

计算机科学视角的社会情感计算

金升菊

(贵州民族大学工程实训中心, 贵州 贵阳 550025)

摘要: 随着计算机科学技术的发展, 社会情感计算成为一个社会科学与自然科学深度融合的一个重要课题。从认知理论、心理学视角, 社会情感是社会人群的思想情感集中表现, 受到自然人的视觉、听觉、触觉和认知能力的影响, 无法通过定量进行计算; 然而从计算机科学的视角, 社会情感是社会人的内心世界表达, 表现形式为人脸表情、语音情感、文字文本情感、生理信息、肢体行为等。本文从计算机科学出发回顾该语音情感计算领域研究成果, 并总结了语音情感计算面临的机遇与挑战。

关键词: 人才培养 教学改革 社会情感计算 大数据

中图分类号: TP391 **文献标识码:** A

0 引言

社会情感是否能进行定量的计算与评价一直以来受到社会学、心理学、认知科学以及计算机科学学界的关注。社会情感是否可计算可以归结为一个哲学终极问题^[1], 其衍生的情感能否计算? 如果不能, 那为什么不能计算? 如果能, 那么计算什么? 怎么计算? 计算效果如何评价? 计算结果价值何在等等一系列问题。1985年人工智能专家 Minsky 教授提出“计算机能够具有情感的能力”开启了情感计算的先河。美国麻省理工大学媒体实验的人工智能专家 Picard 教授在 1995 年进一步总结了情感计算的概念, 并于 1997 年撰写了《Affective Computing》让情感计算的大门正式打开。国际计算机以及人工智能业界领域创办了著名的社会情感计算权威期刊 (IEEE Transaction on Affective Computing) 主要研究和传递社会情感计算前沿技术和理论方法。同时, 两年一届的国际著名社会情感计算会议 Affective Computing and Intelligent Interaction (ACII) 也从学术发展视觉报道社会情感在认知理论、认知实践等方面对研究相关问题进行全面归纳总

结。2018 年 5 月在北京召开了首届亚洲情感计算机与智能交互会议 (ACII Asia 2018)。这些研究活动以及业界的研究成果说明了情感是可以计算的, 计算的对象可以是人脸表情、语音表述、文本文字、生理信息、肢体行为等等信号, 计算方法可以是基于模式识别的特征描述与识别框架 (包括了对应的特征提取以及识别算法), 计算效果评价有查准率、查全率、正确率、R-square、AIC、BIC 以及 F 值等等, 计算结果在舆情控制、教育教学、医疗辅助系统、智能侦查 (测谎系统等)、娱乐游戏、智能驾驶等领域都有着十分广泛的应用。

1 语音情感计算

语音情感计算是社会情感计算中的一个重要课题与方向, 其主要是通过对语音信号特征的统计结合识别算法进行计算分析。总体而言, 该方向研究重点是语音信号的特征分析和识别算法的设计与优化。以下分别介绍语音情感计算中的三个关键因素: 语音情感数据库、语音情感特征以及识别算法。

1.1 语音情感数据库

在计算科学研究中, 国际公开的数据集是基础。经过社会情感计算多年的

研究, 在国内外积累了一定数量的语音情感数据库。语音情感数据库分为离散型和连续型^[2-4], 在离散型中情感被分为高兴、愤怒、恐惧、惊奇、悲伤、厌恶等类型, 然而在连续型中是通过唤醒度、愉悦度、支配度等三维通过回归得出情感评分。离散型语音情感数据库有 Belfast 英语情感数据库、maribor 语音情感数据库、SUSAS 语音情感数据库、WCGS 语音情感数据库、柏林 EMO-DB 德语情感数据库、FAU/AIBO 儿童德语情感数据库、CASIA 汉语情感数据库、IEMOCAP 情感数据库、FERMUS 语音情感数据库、WOZ 语音情感数据库、ACCORPUS 系列汉语情感数据库等, 维度型语音情感计算数据库有 VAM 语音情感计算数据库、Semaine 语音情感计算数据库等。

除了以上的经典语音情感数据库外, 还有始于 2011 年每年一届的 International Audio/Visual+ Emotion Challenge and Workshop (AVEC) 年度竞赛, 每年均有新的语音情感数据库被建立并被使用。受国家社会科学重大项目基金资助的“贵州省少数民族语言资源有声数据库建设”也积累了大量汉语情感语料库, 目前可申请用于研究使用。

基金项目: 贵州民族大学—贵州省科技厅联合基金项目“基于稀疏表示的社会情感计算研究 (黔科合 LH 字 [2014]7392 号)”; 2014 年度国家社科基金重大招标项目“贵州省少数民族语言资源有声数据库建设” (14ZDB104)。

