

基于语义分析的情感计算技术研究进展^{*}

饶元 吴连伟 王一鸣 冯聪

(西安交通大学 软件学院, 社会智能与复杂数据处理实验室, 陕西 西安 710049)

通讯作者: 饶元, E-mail: yuanrao@163.com

摘 要: 随着机器学习和大数据技术的应用发展, 基于语义分析的情感计算与分析技术在研究人类的感知、注意力、记忆、决策、社会交流等诸多方面起着重大作用, 它不仅影响到了人工智能技术的发展, 还影响到了人/机交互的方式, 并受到学术界以及企业界的广泛关注。本文在针对情感定义以及相关 90 多种情感模型分析的基础上, 归纳并提出了目前情感分析领域中存在的 6 项关键性问题与挑战, 其中主要包括: 情感的来源与本质特征的表示问题; 多模态条件下的情感计算问题; 外部因素对情感演化过程的影响度量问题; 情感的个性化度量问题; 情感群体化特征与传播动力学机制问题以及细微情感的表达、算法改进与优化等问题。同时, 本文针对其中的关键问题与技术挑战进行了理论探讨、技术分析、实际应用以及当前工作进展与趋势分析, 从而为深入研究和解决基于语义分析条件下的情感计算提供了新的研究线索与方向。

关键词: 情感语义分析; 人工智能; 情感演化计算; 群体情感; 人工情感生成; 观点挖掘

中图法分类号: TP311

中文引用格式: 饶元, 吴连伟, 王一鸣, 冯聪. 基于语义分析的情感计算技术研究进展. 软件学报.

英文引用格式: Rao Y, Wu LW, Wang YM, Feng C. Research on Emotional Computing Technology Based on Semantic Analysis. Ruan Jian Xue Bao/Journal of Software, 2016 (in Chinese).

Research on Emotional Computing Technology Based on Semantic Analysis

RAO Yuan WU Lian-Wei WANG Yi-Ming FENG Cong

(Lab of Social Intelligence & Complex Data Processing, School of Software Engineer, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: With the development of Machine Learning and the application of Big Data, the semantic-based emotional computing and analysis technology plays a significant role in the research on the human perception, attention, memory, decision-making, social communication and so on. It not only affects the development about artificial intelligence technology, but also affects the human/machine interaction and smart robot technology, which is concerned widespread by the academic community and the business community. Based on the definition of affective and the analysis of more than 90 emotional models, there are five vital problems and challenges about affective computing are summarized and put forward in this paper as follows: where is emotion stem from and how to represent their essential features; how to analyze and compute the emotion under the Multi-model environment; how to measure the influence of external factors on the process of emotional evolution; how to measure individual emotion by various of personalized characteristic; how to measure the crowd psychology and emotion and to analyze the mechanism about propagation dynamics; how to express the subtle emotion and optimize algorithms, and so on. Meanwhile, we put forward some theoretical research, technical analysis, practical application to introduce the current work progress and trend for these technical challenges, which provides new research clues and directions for further study in the field of the semantic-based emotional computing.

Key words: Semantic-based Emotion Computing; Artificial Intelligence; Affective Evolutionary Computing; Crowd Emotion; Emotion Generation

^{*} 基金项目: 国家自然科学基金 (F020807), 教育部“云数融合科教创新”基金项目 (2017B00030), 中央高校基本科研业务费 (ZDYF2017006), 陕西省科技厅协同创新项目 (2015XT-21) 以及陕西省软科学重点项目 (2013KRZ10) 的联合支持

Foundation item: National Natural Science Foundation of China (F020807); "Cloud Number Fusion Science and Education Innovation" Fund Project of the Ministry of Education (2017B00030); Fundamental Research Funds for the Central Universities (ZDYF2017006); Collaborative Innovation Project of Shaanxi Science and Technology Department (2015XT-21); Key Projects of Shaanxi Soft Science (2013KRZ10)

收稿时间: 0000-00-00; 修改时间: 0000-00-00; 采用时间: 0000-00-00; jos 在线出版时间: 0000-00-00

CNKI 在线出版时间: 0000-00-00

1 概述

自从美国 MIT 大学的 Minsky 教授在《The Society of Mind》^[1]一书中首次明确提出了智能机器的情感识别问题后,引起了人们的广泛关注与遐想^[2];随后,情感进一步被认为是由主观体验^[3](Subjective Experience)、外部表情^[4](Emotion)以及生理唤醒^[5](Physical Arousal)三个核心要素组成,其中主观体验反映了个体对不同情感状态的自我感受;而外部表情则反映了不同情感状态下的肢体动作和行为的可量化形式,包括了面部表情、姿态表情和语调表情等;而生理唤醒则是反映了情感产生时对人的生理反应以及特征的激活水平。为了进一步对情感进行有效的度量, Picard (1997) 在《Affective Computing》一书中正式提出了“情感计算”的概念^[6],并指出情感计算是关于人类情感产生、情感识别、情感表示以及影响情感因素度量等方面的计算科学,它利用计算机技术实现了信息载体(如生理特征、文本中的词汇、声音、视频图像等)与人类情感的极性倾向(褒义、贬义或中性)以及强度之间的关系度量,从而促进了情感计算科学的正式诞生,并推动了机器人与机器视觉、人工交互设计、语音识别和文本分析等领域的发展。

近年来,随着人工智能技术、深度学习算法以及脑科学与认知科学的快速发展与进步,人们发现:情感是智能的一部分,而不是与智能相分离的^[7]。因此,如何能够从智能计算中建立情感表示与情感度量之间的关系,并通过赋予计算机识别、理解、表达和适应人的情感的能力来建立一个和谐的人机环境,使计算机具有更高的、全面的智能则成为了目前工业界与学术界共同关注的焦点问题^[8,9]。李飞飞^[10]也进一步指出,情感计算将会是人工智能领域下一步重要的突破与发展方向,但由于情感计算目前面临的巨大技术挑战,本文希望能够通过对近年来情感计算领域中存在的问题以及研究进展进行分析和综述,为本领域内的研究人员廓清情感计算的整体发展状况与趋势,为未来的研究工作梳理新的思路并奠定基础。

2 情感计算的核心要素模型与挑战

2.1 情感的基本概念

情感是人类的一种主观意识,它是人脑对某一种客观存在的主观反映,而情感计算则与情感相关,是源于情感或能够对情感施加影响的计算^[11]。例如,在生活中,人们会随时随地的受到喜、怒、悲、惧等情绪起伏变化的影响,这种源自于情感的变化是人类行为的灵活性、决策自主性和思维创造性的根本来源。与此同时,人们也需要对一些客观的事物或事件进行主观且具有倾向性的评价,从而反映出对客体的情感所施加的影响和作用。因此,目前情感分析主要分为情绪的极性化分析计算(Emotion Analysis)与观点倾向性分析计算(Sentiment Analysis)两个方面的问题。

针对前者,以 MIT 的 Picard 教授为代表^[12,13],其研究主要关注于通过情感信息的采集(如生理应激反映的特征采集)、情感识别分析、情感理解认知以及情感信息的表达来实现对情感的特征分类以及情感极性强度的计算和度量;而针对后者则是以 UIUC 大学的 Bing Liu 教授为代表^[14,15],其研究主要关注于主体对特定客体的观点或倾向特征,其中,客体不仅包括了产品、服务、组织、个人、焦点问题、事件和主题等,同时也包括了某一个客体中的具体属性特征。通过主体对客体的情感、评价、评估、态度和情绪的分析与计算,来反映出主体对客体的反对、支持或中立的态度、立场和主观倾向。下面本文在建立情感基础概念模型的基础上,针对情感计算过程中的一些关键问题、核心技术以及挑战进行深入研究和综述分析。

2.2 基于情绪的情感计算的核心要素模型

根据 Picard 提出的基本概念^[16],情感计算可形式化地采用一个三元组来表示,即:

$$SC = \langle S, C, W \rangle \quad (1)$$

其中, S 表示信息载体 (Semantide) 所形成的集合,即 $S = \{s_1, \dots, s_i, \dots, s_n\}$; 这些信息载体主要包括了: 文本、音频、视频以及图像等不同类型。由于不同信息载体所包含的信息具有不同的数据特征,特别是情感特征,而这些特征的表示方式存在着明显的差异,因此,在传统的情感计算中,往往只针对某一个特定的载体类型(即单模态)来进行内容的情感分析,如文本的情感、音频的情感或者是视频的情感分析等,但是随着人们对情感特征理解的加深,多种信息载体相互融合的多模态条件下,如何进行情感分析与计算则为了目前研究中的一个热点^[17]。

C 表示情感类别或者是不同倾向分类所形成的集合,即 $C = \{c_1, \dots, c_k, \dots, c_v\}$; 该方式可以表示情感的离散特征,例如: 积极与消极,或者是积极、中立和消极,也可以进一步分为喜、忧、哀、乐、爱、恶六种类型等等,甚至可以利用这些基本的情绪组合出一些更复杂的情感。因此,根据不同的应用目标,可以将情感的特征进行二分类或者多分类,从而形成不同的情感分类模型; 这些分类模型直接反映出了人们对情感粒度与特征的基本认知,但是由于忽略了情感特征在时间维度上的连续变化特性,也会存在着一些明显的不足与缺陷,本文在后续章节将会对离散和连续这两类情感模型进行进一步的分析。

W 表示不同情感特征强度所形成的集合, 即 $W = \{w_1, \dots, w_k, \dots, w_v\}$; 一般的强度可以分为高, 中, 低三个维度, 也可以进一步分为极高、高、中、低、极低五个维度等等, 这些强度特征往往与情感特征相结合, 构成了情感计算的核心与基础。

通过上述定义, 情感计算可以表示为通过对信息载体中数据特征的获取、识别和度量, 实现内容情感特征在不同极性维度下的计算, 并形成了对客观存在的事物进行的主观情绪判断与度量。因此, 情感计算可以表示为上述三个要素集合之间的笛卡尔积所形成的状态空间组合, 即:

$$SC = S \times C \times W \quad (2)$$

基于上述情感的三维状态空间, 不同的特征组合将导致不同的情感特征和不同的情感计算模型, 也导致了不同的计算与评价方法。

2.3 情感计算目前面临的核心问题与挑战

由于上述情感特征的分类较为简单, 一方面无法考虑到人类丰富情感的本质; 另一方面, 由于多种因素和问题相互交织和影响, 使得情感计算目前仍然面临着许多核心的问题以及关键的技术挑战, 这些问题与挑战主要如下:

挑战 1: 情感的来源与本质特征表示。传统的情感识别模型主要是基于不同特征的模式分类, 但在情感计算过程中, 不同的情感分类特征均反映出情感的本质来源特征。作为人工智能的一部分, 如何通过脑科学、认知科学、心理学等交叉学科的研究成果, 来深入理解情感的产生来源以及情感认知的本质特征, 已成为情感识别与计算研究的核心问题, 相关的理论突破将会改变、甚至颠覆情感计算的基本方法, 从而成为目前脑科学、认知科学与人工智能等领域中研究的前沿^[18]。

挑战 2: 多模态条件下的情感计算问题。根据式 (2), 传统的情感计算研究, 无论是采用传感器来获取人的生理情感特征, 还是通过文本、视频、图像或者音频等来获取不同的情感特征, 均是在某一种特定的数据类型条件下来开展的, 即属于单模态条件下的情感计算, 在这种模式下, 情感特征的获取方式与特征的表示方式相对较单一。而在实际情况下, 情感往往会涉及到不同的数据来源和各种相关因素的相互影响, 即多模态下的情感计算, 通过对不同模态的情感特征进行描述与融合, 丰富了情感度量的上下文语义以及细微情感表达的复杂程度, 如何从多模态的角度进行情感信息的融合、识别与理解, 实现自然和谐的人机交互已成为了目前一个关键性的技术挑战与研究热点^[19]。

挑战 3: 外部因素对情感演化过程的影响度量问题。由于情感更倾向于人对某种社会需求在欲望上的态度和体验, 而这种体验不仅可以采用离散化的分类方式来表达, 例如 Oatley 曾提出了快乐 (happiness)、焦虑 (anxiety)、悲伤 (sadness)、愤怒 (anger) 和厌恶 (disgust) 五类经典的基本情绪模型^[20], 而且通过这些基础情感进行组合和派生可以生成其它新的情感; 另外, 随着时间以及外部环境的动态变化, 人的情感生理特征在激活、表情展示以及内在体验感受等方面均会产生一些影响和变化, 这种连续的情感变化往往可以采用多维空间模型来进行度量。但是, 如何建立一个符合人类情感规律并适用于机器实现的统一人工情感自动生成模型^[21], 来表达和度量情感与认知随时间或者是外界环境变化而自适应演化的规律则成为了目前一个巨大的挑战。

挑战 4: 情感的个性化度量问题。在情感识别与认知的过程中, 一些跨文化、跨区域、不同性别、不同年龄或者不同教育程度等个性化特征的差异对情感的识别与度量将产生显著影响。因此, 根据个性化的情绪影响因素^[22], 来建立一个个性化的智能推理模型以及持续改进算法, 实现真正意义上的拟人情感推理, 这将对情感计算、基于自动问答的人机交互与机器人等学科的发展起到重大的推动作用。

挑战 5: 情感群体化特征以及传播动力学机制问题。由于每个人都具有社会化的群体性属性和特征, 在群体聚集的环境下, 情感具有非常明显的传染或者是传播的特征, 情感通过传播使得群体的情感趋向于达成某种一致, 并成为了群体的共性特征。而考虑到社会网络环境下的群体涌现行为, 通过对个体情感特征的度量, 如何来实现对群体情感特征的影响, 并使得个体情感在社会化传播过程中的传播动力学机制^[23], 这将对社会计算领域中的新闻传播、舆情监控、广告推送以及金融风险评价等领域带来颠覆性作用。

挑战 6: 基于细微情感表示的情感计算算法持续改进问题。情感计算不同于一般意义上的模式分类问题, 特别是情感信息容易受到外部环境、生理、心理、文化背景、语境、语义等因素的影响, 如何准确地描述、表示以及计算细微的情感特征也是目前情感计算过程中的一个关键的挑战。随着深度学习算法的不断优化, 人们发现深度学习算法不仅需要采集大量的数据信息, 同时计算的过程无法有效理解, 为了实现小样本条件下持续学习方法, 越来越多的人开始关注并利用对抗生成网络 (GAN)^[24]以及迁移学习方法^[25]来改进计算的过程, 同时, 在进行人脑模拟与研究中, CMU 的 Tom Michel 和 UIUC 的刘兵分别提出了“Never Ending Learning”^[26]和“Lifelong Learning”^[27]持续学习算法框架, 这些新算法的不断提出, 为对情感计算中的问题解决提供了新的方法和机遇。

综上, 通过对情感计算领域中存在的情感来源与本质特征进行表示与理解, 本文将围绕上述围绕情感计算领域中的核心问题与关键技术挑战, 特别是针对多模态下的情感特征融合与处理、在特定时空特征条件下的情感演化行为、情感的个性化度量、群体情感特征以及情感传播动力学机制等关键性问题进一步深入分析并综述目前最新的技术成果与研究进展。

3 情感的来源与本质特征的识别与表示

3.1 情感来源研究

人们发现仅依赖于单纯的逻辑思维与推理能力无法适应外部多变且复杂的环境,借助于神经生理学、认知学和脑科学的快速发展,尤其是脑成像技术的运用,纽约大学的 Joseph LeDoux 发现了“情绪短路”的现象^[28],即在通往皮质层的神经通路之外,还有一束小的神经元直接从丘脑通向杏仁核,这条神经通路可将情绪感觉信息输入到杏仁核并在皮质层获取感觉前率先做出反应,这表明一些情绪本身可以直接触发行行为和反应,而不需要意识和认知的参与。这一发现不仅证明了情感也是智能的一部分,同时,也证明了通过情感因素的辅助判断更容易获得最佳的决策结果,即情绪不仅参与到了理性的决策过程之中,并且通过情绪指引方向可以获得复杂问题求解的最佳决策。一旦计算机具有相应的情感处理能力,就则可以大大降低目标求解空间中的复杂程度,并极大地提高计算的效率与能力。

为了进一步对情感来源进行度量,《心理学大辞典》将情感定义为:人针对客观事物是否满足自己的需要而产生的态度体验。这个定义从“需求满足的角度”来对情感进行阐述,并反映出情感的某种价值倾向。在心理学基础理论的基础上,为了对情感进行数学建模与计算机表示,人们先归纳出了情感强度计算的三个关键定律^[29]:

