

DOI: 10.11992/tis.201804029

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.tp.20180927.1124.004.html>

触觉手势情感识别的超限学习方法

魏佳琪¹, 刘华平², 王博文¹, 孙富春²

(1. 河北工业大学 省部共建电工装备可靠性与智能化国家重点实验室, 天津 300130; 2. 清华大学 智能技术与系统国家重点实验室, 北京 100084)

摘要:为了解决声音和图像情感识别的不足, 提出一种新的情感识别方式: 触觉情感识别。对 CoST(corpus of social touch) 数据集进行了一系列触觉情感识别研究, 对 CoST 数据集进行数据预处理, 提出一些关于触觉情感识别的特征。利用极限学习机分类器探究不同手势下的情感识别, 对 14 种手势下的 3 种情感(温柔、正常、暴躁)进行识别, 准确度较高, 且识别速度快识别时间短。结果表明, 手势的不同会影响情感识别的准确率, 其中手势“stroke”的识别效果在不同分类器下的分类精度均为最高, 且有较好的分类精度, 达到 72.07%; 极限学习机作为触觉情感识别的分类器, 具有较好的分类效果, 识别速度快; 有的手势本身对应着某种情感, 从而影响分类结果。

关键词: 触觉; 情感识别; 极限学习机; 特征提取; 触摸手势; 支持向量机; 人机交互; 机器学习

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1673-4785(2019)01-0127-07

中文引用格式: 魏佳琪, 刘华平, 王博文, 等. 触觉手势情感识别的超限学习方法[J]. 智能系统学报, 2019, 14(1): 127-133.

英文引用格式: WEI Jiaqi, LIU Huaping, WANG Bowen, et al. Extreme learning machine for emotion recognition of tactile gestures[J]. CAAI transactions on intelligent systems, 2019, 14(1): 127-133.

Extreme learning machine for emotion recognition of tactile gestures

WEI Jiaqi¹, LIU Huaping², WANG Bowen¹, SUN Fuchun²

(1. State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300130, China; 2. State Key Lab. of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China)

Abstract: To overcome the deficiencies of sound and image emotion recognition, a new emotion recognition method, haptic emotion recognition, is proposed. A series of haptic emotion recognition studies on Corpus of Social Touch (CoST) datasets were performed. First, the CoST data was preprocessed, presenting some features about haptic emotion recognition. Using the extreme learning machine classifier to explore emotion recognition under different gestures, three kinds of emotions, gentle, normal, and irritable, under 14 kinds of gestures, were identified with higher accuracy and a faster recognition speed (0.04 s). The results showed that differences in gestures will affect the accuracy of emotion recognition, wherein the recognition effect of the gesture “stroke” is the highest in classification accuracy under different classifiers. This new method yielded better classification accuracy, reaching 72.07%. As a classifier of haptic emotion recognition, the extreme learning machine had better classification effect and faster recognition speed. Some gestures corresponded to certain emotions, which affected the classification results.

Keywords: haptic; emotion recognition; extreme learning machine; feature extraction; touch gesture; support vector machine; human-computer interaction; machine learning

随着计算机技术的发展, 人机交互的领域愈加广泛, 机器人的情感理解能力逐渐成为衡量智能程度的标准。触觉是最基本交互方式之一, 无

论是对人类还是机器人, 触觉信息都是与周围环境进行交互的重要媒介^[1]。人们逐渐意识到, 触觉可以用来调节情绪, 增加人与人之间情感和信任, 帮助建立良好的人机关系。同时, 越来越多的人使用触摸功能的机器人完成医疗以及远程数学等任务。

收稿日期: 2018-04-18. 网络出版日期: 2018-09-29.

基金项目: 国家自然科学基金重点项目(U1613212); 河北省自然科学基金项目(E2017202035).

通信作者: 刘华平. E-mail: hpliu@tsinghua.edu.cn.