注释: ①部分文献将特征描述称为特征提取, 本文按照严格意义均采用特征描述。

1.2 语音情感特征

语音情感计算数据库文件一般都是视频文件或者音频文件。在这些原始数据的基础上通过特征描述^[5]将这些文件数据数字化为相应的矩阵或者向量。经过多年的发展,语音情感数据特征描述积淀了多种优秀的特征描述算法。语音情感属信号处理领域,因此时域频域处理两种典型框架^[5]。比较规范的是将特征归纳为韵律学特征、基于谱的相关特征、音质特征、融合特征。韵律特征有时长(duration)、基频(pitch)、能量(energy)等,典型的普特征是LPC(linear predictor coefficient), OSALPC(one-sided autocorrelation linear predictor coefficient), LFPC(log-frequency power coefficient)等线性谱特征以及LPCC(linear predictor cepstral coefficient), OSALPCC(cepstral-based OSALPC), MFCC(mel-frequency cepstral coefficient)等倒谱特征,分形维数谱特征是近年来用于语音情感分析的一种统计特征方法^[6]。音质特征则包括了共振峰频率及其带宽、频率微扰和振幅微扰、声门参数等。融合特征是类似于BoW的一种集成方法,其可以将前述三种方法或者更多特征描述方法进行融合以表达更多的语音信号信息。值得一提的是近年来卷积神经网络、循环神经网络、深度信念网络、生成对抗网络等深度学习为语音情感计算中特征描述与生成提供了新的理论、思路和方法。

1.3 情感计算算法

在语音情感计算中最后一个重要的环节就是识别(回归)算法。对于前述的离散型语音情感识别数据库中将采用的是分类识别方法,维度型语音情感计算数据库采用的是回归的方法。从计算机科学视角分类识别与回归算法本质是一致的。两者的区别是离散型情感识别的输出是高兴、恐惧等类别的标识而维度型输出结果则是连续型的具体数值。从统计学的视角二者研究方法和性质是一致的,主要的方法有如下几种。

基于概率论基础的贝叶斯分类或者

回归方法,该类方法以先验概率为基础,困难在于先验概率的获得。K近邻方法在数据点的K个在某种距离测度上最近点选择类别标签最多的类别作为该数据的类别(或者是K个近邻点的平均值)。决策树方法是一种自上而下逐步分解的策略,在分解过程中采用了熵等指标作为分解参考指标。支持向量机是一类以最小化结构误差风险作为目标的方法,其中一般将采用到核方法将在低维空间中线性不可分的数据投影到高维空间中实现线性可分^[6,7]。AdaBoost方法是一种逐步迭代、以指数函数作为损失函数、加权投票(系数)输出的集成学习方法。神经网络是一种由大量的节点(或称神经元)之间通过激励函数构成权值相互联接模拟人脑工作过程的方法,可以用在语音情感分类或者回归中。混合高斯模型是一类用多个高斯通过加权形式表示数据的方法,权值一般是通过EM算法得出。隐马尔科夫链是一种通过转移矩阵进行状态转移的方法,在真实参考矩阵和观察转移矩阵之后逐步调整以适应目标需求。稀疏表示则是在大数据环境下通过设计重构误差和原子稀疏性约束框架下,把语音情感数据特征构成矩阵,在矩阵中计算出一组基,通过求解1-范数条件下整个目标函数最优解,从而得到最稀疏表示系数^[8-9]。本质上,线性回归(Linear Regression)、逻辑回归(Logistic Regression)、多项式回归(Polynomial Regression)、逐步回归(Stepwise Regression)、岭回归(Ridge Regression)、套索回归(Lasso Regression)和ElasticNet回归均可用于语音情感计算的分类或回归问题。

2 语音情感计算面临机遇与挑战

经过学术界大量研究回答了社会情感是否可计算的问题。社会情感可以计算,计算的对象有人脸表情图像、语音表述、文字文本、生理信息信号、肢体姿态行为等等^[10-12]。当认知科学、神经科学、社会科学、心理学、计算机科学、统计学等领域把计算机看到了什么、计

算机听到了什么总结形成与知觉、视觉、听觉、学习、记忆、理解等模仿人体真实行为时,计算机能否知道人体表现出来的情感如何催生了情感计算新兴研究领域。

随着认知科学、神经科学、社会科学、心理学、语言学等学科的发展,同时计算机科学、统计学学科尤其是数据科学学科和人工智能学科深度融合与深入,社会情感计算领域面临诸多发展机遇。这些机遇主要体现在领域数据库的形成和交叉学科界限逐步消除。目前国际公开的Jaffe人脸表情库等人脸表情图像库已经相当成熟,柏林EMO-DB德语情感数据库也称为语音情感计算中的典范,通过网页爬虫技术获取的文字文本信息,通过可穿戴设备获取的生理信息信号以及通过摄像头获得的肢体姿态行为等数据库已经基本完善,研究者也可用数据采集设备、网络技术等建立相应的数据库。同时,计算机科学技术、人工智能技术、机器学习等技术已经开始普及,这些技术已经开始在认知科学、神经科学、社会科学、心理学、语言学中得到应用并取得一些成果。