定律 1: 情感是事物的价值率高差在人的头脑中的主观反映值,且事物对外界刺激产生的情感强度与其价值率的高差成对数函数关系,即:

$$\mu = k_m \log(1 + \Delta p) \quad (3)$$

其中, μ 表示情感强度值, k_m 表示强度系数, Δp 为价值率高差。从上式中可见,价值率高差根本上决定了人对该事物的基本“立场、态度、原则和行为取向”,决定着人对该事物的价值投入方式和投入规模,并形成了具有主观意识特征的情感,但是情感的强度并不与事物的价值率高差成正比,而是存在一种函数关系。

定律 2: 人对事物的情感强度随着人对该事物的作用规模的增长而下降。

该定律反映了情感强度的边际效应递减特征,人对于一般事物的情感强度会随着时间不断增长而不断衰减,并且当情感强度 μ 足够小时,情感强度的衰减速度 $d\mu/dT$ 与情感强度 μ 成正比,即

$$d\mu/dT = -K_i \times \mu \quad (4)$$

定律 3: 情感强度与时间成负指数函数关系,即:

$$\mu = \mu_0 \exp(-K_i T) \quad (5)$$

其中, μ_0 为初始情感强度, T 为时间, K_i 为情感强度衰减系数。

上述三个定律从价值率的角度对情感进行了数学建模和计算,虽然从公式中可以获取情感强度与时间之间的函数关系,但是并没有考虑到情感在“态度表达”上存在的多样性特征,例如:高兴、厌恶、生气、吃惊、害怕和伤心等情感的基本类型^[30],因此,Ortony 等人尝试在对情感进行分类的基础上,通过构造一组情感函数来建立一个 OCC 情感评价模型^[31],并假设 $D(p,e,t)$ 表示对象 p 在时刻 t 想做事件 e 的期望程度,若事件期望产生了有益的结果,则 $D(p,e,t) > 0$,否则, $D(p,e,t) < 0$ 。进一步假设 $I_g(p,e,t)$ 表示情感强度变量的组合, $p_j(p,e,t)$ 表示产生某种情感状态的可能性大小,则产生该情感度量的规则可定义为:

If $D(p,e,t) > 0$;

then $p_j(p,e,t) = F_j(D(p,e,t), I_g(p,e,t))$

其中, $F_j(\)$ 是情感表示函数,当情感发生的概率大于某一阈值 T_j 时,则情感强度 $I_j(p,e,t)$ 有:

If $p_j(p,e,t) > T_j(p,e,t)$

Then $I_j(p,e,t) = p_j(p,e,t) - T_j(p,e,t)$

Else $I_j(p,e,t) = 0$

从该模型可以看到情感的产生存在着一个特定的激活阈值,只有大于这一个阈值时,才有可能实现情感的激活。孟昭兰进一步研究发现^[32]:基本情绪在个体生活中首次显现的时间是不同的,并且存在一定的时间顺序,这种顺序服从于婴儿的生理成熟和适应需求,这也反映出了情感的某种价值性倾向与规律。另外,从实际案例研究中发现,情感的强度也并非总是呈现出随时间相关的负指数衰减关系。例如,恐惧和悲伤等情感在触发事件发生后的情感强度随着时间而快速的衰减,但爱、恨、情、仇等情感的强度却可以持续多年不仅无衰减,甚至会随时间的流逝而增强。

为了进一步探索情感的多样性与维度特征, Picard 等人^[11]通过不同的传感器来获取人的表情、姿态、手势、语音、语调、血压、心率、皮肤电反应、呼吸频率和深度、体内神经化学物质的分泌量或排出量等各种特征数据,并通过当时的环境、语境、情境等上下文信息,来理解和识别人的情感,研究发现不同类型的情感不仅与生命体的基本生理特征息息相关,同时也

与不同的文化传统有关,但目前情感来源的机理仍然没有一个明确而且统一的理论能够进行完美的解释,仍然存在着大量的研究问题与挑战。

3.2 经典的情感模型分类与细微情感表示

Ekman 曾通过跨文化的研究将情感划分为:高兴、厌恶、生气、吃惊、害怕和伤心六种基本类型^[30],并且每类情感都存在一个阈值,仅当情感强度超过该阈值时,情感才会对一些外在行为起作用。在此基础上,人们从不同的角度提出了许多不同的情感模型来实现情感的度量与计算,下面首先从离散与连续性的情感特征表达中,对最具有代表性的模型进行梳理,并从这些模型中来进一步分析细微情感的处理方法。

从离散特征的角度,可以将情感分为主动(趋近)和被动(退避)两大类,并在此基础上来进行细化分析。其中, Fox 将 6 个维度的情感特征分解为 3 个层次共 18 个细分特征^[33]; Ortony 等人从认知评价的角度提出了情感 OCC 模型,并细化出了 22 个细分的情感特征^[31]; Clark Elliot 在考虑到社会关系后,开发了一个基于 Agent 的情感推理器,将 OCC 模型的情感类型扩充到了 26 种^[34];德国人工智能研究中心(DFKI)提出了基于 Sentibank 视觉情感分析框架模型,并建立了一个具有 24 种情感的分类特征^[35]。Plutchik 指出情感识别是对于认知和行为的处理与反馈的一个过程,他将情感定义为 3 组、8 个维度共 24 个情感特征,并且不同维度之间的区域表示两种情感之间的混合的过渡情感,扩展出了 8 个过渡的情感特征,从而形成了一个具有 32 个细分情感特征,并能够对部分连续的情感进行表示与度量的经典模型,图 1 则是在经典的 Plutchik 情感模型基础上进行扩展的具有 72 个细化特征的情感轮模型示意图^[36]。

另外,在人机交互情感网模型中(HUMAINE, Human-Machine Interaction Network on Emotion),通过情感标记表示语言 EARL (Emotion Annotation and Representation Language)将情感分成了 48 个细分的类型^[37]。在 Shaver 等人定义的树状情感结构的基础上, Parrott 进一步将情感定义成了具有三层分类结构共 115 个细分的情感特征的情感模型^[38]。

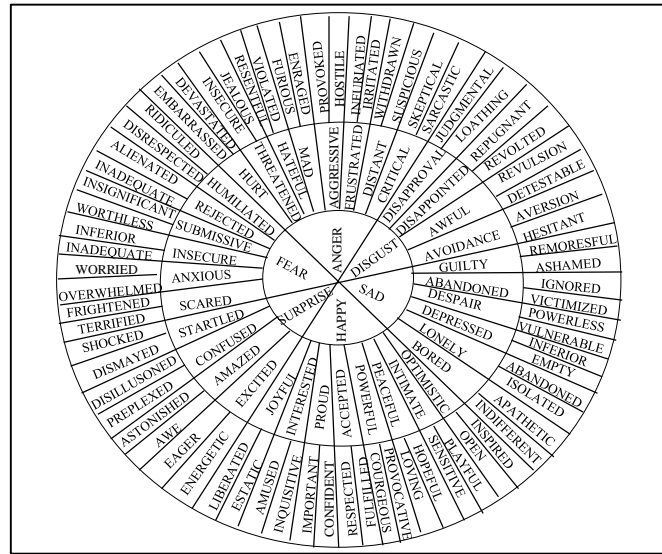


Fig.1 Extended Plutchik emotion wheel model sketch map^[36]

图 1 扩展的 Plutchik 情感轮模型示意图^[36]

Tiffany^[39]进一步列出了 154 个不同的情感细分类型, Steven J. DeRose^[40]通过文献整理提供了 800 多个英文情感词来实现对情感细化分类,上述工作通过对情感词所对应的情感特征的细分来实现了对于复杂情感的分类与测量。为了有效地标记和描述这些离散的情感特征,将人们对多个不同事物或是某一事物中不同属性的情感特征定义为一个情感矢量 M , 并表示为:

$$M = \{\mu_1, \mu_2, \dots, \mu_n\} \quad (6)$$

其中, μ_i 是指人对于第 i 个事物特征所产生的情感。如果当该事物又包含了若干更具体的子事物或者属性时,则用 μ_{ij} 表示某人对第 i 个分类下的第 j 个子事物特征的情感描述,从而可采用一个二维的情感矩阵表示:

$$M = \{\mu_{ij}\}_{n \times m} \quad (7)$$

式中, n 表示分类事物的数目, m 表示每一个分类下的子事物或属性特征的数目。

由于语言所具有的模糊性特征,将情感与语言进行关联时,可以通过语言标签来实现情感的分类和标识。这种离散枚举的方式一方面很难描述和定义一些复杂的情感;另一方面,也无法有效度量每一种情感的强度。特别是在现实生活中,人的

情感突变往往较少发生,情感主要表现为在时间或空间上连续性的变化过程,在这种变化过程中细粒度情感特征的表示较为复杂,因此,需要采用新的方式来进行建模和计算。

为了定义情感的这种连续变化特征,Plutchik 通过两个不同的基础情感特征的融合形成了复杂的情感特征,使得情感的描述具有一定的连续性,但这种方式无法对复杂的组合情感的强度进行有效度量。Thayer 等^[41]则根据压力与能力两个维度将情感定义为:快乐—焦虑和平静—活力四个象限,并对空间中的任意一个点可以进行情感的粗粒度的划分;Whissell^[42]也将情感定义为在评价与激活二维连续空间中的一个点,其中,评价维表示人的正向与负向的情感,而激活维表示人在特定情感状态下主动或被动的行为,并且定义和标记了近 9000 个<评价,激活>词对所形成的情感词典。Li 等^[124]基于一个情感种子词典,通过将 Word Embedding 与支持向量回归算法相结合来学习分布在 Russule 模型中不同效价度和激活度两个维度上的情感词,在此基础上,通过计算一个新词在这两个维度上的度量值大小来实现不同粒度情感词的分类以及种子词典的扩充。

罗素和兰格也认为不同的情感之间存在着关联并且呈连续性变化,并进一步将情感定义为情感空间中的一个连续存在的向量,向量的方向角反映了不同的情感,而向量的模则反映了情感的极性大小。在此基础上,考虑到情感强度、不同用户群体以及时间等因素的影响和作用,情感往往可以采用不同维度的空间模型来进行表征。Mehrabian 等^[43]采用愉悦度(即极性 P)、激活度(即生理激活水平 A)和优势度(即个体对情景和他人的控制状态 D)3 个维度组成的 PAD 空间模型来测量和表示具体的情感,并且每个维度用-1~+1 来表示情感极性的值,例如愤怒的情感坐标值为(-0.51,0.59,0.25)。由于 PAD 模型中的三个维度对情感的主观体验、外部表现和生理唤醒这三个特征都有一定的表征,因此,得到了较为广泛的应用。例如,Gao 等^[125]发现微博中有许多词表达着不止一种的情感,因此,不能只采用一种情感标签来标记,于是利用 PAD 模型将情感词投影到三维的向量空间中,通过极性强弱来实现细粒度情感的区分与表示。

为了进一步描述情感连续变化特征,人们从低维情感空间逐步向多维情感空间模型的方向发展。在经典的<激活度、效价、优势度>、<评价、激活、能量>和<愉悦度,激活度,优势度>三维情感空间^[43]的基础上,Fontaine 等人提出了一个<效价,潜能、激活度、不可预测性>四维的情感空间^[44]。通过维度空间不仅可以表达出更为有效的情感追踪,同时,对随时间变化的连续情感状态的处理也将更加有效。但是由于情感维度空间中采用情感词之间的空间距离来进行计算比较,这对于一些复杂的情感词而言,无法实现有效的度量。为了解决这一问题,Erik^[45]在 Plutchik 模型的基础上,提出了一个基于<愉悦度,关注度,倾向度,敏感度>四个维度、三个极性强度,且具有 24 个细化特征的沙漏情感模型(Hourglass of emotion),其中,沙漏的竖轴表示了情感的极性强度大小,并且情感的衰减过程满足于指数分布(如图 2 所示)。

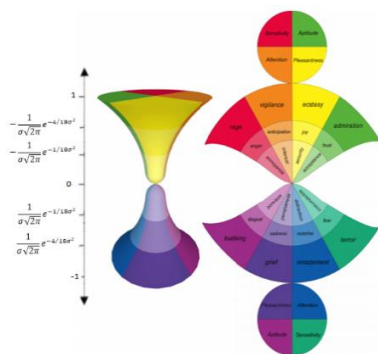


Fig.2 Schematic diagram of Hourglass model structure

图 2 Hourglass 模型结构示意图

从上述不同情感模型的演化过程可以发现,情感的表示一方面从简单的二元分类模型(积极与消极)逐步演化向具有更高维度特征的复杂细粒度情感表示方向发展,另一方面也正在从离散的情感特征标签向连续的情感度量方向深化;特别是在人机交互以及情感自动生成的过程中,如何实时准确地识别和区分出不同细粒度情感之间的差异,则需要在上述模型基础上,探索新的有效识别方法。例如,Wang^[126]等利用深度学习的方法来研究图像中细微情感的识别与分类,在人工标注的情感图像训练集的基础上,将形容词与名词组合成一个 DCAN 神经网络来训练图像中细微情感的分类,然后,通过预测网络来识别细粒度情感分类下的差异。

与细粒度情感相对应,人们还存在着一个微表情(Micro-expression),即一种持续时间在 0.5s 以内的快速的面部表情^[140],它能够提示出人们在面对高风险、尤其是带有一定恶意性质的事件时试图隐藏真实情绪或意图的行为,因此,微表情能够在测谎以及危险行为检测等诸多应用领域中提供必要的线索,由于微表情持续时间短、强度低、动作区域少,难于肉眼

识别, 高强度训练的人也只有 47% 的识别率, 目前大量的研究工作聚焦到如何在传统算法的基础上, 结合深度学习的方法对微表情的特征差异进行识别、度量以及表达, 相关工作仍然需要进一步地深入研究^[14]。

综上, 根据离散与连续的情感模型中存在的差异, 以及不同粒度下的情感特征进行比较的结果如表 1 所示。

Table 1 Comparison of characteristics of different emotional models

表 1 不同情感模型的特征比较

分类	模型	情感特征
离散型	Ekman 模型	高兴, 厌恶, 生气, 吃惊, 害怕, 伤心
	Fox 模型	趋近型: 愉快 (骄傲、极乐), 兴趣 (关心、责任感), 愤怒 (敌意、嫉妒) 回避型: 忧伤 (痛苦、剧痛), 厌恶 (轻蔑、怨恨), 害怕 (惊恐、焦虑)
	OCC 模型	欣喜, 不平, 庆幸, 同情, 希望, 担心, 满足, 恐惧, 释怀, 失望, 高兴, 悲伤, 自豪, 自责, 赞赏, 责备, 感激, 愤怒, 满意, 悔恨, 爱, 厌恶
	Elliot 模型	在 OCC 基础上加入: 羡慕, 嫉妒, 喜欢, 不喜欢
	Plutchik 模型	第一层: 狂喜, 钦佩, 恐惧, 惊奇, 悲痛, 厌恶, 愤怒, 警惕 第二层: 快乐, 信任, 害怕, 惊奇, 哀伤, 嫌恶, 气忿, 预期 第三层: 宁静, 接受, 忧虑, 错乱, 沉思, 无聊, 烦恼, 兴趣 混合情感: 乐观, 热爱, 顺从, 敬畏, 不赞同, 悔恨, 蔑视, 侵略性
	Parrott 模型	第一层: 喜欢, 快乐, 惊奇, 生气, 悲哀, 恐惧 第二层: 喜爱, 欲望, 渴望; 欢呼, 热心, 满足, 骄傲, 乐观, 沉迷, 安慰; 惊奇; 易怒, 激怒, 愤怒, 厌恶, 痛苦; 受难, 悲哀, 失望, 羞耻, 忽视, 同情; 惊骇, 紧张不安; 第三层: 崇拜, 喜爱, 喜欢, 吸引力, 关怀, 亲切, 同情, 多愁善感; 希望, 激情, 迷恋; 渴望; 娱乐, 极乐, 快乐, 欢喜, 高兴, 快活, 喜悦, 欢庆, 得意, 狂喜, 欢快; 热衷, 热情, 兴奋, 激动, 愉快; 满意; 得意; 热切, 希望; 沉迷, 狂喜; 宽慰; 惊奇, 惊讶; 恼怒, 烦恼, 抱怨, 粗暴, 暴躁; 挫败; 生气, 愤怒, 暴怒, 激怒, 敌对, 残暴, 痛苦, 仇恨, 蔑视, 怨恨, 报复, 讨厌, 愤恨; 厌恶, 蔑视, 憎恨; 嫉妒; 痛苦; 烦闷, 极其痛苦, 伤害; 抑郁, 绝望, 沮丧, 忧郁, 不快乐, 悲伤, 悲痛, 悲哀, 不幸, 忧愁; 沮丧, 不悦; 内疚, 后悔, 悔恨; 疏远, 挫折, 忧郁, 尴尬, 乡愁, 羞辱, 不安全, 侮辱, 孤立, 寂寞, 拒绝; 怜悯, 哀叹, 同情; 惊恐, 震惊, 害怕, 恐惧, 可怕, 恐怖, 恐慌, 歇斯底里, 羞愧; 焦虑, 悬念, 拘束, 害怕, 担心, 悲痛, 恐惧
连续型	Thayer 模型	压力: 快乐-焦虑; 能量: 平静-有活力
	罗素兰格情感模型	罗素情感: 唤醒, 兴奋, 愉快, 满足, 困乏, 沮丧, 痛苦, 忧虑 兰格模型: 喜欢, 愉快, 兴奋, 恐惧, 愤怒, 憎恨, 悲伤, 沮丧, 无聊, 放松
	PAD 模型	愉悦度, 激活度, 优势度
	Schlosberg 模型	快乐-不快乐维特征: 快乐, 爱, 愉快, 幸福, 不快乐, 愤怒, 伤心 注意-拒绝维特征: 惊奇, 痛苦, 注意, 拒绝, 反感, 厌恶 激活水平维: 强度
	Hourglass 模型	四个维度: 愉悦度, 关注度, 倾向度, 敏感度 主要特征: 愤怒, 气愤, 恼火; 警觉, 预期, 兴趣; 狂喜, 高兴, 安全; 羡慕, 信任, 接受; 恐怖, 害怕, 恐惧; 惊奇, 奇怪, 迷惑; 悲痛, 悲哀, 哀愁; 厌恶, 厌烦, 厌倦