图像和声音一直以来是情感状态识别的主流方式,但有些时候图像和语音并不能真实地反映出用户当时的情感状态。图像识别在计算量上特别大,而且非常依赖光照条件。比如在赛车时,由于头盔限制了图像的采集,并不能有效地利用图像识别出驾驶员的情感信号;语音识别同时也存在着缺陷,用户通过语音来表达情感时,常常伴随着语义的理解问题,需要通过语料的内容才能正确地判断说话人的情绪。因此,发展一种新的情感识别研究方向成为趋势。

触觉在复杂的社会信息中扮演着重要角色,情感识别正逐步成为人机交互领域研究的热点和发展方向^[2]。人机情感交互让机器人更加了解人类的情绪,并做出相应的反馈。这也让触觉情感识别走进了人类生活。例如:智能家居能够通过触摸来识别出用户的情感,进而做出反应调整灯光或音乐;在线学习的用户通过触摸屏幕传递情感信息,当用户面对学习系统表现出烦躁或者焦虑情绪时,系统能够及时地给用户鼓励信息,使用户的注意力重新转移到学习中;在医疗方面还可以用情感识别来治疗自闭症患者和对养老院老人进行心理疏导。

1 相关工作

Morrison 等^[3]用实验证明触摸是传递亲密情感的主要方式;Debrot 等^[4]证实了情侣之间的积极接触增强了他们的情感状态;Park 等^[5]设计了触觉电话,证实触觉在面对面交流中的作用;Guest 等^[6]研究触觉与情感之间的联系,并建立了一份英语中最常用的形容词,用来描述触摸体验;Kim 等^[7]提出一种基于神经网络的纺织品-情绪索引方法,研究情感与图案之间的联系;Huang 等^[8]利用神经网络得到从图像特征空间到心理空间的映射函数,确定图像与情感之间的联系;Jung 等^[9]建立 CoST(corpus of social touch) 触觉数据集,对 14 种触觉手势进行识别分类,准确率达到 60%;周楠等^[10]利用 3D 卷积神经网络对触觉手势识别进行了改进;Hughes 等^[11]设计了一种新型传感器,将距离信号和力信号用于手势识别,并证明与单独的力信号相比,近距离感知信号增加了手势识别的准确性;Maramis 等^[12]利用安卓手机设计开发了一款 APP 来引导用户自然地表达情感并进行触觉情感识别;Gao 等^[13]在 ipod 上通过游戏探究触觉行为是否反应玩家的情绪,并用机器学习完成 4 个情绪状态的分类。但是触摸屏在硬件方面存在一定的局限性,并不能像压阻式传感器一样具有很好的延展性和可弯曲性^[14],且仅能

通过在触摸屏上滑动来采集数据,形式较为单一。本文提出了一种基于阵列式触觉传感器的情感识别方式,并在手势识别数据集上进行了实验,探究不同手势下的情感识别。

2 情感识别

基于触觉手势的情感识别技术主要分为数据采集、数据预处理、特征提取、情感识别以及结果分析 5 个部分,如图 1 所示。数据采集通过阵列式触觉传感器固定在载体上采集不同参与者的压力值数据;数据预处理是将已采集的压力值数据随时间变化的数据进行处理,比如除掉错误的帧或者错误的样本数据;特征提取是对处理后的压力值数据进行分析,并提出情感识别相关的特征进行提取;情感识别是将提取的情感识别特征输入到分类器中进行训练;最后将情感识别的结果进行分析。

