

# 面向可穿戴多模生物信息传感网络的栈式 自编码器优化情绪识别

戴逸翔 王 雪 戴 鹏 张蔚航 张鹏博

(清华大学精密仪器系精密测试技术及仪器国家重点实验室 北京 100084)

**摘 要** 情绪识别是指采用无生命的传感器和计算机感知测量识别人类情绪状态,其主要环节包括情绪相关信号获取、特征提取以及分类识别.情绪识别可为人类情绪健康监测乃至情绪相关心理精神疾病的初筛提供科学依据.该文构建了多模可穿戴生物信息传感网络被测个体的多模情绪相关信号(脑电、脉搏以及血压),经由身体主站将信号传输至远程网络数据中心,并将情绪识别的结果进行网络发布,简化了测量结构,使得被测个体日常情绪监测和远程监控成为可能.由于信号测量和特征提取过程中存在不确定性,该文提出了栈式自编码器(基于深度学习理论)优化的情绪识别算法.71天时间跨度的实验结果表明,栈式自编码器预学习后的特征向量具有更高的一致性与可分性,情绪识别率较相关研究提高了约5%.

**关键词** 情绪识别;多模感知测量;可穿戴生物信息传感网络;栈式自编码器;深度学习;物联网;传感器网络  
**中图法分类号** TP391 **DOI号** 10.11897/SP.J.1016.2017.01750

## Stacked Auto-Encoder Optimized Emotion Recognition in Multimodal Wearable Biosensor Network

DAI Yi-Xiang WANG Xue DAI Peng ZHANG Wei-Hang ZHANG Peng-Bo

(Department of Precision Instrument, State Key Laboratory of Precision Measurement Technology and Instruments,  
Tsinghua University, Beijing 100084)

**Abstract** Emotional health draws great concern with the enhancement of public health consciousness. Emotional health is closely related to the quality of personal life. Even for some special groups of people, like pilots, soldiers, etc., their emotional states will have impacts on the stability of communities. Traditionally, to evaluate emotional states of human beings relies on the doctors or psychologists to communicate with subjects and give scores based on various questionnaires. This approach is not scientific enough and leads to the difficulties in the emotional health monitoring in daily-life. Emotion recognition enables lifeless sensors and computers to measure and interpret human emotions. It is a procedure consisting of emotion-related bio signals recording, features extraction and emotional states classification, providing scientific evidence for emotional health monitoring and primary diagnosis of potential mental diseases. In the related works concerning emotion recognition, the application scenarios are usually restricted in the hospitals or labs and the common-used classifiers are not suitable for the daily emotion recognition data set. This paper develops a multimodal biosensor network to simplify the sensing framework so that it can finish emotion recognition tasks when subjects are participating in daily tasks without so

收稿日期:2016-12-18;在线出版日期:2017-03-16. 本课题得到国家自然科学基金(61472216)、教育部博士点基金(20120002110067)、精密测试及仪器国家重点实验室研究项目资助.戴逸翔,男,1990年生,博士研究生,主要研究方向为生物信息感知网络与认知信号测量解析. E-mail: daiyx12@mails.tsinghua.edu.cn.王雪(通信作者),男,1963年生,博士,教授,博士生导师,主要研究领域为精密测试传感技术和无线传感网络测量. E-mail: wangxue@mail.tsinghua.edu.cn.戴鹏,男,1990年生,博士研究生,主要研究方向为视觉传感网络与机器视觉.张蔚航,男,1992年生,博士研究生,主要研究方向为医学图像处理与生理信号感知.张鹏博,男,1992年生,博士研究生,主要研究方向为生理信号测量与分析.