3.3 基于观点挖掘的情感计算分类

除了经典的情感特征分类外, Kim 和 Hovy^[47]曾将情感计算定义为观点识别或者意见挖掘, 即人们针对诸如产品、服务、组织、个体、事件和主题等实体所产生的观点、情感、评论和态度等情感倾向进行挖掘、归纳和推理的技术。其中, 观点可由主题(Topic)、持有者(Holder)、陈述(Claim)、情感(Sentiment)等要素组成的四元组来形式化定义, 即:

$$Option = \langle Topic, Holder, Claim, Sentiment \rangle \quad (8)$$

考虑到随时间的变化, 持有者对实体对象的认知和情感倾向也会发生动态的变化, Jindal 和 Liu 等人^[47]进一步将观点分为直接观点和比较观点两类, 其中:

定义 1: 直接观点 (Direct Opinion) 是对目标对象情感的直接表达, 它可以用如下五元组来进行表示:

$$DO = \langle O_j, f_{jk}, h_i, t_l, SO_{ijkl} \rangle \quad (9)$$

其中, O_j 表示一个目标对象, f_{jk} 表示对象 O_j 的第 k 个特征, h_i 表示对特征 f_{jk} 持有特定观点的第 i 个人, t_l 表示表达该观点的时间, SO_{ijkl} 则表示在时间 t_l 的情况下, 针对一个客观对象 O_j 中的某一个特征 f_{jk} , 观点持有人 h_i 所表达的情感倾向值, 并且 SO_{ijkl} 值可以表示为正向、负向或者中立。在此基础上, Ding & Liu^[48]提出了一个情感词聚合的算法来对观点倾向进行计算, 其过程如下:

输入: $\text{Pair}(f, s)$ 表示文本中所包含的对象 O 的特征 f_{ik} 以及相关句子 S 所形成的点对集合;

输出: $SO(f)$ 的值, 即表示特征 f 的观点倾向;

Step 1: 根据转折词 (但是、反而、除非等) 将句子进行切分, 并进行过滤后, 使得每一个句子 s_f 均包含相应的特征 f ;
//初始化并构造 $\text{Pair}(f, s)$ 集合;

Step 2: 对句子中的观点或情感词进行分词, 即 $s_f = \{w_1, \dots, w_n\}$, 且 $SO(s_f) = \sum_{i=1}^n SO(w_i)$;

Step 3: 对文档中包含特征 f 所有的句子 $s_j^{(f)}$ 进行倾向聚合, 即 $SO(f) = \sum_{j=1}^m SO(s_j^{(f)})$;
且当 $SO(f) < 0$ 时, 情感倾向为负; 当 $SO(f) = 0$ 时, 情感倾向为中立; 当 $SO(f) > 0$ 时, 情感倾向为正;

Step 4: 输出最终 $SO(f)$ 的情感值

随后, Ding^[49]进一步对该算法进行了改进, 并针对特征 f 的观点倾向采用下式来计算:

$$SO(f) = \sum_{i=1}^m \frac{w_i O}{d(w_i, f)} \quad (10)$$

其中, $d(w_i, f)$ 表示从特征 f 到情感词 w_i 的距离, 而 $w_i O$ 表示对词 w_i 的情感倾向。

由于直接观点是针对于特定对象或其特征来开展的一种主观性评价, 它具有集群性与时间演化性的特点。集群性反映在针对特定对象或特征所具有相同观点的持有者组成的集合, 由于这些观点持有者本身也具有相同或相似的感情特征, 因此, 在社会网络环境下, 这些观点持有者通过共同的观点来形成一个临时社区, 并通过观点的传播来影响更多的人。而时间演化性是指同一个观点的持有者也会随着时间的变化, 对特定对象或特征的观点发生一定的变化, 这种变化所反映出来的观点动态演化行为, 为情感演化动力机制的深入研究提供了一个有效的研究场景。

定义 2. 比较观点(Comparative Opinions): 是针对两个或多个对象或者对象间的特征进行比较时所表达出来的情感倾向, 例如: “张三的考试成绩比李四高”。因此, 可以形式化表示为一个六元组:

$$\langle O_j, O_c, f_{cjk}, PO_{icjkl}, h_i, t_l \rangle \quad (11)$$

其中: O_c 和 O_j 分别表示对象集合中具有相同特征 f_{cjk} 的两个不同对象; h_i 表示对特征 f_{cjk} 持有特定观点的第 i 个人, t_l 表示表达该观点的时间; PO_{icjkl} 表示对比观点的持有者 h_i 优先倾向选择的对象 (preferred object), 它所表示的不再是正负的观点, 而是等价或者是比较的差异, 例如 “ O_c 的特征 f_{cjk} 比 O_c 高/好/坏/...或者相同/相等”; 也可能是一个比较的最高级, 例如: “该书是本领域中最好的一本” 等。

综上, 这两类观点计算的核心要素与特征以及计算方法存在着一定的差异。但是, 在观点挖掘过程中也具有相似的关键性要素与任务, 即: 对象的识别、对象特征的抽取与主题的识别、观点持有者的识别与抽取、观点分类以及情感计算与度量等, 这些任务也是观点挖掘计算过程中迫切需要解决的关键问题与核心任务。本文后续内容暂不考虑这两类观点的识别处理, 而是重点讨论基于内容信息的情感特征计算和评价。

4 基于单模态与多模态条件下的情感计算

根据 2.1 节式 (1) 中的情感定义, 情感具有的不同载体, 例如: 基于语言的文字与声调, 基于视觉的面部表情与肢体行为, 基于生理的特征信号与应激反应等等, 这些不同的特征对情感的表达与特征的描述均存在着差异, 传统的研究方法往往是从某一个特定的情感载体的角度来展开研究, 并从中获得情感的度量; 但是这种方式并不能够满足实际情况, 现实的情感计算往往需要考虑到多个载体所反映的不同特征, 并在此基础上进行融合计算后并获得最终的结果。本文从单一模态下的情感计算出发, 通过对不同载体条件下的情感特征分析与比较, 进而对多模态条件下的情感计算最新研究进展进行分析与综述。

4.1 基于单模态的情感计算方法

1) 基于文本的情感计算

基于自然语言处理的文本内容情感计算主要集中在对词汇以及词汇的上下文语义特征的分析上, 目前仍存在着许多困难和挑战^[50]。一方面, 新闻、博客、客户评论等不同类型与领域的信息主要还是采用文本的方式来进行描述, 这不仅需要对语言表达过程中的隐式与显式、规则与不规则, 语法与语义等方面的特征进行深入的研究, 同时, 还需要考虑到句子的浅语义特征, 即通过极性化术语、情感词以及其共现频率等方式来实现情感的计算。另一方面, 产生情感的环境及刺激因素等都是通过语言来表述的, 而非听到、看到和触摸到的真实刺激, 这也使得不同个性偏好与背景的人员容易对文本内容的理解产生差异。因此, 在针对文本句子语法与语义结构分析的基础上, 准确实现对自然语言的理解仍然是基于文本情感计算的关键。

传统基于文本的情感计算主要采用两种方式:一是基于情感字典、词典、关系表等情感词袋(BoW)进行规则匹配策略,例如,Hatzivassiloglou等^[51]根据连词所连接的两个形容词的情感倾向之间存在的关联性来挖掘出一个基于形容词的情感评价词库,并基于词库来实现情感极性的匹配与度量;二是在海量语料库的基础上,采用基于统计的机器学习方法,即通过建立大规模的情感语料库,并通过情感和极性标注来实现对样本内容的训练。例如,Pang等^[52]人针对11,855个电影评论中提取的单句构建了一个斯坦福大学情感数据库,并通过词语解析树分析并构建了总共215154个独特的短语,并且每个词语由三个人来进行人工标注。在此基础上,采用并对比了几种不同的有监督机器学习算法针对复杂的语言现象来实现文本情感分类,精确率达到了82.9%, Socher等人^[53]进一步采用RNTN递归神经张量网络,在相同数据集下实现情感分类的精确率提高到了85%。

由于BoW模型只考虑到了文本中的词而没有考虑到词在文本中的上下文特征以及文本中存在的概念,即忽视了文本中存在的语义特征。为了考虑上下文情感的影响和作用,傅向华等^[54]以动态的滑动窗口为处理单位,采用了将LDA模型与情感词典相结合的话题情感分析方法实现了话题识别与情感倾向分类。林鸿飞等人^[55]则从认知学角度,结合了上下文的情感内容提出了一个文本情感认知模型:

$$E = (S, TR, P, CC) \quad (12)$$

其中, E表示当前句子的情感; S表示上一句的情感,即当前句子的情感基础; TR表示情感迁移的规则,即是指从语料库中统计出来的一些情感迁移规律和规则; P表示了情感的极性 or 强度大小; CC则表示了当前句子的认知语境,且CC可以进一步表示为:

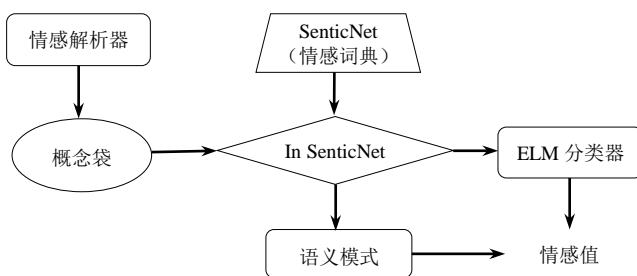
$$CC = (W, K, L) \quad (13)$$

式中, W表示当前句子的词汇信息,包括了情感词、否定词、标记词等具体词汇,也包括了句法和语法等信息; K是百科知识,在文本情感识别中是指与情感相关的各种图式; L是逻辑信息,即将词汇信息和百科知识相联系的能力,逻辑信息决定该概念与其它概念的联系方式。

由于词汇在文本情感分析过程中的重要性,Jin等人^[56]在定义一个生气、开心、悲伤和中立四维情感模型的基础上,为每个情感维度构建了一个能够表达特定情感倾向和权重的特征词向量eVector,该向量包括具有不同权重的情感词、经常出现但不表达情感的普通词和既不是情感词也不是经常出现的普通词,采用该向量实现的词的组合结构来反映出每一个情感维度的表达特征和方式。实验结果表明,采用情感词向量的性能远优于采用词袋模型(BoW)所获取的200个维度所对应的性能。进一步考虑到信息的发布者、接收者以及旁观者所存在的个体特征在情感表达与认知过程中的差异,徐睿峰^[127]等提出了一个将情绪表达和情绪认知分离的情绪词典构建方法,将不同角度对HowNet中所挑选出来的1259个正负情感词条的情绪状态、认知结果及其强度进行区分标注,从而提高了情绪词典的适用性程度以及情感计算精确程度。

针对句子中存在的概念语义特征,Cambria等人^[57]提出了一个基于概念袋模型(BoC)的SenticNet情感分析资源库,且SenticNet资源库中包含了50,000个概念,即通过从句子中抽取词的高层语义概念来替代候选词,从而提高了句子中情感识别的准确率。目前,该模型库已广泛应用于观点挖掘、知识表示、用户画像、情感的可视化和短文本的分类等多个领域。考虑到文本的语法结构、语义以及统计学特征。

为了降低传统算法中人工标注的工作量,梁军等^[59]采用RNN网络来发现特征,并引入情感极性转移模型来加强对文本关联性的捕获。该方法采用了非监督模式来识别观点的持有者与主题,并结合句子的依存语法树来进行语义解析,随后再将所获得的特征作为深度学习算法的输入,同时,结合句子的语义模式(Sentic Patterns,简称SP)来修正文本句子的情感权重,从而计算出句子的情感极性强度的大小。例如:“这辆车很旧但是不贵”和“这辆车很旧尽管它不贵”两句话表达了完全不同的观点倾向,但在句子结构上其区别仅在于连词所反应的语义差异,因此,利用SP考虑整个句子的情感语义模式有助于提高情感识别的准确率。Poria^[60]进一步利用电路中逻辑门的方式来反映语言学的基本模式,并在保持句子结构特征一致性的基础上实现对语句中情感特征的挖掘,并提出一种基于语义的混合计算框架如图3。此外,考虑到语言学特征在文本情感计算中的应用,Qian^[128]等把语言学规则(情感词典,否定词和程度副词)融入到现有的句子级LSTM情感分类模型中,考虑当前词和它相邻位置词的情感分布,定义了四种语言学规则来优化了损失函数,结合单向LSTM和双向LSTM两种模型进行性能实验的比较发现:基于语言学规则的双向LSTM方法具有最佳的情感分类准确率。

Fig.3 A semantic based hybrid computing framework^[60]图3 基于语义的混合计算框架^[60]

当涉及到不同上下文或领域中的文本内容时,一些情感词所包含的语义存在着一定的差异,这种情感词的语义模糊性特征一方面导致了情感分析的二义性;另一方面也导致了许多精细化的情感特征无法进行明确的定义,为了表示文本内容的深层情感语义,近年来越来越多的研究采用了基于 Word Embedding 的表示学习方法, Tang 等^[129]采用了三种神经网络来构建学习特殊情感词的嵌入方式(SSWE),将情感信息编码为词的连续表示,并实现了区分“好”和“坏”两个极性相反的情感在词向量空间中的表示。在此基础上,利用 SSWE 表示的情感特征作为监督学习框架中的输入,对 Twitter 内容的情感分类准确率达到了 86.58%。Bollegala^[130]等将情感的词嵌入方法应用到不同领域的情感分类中,通过构建三种不同的目标函数来分别获取在源领域和目标领域同时出现的情感特征、在源领域文档中的约束标签以及源领域和目标领域中均未被标记的文档映射到嵌入空间中的几何属性特征,将源域和目标域的特征以及文档投影到低维词嵌入空间中,实验结果表明:通过情感词嵌入将上述三种目标函数进行融合后,可获得最佳的情感分类性能。另外,考虑到上下文的影响, Ren^[131]等构建了两个神经网络模型分别处理当前 Tweet 的特征以及与其相关的 Tweet 特征,即将文本内容转化为词嵌入向量,一方面考虑到在两种神经网络上使用共享的词向量表示,另一方面考虑到共享的词向量可能无法获取本条 Tweet 的内容情感特征与上下文特征之间的差异。因此,在采用不同的词嵌入矩阵来扩大模型参数空间的基础上,利用神经网络的池化方法来自动抽取有效的特征来实现两种网络非线性融合,从而获得最终的情感分类结果。另外,由于自然语言的语义模糊性、歧义性、文本内容的多尺度特征,特别是讽刺、暗喻、礼貌、语言文化特征与风格等不同特征之间的互相影响,导致了精确描述和定义混合的细分情感仍然存在一些关键性的技术挑战。