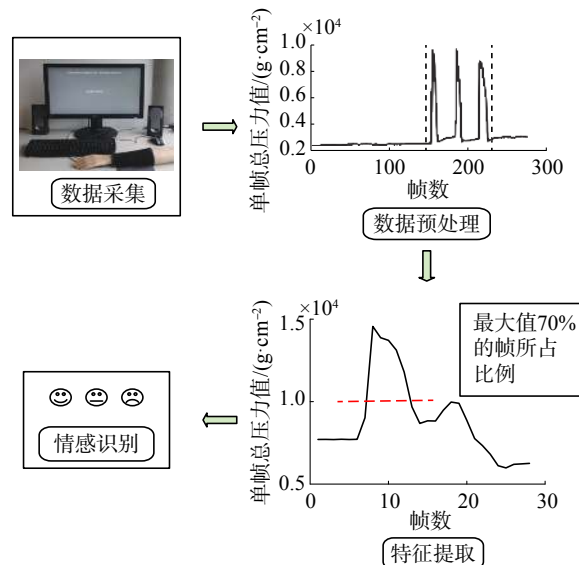


图1 情感识别流程图

Fig. 1 The flow chart of emotion recognition

本文研究的基于触觉手势的情感识别技术是对多种手势进行情感识别,目的是探究手势对情感识别结果产生的影响以及情感之间的联系。首先,对采集到的数据进行数据预处理,将“不规则”的帧排除;将预处理好的数据按照手势的不同拆分;再对其中每一帧的数据进行分析,提出情感识别相关的特征;将每个特征集按照参与者划分为训练集和测试集;最后用极限学习机 (ELM) 分类器测试实验结果,并对实验结果进行对比分析。

ELM 是新加坡南洋理工大学的黄广斌教授提出来的针对单个隐含层的前馈型神经网络 (SLFNs) 的监督型学习算法^[15],其主要思想是:输入层与隐含层之间的权值参数以及隐藏层上的偏置向量参数,是不需要像其他基于梯度的学习算法一样通

过迭代反复调整刷新, 而是直接求解一个最小范数最小二乘问题, 最终化归成一个矩阵的广义逆问题。ELM 的原理如图 2 所示。

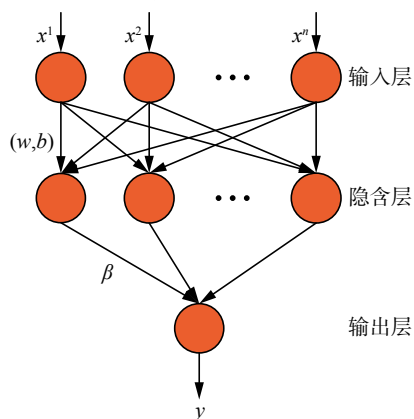


图 2 ELM 的网络结构

Fig. 2 The model structure of ELM

3 数据处理及特征提取方法

在以往的研究中, 大部分学者把重心放在手势识别上^[16], 本文主要对不同的手势进行情感识别。对数据集的处理一般包括 3 部分: 数据预处理、特征提取、特征选择。因为特征维度并不大, 所以在本文中并没有使用常用的随机森林模型进行特征选择。

文献^[16]中提出的 CoST 数据集包含 31 个参与者, 每个参与者根据要求分别执行 14 个手势, 每个手势重复 6 次, 每个动作分别以 3 种情感(温柔、普通、暴躁)来执行。也就是说, 每个参与者都会执行 252 次动作。手势包括 grab、hit、massage、pat、pinch、poke、press、rub、scratch、slap、squeeze、stroke、tap、tickle。表 1 是对 14 个手势的定义^[17]。

表 1 手势的定义

Table 1 Gesture definition

手势名称	手势定义
grab	突然抓住手臂
hit	对手臂进行打击
massage	按摩手臂
pat	轻轻的拍手臂
pinch	拧手臂
poke	戳手臂
press	按压手臂
rub	用不变的压力搓手臂
scratch	用指甲挠手臂
slap	扇手臂
squeeze	挤手臂
stroke	手在手臂上来回划动
tap	敲手臂
tickle	用手指轻触手臂

该数据集的采集由安装在假肢手臂上的 8×8 阵列式触觉传感器以 135 f/s 的频率完成。每一个手势均由多个帧的数据构成, 数据集的每一帧都由 64 个传感器通道的压力值组成, 压力值单位为^[18] g/cm^2 。图 3 为传感器的 8×8 压力通道一帧的灰度图, 白色部分对应压力值最高的通道, 黑色部分对应压力值最低的通道。