many disturbances. Several wearable biosensor nodes record multimodal bio signals (electroencephalography, pulse and blood pressure) and transmit them to a body station employing Arduino UNO3 and its expansion boards. The body station works as a web client connecting to a web data center on the Internet by wireless routers or personal hotspots. The web data center is established on NI-PXI 1065 with a static public IP address. The recognition algorithm is carried out in the data center and the results are displayed for authorized web terminals with the assistance of web publishing service supported by LabVIEW. The multimodal wearable biosensor network can provide emotion-related bio signals from which typical features are extracted based on the existing theories. In particular, due to the uncertainties in signal acquisition and feature extraction, a stacked auto-encoder (based on the deep learning theory) optimized emotion recognition method is proposed to improve the recognition process. The stacked auto-encoder helps to pre-learn the feature vectors and with the fine tuning it generates a better scheme for emotion classification phase. There are 9 emotional states for classification according to the Valence-Arousal dimensional model. A two-layer stacked neural network with a softmax classifier is designed to finish the final classification tasks. The experiment convinces that the feature vectors pre-learned by stacked auto-encoder are of higher quality both in centrality and distinguishability based on the similarity evaluation theory. The final recognition rate is also improved approximately 5% compared to related works. The main contributions of this paper are the wearable network-based sensing structure, the stacked auto-encoder optimized multimodal emotion recognition algorithm and the quantitative analysis on 71-day experimental data. This is a novel system for daily emotional health monitoring and can provide scientific suggestions for doctors or guardians. However, in the future, large scale of data should be accumulated. Moreover, the dynamic performance and energy efficiency also need improving.

**Keywords** emotion recognition; multimodal sensing; wearable biosensor network; stacked auto-encoder; deep learning; Internet of Things; sensor networks

## 1 引言

随着全民健康意识的普遍增强,情绪健康成为大众日益关注的问题。不良情绪状态不仅影响个人生活质量,也极有可能是严重精神心理疾病的前兆或成因。对于某些特殊群体,如驾驶员、飞行员、军人等,该类人群的情绪健康甚至会影响到公共安全和稳定<sup>[1]</sup>。所以,对情绪健康的定量分析与监测识别成为近年来的研究热点<sup>[2]</sup>。传统的情绪状态评估是由医生或者心理学工作者与被测对象进行互动,通过包含特定任务的量表对被测对象进行打分评测,主观性较强且缺乏远距离实时监控的手段。即便在医院病房,例如近年来开展的“阳光病房”等试点项目,医生为了满足对病人情绪监控的需求,也只能借由查房、问诊等手段实现。因而亟需发展应用于日常生活情绪健康评估的有效手段。面向情绪识别,采用传感器、计算机等非生命体观察被测对象的情绪

特征信号进行观测记录、特征提取和计算分类<sup>[3]</sup>,能够实现情绪状态的实时客观数据分析,从而为个体情绪健康状态的综合评估乃至精神心理疾病的日常初步筛查提供科学依据。

个体情绪的定量描述较为复杂,依据近一个世纪应用心理学家普拉切克、罗素等人的研究成果,目前对情绪进行描述和分类主要依据两类模型:离散模型和维度模型<sup>[4]</sup>。离散模型将人的个体情绪分为基本情绪和复合情绪,复合情绪由基本情绪依据各种规则组合叠加而成,但基本情绪的界定仍然存在争议。维度模型中的情绪环状模式是目前采用最为广泛的情绪描述分类模型<sup>[4]</sup>。

如图1所示,依据维度模型,以效价(Valence)和唤醒度(Arousal)两个维度描述各类情绪。效价度量情绪的极性,从负面情绪过渡到正面情绪,效价也随之增大。唤醒度度量情绪的唤醒程度,随着唤醒度的提高,被测个体的警觉性也随之增加。两者经由2-D SAM(Self-Assessment Manikins)量表实现标

定<sup>[5]</sup>. 该效价-唤醒度 Circumplex 模型足以描述大部分情绪, 在已有的研究成果中被广泛使用.