2) 基于音频的情感计算

基于语音的情感计算往往将韵律学、语音学以及声学特征用来进行情感特征的识别,这些特征不仅包括:音高、音强、音调、带宽、语速、持续时间以及基音过零率等,同时还包括了一些与能量相关的特征,如共振峰、Mel 倒谱系数(MFCC)和 LPC 倒谱系数、短时能量及短时平均幅度等,因此,传统语音情感特征包括了时间构造、振幅构造、基频构造、共振峰构造、MFCCs 和 Mel 频谱能量动态系数等共 101 维特征来表征语音的情感^[61]。同时,通过基于开源情感识别工具 openEAR^[62](open-source Emotion and Affect Recognition)的经典 OpenSMILE 工具包^[63]可以实现对上述音频特征的自动抽取。

另一方面,由于这些音频特征可以分为局部性特征或者是全局性特征;其中,局部特征主要是将整个音频文件按照时间间隔或者是发音间隔进行切片后所具有的特征,而全局特征则是对所有音频片断的局部特征进行汇总后的平均值、方差等。因此,在<激活度、效价、优势度>三维空间模型中,全局特征对高激活度的情感特征分类具有敏感度,例如:生气与厌恶;但全局特征并不适用于区分低激活度的情感,如生气与快乐。赵力等^[64]针对含有欢快、愤怒、惊奇、悲伤等 4 种情感的语音片段,利用全局特征和各元音时序结构与韵律特征组合成的情感识别特征,在分析了它们的时间、振幅、基频和共振峰结构特征的基础上,结合句子中元音个数的不同,发现不同类型的情感语句中情感信号特征具有如下的分布特征:一般的情感句子在句首部分的情感信号特征分布较高;而惊奇类型的句子其情感信号特征主要分布在句尾。进一步的研究表明,情感不仅与声音特征有关,同时也与说话者的性别、年龄、性格、教育背景、语言、周围的环境等情感上下文信息紧密相关^[65]。

通过进一步研究声音特征与情感特征之间存在的关联, Hevner^[66]针对音乐通过建立一个情感环模型来表示向往、抒情、轻盈、欢快、热情、生机、神圣和悲伤等八大类以及细分的情感分类。而更多的研究工作则是从一些普通的视频、日常对话与采访或者是呼叫中心的录音中来获取相关的音频信息,其中发现与说话者相关的方法比与说话者无关的方法具有更高的性能。针对前者, Navas^[67]采用高斯混合模型(GMM)针对大量发言者的韵律、声音质量以及 MFCC 等音频特征进行分类,并获得了 98% 的识别准确率,但这种方法的最大困难在于需要获取海量的用户数据。而针对后者, Park 等人^[68]用动态循环神经网络(DRNN)模拟人脑的非线性动态特性,通过采用基音特征以及包含一个输入节点、两个隐层节点和 4 个输出节点的 RNN

网络来对 4 种不同的情感进行分析和识别,在不同噪音的影响下平均识别准确率达到了 77%。因此,在音频特征分类之前对于噪音的预处理则成为了提高音频特征抽取质量的关键,这不仅需要通过语音增强来发现抗噪性能比较好的语音特征,同时也需要采用模型补偿等手段来进行完善。

3) 基于视觉模态的情感计算

在视觉模态下,人的面部表情特征以及人体的姿势特征的识别对于情感分析具有非常重要的作用,但是由于不同年龄与性格的人在会话过程中,面部表情也存在着差异,很难用一种统一的方式来度量,Ekman 等人率先从人的面部表情中抽象出来了喜、怒、哀、恶、惊等 5 种基本的情感特征,在此基础上,他与 Friesen 进一步开发了一个基于活动单元(Action Unit, AU)的面部表情编码系统(Facial Action Coding System, FACS),这些活动单元(AU)通过人脸面部特定的 64 个独立的肌肉动作来表达相应的表情特征,并且每个动作都有其文字解释和照片比对,同时,不同的 AU 组合也可以用来定义和表示细化的脸部表情^[69](如图 4 所示)。




Upper Face Action Units					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
					
Inner Brow Raiser	Outer Brow Raiser	Brow Lowerer	Upper Lid Raiser	Cheek Raiser	Lid Tightener
*AU 41	*AU 42	*AU 43	AU 44	AU 45	AU 46
					
Lid Droop	Slit	Eyes Closed	Squint	Blink	Wink
Lower Face Action Units					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
					
Nose Wrinkler	Upper Lip Raiser	Nasolabial Deepener	Lip Corner Puller	Cheek Puffer	Dimpler
AU 15	AU 16	AU 17	AU 18	AU 20	AU 22
					
Lip Corner Depressor	Lower Lip Depressor	Chin Raiser	Lip Pucker	Lip Stretcher	Lip Funneler
AU 23	AU 24	*AU 25	*AU 26	*AU 27	AU 28
					
Lip Tightener	Lip Pressor	Lips Part	Jaw Drop	Mouth Stretch	Lip Suck

Fig.4 AU combination schematic in facial expression coding system^[70]

图 4 面部表情编码系统中 AU 组合示意图^[70]

在 FACS 模型基础上, Tian 等人^[71]提出了一个基于多状态的脸部组件模型(MSFCM),该模型不仅包括了不同年龄条件下眼睛与嘴唇开闭、瞳孔位置以及眉毛与下颚的区域等人脸局部区域的静态特征,这些区域形状、纹理、边缘轮廓等静态特征在不同环境下,具有相对的稳定性;另外,也包括了脸部皱纹与眉毛之间的距离与形态变化等动态特征。除了上述的面部表情特征外,为了更好地表示情感,Ekman 等将头部的运动特征、眼睛的位置特征以及身体的姿势等局部区域特征也加入到了情感特征定义的范围,从而为人工情感与和谐的人机交互应用的发展奠定了重要的研究基础^[72]。

由于上述方法主要基于实验室环境下的情感识别可测试数据集(如 Cohn-Kanade+, Semaine, MMI, GEMEP-FERA 等^[73]),这些数据集主要依赖于实验室条件下的数据标记,一旦应用于真实环境,则会受到室内与室外不同色彩背景、封闭环境中的杂乱背景等因素的影响,经典的计算方法将面临着新的挑战,为此, Dhall 等人^[74]在面向真实场景下的情感识别挑战赛中,针对所提供的真实环境下的视觉数据(如: HAPPEI, GENKI, AFEW 和 SFEW 等)来进行情感的识别,这不仅需要考虑到人脸检测方法的健壮性与精度,同时还需要考虑到姿势、光照、运动、遮挡和背景等多个不同的外部因素,如何创建和扩展新的算法以及算法验证成为了目前重要的研究任务。

4) 基于生理特征的情感计算

文本、语音、表情和势态均可以通过主观意识的控制来隐藏实际的情感,但是生理信号特征主要来自于神经系统和内分泌系统,而不受主观意识的控制,从而能够更加真实地反映出人的情感特征。因此, Picard 等人^[13]在对情感进行研究时,采用了不同图片来诱发人产生愤怒、憎恶、悲伤、柏拉图式的爱、罗曼蒂克的爱、高兴和崇敬等 7 种情感的同时,对实验者的肌电、脉搏、皮肤电导、呼吸等 4 种生理信号进行了测量,并建立了首个情感生理反应样本库。进一步的研究发现^[75],情感生理特征的特异性与可区分性是情感计算和人机交互研究的关键,由于存在着情感特异性自主神经系统的活动,因此可以通

过对用户的某些生理特征,如:皮肤电导、呼吸频率、心率、脉搏、体温以及脑电波等特征的识别、跟踪与分类来分析和识别其所对应的情感,这对改善人机交互的环境具有重要的作用。

Kim^[76]和 Song^[77]等进一步采用了视频和音乐以及电影所诱发的不同情感组合,如悲伤、压抑、愤怒;惊奇、高兴、悲伤、愤怒和愉悦;以及恐惧、快乐和轻松等,通过心电、皮肤温度、皮肤电导和呼吸这 4 种信号的测试结果,建立了一个相应的本体特征库。温万惠^[78]则采用高收视率电影剪辑作为唤起材料,以激发高兴、惊奇、厌恶、悲伤、愤怒和恐惧 6 种情感,记录被测试者观看电影并唤起情感时的皮肤电导、心率、脉搏、心电、呼吸、面部肌电和额叶的两路脑电信号特征,从而获得了 300 名普通大学生的情感生理反应样本库,通过进一步采用 Fisher 映射对情感生理反应样本进行识别和分类,实现了情感的预测。其中,由于发现了不同生理信号各自的相关矩阵最大特征值和对应的特征向量分布背离了随机矩阵理论的预测,表现出集体的两两相关,揭示出由情感因素的影响所产生的相关生理反应。李昕等^[79]在 MIT 大学和 Augsburg 大学提供的两个公开情感数据库基础上,设计了一个面向高校人群生理参数的压力情感数据库。其中,利用 ECG 和 EMG 信号数据以及 EEG 信号数据对压力识别的平均正确率分别达到了 70.62% 和 75.87%,从而为在压力环境下的情感计算提供了参考。

为了更有效地将生理的情感特征与业务上下文场景相关,一些基于特征传感网络环境下可穿戴的场景感知设备,通过各种传感器获取和收集人的情感状态所引发的生理以及行为的特征信号^[80],从而创建出具有感知、识别和理解人类情感能力的情感计算模型,并使之能够针对个体的情感做出智能、灵敏和友好的反应。

综上,基于单模态的情感分析与计算从不同的数据对象的情感特征出发,通过对这些情感特征的识别与分类,来实现情感有效计算与预测,尽管在单模态下,情感的特征存在着一定程度的语义缺失,但是这些特征的处理为多模态下的数据融合与情感计算奠定了坚实的基础。针对不同模态条件下情感特征的获取指标与线索分类如表 2 所示。

Table 2 Emotional characteristics under different single modal conditions

表 2 不同单模态条件下的情感特性

分类	指标与线索
文本特征	基本特征: 词性、极性词、词频、位置 语义特征: 内容专业性、可理解性、合理性, 价值性, 有用性、可靠度、倾向度
音频特征	韵律特征 (如音高、音强、带宽、能量、短时平均幅度、持续时间、高阶共振峰、基音过零率、语速)、频谱特征 (谱矩心、谱截止频率、相关密度、梅尔频率能量)、倒谱特征 (线性预测倒谱系数、梅尔频率倒谱系数、感知线性预测)
视觉特征	面部表情特征: 几何特征 (ERS: 33 个几何特征) 局部区域特征: 纹理特征; 姿态特征
生理特征	生理特征: 皮肤电反应; 呼吸; 心率; 体温; 脑电波; ECG 信号

4.2 基于多模态条件下的情感计算

在日常生活中,人们在交流过程中对情感的感知主要基于多模态的方式,并通过多种模态特征的融合来实现情感的判断^[81]。例如,人们在交流时,如果在倾听到内容的同时,也观察到了发言者的语气和语调以及面部表情和肢体行为,则会更加准确地帮助我们理解和判断出发言者说话的动机、意图以及相应的情感状态。另外,随着移动应用技术的快速发展与进步,大量的智能终端通过社会化媒体的应用,形成了混杂海量视频、音频、文本等异构内容的大数据集合,针对这些异构数据的处理与应用,直接推动了情感计算技术从单模态向多模态方向的发展和演化,通过对多模态数据的情感特征分析,更好地实现情感特征的识别^[82]。因此,针对多模态下的情感识别与计算研究主要是从视觉 (包含图像和视频等)、听觉、触觉、文本以及视听混合等不同类型的多模态数据出发,模拟人脑对信息的融合与处理方式,来获取和实现情感的计算与度量。

一般地,多模态的情感计算主要分为两个步骤:一是基于单模态的数据处理和情感计算;二是进行多模态之间的情感融合。这两步均非常重要,一方面单模态条件下的分析质量将会直接影响到多模态条件下的计算结果;另一方面,较差的融合质量将会降低多模态系统计算结果的稳定性。通过对收集的样本数据进行算法比较发现:85%的基于多模态的情感分析系统比单模态系统具有更高的准确率,且性能平均提升了 9.83% (中位数达到 6.60%)^[19]。D'Mello 等人^[83]进一步的研究发现,当样本数据为自然视频数据集 (Natural Video Data) 时的训练结果仅比单模态分析结果提高了 4.59%,而样本数据为人为加工数据 (Acted Video Data) 时,其训练结果则比单模态提高了约 12.7%。Pérez-Rosas 等人^[84]在开发具有视频、音频与文本三种模态组合的 MOUD 样本数据集中,同时考虑到了包括韵律、能量、声音概率、频谱等 28 种多模态的情感特征,情感识别的错误率结果相比最佳的单模态处理结果降低了 10.5%。

Poria^[85]在进行文本、图像以及音频三模态的特征融合时,首先采用了一个改进的7层CNN卷积网络对最长为69个字的短文本进行分析,且输出层采用了SVM分类器。文本中的词通过word2vec来构建一个300维的词向量,并结合6种词性特征作为输入,共生成了21114个特征作为输入参数。实验结果表明:单纯采用CNN分类的准确度达到75.5%,而采用SVM改进后得到的分类准确度达到了79.77%,改进后的方法对分类结果有显著的提高。其次,在针对图像特征时,采用CLM-Z面部识别库提取68个面部特征点,并计算每两个点之间的距离;得到2278个距离,然后采用GAVAM软件提取6个面部位置坐标,并计算这些数据的均值和标准差,共得到4568个面部特征。第三,针对音频特征,使用openSMILE软件抓取与音高、音强等相关的音频特征,该软件可以抓取各种低层描述符号,由此得到6373种音频特征。

结合上述模态提取的特征,进一步使用基于特征层或者基于决策层两种策略来实现多模态的融合。其中,针对特征层融合策略是指将多模态的特征先进行融合后,再进行分类与计算。在基于特征层的情感融合过程中,针对单模态、双模态两两融合、三模态融合的实验结果表明:文本模态对结果影响最大,且融合后的结果准确率显著高于单模态。Poria将三种模态分别利用多核学习算法(MKL)来实现分类,并采用SPF-GMKL方法来达到不同模态下的特征融合的最佳效果^[86]。Chao^[87]采用池化建模的方法分别抽取音频和视频等5种不同的时序动态信息的特征进行池化,且对每一组特征建立一个独立的预测系统,并将结果进行融合后实现情感的分类。而基于决策层融合策略的优点在于:不同模态下的决策数据特征可以很容易地进行比较与融合,同时,三种模态同时在决策层进行融合的策略与将两种模态进行特征融合后,再进一步通过决策层来实现与另一种模态融合策略相比,后一策略的分类准确率存在着明显的提升。综合上述的实验结果表明,尽管决策层融合的效果略低于特征层融合,但特征融合处理的速度却得到了明显的提升。

除了上述两种基本的融合策略外,在实际应用中还存在一些常用的融合策略,如:混合融合策略、基于模型的融合策略、基于规则的融合策略、基于分类的融合策略以及基于估计的融合策略等,这些不同的特征融合策略在不同的数据特征条件下也具有不同的应用效果。例如,Wollmer等人^[88]针对YouTube中的音视频多模态数据进行特征抽取,并采用BLSTM来进行音频与视频特征的融合,同时,将初步融合的结果与基于内容的语言学特征来进行决策融合,从而获得最终的情感计算与预测,上述混合特征融合过程如图5所示。

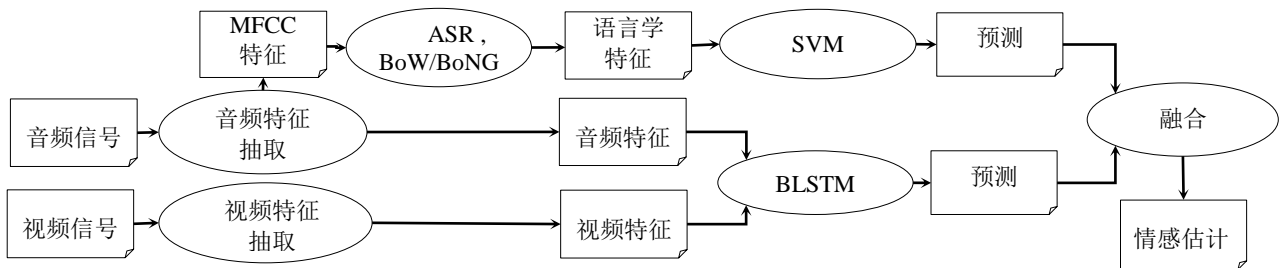


Fig.5 A hybrid fusion method based on multimodal emotion analysis in YouTube video^[88]