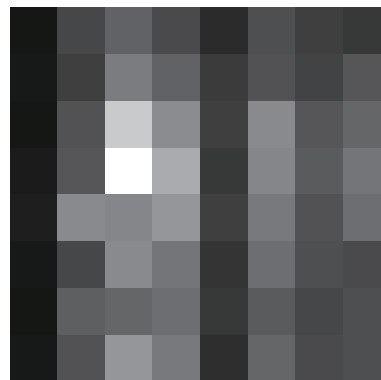


图 3 传感器的灰度图

Fig. 3 Gray scale image of sensor

3.1 数据预处理

首先, 对部分数据进行观察, 画出 CoST 数据集帧数与每一帧压力总和之间的关系图, 如图 4 所示, 但手势的前半部分(第 1 个虚线前)属于手势执行前的噪声, 虚线内的部分才是手势执行的过程, 所以虚线前并无实际意义, 需要删除此部分。因此对整体的 CoST 数据集进行分割, 以此来删除掉手势中无意义的帧。这样就能保证计算的时间序列特征更有意义。

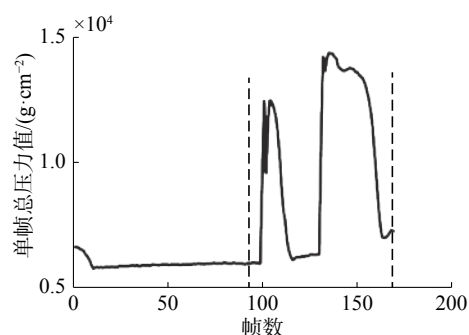


图 4 对“暴躁的 massage”的分割图

Fig. 4 Segmentation of a “rough massage”

3.2 特征提取

不同的情感环境下, 最主要的 3 个影响因素分别为压力值、传感器的接触面积和触摸的有效时间。显然, 不同情感的平均压力值和最大压力值是判别情感程度的初步特征, 压力值由小到大分别为: 温柔的动作压力值、正常的动作压力值、暴躁的动作压力值。光有这两个特征远远不够。