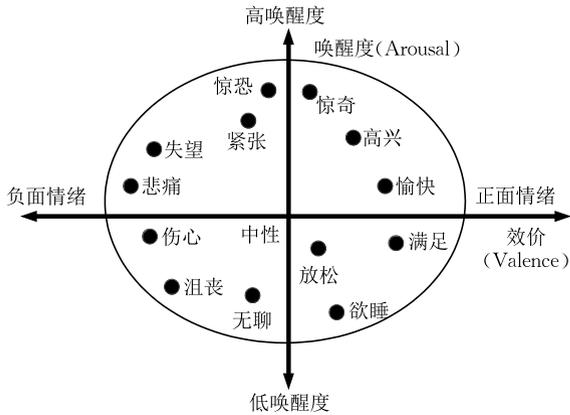


图 1 情绪识别维度环状理论模型

目前已有多种模态的生物指征参数(如脑电、心电、呼吸、皮肤阻抗、表情、声纹、词频等)均被证明与情绪变化紧密关联<sup>[3,6-8]</sup>. 大脑是人类情绪思维的中心, 脑电信号可以直接反映人类的情绪变化. 国内外研究者也基于此展开了面向情绪监测识别的相关研究. Chanel 等先后选择包含情绪色彩的图片、回忆以及游戏作为情绪刺激源, 提取脑电 (Electroencephalography, EEG) 功率谱密度 (Power Spectrum Density, PSD) 和各波段能量谱特征, 先后采用朴素贝叶斯 (Naïve Bayes, NB)、支持向量机 (Support Vector Machine, SVM) 等分类器实现了 3 类情绪的分类, 识别率为 56%~63%<sup>[1,9-10]</sup>. Lin 等人<sup>[11]</sup> 基于 SVM 分类器, 采用高精度医疗设备实现了医疗级 4 类情绪的高精度分类, 识别率为 83%. Murugappan 等人<sup>[12]</sup> 结合功率谱与小波特征, 采用  $K$  最近邻分类器 ( $K$  Nearest Neighbor, KNN) 实现情绪分类.

情绪监测识别的最新研究趋势是对多模情绪相关特征信号进行融合分析. 多模融合情绪识别研究的先行者麻省理工学院 Media Lab 的 Piscard 等率先采用心电、血压、皮肤温度、脉搏、血容量、呼吸、肌电等多模信号实现了对 8 种情绪的分类识别, 从多模生理信号中提取了均值、方差、一阶二阶差分等 40 余种特征, 最终对愤怒、高兴、悲伤三类情绪分类达到 80% 以上的识别率, 但是 8 类情绪分类识别率仅有 40.6%<sup>[13]</sup>. 该研究组发布了最早的多模情绪识别数据库之一, 共有 20 天时间跨度的数据用于 8 类情绪(愤怒、厌烦、悲伤、兴奋、平和、平静、敬畏、喜悦)的识别. 随后, 该研究组又进一步开发出了用于真实驾驶环境的心电、皮肤阻抗、呼吸等多模信号情绪识别系统, 用于驾驶员的压力状态监测<sup>[14]</sup>. 近十

年来, 该研究组又针对自闭症患者的情绪监测识别应用, 研制了面向无线网络与手机应用的情绪监测系统, 并不断推进情绪监测识别任务从实验室与医院转向日常生活领域<sup>[15]</sup>. 德国奥格斯堡大学的 Wagner 等人也发布了一个单个被测对象 25 天的 4 类多模情绪识别数据库<sup>[16]</sup>. Soleymani 等人<sup>[8]</sup> 和 Koelstra 等人<sup>[17]</sup> 引入脑电、心电、皮肤温度、表情、呼吸等多模情绪相关信号, 构建了多模情绪数据库 DEAP, 该数据库是目前广泛研究的多模情绪数据库, 依据已有的人类情感认知的医学结论, 相关研究组进一步从 DEAP 数据库中提取脑电(功率谱特征、能量分布特征)、脉搏(心率变异性特征、频带内峰值特征)、心率(时域统计特征)、血压(时域统计特征)、呼吸(时域统计特征、频带内峰值特征)等特征, 并面向情绪维度模型, 实现了 LVLA、HVLA、LVHA、HVHA 四类维度情绪标签的识别. Choi 等人<sup>[18]</sup> 提出了多模信息融合恐惧监测方法. 随着深度学习理论的发展, Jirayuchareonsak 等人<sup>[19]</sup> 提出了面向 DEAP 数据库的深度神经网络情绪识别方法. Khosrowabadi 等人<sup>[20]</sup> 提出了基于 Emocog 的新型情绪识别神经网络. 此外, 一些最新的分类器(如信誉优化分类器、模糊粗糙近邻分类器等)也被广泛应用于情绪识别<sup>[21-22]</sup>. 总体而言, 当前的研究主要集中在实验室与医院场景、采用繁复的设备进行测量, 无法满足日常情绪监测的需求. 同时对多模特征提取融合能力有限, 多采用 NB、KNN、SVM 等传统分类器<sup>[4]</sup>, 分类识别率仍需提高.

本文针对日常情绪健康状况的监测识别, 搭建可穿戴传感网络测量系统测量被测个体的脑电、脉搏、血压信号, 减少测量硬件对被测个体的影响, 实现在被测个体完成日常生活任务的过程中监测识别其情绪. 同时, 采用栈式自编码器深度神经网络分类算法 (Stacked Auto-Encoder, SAE) 实现对多模情绪信息的融合分析, 并最终优化识别率. 本文提出的面向可穿戴多模生物信息传感网络的栈式自编码器优化情绪监测识别方法, 可为日常情绪健康监测与评估提供可靠方法, 为精神心理疾病的初筛提供科学依据.