图5 基于 YouTube 视频中多模态情感分析的混合融合方法示意^[88]

5 基于时空等外部条件下的情感演化计算

情感演化研究的是复杂时空环境与交互式条件下情感随时间的动态变化规律。研究表明^[57]:时空建模方法在基于多维度的情感识别与计算中具有重要的作用。这种情感的时空特征一般具有两类基本特征:情感本身所具有的时序演化特征;以及用于情感分析的言语、面部表情、肢体行为、生理特征等因素在动态交互过程中的时序特征。前者反映了情感的内在变化规律,后者反映了情感外在观测因素的变化规律。因此,如何基于情感的识别、表示与建模,在特定的时空环境约束下,通过对情感的度量和计算来实现情感动态演化过程的预测以及情感交互能力的分析,则成为了重要的研究任务^[89]。

5.1 基于时空条件下的情感演化计算与预测

情感认知理论认为情绪的产生是环境事件、生理状态和认知过程三个方面共同作用的结果,其中认知过程是决定情绪的关键因素,但是环境事件因素往往也起了重要的影响作用。达尔文根据世界不同区域的人出生后存在的一些共同的天然情感特征提出了基本情感理论的假设:一方面这些情感可以通过一些情感词来标记,如恐惧、爱、悲伤、欢喜等;另一方面,随着人在成长过程中的环境、生理状态以及认知能力的变化,不同的情感特征也会出现不断细分和动态演化的行为过程。例如,每个人都存在着一个情感的日周期特征,即一天的不同时刻下人的情感存在着明显的差异与波动,这种周期性的情感变化取决于人脑对于生理以及外界环境价值关系的一种周期性变化^[90]。由于根据情感所具有的动态、连续和内在的特征,用户当前情绪不仅与外部事件或信息相关,同时也与情感的历史状态相关联。因此,通过不同数据模型将用户过去的情感关联起来作

为已知信息,并结合当前感知的情感状态信息,不仅可以预测用户的未来情感,而且可以推断用户的情感结构模型,该模型如图6所示。

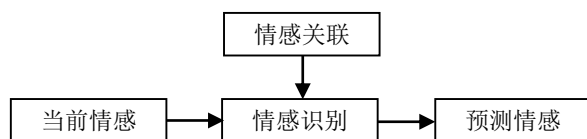


Fig.6 Emotion prediction model based on State Association

图6 基于状态关联的情感预测模型

Cunningham 等人^[91]进一步考虑到当前的情感状态不仅受到此刻环境的影响,同时也与上一个时刻的情感轨迹有关,即在 T_{n+1} 时刻的个体情感状态,由当前环境 S_{n+1} 和 T_n 时刻的情感状态 A_n 以及在 T_n 时刻所预测的情感状态 A'_{n+1} 所共同决定,从而形成了如图7所示的情感迭代再生模型。

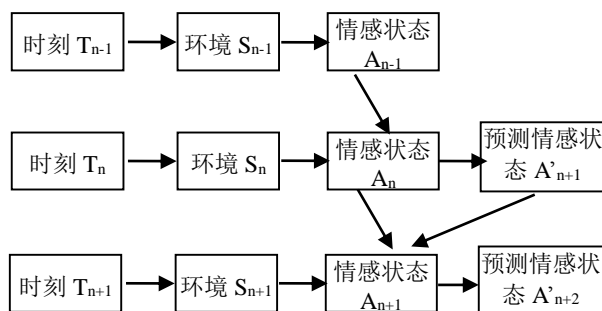


Fig.7 Schematic diagram of emotion iterative regeneration model

图7 情感迭代再生模型示意图

为了进一步研究情感的时序或者周期性特征,人们将情感作为情绪过程的一个信息观察序列来进行表征,即情绪过程可以假设为一种马尔可夫过程,且状态转移概率矩阵中各要素的值由多种因素来共同决定,这不仅与该个体的性格特征有关,同时也与意识刺激的类型相关^[92],该方法多用于有限、离散状态下的情感识别,但是对于情感特征连续变化的描述能力较弱。Zeng 等人^[93]则采用了一种基于最大熵和最大互信息的多流融合隐马尔可夫模型(MFHMM)来建立各种流数据之间的最优连接,并实现了音视频内容的情感识别;Yang 等人^[94]针对基于音频与视频流的时序特性采用了耦合隐马尔可夫模型(HMM)的多模态融合方法来实现对异常情绪的自动检测;Hirose^[95]则采用基于 HMM 和 DNN 的方法来提高语音自动合成过程中的质量。Wöllmer^[96]针对人类情感中存在的复杂性、连续性、精细化和动态化的问题,采用 LSTM 来对实际场景中的情感特征数据进行时序建模,通过采用遗忘门、输入门和输出门的控制来实现了短期情感状态记忆,神经元的内部结构以及神经元之间的情感状态的变化与输出如图8所示。在此基础上,利用 LSTM 结构将不同时刻下引发情感的外部因素、环境因素以及时序过程中的情感状态的变化等因素进行了综合,同时将音频与视频的特征分别抽取后再进行特征融合,提升了情感预测的准确率。

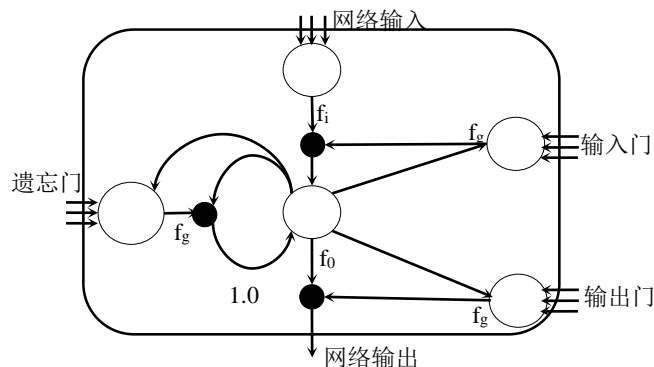


Fig.8 The basic structure of Wöllmer by LSTM

图8 Wöllmer 采用的 LSTM 的基本结构特征

Metallinou 等人^[97]在 LSTM 模型的基础上进一步提出一个基于双向长短时记忆(Bidirectional Long-Short-Term Memory, BLSTM) 神经网络的对话系统, 通过基于视觉与语音两个通道的上下文情感识别模型, 来提高情感识别的准确率。另外, 在多模态条件下的时序状态融合过程中, 通过将 RNN 与 LSTM 网络相结合^[98], 将来自各通道在不同时间片段上的情感特征与决策信息按序融合, 实现了场景复原, 并提高了情感预测的准确率。Tang^[122]提出了一种自下向上的文本向量表示模型, 该模型首先采用 CNN 或 LSTM 对句子的内容表示建模, 随后采用双向 GRU 模型对句子表示进行编码、获取文档表示并发现句子间的关联关系 (如图 9 所示):

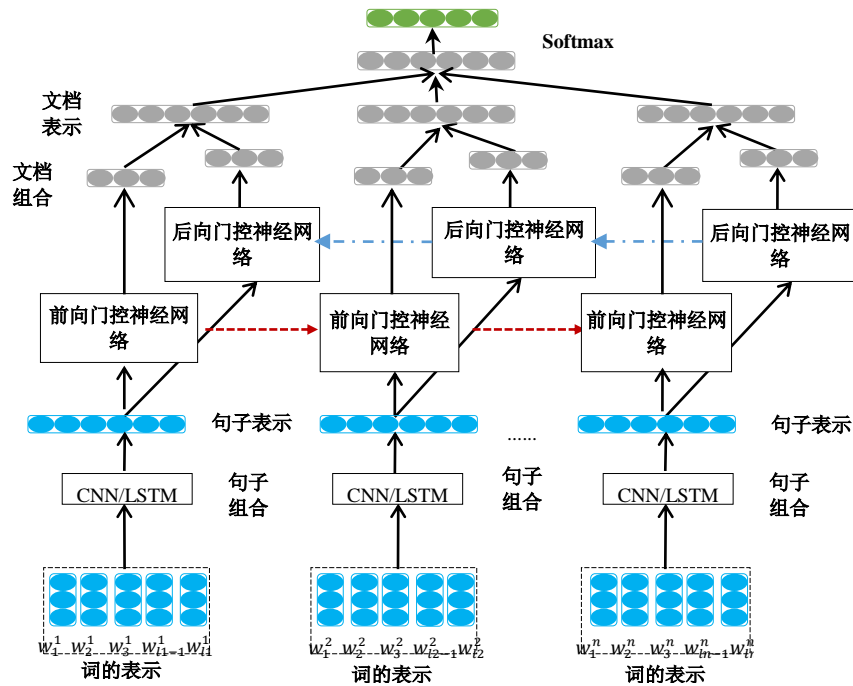


Fig.9 Text emotional representation model based on CNN/LSTM/GRU

图 9 基于 CNN/LSTM/GRU 的文本情感表示模型

Yang 在此基础上进一步考虑到文档的分层结构和不同单词与语句的信息量大小, 提出了一个 HAN(Hierarchical Attention Network) 模型^[123] (如图 10 所示), 即在分层网络的构建时加入了两个 Attention 层来分别实现对句子和文档中的词和句子重要性的评价与建模, 从而在特定时空与上下文语境下, 提高内容与情感识别精度的同时, 还可以对单词和句子的重要性进行分析与可视化。

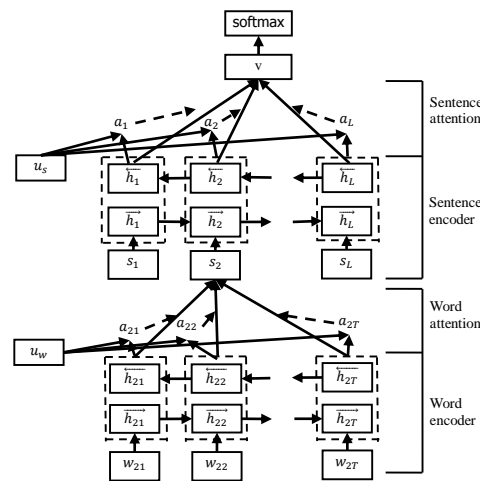


Fig.10 Hierarchical concern network structure (HAN) model

图 10 分层的关注网络结构 (HAN) 模型

5.2 基于特定时空条件下的情感交互计算

除了受到生理因素的影响外,情感也会受到社会文化经验以及个性化特征的制约,并体现出在特定时空环境下,人的社会决策与情感之间的交互作用过程。人们的交互行为主要存在着2个关键的动机:一是为了获取和发布信息来实现信息的有效沟通;二是为了在获得支持或者是情感宣泄过程中得到情感的满足。在不同的时空环境下,情感常能够受到交互环境以及其它外界因素的影响,并呈现出对事件触发过程中的突发性、起效快、持续时间短、自动评价、具有连续和一致性反应等共性特征,这些特征也反映到人/人之间的交互以及人/机之间的交互过程中。卜湛等^[99]将网络论坛中不同用户角色之间关于特定主题事件T进行文本交互评论的过程看作是一个情感的博弈过程,具体表现为用户*i*收到了用户*j*支持或反对的回复时,自身满意度也会随之上下波动,用户*i*在追求自身满意度最大化的基础上,根据双方历史情感交互行为做出最合理的决策,并预测下一个时刻用户反馈评论属于积极或消极情感的可能性,即已知 E_{ijk-1}^T 的值,则 $E_{ijk}^T=?$,用户*i*和用户*j*之间的交互预测过程如图11所示。

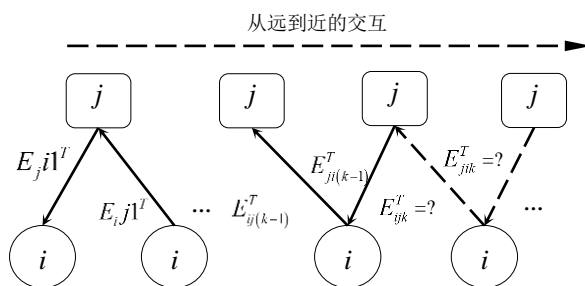


Fig.11 Interuser interaction and prediction based on topic T

图 11 不用户间基于主题 T 的交互与预测过程

这种情感的交互预测不仅在社会交往过程中具有重要的作用,而且在人机交互过程中也起到了关键作用。如何赋予计算机类似于人一样的观察、理解和生成各种情感特征的能力,并实现人/机对话过程中自然、和谐的情感应答与交互则是人工智能 2.0 领域发展的重要方向。

例如,在传统的 e-learning 网络教学中,由于在学习交互过程中的“情感缺失”问题直接导致了学习者无法持续有效地学习。因此,实时判断和预测学习者当前的情感特征以及情感变化,动态且自适应地调整学习内容,已成为构建一个智慧学习环境的关键^[100]。考虑到在线交互过程中,情感的变化也是一个连续的过程,为了进一步实现交互过程中的情感预测,Liscombe^[101]将对话系统中说话者的历史行为和历史话语中的词汇变化信息作为情感上下文,结合话语中的韵律和词汇的情感特征一起实现了对 7 种情感状态的识别,且识别准确率平均提高了 2.6%。在此基础上,一些研究工作主要从以下两个角度来提高交互过程中情感识别的准确率:一是获取并融合更多的情感上下文特征;二是考虑到交互过程中连续语音在时间序列上相关性对情感分析的影响。针对前者,国玉晶等^[102]将交互对话过程中存在的上下文环境、动态环境和基于语句的全局环境共 3 类 5 种环境特征融入到了情感分析的过程中;针对后者,白李娟等^[103]则将上下文语音情感特征定义为:上下文动态特征(CEF1),上下文差分特征(CEF2),上下文边缘动态特征(CEF3)和上下文边缘差分特征(CEF4)等四类,并在特征获取的基础上,结合对话系统的情感交互矩阵,制定出基于情感上下文的情感推理规则,在自行录制的包含 6 种基本情感的数据集上,将语音情感识别率较传统的基于声学特征方法相比提高了 12.17%。

5.3 基于特定时空条件下的人工情感自动生成

情感一方面具有信号、生理特征和先验事件所反应出的独特特征;另一方面也与环境等外部的关联因素相关,因此,情感作为连接人的内部心理状态与外部环境状态的桥梁,影响着信息处理的全过程。在人机交互以及机器人等多个领域的研究中,如何利用信息科学与认知科学的手段,对人的情感、意志、性格等心理活动通过计算机或者算法模型进行模拟或重现,并在情感的自动识别、生成与表征的基础上,实现具有心理学广义层次上的人工情感,从而使得机器不仅具有人工心理与情绪的表达能力,同时,还具有认知、目标和意图的识别与分析能力,并已成为了人工智能领域研究的重要挑战之一^[1]。

一般地,人工情感自动生成主要包含了以下四个阶段:一是对特定时空约束条件、环境或者场景进行高层的情感建模;二是实时地对具体的交互行为、对话或者生理特征进行情感识别、特征抽取;三是对要反馈和输出的情感进行推理和决策;四是根据推理的结果,采用不同类型的情感词典、行为库或表情库来进行显式的情感表达。因此,人工情感生成的整体逻辑过程与框架如图 12 所示。

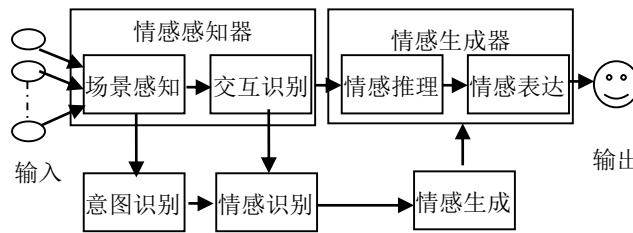


Fig.12 The overall logical framework of artificial emotion generation

图 12 人工情感生成整体逻辑框架

其中,情感感知器所涉及到的场景感知以及交互过程中的情感状态与特征的识别均可以采用前文的研究工作来实现;而基于情感上下文的情感推理与情感表达则成为了情感生成器中的核心与关键。例如:Hao Zhou^[104]等利用基于内部与外部的记忆模式开发了一个情感聊天机器,该机器将高层的情感语义抽象出来并嵌入到情感分类之中,同时结合所获取的内部显性变化的情感特征,利用外部的情感词汇来进行情感语句的生成与表达,从而实现了一个具有情感会话生成与表达能力的自动聊天工具。而 Oh^[132]等提出一种基于深层语义的自然语言理解(NLU)会话服务,它采用 Pointer Network Model 神经网络模型来学习句子中的关键词,并从句子语料库中映射获得相应特征来学习句子中的关注点,并结合 LSTM 通过短暂时空条件下的上下文分析来判断用户的情感,实现了对用户情感变化的动态连续监测与用户关注内容的理解。此外,Shen^[133]等构建了一个分层的递归编码-解码器(SPHRED),利用深层的情感语义特征实现了在特定条件下的情感语句的自动生成。