情感的准确辨别,至少需要2种行为特征:触觉行为的持续时间和强度^[19]。因此我们对其他特征进行了探究。

首先对 CoST 数据集的 poke 手势进行了分析,如图5、图6所示,通过对手势的压力值大小与时间关系的二维图、三维图的分析,进而推断出影响情感的因素。

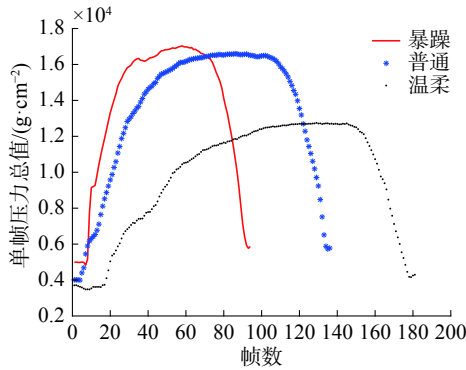


图5 3种情感下的二维图

Fig. 5 A two-dimensional comparison of three emotions

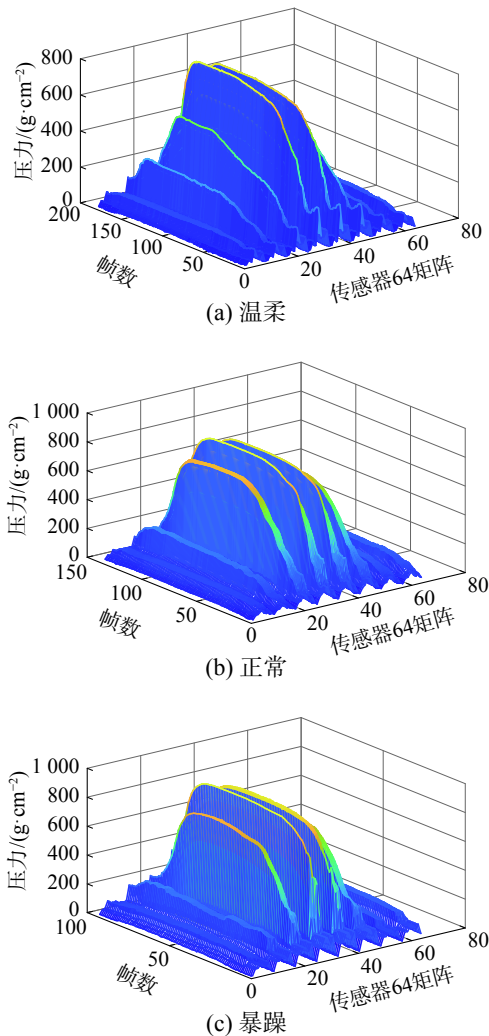


图6 3种情感下三维图

Fig. 6 A 3D comparison of three emotions

- 1) 一帧中阵列传感器所有通道的平均值。
- 2) 所有帧所有通道的平均值和最大值。
- 3) 该动作的所有帧当中,无信号的百分比。如果这一帧的64列的平均压力值小于 60 g/cm^2 ,那么就说明这个帧是没有信号的。
- 4) 所有帧当中每一列的平均压力。
- 5) 所有帧当中每一行的平均压力。
- 6) 每一帧的总压力,然后求帧与帧之间总压力差的绝对值。此特征显示该情感下不同时刻间力的大小变化。
- 7) 取特征2)中的平均值,计算达到平均值的帧数所占该动作总帧数百分比。
- 8) 求出每一帧中64个通道的最大值,求当前帧下,达到50%最大值的数量。
- 9) 接触面积。
- 10) 压力峰值的间距。每一帧的压力值总和随着帧数变化,其中极大值之间的时间差。
- 11) 每一个动作下不同情感的有效接触时间。即求出每个动作各个情感的总有效帧数。

由此得到数据特征集。边缘检测是图像处理 and 计算机视觉中的常用手段,目的是标识图像中亮度变化明显的点^[20]。图像属性中的显著变化通常反应了属性的重要事件和变化。用于视频分类的特征可以用于 CoST 数据集,因为 CoST 数据是以固定速率更新的网格形式的压力值,其类似于低分辨率灰度视频。利用 sobel 算子,通过锐化对比度,构建了第2组数据帧,使用3.2节中相同的提取特征程序,获得一组新的数据特征集。将两组数据特征集合整合到一起。

3.3 分类器实验及结果

本文采用了3种分类器:支持向量机(SVM)、极限学习机(ELM)和随机森林(RF)。本文的支持向量机分类器使用了 libsvm 工具箱的 RBF 核函数来训练样本;本文的 ELM 分类器对参数最优值进行了研究,如图7所示,激活函数设置为 sine,隐含层数量设置为100时,ELM 的分类效果最好。随机森林是一种包含多个决策树的分类器,并且其输出的类别是由个别树输出的类别的众数而定,决策树数量设置为300时,随机森林的分类效果最好。