本文第 2 节介绍可穿戴生物信息传感网络的节点开发过程与网络结构设计方案, 并分析该传感网络的结构特点与日常情绪监测识别的对应关系; 第 3 节提出适用于可穿戴传感网络测量环境与数据集特点的多模融合栈式自编码器情绪识别优化分类方法; 第 4 节基于上述测量系统与多模情绪数据处理方法进行实验, 并对实验结果进行分析与对比; 第 5

节对全文内容进行总结并展望未来工作。

## 2 可穿戴生物信息传感网络情绪数据监测

随着以先进测量传感技术、计算机技术为核心的传感网络技术(如物联网等)的飞速发展,传感网络远程分布式测量技术在健康监护领域的应用日益广泛,Mobile Health Sensing/Telemedicine 等概念及其应用成果不断涌现<sup>[23]</sup>. 医疗健康监护领域逐渐形成包括医院信息系统、公共卫生信息系统、远程医疗、家庭护理和区域协同医疗等新型医疗健康体系,得到了学术界和业界的广泛重视<sup>[24]</sup>.

本文设计的可穿戴多模生物信息传感网络是无线传感网络在生物信息测量感知领域的应用实例. 该类型的传感网络采用可穿戴、小型化、便携式的传感节点与身体主站,基于分布式远程网络测量技术,实现多种模态情绪信号的测量感知. 本文提出的可穿戴多模生物信息传感网络拓展了情绪监测识别应用领域,使其不再局限于实验室或医院场景.

### 2.1 可穿戴传感网络测量节点与身体主站

脑电测量一般依据 10~20 系统布置多电极,并采用涂抹导电膏的方式降低接触阻抗. 传统的心电测量也需要在左右臂和腿上布置 RA、LA、RL 等多个电极夹. 本文搭建的可穿戴生物信息传感网络简化了脑电、脉搏、血压三类情绪相关信号的测量过程,采用稀疏测量架构<sup>[25]</sup>,为情绪健康监测提供硬件基础.

如图 2 所示,采用 TGAM 芯片组与干电极构成可穿戴脑电测量节点,对 Fp1 前额位置的脑电进行测量,该位置的脑电被证明与情绪的纵轴(Arousal)高度相关. 脑电原始波形在 TGAM 芯片组中实现放大和整流,经由 1 Hz~100 Hz 的硬件滤

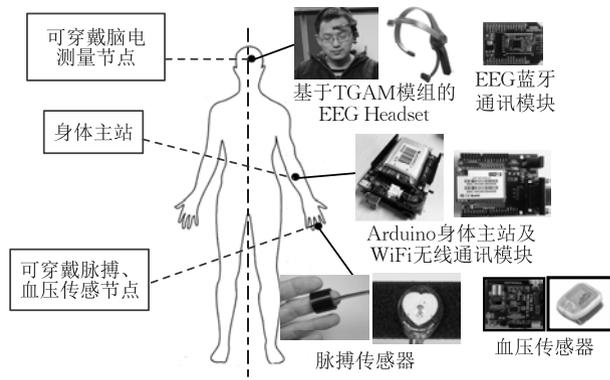


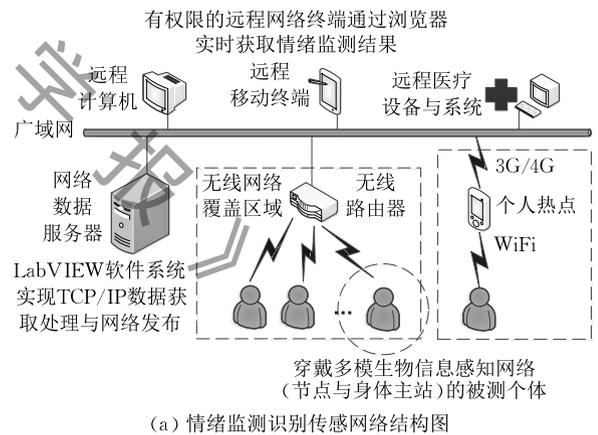
图 2 可穿戴传感网络测量节点与 Arduino 身体主站

波电路,由 AD 芯片转化为数字信号,并采用蓝牙接口与 Arduino 身体主站实现通讯. 可穿戴脉搏测量传感节点通过尼龙扣固定在中指指腹部位,采用光电容积法测量脉搏信号. 血压测量节点基于 Han-back HBE-Zigbex II Platform 开发.

