把目光投向了人类的内心世界。“认识你自己”从来就是一个巨大的难题,近代科学发展的大部分成果是在认识和改造外部世界中取得的,而对人类本身的认识还面临着许多难解之谜,这也正是情感计算所面临的困难。人类追求和谐人机交互的巨大需求和不畏惧任何挑战的决心则是推动情感计算发展的巨大动力。

心理学、生理学、认知科学以及计算机科学相关领域的研究为情感计算的研究提供了理论基础和工具方法,情感计算的提出也为这些学科的发展提供了很好的机遇。

情感计算一提出,就受到了学术界和企业界的广泛关注。瑞士政府成立了情感科学国家研究中心<sup>[9]</sup>。心理学、神经科学、哲学、历史学、经济、社会、法律等多学科合作开展情感计算的研究与应用。MIT 情感计算研究小组开展了多方面的研究<sup>[10]</sup>。例如,开发可穿戴设备,识别真实情境中的人类情感,研究人机交互中的情感反馈机制,研制能够用肢体语言表达情感的机器人。英国伯明翰大学开展了“认知与情感的研究”<sup>[11]</sup>。IBM 公司实施的“蓝眼计划”<sup>[12]</sup>,可以使计算机未卜先知。当你看地图时,计算机可以找到你感兴趣的地点,并主动介绍其风土人情。

情感计算的研究在很多方面取得了进展,尤其是在改善人机交互的和谐性上,得到了用户的认可。MIT 的情感计算

小组研究了情感支持对用户的影响<sup>[13]</sup>。实验结果表明,在用户产生挫败感时,计算机如能给予情感支持(例如主动倾听、对用户的遭遇表示同情甚至歉意),将降低用户的挫败感。以此为基础,MIT 情感计算小组设计了一个情感对话系统:虚拟人“Laura”可以通过文字界面与用户交流锻炼身体的心得。锻炼者完成目标时,Laura 会赞扬他;若锻炼者未完成目标,Laura 则会鼓励他。经过一段时间的“相处”后,大多数用户都更主动地锻炼并愿意与 Laura 继续交流,甚至有人认为 Laura 非常有魅力。

情感计算的进步为我们展现了无限的前景。在可预见的未来,情感计算在教育、娱乐、艺术等领域将得到广泛的应用。学生在学习过程中会体验到兴趣、困惑、沮丧、快乐、厌倦等情绪。有经验的教师会观察学生的情绪,及时调整学习内容并鼓励和引导学生克服困难。如果计算机能够识别学生的情绪,并做出适当的反馈,它也能成为一位循循善诱的老师。在电子游戏领域,通过可穿戴设备和游戏手柄等外设,游戏软件可以感知游戏者的情感状态。这些信息可以帮助设计者了解游戏者的偏好。同时,游戏者的化身可以通过一些动作来表现这种情感变化,使游戏更逼真。例如,游戏者吃惊时,化身会猛然向后跳去。人们往往很难明确说出自己的喜好,但在看到自己心仪的设计时,情感会准确无误地告诉我们“这就是我想要的”。如果计算机能够在人们浏览设计方案时,记录其情感反应,自动分析用户的偏好,无疑将大大提高设计的效率和满意度。随着情感计算研究的深入,将不断有新的应用被发掘出来,人们将在日常生活中享受情感计算带来的便利和体贴。

情感计算的发展必将极大地改变人类的生活,自然和谐人机交互的梦想离我们会越来越近。

## 参考文献

- [1] Minsky M. The society of mind. Simon & Schuster, 1985
- [2] Picard R W. Affective computing. Cambridge, MA: MIT Press, 1997
- [3] Goleman D. 情感智商. 耿文秀,查波,译. 上海:上海科学技术出版社, 1997
- [4] Johnston VS. 情感之源. 翁恩琪,等译. 上海:上海科学技术出版社, 2002
- [5] 孟昭兰. 人类情绪. 上海:上海人民出版社, 1989
- [6] Ekman P, Friesen W. Facial action coding system (FACS): manual. Palo Alto: Consulting Psychologists Press, 1978
- [7] Mota S. Automatic posture analysis for detecting learner's affective state [D]. Cambridge, MA: MIT, 2002
- [8] <http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html>
- [9] <http://www.unige.ch/fapse/emotion/>
- [10] <http://affect.media.mit.edu>
- [11] <http://www.cs.bham.ac.uk/research/projects/cogaff/cogaff.html>
- [12] <http://www.almaden.ibm.com/cs/BlueEyes/index.html>
- [13] Bickmore T. Relational agents: effecting change through human-computer relationships [D]. Cambridge, MA: MIT, 2003