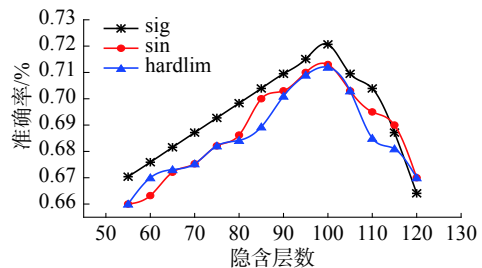


图7 ELM 参数灵敏图

Fig. 7 Parameter-sensitive map of ELM

将 31 个参与者中随机抽取的 21 个人的所有特征数据当作训练集, 另外 10 个人的所有特征数据当作测试集。将训练集和测试集再按照手势分成 14 组, 依次进行分类识别。分类结果如表 2 所示, 从 14 个手势的整体来看, SVM 和随机森林的分类准确率较为相似, ELM 无论是从准确度上还是分类时间上都要优于其他 2 个分类器。图 8 中, 雷达图的最外层代表 80% 的准确率, 最内层代表 30%。从图 8 可以看出, 14 个手势中 stroke 手势的情感分类效果最好, 在 3 个分类器的分类结果中均为最高, 在触觉情感识别中, ELM 的输入层与隐含层之间的权值参数以及隐藏层上的偏置向量参数是不需要像其他基于梯度的学习算法一样通过迭代反复调整刷新, 而是直接利用情感特征求解一个最小范数最小二乘问题。所以 ELM 在情感识别中效果较优。其中, ELM 分类器对 stroke 手势的情感识别分类精度达到 70.95%。图 9 为 SVM 分类器下的混淆矩阵图, 从图中可以看出 stroke 手势在 3 种不同的情感状态下均有较好的分类效果; 手势和情感之间存在着某些联系, 不同的手势本身对应着不同情感, 会对实验结果产生影响。比如当人们执行 pinch 手势时, 更倾向于暴躁情绪, 同时暴躁情感的识别准确度也最高, 也更容易将其他 2 种情感混为暴躁情绪; 而 press 手势则被容易理解为一种温柔情感下的动作, 其在执行和识别的过程中更倾向于温柔情绪。

表 2 情感识别准确率

Table 2 Accuracy of emotion classification %

分类结果	SVM	ELM	随机森林
grab	58.89	63.33	59.44
hit	51.11	59.44	53.89
massage	50.56	58.33	51.11
pat	57.78	61.11	54.44
pinch	58.89	64.44	57.78
poke	47.22	56.67	51.67
press	55.56	57.78	55.00
rub	65.56	65.00	59.64
scratch	64.44	62.78	65.56
slap	52.78	57.78	52.22
squeeze	50.00	56.67	51.67
stroke	70.95	72.07	67.60
tap	51.67	58.89	54.44
tickle	62.22	62.22	52.22
平均	56.97	61.18	56.19

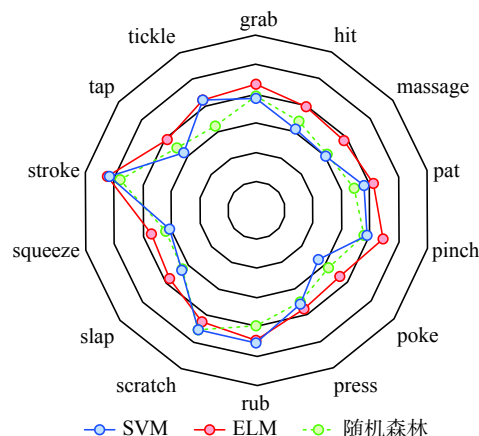
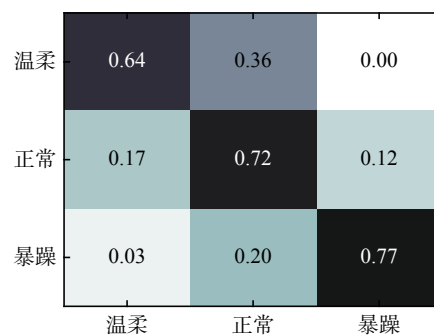
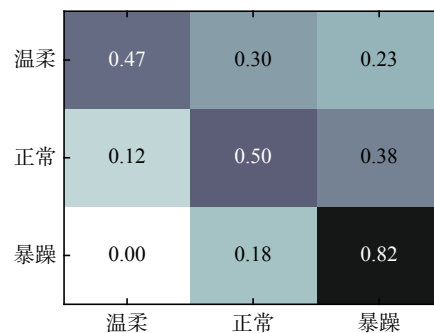


图 8 情感识别雷达图

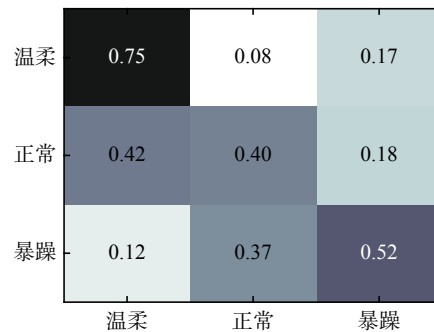
Fig. 8 Radar map of emotion classification