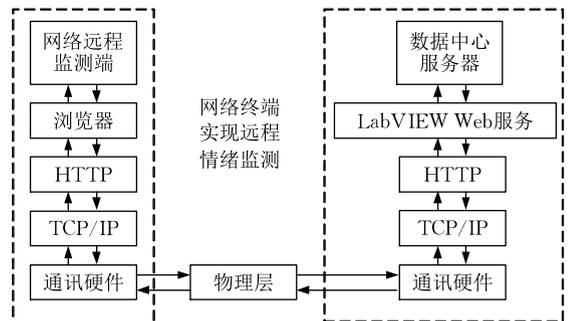
Arduino 是一款开源嵌入式平台,具有便捷灵活、易于搭建测量与数据处理系统的特点,主要组成部分包含 Arduino 开发板、扩展板和软件(Arduino IDE). Arduino 身体主站基于 Arduino UNO 3 及电池、无线通讯模块叠插搭建,固定在上臂上. 各可穿戴传感节点的数据经由各个分立的通讯以及数据预处理模块传输至身体主站,然后身体主站通过 WiFi 模块,作为远程网络的一个客户端,将情绪相关的信号数据传输到远程网络数据中心完成情绪识别评估过程.

### 2.2 情绪监测识别可穿戴传感网的网络结构设计

穿戴测量节点以及身体主站的被测个体一般处于不同的测量环境中. 对于无线 WiFi 信号覆盖的区域,可直接选择无线路由器作为远程网络的接入点. 对于无线信号并不覆盖的区域,可选择建立手机热点,通过 3G/4G 信号接入远程网络,如图 3(a)所示.



(a) 情绪监测识别传感网络结构图



(b) 网络发布服务实现过程

图 3 情绪监测识别传感网络设计

图 3(a)所示网络结构中,多模情绪信号数据经

由远程网络传输至具有固定公网 IP 的网络数据服务器(Web Data Server). 该网络服务器基于 NI-PXI 1065 硬件平台, 采用 LabVIEW 与 MATLAB 混编实现情绪数据处理和状态识别算法. 假设该服务器具有足够的计算资源, 因而本文暂不讨论识别算法的实时性. 网络数据服务器通过网络发布(Web Publish)功能将情绪识别的结果发布到公网上. 基于软件系统的网络发布服务, 如图 3(b)所示, 远程网络终端可以通过浏览器软件实时观察数据中心服务器的数据信息, 从而实现远程情绪监测. 有权限的网络控制终端, 如远程计算机、手机等即可访问该服务器提供的数据处理结果, 实现对被测个体情绪的网络化实时监测识别.

基于以上过程, 面向可穿戴多模生物信息传感网络情绪监测识别的传感测量系统已经构建完毕. 该可穿戴多模生物信息感知网络在测量端具有可穿戴、无线网络接入等特点, 符合日常场景下情绪识别的应用背景; 同时该网络能够在数据处理端实现情绪识别结果的远程网络共享与监控, 为情绪数据的积累与情绪心理健康的长期监测提供了解决方案.

下文将根据情绪监测识别的问题以及该测量系统的特点提出多模融合栈式自编码器优化处理算法, 进一步提高情绪监测识别的识别率.

### 3 多模融合栈式自编码器优化情绪识别

情绪识别本身是一个复杂度、抽象度较高的问题, 情绪状态的描述、情绪特征的提取都存在不确定性. 深度学习(Deep Learning), 可以实现对复杂模型的特征提取和学习分类<sup>[19]</sup>, 优化情绪数据的挖掘和情绪状态的识别效果. 栈式自编码器深度神经网络优化情绪识别算法即是采用深度学习栈式自编码器对提取的特征进行预学习, 并通过分类器进一步修正预学习过程, 最终达到优化识别率的目标, 算法具体流程如下.