虽然,语音情感计算取得了一些成绩,在很多机遇面前也同样面临诸多挑战。例如,拥有自适应能力语言情感信息的获取与建模问题,情感深度理解问题,智能人性化的实现问题等。在目前的语音情感数据库上数据样本均为实验环境下的样本,那么在自然环境下,尤其是在复杂环境中如何对语音情感数据采集尤其重要,对语音信号建模是核心关键。在现有的语音情感研究框架下,局限于分类或者是回归的问题,如何深度理解语音中潜在的情感问题是未来的研究方向之一。研究成果产品化是研究的目标之一,通过语音情感计算后如何实现智能任性的产品对于社会情感计算中具有挑战性的问题。

3 结束语

从计算机视觉出发探讨了语音情

(下转第64页)

协调开展好各类文体活动,让干部职工充分参与、亲身经历、实际获益,提高干部职工的本位意识和归属感。

拓展思想政治工作的载体。“工欲善其事,必先利其器。”思想政治工作的载体选择在很大程度上影响着方法的选择和实际的效率。一方面,对在实践中证明仍然有效的方法和载体,继续坚持并不断深化,努力创新内容形式,在互动性、吸引力上强效果,充分发挥传统方式方法灌输正确理论、强化主流意识、预防不良苗头的作用。另一方面,积极探索思想政治工作的新领域和新方法,与时俱进,利用干部职工们喜闻乐见、易于接受的新载

体开展思想政治工作,让新媒体成为实现“实施教育有目的、接受教育无意识”要求的有效手段。

打破思想政治工作的“单一性”,向“立体性”转变。思想政治工作并不是单一的、直向性的,而是立体的,交互性的。根据思想政治工作这一特性,应打破政工人员做思想政治工作的“专利”和思想政治工作的“单一性”,在组织管理上形成交叉管理,使政工人员和行政管理人员共同承担起思想政治工作和央行履职工作的双重职责,形成思想政治工作的辐射功能和群体优势,能更好地鼓励每位干部职工在各自的岗位上最大限度地发挥自己的

聪明才智,让广大员工达到思想上解惑、精神上解闷、文化上解渴、心理上解压的新要求,从而逐步形成大政工的思想政治工作格局。

参考文献:

- [1] 赵艳.新时期思想政治工作的创新路径探究[J].东方企业文化,2015(16):102.
- [2] 王志宽,李树彪,苑贵斌.新时期思想政治工作的创新理念研究[J].人才资源开发,2015(10):30.
- [3] 李春.新时期人本理念下企业思想政治工作方法创新初探[J].中国电力教育,2012(9):92,96.

(上接第61页)

感识别的问题。语音情感是可计算问题,其中包括了语音情感数据库、语音情感特征描述和语音情感计算的方法等。还探索了语音情感计算面临的基于与挑战。社会情感计算是一个充满生机的课题,语音情感计算是其中一个重要的分支。除了本文涉及的问题之外,还有诸如对高维语音信号的降维、无监督的语音信号聚类、识别算法中的优化、深度学习在语音情感识别中的应用等等均值得探讨与研究。

参考文献:

- [1] 李佳源.情感计算的研究现状与认知困境[J].自然辩证法通讯,2012,34(2):23-28,125.
- [2] 韩文静,李海峰,阮华斌,马琳.语

音情感识别研究进展综述[J].软件学报,2014,25(1):37-50.

- [3] 赵腊生,张强,魏小鹏.语音情感识别研究进展[J].计算机应用研究,2009,26(2):428-432.
- [4] 林奕琳,韦岗,杨康才.语音情感识别的研究进展[J].电路与系统学报,2007(1):90-98.
- [5] 王薇,杨丽萍,魏丽,刘艳.语音情感特征的提取与分析[J].实验室研究与探索,2013,32(7):91-94,191.
- [6] 李书玲,刘蓉,张鑒钦,刘红.基于改进型SVM算法的语音情感识别[J].计算机应用,2013,33(7):1938-1941.
- [7] 杨永健,聂瑜,吴洋,孙广志,杨

仲尧.基于SVM新的情感计算方法[J].吉林大学学报(信息科学版),

- 2017,35(4):438-442.
- [8] 李航,统计学习方法[M],北京:清华大学出版社,2015.
- [9] 王志良,解仑,董平.情感计算数学模型的研究初探[J].计算机工程,2004(21):33-34,167.
- [10] 张瑞.文本情感计算研究综述[J].管理观察,2017(13):28-30.
- [11] 赵思成,姚鸿勋.图像情感计算综述[J].智能计算机与应用,2017,7(1):1-5.
- [12] 张迎辉,林学闇.情感可以计算——情感计算综述[J].计算机科学,2008(5):5-8.