常规的情感推理多采用基于定义的情感规则的匹配或映射来实现,例如,Shah^[105]等通过情感感知器分析出图像中使用最为频繁的情绪标签,并在时间受限的情况下,与已标记的音乐情感特征进行匹配,根据匹配的规则生成一个包含音乐的电子相册。而邵曦^[106]等提取了音乐中 MFCC 和相关谱感知线性预测(RASTA-PLP)特征,使用核化典型相关分析(KAAC)方法来分析音乐情感特征和图像情感特征的相关性,从而在图像播放的过程中为其匹配对应情感的音乐。毛启容^[103]等也分别通过情感词和情感上下文分析法来进行情感分类,并根据情感分类结果与已进行标记的图像建立映射,从而生成一个具有相应情感的图像来实现对文本的情感表达。在这种基于规则匹配的情感分析基础上,社交机器人则在通过强化信号来获取心理状态外,一方面需要考虑到外界环境的特征,并辅助目标意图的识别以及行为动机的推理等控制机制;另一方面需要考虑到多模态条件下的情感特征,最终通过多类型和多模态的情感特征融合来实现情感的自动生成。

尽管上述工作取得了一些初步进展,但是情感的自动生成目前还存在大量的技术挑战,例如在实现人机交互过程中,不仅需要进行实时、准确的情感判断与恰当的响应,还需要实现对细微情感的特征定义与计算,如将“微笑、笑、大笑、狂笑”之间的特征距离进行细分,自动做出不同程度的“笑”来进行响应,要达到这一理想的人/机和谐交互的目标,还存在着许多需要技术突破的困难与挑战。

6 情感群体化特征以及传播动力学机制

传统的情感研究主要依赖于数据样本中的单一主题,特别是针对个人的情感来展开分析。然而,由于每个人都具有一定的社会化群体性属性和特征,且随着网络社交媒体的发展,针对某一个社会事件,不同的用户不仅可以上传个性化的音频、视频以及文本内容来表达个人的情感或倾向,同时也可以利用网络来传播和分享这种情感,引发个人的情感或情绪在网络中快速地扩散与传播,并逐步形成了群体共同的情感或情绪,甚至通过这种群体的情绪来影响人们对事物的观点和态度的认知。例如:通过群体情感计算与分析,可以广泛地应用于新闻舆情、谣言信息的散播能力、产品口碑与推广的效率和质量以及金融市场的价格波动等多个领域,其中所反应出的情感群体化特征、情感的传播动力学机制以及社会网络环境下的情感涌现行为等关键问题已成为目前研究新的热点。

6.1 多因素的情感个性化度量

上述的研究主要针对着抽象个体的情感来展开,并没有考虑到实际情况下每一个个体的性格、教育程度、文化背景、宗教信仰、年龄、性别、从事行业 and 习惯等属性对个性化价值判断与情感倾向产生的影响^[22],2002 年 Kshirsagar^[134]发现人的

性格会影响心情的变化,而心情的变化又会影响到情感状态的变迁和面部表情改变,于是提出了一个“性格—心情—情感—表情”多层情感模型,将人的性格特征与情感进行了关联,并用于虚拟人面部表情合成中,但是由于该模型将心情只是粗略地分为:好、坏和中性三种心情,并不能充细化地描述复杂的心情;随后,Gebhard^[135]提出了ALMA(A Layered Model of Affect)模型,通过引入PAD空间模型来定义心情,从而细化了“性格—心情—情感”三层相互映射关系。从认知学的角度上看,个体的认知评价能力对于激活和区分情感起到了关键性的作用,即由于文化的特异性价值体系使得不同文化的成员在表达情感的过程中存在着明显差异,并导致在认知评价上的差异^[30]。陈师哲等^[136]并针对在多文化场景下的多模态情感识别问题进行了研究,通过利用语音与面部表情等多模态情感特征的融合方法,结合分文化选择模型、多文化联合训练以及基于共同情感空间的多文化联合训练3种训练策略,在CHEAVD和AFEW中/英文多模态情感数据集下进行对比实验结果表明:文化因素对于情感识别具有重要的影响;且第三种策略在两种数据集下的MAP与MF1的值分别达到了66.63,65.60和72.12,70.60。

另外,由于发布信息的上下文背景以及信息发布者属性特征的差异,其情感含义与极性可能会完全不同。Song^[137]针对微博用户,提出了一种基于潜在因子模型(LFM)的个性化情感分类方法,将用户和微博内容特征映射到低维因子空间中来获取情感的变化与用户之间的关联关系,从而挖掘出同类用户对情感的表达方式。宋晓勇^[138]等则将用户先按性别,再按年龄阶段(如青年、中青年、中老年)组合成6种类型,来研究针对单一主题的在线评论数据,不同类型用户的个性化特征对情感分类器的质量影响,通过将全局的情感分类器与特定用户类型的情感分类器组合使用,从而构建了一种个性化的在线评论情感分类模型(PORSC),实证结果表明该方法比起LFM方法更适合个性化的情感分类。随着计算与存储能力的不断提高,利用个性化的情感特征来进行用户特征画像以及用户行为和意图的识别与预测,甚至是拟人情感推理,已在许多基于情感的个性化的智能推荐电商平台以及情感机器人领域得到了广泛的应用^[139]。

6.2 情感群体化特征

Edward Titchener很早就提出了移情(empathy)的概念,即人不仅能看到他人的情感,且还能通过条件反射、联想、模仿和角色扮演等方法来感受到他人的情感,从而实现了情感的迁移与传播^[107]。孟昭兰也认为情绪可以发动、干涉、组织或破坏认知过程和行为,同时,认知对事物的评价则可以发动、转移或改变情绪的反应和体验^[32]。而当人们通过某种方式或者某个共同的目标聚集在一起并形成社群或社区时,相互之间的交流互动进一步促进了移情作用,逐渐形成了与个体情感特征不同的群体心理与群体情感特征,这种群体性情感反过来直接影响到人们对事物的观点、态度以及认知。因此,Averill认为情感不仅仅是一个认知或者心理学的产物,而且是一种基本的社会结构^[108],例如:在社交网络中不同社群的风格、词语特征以及情感趋向均具有明显的差异,甚至于在不同民族语言进行翻译的过程中,也存在着某些特定的感情词(如:恐惧、怜悯等)没有对应词的现象,这些现象也都反映出了情感的社会化本质特征。

一般地,群体主要具有以下三个特征:1)群体是个体的集合;2)群体是个体通过某种方式、某种原因而聚集成,且具有特定的目标或兴趣的相对稳定的集合;3)群体具有区别于个体,且不为个体所拥有的独特的群体特征。因此,目前在微博等网络社交媒体中进行基于事件的群体情感分析时,主要采用以下两步:一是情感分类,即先将微博中文本强度最高的情感作为该微博文本的情感类型;二是群体情感强度计算,即采用算术平均的方式按不同情感类型来进行强度的计算^[109]。但是这种计算方式并没有充分考虑到群体的第三个特征,即群体与个体之间的差异性特征。例如,在网络社交媒体中的“人肉搜索”现象充分反映出网络群体目标与行为特征在群体内部成员互动过程中所具有的极其不稳定的情感倾向,这种不稳定的情感倾向来源于以下三个方面:1)群体中由于采用了匿名方式而可以充分表达和反映出个人的情感和本能,而无需担责;2)网络群体通过“移情”的方式,将个体的情感与行为在群体的范围内快速传染与传播,并逐步形成了网络中稳定存在的群体情感;3)群体的情感易受外界信息或事件的刺激与影响而发生变化^[110]。

因此,由于受到传染性、社会性与随机性的影响,群体的情感往往具有极端情绪化、易感染性和易变性的不确定特征,这一方面表现为群体难以接受不同意见和观点而呈现出的集体的专横与偏执;另一方面,群体情感又特别容易受到外界刺激因素的影响而产生变化,导致了群体情感与群体行为的不确定性,并表现为群体情感非线性传播过程中存在着复杂的“涌现”的现象,该现象是通过自组织的方式,将具有共同关注点、兴趣或者是相同情感的人进行聚合,并通过相互作用、相互影响而达成集体行为的动态过程,这种集体行为是一种从个人行为出现、叠加到相互传染和互动,形成了一种具有不稳定性与非线性演化特征的动态行为。特别是当一些非线性的微小局部变化积累达到某一个特定的阈值时,会引起整个系统产生临界相变,并导致了网络群体、网络结构、群体情感与行为以及信息扩散数量上的涌现^[111]。例如,受到群体中一些意见领袖的鼓动和控制影响,可能会导致群体的情绪爆发,并引发一定的破坏性行为。

利用这些特征, Robert Shiller 提出了行为金融学的概念与理论^[112], 指出尽管传统的金融学理论已证明市场是不可预测的, 但是考虑到人的投资行为往往受到个人情感以及整个市场情绪的影响, 一旦能够对当前市场的情绪进行度量, 则在一定程度上就可以通过情绪对投资行为的影响来实现对市场动态变化的预测^[113]。同样, 在利用微博和微信等社交网络媒体针对特定的网络事件进行舆情分析与群体情感预测时^[114], 如何结合个体的情感特征来度量和分析群体情感的动态变化因素以及情感在个体之间的传播动力学机制, 不仅对新闻、舆情、广告等信息的接受与传播机制具有重要的研究价值和作用, 同时, 对于深入理解群体心理与群体情感的特征、激励机制以及演化过程均具有重要的研究意义。

6.3 情感传播模式与动力学机制

Mark Granovetter 曾根据人们之间的互动频率、情感力量、亲密性程度以及互惠交换能力四个要素来度量人际关系在社会网络中的传播能力^[115], 特别是将互惠交换能力和情感力量定义为人际网络中进行交互传播的一个重要动机来源, 即分享与传播信息的目的不仅是为了更好地获得对等的信息资源, 同时也为了进行情感的发泄与交流。因此, 当某一个人在网络中发布消息时, 该消息被传播至每一个关注它的用户, 而其他用户评论或转发也会以同样的方式传播下去, 其中, 信息传播的效果不仅取决于传播者的情感与内容对信息接受者的固有信仰、观点、态度的威胁或强化程度, 同时也取决于群体情感和行为对其个体成员的直接影响。

这种情感的快速扩散方式与传染病的传播模式存在着高相似性, 通过“感染”、“免疫”和“抗体”等方法激发个体向群体的趋同。在快速的传播过程中, 用户个体与网络群体之间的情感和观点会不断交织和重新塑造。尤其是通过社会媒体容易引发负面情感信息的快速扩散与传播, Berger 和 Milkman^[116]针对纽约时报中的文章进行研究后发现, 信息内容情感特征与传播能力相关, 具有积极和正能量的文章容易被分享, 特别是那些能够刺激读者产生愤怒或惊愕情感的文章会更加易于传播; Bessi 在针对 Facebook 的研究中也证明了类似的结果^[117], 即用户选择转发或传播科学类或谣言类信息的概率, 与其观点相同的好友数目成正比。周东浩则将微博看作是一个融合了社交图谱和兴趣图谱的关系网络^[118], 每一个网络节点均融合了关系特征和具有情感的偏好特征, 并且网络舆情信息传播的概率与网络节点之间的结构相似度以及节点用户对信息内容的情感偏好具有强相关性, 同时, 用户情感特征在一定条件下也容易转化为信息传播的动机和行为。为了进一步梳理用户情感与传播动机的关系, 钱玉霞针对国内虚拟品牌社区中的信息传播过程抽象出了积极、消极和中立三大类 13 项传播动机的核心要素^[119]。其中, 积极动机包括分享快乐、提升影响力、奖励动机、帮助企业和帮助他人动机; 消极动机包括负面情感发泄、心理安慰、警示他人、寻求补偿和报复; 中立动机包括参与互动、支持社区发展和习惯性动机等, 这些动机的划分为用户情感的建模与传播动机和意图之间建立了映射关系。

上述网络群体行为特征与传播能力的研究成果与经典的传播效果心理一致性模型相似, 本质上反映了用户情感从分散到一致的同步行为^[120], 即通过情感的传播与演化而达成群体心理一致的复杂非线性的同步涌现过程, 该过程包含了情感传播过程中网络拓扑结构的动态演化, 且当网络中的从众性指标大于特定的从众性阈值时, 网民情感呈现出明显的同步收敛的现象, 否则将会形成并保持多种不同情感共存。为了针对网络群体行为进行评价, 王元卓^[121]提出了一个网络群体行为分析评价的模型框架, 该模型将网络群体行为评价指标以及行为分析模型与网络的群体事件模型相结合, 从而可以根据网络的不同结构与行为来实现对群体中各种特征的分析, 但是如何利用该框架来协助计算群体情感对网络群体行为的影响并进行指标评价将是进一步需要研究的问题, 整个框架如图 13 所示。

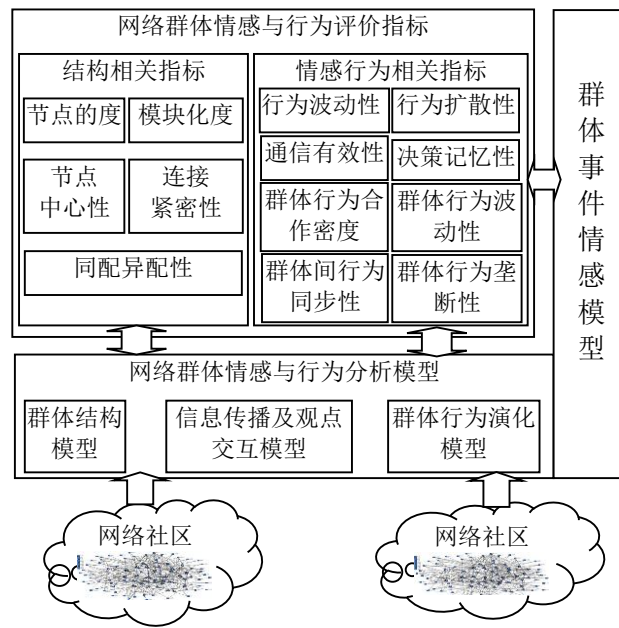


Fig.13 Behavior analysis and evaluation framework based on group event characteristics

图 13 基于群体事件特征的行为分析与评价框架

7 总结

情感语义分析与计算是一个多学科高度交叉综合化的研究领域，它包括了多成分、多维度和多水平的结合，并融合了心理和生理等诸多因素特征，在人类的感知、注意力、记忆、决策、社会交流等诸多方面起着重大作用，它不仅影响到了人工智能的实现，还影响到了人/机交互的方式，并赋予计算机合适的情感表达能力来实现人机之间的自然交往。因此，情感计算已成为人工智能领域中至关重要的一个研究领域，受到学术界以及企业界的持续关注。

本文在针对情感定义以及相关 90 多种情感模型分析的基础上^[57]，归纳并提出了目前情感分析领域中存在的 6 项关键性问题与挑战，即：情感的来源与本质特征的表示问题；多模态条件下的情感计算问题；外部因素对情感演化过程的影响度量问题；情感的个性化度量问题；情感群体化特征以及传播动力学机制问题；细微情感的表达与高精度度的情感计算算法的改进与优化等问题。随后，本文进一步针对其中的一些关键挑战进行详细的综述，深入介绍的领域中的相关问题、工作思路以及目前的工作进展，尝试着为研究人员建立一个完整的领域研究视图。随着深度学习以及强化学习等新方法的深入，将会进一步推动情感计算的理论突破以及在更多领域的实际应用。其中，最具有突破和影响的研究工作将会在以下几个方面展开：

- 1) 在脑科学以及认知科学的研究基础上，通过多模态的情感感知、分析、计算和预测，以及在此基础上的人工情感生成，该领域的研究突破将会极大的促进人工智能、人机交互以及服务机器人等领域的研究发展；
- 2) 细微情感的表示与高准确率的情感计算算法优化，传统的情感模型多是建立在单一模态条件下，对于情感的表达与特征建模存在着一些缺陷，而基于多模态条件下，考虑到更多的特征与不同类型的数据，可以获取更加细微的情感变化；
- 3) 在个性化情感研究的基础上，进一步考虑到不同社会环境、文化背景、种族与宗教等社会化因素下所聚合形成的群体情感以及非线性的“涌现”特征，实现对群体情感的形成、演化特征以及传播动力学机制的建模与度量，这将会对多机器的协同、新闻传播、舆情监控、广告推送等领域的应用产生重要的影响。