(a) 划



(b) 拧



(c) 按

图 9 手势的情感识别混淆矩阵

Fig. 9 Emotion recognition confusion matrix of the gestures

4 结束语

本文提出了基于触觉手势的情感识别技术,

通过对 CoST 数据集的 7 805 个手势进行分析, 对数据集进行数据预处理, 剔除部分错误的手势, 使数据集更加完善。使用 MATLAB 等软件对数据进行整理, 并通过相关论文引证, 提出一系列情感识别的特征。使用多种分类器进行分类对比, 整体来看 14 种手势中 SVM 和随机森林的情感识别效果不相伯仲。不同分类器下的 stroke 手势的情感识别效果均为最高, 但 ELM 的 stroke 手势的情感识别效果要比基于 RBF 核函数的 SVM 分类器的效果好 (SVM=70.95%, ELM=72.07%)。并且 ELM 的识别时间要明显短于 SVM (SVM=0.33 s, ELM=0.04 s)。本文针对 CoST 数据集进行了一系列研究实验, 得到了 3 点结论: stroke 手势具有最好的情感识别效果, 且分类精度较高; ELM 作为触觉情感识别的分类器具有较好的表现, 识别精度高且识别速度快; 有的手势 (比如 pinch、press) 本身对应着某种情感, 会影响情感识别的结果。因此, 在设计触觉情感识别传感器时, 可以诱导用户做出情感识别率最高的手势, 从而间接地提高分类精度。这为以后的触觉情感识别系统的设计提供了思路, 也为建立新的情感识别数据集奠定了基础。

参考文献:

- [1] 马蕊, 刘华平, 孙富春, 等. 基于触觉序列的物体分类[J]. 智能系统学报, 2015, 10(3): 362–368.
MA Rui, LIU Huaping, SUN Fuchun, et al. Object classification based on the tactile sequence[J]. CAAI transactions on intelligent systems, 2015, 10(3): 362–368.
- [2] 郝敏, 刘光远, 温万惠. 基于进化策略的生理信号情感识别[J]. 智能系统学报, 2009, 4(4): 352–356.
HAO Min, LIU Guangyuan, WEN Wanhui. Recognition of emotion in physiological signals using evolutionary strategies[J]. CAAI transactions on intelligent systems, 2009, 4(4): 352–356.
- [3] MORRISON I, LÖKEN L S, OLAUSSON H. The skin as a social organ[J]. *Experimental brain research*, 2010, 204(3): 305–314.
- [4] DEBROT A, SCHOEBI D, PERREZ M, et al. Touch as an interpersonal emotion regulation process in couples' daily lives: the mediating role of psychological intimacy[J]. *Personality and social psychology bulletin*, 2013, 39(10): 1373–1385.
- [5] PARK Y W, BAEK K M, NAM T J. The roles of touch during phone conversations: long-distance couples' use of POKE in their homes[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Paris, France, 2013: 1679–1688.
- [6] GUEST S, DESSIRIER J M, MEHRABYAN A, et al. The development and validation of sensory and emotional scales of touch perception[J]. *Attention, perception, & psychophysics*, 2011, 73(2): 531–550.
- [7] KIM N Y, SHIN Y, KIM E Y. Emotion-based textile indexing using neural networks[C]//12th International Conference on Human-Computer Interaction. Beijing, China, 2007: 349–357.
- [8] HUANG Xinyin, SOBUE S, KANDA T, et al. Linking KANSAI and image features by multi-layer neural networks[C]//11th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems. Vietri sul Mare, Italy, 2007: 318–325.
- [9] JUNG M M, CANG X L, POEL M, et al. Touch challenge'15: recognizing social touch gestures[C]//Proceedings of the 2015 ACM on International Conference on Multimodal Interaction. Seattle, Washington, USA, 2015: 387–390.
- [10] ZHOU Nan, DU Jun. Recognition of social touch gestures using 3D convolutional neural networks[C]//Chinese Conference on Pattern Recognition. Chengdu, China, 2016: 164–173.
- [11] HUGHES D, LAMMIE J, CORRELL N. A Robotic skin for collision avoidance and affective touch recognition[J]. *IEEE robotics and automation letters*, 2018, 3(3): 1386–1393.
- [12] MARAMIS C, STEFANOPOULOS L, CHOUVARDA I, et al. Emotion recognition from haptic touch on android device screens[M]//MAGLAVERAS N, CHOUVARDA I, DE CARVALHO P. Precision Medicine Powered by pHealth and Connected Health. Singapore: Springer, 2018: 205–209.
- [13] GAO Yuan, BIANCHI-BERTHOUEZ N, MENG Hongying. What does touch tell us about emotions in touch-screen-based gameplay?[J]. *ACM transactions on computer-human interaction*, 2012, 19(4): 31.
- [14] 林连冬, 李思奇, 陈春雨, 等. 触觉传感器非线性补偿仿生算法[J]. 哈尔滨工程大学学报, 2017, 38(2): 288–292.
LIN Liandong, LI Siqi, CHEN Chunyu, et al. Bionic algorithm for nonlinear compensation of tactile sensors[J]. *Journal of Harbin engineering university*, 2017, 38(2): 288–292.
- [15] HUANG Guangbin, WANG Dianhui, LAN Yuan. Extreme learning machines: a survey[J]. *International journal of machine learning and cybernetics*, 2011, 2(2): 107–122.
- [16] JUNG M M, POPPE R, POEL M, et al. Touching the void--introducing CoST: corpus of social touch[C]//Pro-