综上，随着技术的进步及对人脑功能与结构理解的深入，情感计算将极大地促进传统计算的决策速度与效率，从而对人和机器的认知、智能以及交互产生重要的影响。未来，正如 Minsky 所言：没有情感的机器怎么能是智能的呢？

致谢：本文得到了国家自然科学基金（F020807），教育部“云数融合科教创新”基金项目（2017B00030），中央高校基本科研业务费（ZDYF2017006），陕西省科技厅协同创新项目（2015XT-21）以及陕西省软科学重点项目（2013KRZ10）的联合支持，同时感谢哈工大秦兵教授在 SMP2017 会议报告中的启发，在此一并表示衷心感谢。

References:

- [1] Minsky, Marvin (1986). *The society of mind*. New York: Simon & Schuster. ISBN 0-671-60740-5.
- [2] Boehner K, DePaula R, Dourish P, Sengers P. Affect: from information to interaction. In: OW Bertelsen, Bouvin NO, Krogh PG, Kyng M. eds. *Decennial Conference on Critical Computing: Between Sense and Sensibility*. Aarhus, Denmark, DBLP, 2005.59-68.
- [3] Leahu L, Schwenk S, Sengers P. Subjective objectivity: negotiating emotional meaning. In: CVDS Johann, G Marsden eds. *ACM Conference on Designing Interactive Systems*. New York: ACM, 2008.425-434.
- [4] Binali H, Potdar V. Emotion detection state of the art. In: Potdar V. ed. *Cube International Information Technology Conference*. New York: ACM, 2012.501-507.
- [5] Ashbarray L, Geelan B, Salas KD, Lewis L. Blood and violence: exploring the impact of gore in violent video games. In: *Proc. Symposium on Computer-Human Interaction in Play*. New York: ACM, 2016.44-52.
- [6] Picard RW. *Affective computing*. Cambridge: MIT Press, 1997.
- [7] Wang ZL. Artificial psychology and artificial emotion. *CAAI Transactions on Intelligence Systems*. 2006,1(1):1-6.
- [8] Pu MM, Xu B, Tan TN. Brain science and brain-inspired intelligence technology-an overview. *Bulletin of Chinese Academy of Science*, 2016(7):725-736.
- [9] Zeng Y, Liu CL, Tan TN. Retrospect and outlook of brain-inspired intelligence research. *Chinese Journal of Computers*. 2016(1):725-736.
- [10] Asha McLean, Emotional intelligence is the future of artificial intelligence: Fjord [OL/EB], Feb. 20, 2017, <http://www.zdnet.com/article/emotional-intelligence-is-the-future-of-artificial-intelligence-fjord/>.
- [11] Picard RW, Vyzas E, Healey J. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE transactions on pattern analysis and machine intelligence*, 2001,23(10):1175-1191.
- [12] Jaques N, Rudovic O, Taylor S, Sano A, and Picard R, Predicting tomorrow's mood, health, and stress level using personalized multitask learning and domain adaptation. In: *Proc. of Machine Learning Research*, 2017,48:17-33.
- [13] Jaques N, Taylor S, Nosakhare E, Sano A, Picard RW, Multi-task learning for predicting health, stress, and happiness. In: *Proc. of NIPS Workshop on Machine Learning for Healthcare*, Barcelona, Spain. 2016.
- [14] Hu M, Liu B. Mining and summarizing customer reviews. In: Kim W, Kohavi R, Gehrke J, Dumouchel W.eds. *Tenth Acm Sigkdd International Conference on Knowledge Discovery & Data Mining*, New York: ACM. 2004: 168-177.
- [15] Liu B. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 2012,5(1):1-10.
- [16] Picard RW, Affective perception, *Communications of the ACM*, 2000,43(3):50-51.
- [17] Perezgaspar L, Caballeromorales S, Trujilloromero F. Multimodal emotion recognition with evolutionary computation for human-robot interaction. *Expert Systems with Applications*, 2016,42-61.
- [18] Ivanović M, Radovanović M, Budimac Z, Mitrović D, Kurbalija V, Dai WH, Zhao WD. Emotional intelligence and agents: survey and possible applications. *ACM*, 2014.1-7.
- [19] Poria S, Cambria E, Bajpai R, Hussain A. A review of affective computing: from unimodal analysis to multimodal fusion, *INFORMATION FUSION*, 2017,37: 98-125.
- [20] Oatley K, Johnsonlaird PN. Towards a cognitive theory of emotions. *Cognition & Emotion*, 1987,1(1):29-50.
- [21] Belhaj M, Kebair F, Said LB. An emotional agent model for the simulation of realistic civilian behaviors during emergency situations. In: *Proc. of IEEE/WIC/ACM International Joint Conferences on Web Intelligence*. Piscataway: IEEE, 2014:262-269.
- [22] Sagha H, Coutinho E. Exploring the importance of individual differences to the automatic estimation of emotions induced by music. In: Ringeval F, Schuller B, Valstar M, Cowie R, Pantic M. *International Workshop on Audio/visual Emotion Challenge*. New York: ACM, 2015:57-63.
- [23] Mell J. Toward social-emotional virtual humans. In: *Proc. of International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems. New York: ACM. 2015:2015-2016.
- [24] Wang KF, Gou C, Duan YJ, Lin YL, Zheng XH, Wang FY. Generative adversarial networks: The State of the Art and Beyond. *ACTA AUTOMATICA SINICA*, 2017,43(3):321-332.
- [25] Zhang B, Shi ZZ, Zhao XF, Zhang JH. A transfer learning based on canonical correlation analysis across different domains. *Chinese Journal of Computers*, 2015(7):1326-1336.
- [26] Mitchell TM, Cohen WW, Hruschka E, et al. Never-ending learning. In: *Proc. of National conference on artificial intelligence*, 2015: 2302-2310.
- [27] Chen Z, Ma N, Liu B. Lifelong learning for sentiment classification. In: Fomichov VA. *Meeting of the Association for Computational Linguistics*, Stroudsburg: ACL.2015:750-756.
- [28] Ledoux JE. Emotion circuits in the brain. *Annual Review of Neuroscience*, 2000,23(1):155-184.
- [29] Frijda NH. The laws of emotion. *American Psychologist*, 1988,43(5): 349-358.
- [30] Ekman P, Dalglish T, Power M. *Handbook of cognition and emotion*. Wiley, Chichester, 1999.
- [31] Naseem A. OCC model: application and comparison to the dimensional model of emotion. In: *Proc. of International Conference on Kansei Engineering and Emotion Research (KEER 2014)*. LINKÖPING, 2014:607-617.
- [32] Meng ZL, *Emotion psychology*. Peking University Press, 2005.
- [33] Fox, Elaine. *Emotion science: an integration of cognitive and neuroscientific approaches*. Palgrave MacMillan. 2008.
- [34] Elliott C. Using the affective reasoner to support social simulations. In: *Proc. of International Joint Conference on Artificial Intelligence*. Morgan Kaufmann Publishers Inc. 1993:194-200.

- [35] Borth D, Chen T, Ji R, Chang S. SentiBank: large-scale ontology and classifiers for detecting sentiment and emotions in visual content. In: Proc. of ACM International Conference on Multimedia. ACM, 2013:459-460.
- [36] Ginny, The wheel of emotion—helping you understand what you feel. [EB/OL], 2016, Jan. 26, <http://westendcounselling.co.uk/emotions/wheel-of-emotion/>.
- [37] Sheelagh McG., Peter G, Ian S, Paolo P. HUMAINE: Human-Machine interaction network on emotions. 2006. <http://pdfs.semanticscholar.org/2c84/a9fc10a2e40e542e79d09a68c9c66f75c0ec.pdf>.
- [38] Parrott WG. Ur-Emotions and your emotions: reconceptualizing basic emotion. *Emotion Review*, 2010,2(1):14-21.
- [39] Tiffany WS. The book of human emotions: an encyclopedia of feeling from anger to wanderlust.2015.
- [40] Steven JD. The compass derose guide to emotion words[EB/OL].<http://www.deroose.net/steve/resources/emotionwords/ewords.html>, 2015.
- [41] Thayer R. The biopsychology of mood and arousal. United States, Oxford University Press, 1989.101-102.
- [42] Whissell C. Using the revised dictionary of affect in language to quantify the emotional undertones of samples of natural language. *Psychological Reports*, 2009,105(2):509-21.
- [43] Mehrabian A. Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. *Current Psychology*, 1996,14(4):261-292.
- [44] Fontaine J, Scherer K, Roesch E, Ellsworth P. The world of emotions is not two-dimensional. *Psychological Science*, 2007,18(12):1050-1057.
- [45] Cambria E, Livingston A, Hussian A, The hourglass of emotion. In: Esposito A, ed. *Cognitive Behavioral Systems*. Berlin Heidelberg: Springer-Verlag, 2012.144-157.
- [46] Kim S M, Hovy E. Extracting opinions, opinion holders, and topics expressed in online news media text. In Michael G, Anthony A. eds. *ACL Workshop on Sentiment and Subjectivity in Text*. Stroudsburg: AC. 2006:1-8.
- [47] Jindal N, Liu B. Identifying comparative sentences in text documents. In: Efthimiadis EN, Dumais S, Hawking D, J P, Rvelin K. eds. *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. Seattle, Washington, USA: International Acm Sigir Conference on Research & Development in Information Retrieval, 2006:244-251.
- [48] Ding X, Liu B, Zhang L. Entity discovery and assignment for opinion mining applications. In: xpontus, benhdj, Peterk, waelemara, nlarusso, eds. *International Conference on Knowledge Discovery and Data Mining (SIGKDD 2009)*. New York: ACM, 2009:1125-1134.
- [49] Ding X, Liu B. Resolving object and attribute coreference in opinion mining. In: Joshi AK, Huang CR, Jurafsky D, eds. *International Conference on Computational Linguistics*. Shanghai:Trans Tech Publications. 2010:268-276.
- [50] E. Cambria, Affective computing and sentiment analysis, *IEEE Intelligence System*, 2016,31(2):102-107.
- [51] Hatzivassiloglou V, Mckeown KR. Predicting the semantic orientation of adjectives. In: Cohen PR, Wahlster W, eds. *Meeting of the Association for Computational Linguistics & the Eighth Conference of the European Chapter of the Association for Computational Linguistics*, 1997:174-181.
- [52] Pang B, Lee L, Vaithyanathan S. Thumbs up?: sentiment classification using machine learning techniques. In: Proc. of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, Association for Computational Linguistics. Stroudsburg: ACL. 2002.79-86.
- [53] Socher R, Perelygin A, Wu JY, Chuang J. Recursive deep models for semantic compositionality over a sentiment tree- bank. In: Proc. of the conference on empirical methods in natural language processing (EMNLP). 2013,1631.
- [54] Fu XH, Liu G, Guo YY, Guo WB. Multi-aspect topic sentiment analysis of chinese blog. *Journal of Chinese Information Processing*, 2013,27(1):47-55.
- [55] Xu LH, Lin HF. Text affective computing from cognitive perspective, *COMPUTER SCIENCE*, 2010,37(12):182-185.
- [56] Jin Q, Li C, Chen S, Wu H. Speech emotion recognition with acoustic and lexical features. In: Proc. of 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2015. 4749-4753.
- [57] Cambria E, Poria S, Bajpai R, Schuller B. SenticNet 4: a semantic resource for sentiment analysis based on conceptual primitives. In: Proc. of COLING, 2016. 2666-2677.
- [58] A. Gangemi, V. Presutti, D. Reforgiato Recupero. Frame-based detection of opinion holders and topics: a model and a tool. *Comput. Intell. Mag. IEEE*, 2014,9(1):20-30.doi: 10.1109/ MCI.2013.2291688.
- [59] Liang J, Chai YM, Yuan HB, Gao ML, Jiu HY. Polarity shifting and LSTM based recursive networks for sentiment analysis. *Journal of Chinese Information Processing*. 2015,29(5): 152-159.
- [60] Poria S, Cambria E, Winterstein G, Huang GB. Sentic patterns: dependency-based rules for concept-level sentiment analysis. *Knowledge-Based Systems*, 2014,69(1):45-63.
- [61] Abdelwahab M, Busso C. Ensemble feature selection for domain adaptation in speech emotion recognition. In Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2017.5000-5004.
- [62] Florian E, Martin W, Bjor S. Openeara_A_T, Introducing the munich open-source emotion and affect recognition toolkit. In: Proc. of In Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on, IEEE, 2009. p.1-6.
- [63] James CA. OpenSMILES specification (V. 1.0) [EB/OL]. 2016-05-15, <http://opensmiles.org/opensmiles.html>.
- [64] Zhao L, Wang ZP, Lu W, Zou CR, Wu ZY. Speech emotional recognition using global and time sequence structure feature. *Acta Automatica Sinica*, 2004, 30(3):423-429.
- [65] Tawari A, Trivedi M. Speech emotion analysis: exploring the role of context. *IEEE Trans on Multimedia*, 2010,12(6): 502-509.
- [66] Hevner K. Expression in music: a discussion of experimental studies and theories. *Psychological Review*, 1935, 42(2):186-204.
- [67] Erro D, Alonso A, Serrano L, Navas E, Hernaez I. Interpretable parametric voice conversion functions based on gaussian mixture models and constrained transformations. *Computer Speech & Language*, 2015, 30(1):3-15.

- [68] Park CH, Sim KB. Emotion recognition and acoustic analysis from speech signal, In: Proc. of International Joint Conference on Neural Networks. IEEE, 2003. 2594-2598.
- [69] Ekman P. Facial expressions of emotion: New Findings, New Questions. *Psychological Science*, 2010, 3(1):34-38.
- [70] Ekman P, Friesen W V. Facial action coding system (FACS): a technique for the measurement of facial actions. *Rivista Di Psichiatria*, 1978, 47(2):126-38.
- [71] Tian YL, Kanade T, Cohn J F. Recognizing action units for facial expression analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2001, Vol.23(2):97-115.
- [72] Matsumoto D, Hwang HC, Ekman P. Subjective experience and the expression of emotion in humans. *Encyclopedia of Behavioral Neuroscience*, 2017:342-348.
- [73] Zhong L, Liu QS, Yang P, Huang JZ, Metaxas DN. Learning multiscale active facial patches for expression analysis. *IEEE Transactions on Cybernetics*, 2015, 45(8):1499-1510.
- [74] Dhall A, Goecke R, Gedeon T, Sebe N. Emotion recognition in the wild. *Journal on Multimodal User Interfaces*, 2016, 10(2):1-3.
- [75] Chen YF, Cui YL, Wang SX. Review of emotion recognition based on physiological signals. *System Simulation Technology*, 2017, 13(1): 1-5.
- [76] Shin KH, Kim HR, Lee IK. Automated music video generation using emotion synchronization. In: Proc. of IEEE International Conference on Systems, Man, and Cybernetics. IEEE, 2017.002594-002597.
- [77] Song YW, Gu KW, Li HK, Sun GZ. A lexical updating algorithm for sentiment analysis on Chinese movie reviews. In: Proc. of Fifth International Conference on Advanced Cloud and Big Data. IEEE Computer Society, 2017.188-193.
- [78] Wen WH, Qiu YH, Liu GY, Chen NP, Huang XT. Construction and cross-correlation analysis of the affective physiological response database. *SCIENTIA SINICA Informationis*. 2011, 41(1): 77-89.
- [79] Li X Du XJ, Zhang YP, Ying LJ, Li CW. Research on the performance comparing and building of affective computing database based on physiological parameters. *Journal of Biomedical Engineering*. 2014(4):782-787.
- [80] Dai Y, Wang X, Li XP, Zhang PB. Reputation-driven multimodal emotion recognition in wearable biosensor network. In: Proc. of Instrumentation and Measurement Technology Conference. IEEE, 2015.1747-1752.
- [81] Veni S, Thushara S. Multimodal approach to emotion recognition for enhancing human machine interaction-a survey. *International Journal on Advanced Science Engineering Information Technology*, 2017,7(4):1428-1433.
- [82] Zhang Z, Yang H, Cohn JF, et al. Multimodal spontaneous emotion corpus for human behavior analysis. *Computer Vision and Pattern Recognition*, IEEE, 2016,3438-3446.
- [83] D'mello SK, Kory J. A review and meta-analysis of multimodal affect detection systems. *ACM Comput. Surv*, 2015,47(3):43-79.
- [84] Pérez-Rosca V, Mihalcea R, Morency LP. Utterance-level multimodal sentiment analysis. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*. 2013. 973-982.
- [85] Poria S, Cambria E, Gelbukh A. Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis. In: Proc. of Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015. 2539-2544.
- [86] Poria S, Chaturvedi I, Cambria E, Hussain A. Convolutional MKL based multimodal emotion recognition and sentiment analysis. In: Proc. of International Conference on Data Mining. IEEE. 2017. 439-448.
- [87] Chao L, Tao J, Yang M, Li Y, Wen ZQ. Long short term memory recurrent neural network based multimodal dimensional emotion recognition. *International Workshop on Audio/Visual Emotion Challenge*, ACM, 2015,65-72.
- [88] Wollmer M, Weninger F, Knaup T, Schuller B, Sun C, Sagae K, Morency LP. Youtube movie reviews: sentiment analysis in an audio-visual context, *Intell. Syst*. 2013,28(3):46-53.
- [89] Zhu H, Zhu H, Ge Y, Chen E, Liu Q, Xu T, Xiong H. Tracking the evolution of social emotions with topic models. *Knowledge & Information Systems*, 2016,47(3):517-544.
- [90] Adan A, Archer SN, Hidalgo MP, Di Milia L, Natale R. Circadian typology: a comprehensive review. *Chronobiology International*, 2012,299:1153-1175.
- [91] Cunningham WA, Dunfield KA, Stillman PE. Emotional states from affective dynamics. *Emotion Review*, 2013,5(4):344-355.
- [92] Krell G, Glodek M, Panning A. Fusion of fragmentary classifier decisions for affective state recognition. *Multimodal Pattern Recognition of Social Signals in Human-Computer-Interaction*, MPRSS 2012. 2012,7742:116-130.
- [93] Zeng Z, Hu Y, Liu M, Fu Y, Huang TS. Training combination strategy of multi-stream fused hidden markov model for audio-visual affect recognition. In: Proc. of the 14th annual ACM international conference on Multimedia. ACM, 2006. 65-68.
- [94] Yang TH, Wu CH, Huang KY, Su MH. Coupled HMM-based multimodal fusion for mood disorder detection through elicited audio-visual signals. *Journal of Ambient Intelligence & Humanized Computing*, 2016,1-12.
- [95] Hirose K, Hashimoto H, Saito D, Minematsu N. Modeling of fundamental frequency contours for HMM-based speech synthesis: Representation of fundamental frequency contours for statistical speech synthesis. In: Proc. of International Conference on Signal Processing. IEEE. 2017:771-775.
- [96] Wöllmer M, Kaiser M, Eyben F, Schuller B, Rigoll G. LSTM-modeling of continuous emotions in an audiovisual affect recognition framework. *Image & Vision Computing*, 2013,31(2):153-163.
- [97] Metallinou A, Wöllmer M, Katsamanis A, Eyben F, Schuller B, Narayanan S. Context-sensitive learning for enhanced audiovisual emotion classification. *IEEE Trans on Affective Computing*, 2012,3(2):184-198.
- [98] Sun L, Du J, Dai LR, Lee CH. Multiple-target deep learning for LSTM-RNN based speech enhancement. *Hands-Free Speech Communications and Microphone Arrays*, IEEE, 2017,136-140.

- [99] Bu Z, Wu ZA, Can J, Zhu GX. Affective computing and game theory based prediction for online reviews. *Acta Electronica Sinica*, 2015,43(12):2530-2535.
- [100] Feidakis M, Daradoumis T, Caballé S, Conesa J. Embedding emotion awareness into e-learning environments. *International Journal of Emerging Technologies in Learning*, 2014,9(7):39-46.
- [101] Liscombe J, Riccardi G, Hakkani-Tür D. Using context to improve emotion detection in spoken dialog systems. In: *Proc. of the 9th European Conference on Speech Communication and Technology*, Lisbon, Portugal, 2005,1845-1848.
- [102] Guo YJ, Liu G, Liu J, Guo J. Environmental features based confidence measure for speech recognition. *Journal of Tsinghua University:Science & Technology*, 2009,49(S1): 1388-1392.
- [103] Bai LJ, Zhao XL, Mao QR, Wu BF. Speech emotion feature based on acoustic context extraction and analysis. *Journal of Chinese Computer Systems*, 2013,34(6):1451-1456.
- [104] Zhou H, Huang M, Zhang T, Zhu X, Liu B. Emotional chatting machine emotional conversation generation with internal and external memory. 2017,14.
- [105] Shah R, Yu Y, Verma A, Tang S, Shaikh A, Zimmermann R. Leveraging multimodal information for event summarization and concept-level sentiment analysis. *Knowl-Based Syst*, 2016,108:102-109.
- [106] Shao X, Tao KY. Research on music emotion classification based on music content and lyrics. *Computing Technology and Development*, 2015,184-187+193.
- [107] Preston SD, De Waal FBM. Empathy: its ultimate and proximate bases. *Behavioral & Brain Sciences*, 2002,25(1):1-72.
- [108] Averill JR. The future of social constructionism: introduction to a special section of emotion review. *Emotion Review*, 2012,4(3):215-220.
- [109] Liu CJ, Liu Z, Chai YJ, Fang H, Liu LP. Research on the visualization method of social crowd emotion based on microblog text data analysis. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2016,52(1):178-186.
- [110] Rao Y, Wu LW, Zhang JY. A survey of information propaganda mechanism under the cross-medium. *SCIENTIA SINICA Informationis*, 2017,44:1-23.
- [111] Wang YZ, Yu JY, Qiu W, Shen HW, Cheng XQ, Lin C. Evolutionary game model and analysis methods for network group behavior. *Chinese Journal of Computers*, 2015,38(2):282-300.
- [112] Shiller RJ. From efficient market theory to behavioral finance. *Cowles Foundation Discussion Papers*, 2002,17(1):83-104.
- [113] Ahuja R, Rastogi H, Choudhuri A, Garg B. Stock market forecast using sentiment analysis. In: *Proc. of International Conference on Computing for Sustainable Global Development*, IEEE, 2015,1008-1010.
- [114] He YX, Sun ST, Niu FF, Li F. A deep learning model enhanced with emotion semantics for microblog sentiment analysis. *Chinese Journal of Computers*, 2017,40(4):773-790.
- [115] Granovetter MS. The strength of weak ties. *The American Journal of Sociology*, 1973,78(6):1360-1380.
- [116] Berger J, Milkman KL. What makes online content viral? *Journal of Marketing Research*, 2012,49(2):192~205.
- [117] Bessi A, Coletto M, Davidescu GA, Scala A, Caldarelli G, Quattrociocchi W. Science Vs. Conspiracy: collective narratives in the age of misinformation, *PLOS ONE*, 2015,10(2):1~17.
- [118] Zhou DH, Han WB, Wang YJ. A fine-grained information diffusion model based on node attributes and content features. *Journal of Computing Research and Development*. 2015,52(1):156~166.
- [119] Qian YX. An experimental research on motives of EWOM via domestic virtual brand community. University of Science and Technology of China: A Dissertation for Master's Degree, 2014.
- [120] Zhang YF, Xiao RB. Emergence mechanism of consensus synchronization in internet collective behavior based on cellular automata. *Systems Engineering - Theory & Practice*. 2014,34(10): 2600-2008.
- [121] Wang YZ, Yu YJ, Qiu W, Shen HW, Cheng XQ, Lin C. Evolutionary game modal and analysis methods for network group behavior. *Chinese Journal of Computers*. 2015,38(2):282-300.
- [122] Tang D, Qin B, Liu T. Document modeling with gated recurrent neural network for sentiment classification. In: *Proc. of Conference on Empirical Methods in Natural Language Processing*, 2015:1422-1432.
- [123] Yang Z, Yang D, Dyer C, He XD, Smola A, Hovy E. Hierarchical attention networks for document classification. In: *Proc. of Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2017:1480-1489.
- [124] Li B. Learning dimensional sentiment of traditional Chinese words with word embedding and support vector regression. In: *Proc. of International Conference on Asian Language Processing*. IEEE, 2017:324-327.
- [125] Gao F, Sun, X, Wang, K, Ren, F. Chinese micro-blog sentiment analysis based on semantic features and PAD model. In: *Proc. of 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, 2016:1-5.
- [126] Wang JW, Fu JL, Xu Y, Mei T. Beyond object recognition: visual sentiment analysis with deep coupled adjective and noun neural networks. In: *Proc. of International Joint Conference on Artificial Intelligence*. AAAI Press, 2016:3484-3490.
- [127] Xu RF, Zou CT, Zheng YZ., Xu J, Gui L, Liu B, Wang XL. A new emotion dictionary based on the distinguish of emotion expression and emotion cognition. *Journal of Chinese Information Processing*, 2013,27(6):82~89.
- [128] Qian Q, Huang M, Lei JH, Zhu XY. Linguistically regularized LSTMs for sentiment classification. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics(ACL2017)*. Vancouver, Canada. 2017.
- [129] Tang DY, Wei FR, Yang N, Zhou M, Liu T, Qin B. Learning sentiment-specific word embedding for twitter sentiment classification. *Meeting of the Association for Computational Linguistics*, 2014:1555-1565.

- [130] Bollegala D, Mu T, Goulermas JY. Cross-domain sentiment classification using sentiment sensitive embeddings. *IEEE Transactions on Knowledge & Data Engineering*, 2016,28(2):398-410.
- [131] Ren Y, Zhang Y, Zhang M, Ji D. Context-sensitive twitter sentiment classification using neural network. In: *Proc. of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*:215~221.
- [132] Oh KJ, Lee D, Ko B, Choi HJ. A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. In: *Proc. of IEEE International Conference on Mobile Data Management*. IEEE, 2017:371-375.
- [133] Shen XY, Su H, Li YR, Niu S, Zhao Y. A conditional variational framework for dialog generation. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL2017)*, Vancouver, Canada. 2017:504-509.
- [134] Kshirsagar S. A multilayer personality model, In: *Proc. of SMARTGRAPH '02 Proceedings of the 2nd International Symposium on Smart Graphics*, New York, USA. 2002:107-115
- [135] Gebhard P. ALMA - A layered model of affect. In: *Proc. of The Forth International Joint Conference on Autonomous Agents & Multi Agent Systems*, Utrecht, Netherlands, July 25 to 29, 2005:177~184.
- [136] Chen S, Wang S, Jin Q. Multimodal emotion recognition in multi-cultural conditions. *Ruan Jian XueBao/Journal of Software*, 2018,29(4) (in Chinese).
- [137] Song KS, Feng S, Gao W, Wang DL, Yu G, Wong KF. Personalized sentiment classification based on latent individuality of microblog users. In: *Proc. of International Conference on Artificial Intelligence*. AAAI Press, 2015:2277-2283.
- [138] Song XY, Lv P, Chen NS. PORSC: a sentiment classification model integrating user personality for online reviews. *Journal of Fudan University(Natural Science)*, 2017,56(03):359-365+373.
- [139] Lim H, Kim HJ. Item recommendation using tag emotion in social cataloging services. *Expert Systems with Applications*, December 2017,89(15):179-187.
- [140] Liang J, Yan WJ, Wu Q, Shen XB, Wang SJ, Fu XL. Recent advances and future trends in mirco-expression research. *Bulletin of National Natural Science Foundation of China*, 2013,2:75-78.
- [141] Wang SJ, Ben XY, Yan WJ, Fu XL. The advance and trends of mirco-expression recognition and detection. *Report of the Development Trends about Computer Science and technology from 2016 to 2017*, 2017. 286-313.

附中文参考文献:

- [7] 王志良.人工心理与人工情感.智能系统学报,2006,1(1):1-6.
- [8] 蒲慕明,徐波,谭铁牛.脑科学与类脑研究概述.中国科学院院刊, 2016,31(7):725-736.
- [9] 曾毅,刘成林,谭铁牛.类脑智能研究的回顾与展望.计算机学报,2016,1:212-222.
- [24] 王坤峰,苟超,段艳杰,林懿伦,郑心湖,王飞跃.生成式对抗网络 GAN 的研究进展与展望.自动化学报,2017,43(3):321-332.
- [25] 张博,史忠植,赵晓非,张建华.一种基于跨领域典型相关性分析的迁移学习方法.计算机学报,2015(7):1326-1336.
- [32] 孟昭兰.情绪心理学.北京:北京大学出版社,2005.
- [54] 傅向华,刘国,郭岩岩,郭虎彪.中文博客多方面话题情感分析研究.中文信息学报,2013,27(1):47-55.
- [55] 徐琳宏,林鸿飞.认知视角下的文本情感计算.计算机科学,2010,37(12):182-185
- [59] 梁军,柴玉梅,原慧斌,高明磊,咎红英.基于极性转移和 LSTM 递归网络的情感分析.中文信息学报,2015,29(5):152-159.
- [64] 赵力,王治平,卢书,邹采荣,吴镇扬.全局和时序结构特征并用的语音信号情感特征识别方法.自动化学报,2004,30(3):423-429.
- [75] 陈月芬,崔跃利,王三秀.基于生理信号的情感识别技术综述.系统仿真技术,2017,13(1):1-5.
- [78] 温万惠,邱玉辉,刘光远,程南璞,黄希庭.情感生理反应样本库的建立与数据相关性分析.中国科学:信息科学,2011,41(1):77-89.
- [79] 李昕,杜笑娟,张云鹏,林立娟.基于生理信号的压力情感数据库的建立及分析.生物医学工程学杂志,2014(4):782-787.
- [99] 卜湛,伍之昂,曹杰,朱桂祥.在线评论情感计算与博弈预测.电子学报,2015,43(12):2530-2535.
- [102] 国玉晶,刘刚,刘健,郭军.基于环境特征的语音识别置信度研究.清华大学学报:自然科学版,2009,49(S1):1388-1392.
- [103] 白李娟,赵小蕾,毛启容,吴宝凤.基于声学上下文的语音情感特征提取与分析.小型微型计算机系统,2013,34(6):1451-1456.
- [106] 邵曦,陶凯云.基于音乐内容和歌词的音乐情感分类研究.计算机技术与发展,2015,08:184-187+193.
- [109] 刘翠娟,刘箴,柴艳杰,方昊,刘良平.基于微博文本数据分析的社会群体情感可视计算方法研究.北京大学学报(自然科学版),2016,52(1):178-186.
- [110] 饶元,吴连伟,张君毅.跨媒介舆情网络环境下信息传播机制研究与进展.中国科学:信息科学,2017,44:1-23.
- [111] 王元卓,于建业,邱雯,沈华伟,程学旗,林闯.网络群体行为的演化博弈模型与分析方法.计算机学报,2015,38(2):282-300.
- [114] 何炎祥,孙松涛,牛菲菲,李飞.用于微博情感分析的一种情感语义增强的深度学习模型.计算机学报,2017,40(4):773-790.
- [118] 周东浩,韩文报,王勇军.基于节点和信息特征的社会网络信息传播模.计算机研究与发展,2015,52(1):156-166.
- [119] 钱玉霞.国内虚拟品牌社区中网络口碑传播动机实证研究[硕士学位论文].合肥:中国科学技术大学,2014.
- [120] 张耀峰,肖人彬.基于元胞自动机的网络群体事件舆论同步的涌现机制.系统工程理论与实践,2014,34(10):2600-2008.
- [121] 王元卓,于建业,邱雯,沈华伟,程学旗,林闯.网络群体行为的演化博弈模型与分析方法.计算机学报,2015,38(2):282-300.
- [127] 徐睿峰,邹承天,郑燕珍,徐军,桂林,刘滨,王晓龙.一种基于情绪表达与情绪认知分离的新型情绪词典.中文信息学报,2013,27(6):82-89.
- [136] 陈师哲,王帅,金琴,多文化场景下的多模态情感识别.软件学报,2018,29(4).
- [138] 宋晓勇,吕品,陈年生.PORSC:融合用户个性化特征的在线评论情感分类模型.复旦学报(自然科学版),2017,56(03):359-365+373.
- [140] 梁静,颜文靖,吴奇,申寻兵,王甦菁,傅小兰.微表情研究的进展与展望.中国科学基金,2013(2):75-78.
- [141] 王甦菁,袁晔焯,颜文靖,傅小兰.微表情检测与识别的研究进展与趋势.CCF2016-2017 中国计算机科学技术发展报告(R),2017.286-313.