ceedings of the 16th International Conference on Multimodal Interaction. Istanbul, Turkey, 2014: 120–127.

- [17] TA V C, JOHAL W, PORTAZ M, et al. The Grenoble system for the social touch challenge at ICMI 2015[C]// Proceedings of the 2015 ACM on International Conference on Multimodal Interaction. Seattle, Washington, USA, 2015: 391–398.
- [18] JUNG M M, POEL M, POPPE R, et al. Automatic recognition of touch gestures in the corpus of social touch[J]. *Journal on multimodal user interfaces*, 2017, 11(1): 81–96.
- [19] YOHANAN S, MACLEAN K E. The Role of affective touch in human-robot interaction: human intent and expectations in touching the haptic creature[J]. *International journal of social robotics*, 2012, 4(2): 163–180.
- [20] 顾海巍, 樊绍巍, 金明河, 等. 基于灵巧手触觉信息的未知物体类人探索策略[J]. 哈尔滨工程大学学报, 2016, 37(10): 1400–1407.
- GU Haiwei, FAN Shaowei, JIN Minghe, et al. An anthropomorphic exploration strategy of unknown object based on haptic information of dexterous robot hand[J]. *Journal of Harbin engineering university*, 2016, 37(10): 1400–1407.

作者简介:



魏佳琪, 男, 1995 年生, 硕士研究生, 主要研究方向为新型磁性材料与器件、触觉交互。



刘华平, 男, 1976 年生, 副教授, 博士生导师, IEEE Senior Member、中国人工智能学会理事, 中国人工智能学会认知系统与信息处理专业委员会秘书长, 主要研究方向为机器人感知、学习与控制、多模态信息融合。在 IEEE Trans. On Automatic Control、IEEE Trans. on Circuits and Systems II 以及 Automatica 等国际期刊, 以及 ICRA、IROS 等国际会议中发表论文十余篇。



王博文, 男, 1956 年生, 教授, 博士生导师, 主要研究方向为磁致伸缩材料与器件、振动发电技术、磁特性测试技术。发表学术论文 200 余篇, 被 SCI 收录 100 余篇。