#### 3.1 数据预处理与时间窗口平移

可穿戴传感网络获取的多模情绪信号包含了较多的噪声信号. 本文主要采用情绪识别常用的多模信号滤波与时间窗口平移方法对脑电、脉搏、血压信号进行预处理<sup>[19]</sup>(去噪并不是本文的主要工作, 优化的去噪算法将作为未来重要的研究内容). 1 Hz~40 Hz 带通滤波器用于 Fp1 脑电信号滤波, 50 Hz 陷波器用于其他信号滤波. 采用时间窗口平移的方法截

取滤波后的信号, 将其作为后续特征提取与训练分类的样本数据段. 以上过程可表示为

$$S_i^n = f_T(n) \sum_{k=1}^n F_i^k R_i^{n-k} \quad (1)$$

$$f_T(n) = \begin{cases} 1, & n_0 \leq n \leq N + n_0 \\ 0, & n < n_0 \text{ 或 } n > N + n_0 \end{cases} \quad (2)$$

其中,  $R_i^n$  表征第  $i$  个模态的第  $n$  个原始数据波形,  $F_i^n$  表征上述滤波器,  $f_T(n)$  是平移的时间窗口,  $n_0$  是时间窗口起点,  $N$  是时间窗口跨度. 经过预处理和时间窗口平移截取后, 得到多模生物信号数据段  $S_i^n$  作为特征提取的输入. 时间窗口的设置过程如图 4 所示. 将情绪任务开始时间点设置为首个时间窗口的起点, 通过实验设置固定窗口长度, 随着时间窗口的向后平移, 实质上多模数据段被等分为若干小段, 最末端的时间窗口与情绪任务结束的时间点保持一致.

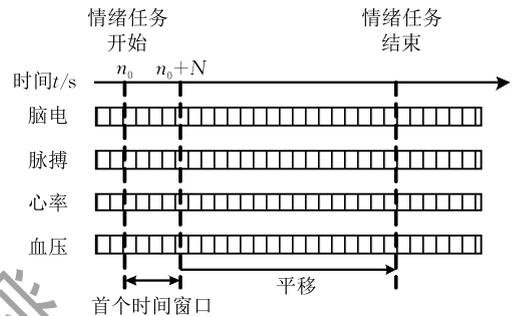


图 4 时间窗口数据预处理

由于情绪识别任务无需在毫秒级或较短的时间跨度内给出结果, 时间窗口预处理本质是一种时间尺度上的平均, 该方法针对特定时间跨度内的多模生物信号进行特征提取与识别, 可减小瞬态信号变化造成的特征变化过快, 减小多模生物信息传感网络数据同步对情绪监测识别的影响.

#### 3.2 多模情绪特征提取与类标签确定

依据已有研究成果, 特别是广泛研究的 DEAP 多模情绪数据库涉及的特征提取列表, 本文从多模生物信号数据段提取特征如下<sup>[8,13,19]</sup>:

脑电信号—— $\alpha$  波 (8 Hz~13 Hz)、 $\beta$  波 (13 Hz~31 Hz)、 $\gamma$  波 (31 Hz~100 Hz)、 $\theta$  波 (4 Hz~8 Hz) 和  $\delta$  波 (<4 Hz) 的功率谱分布、能量百分比; 功率谱峰值点对应的频率值; 基于 TGAM 脑电模组及其配套 eSense 脑波测量技术对被测对象专注程度和冥想深度的测量平均值和标准差; 脑电波的时域标准差、均值分析. 脉搏信号——心率时域均值和标准差; 心跳间隔时间; HRV 心率变异性分析, 包括全部窦性心搏 RR 间期 (简称 NN 间期) 的标准差、RR 间期平

均值标准差、相邻 RR 间期差值的均方根、血压——时域均值和方差。依据已有结论, 以上列出的部分特征可作为单一指标, 与情绪的特定维度关系密切(例如专注度指标与情绪纵轴 Arousal 相关等); 部分特征作为单一指标与情绪类别并没有明确关联, 但可以通过提取联合特征(例如  $\alpha$  波与  $\beta$  波的比值也可作为 Arousal 值的识别特征); 其余特征间的关联性并不明确, 需要采用下文涉及的特征预学习方法挖掘特征间关联性, 从而为分类器实现情绪多类分类任务提供依据。以上提取的特征在后文实验中将做示范性分析。

提取的特征经过尺度变换归一化, 得到数据样本的特征向量  $V = [v_1, v_2, \dots, v_n]$ 。被测个体依据 2D-SAM 量表, 完成 1~9 的 Valence 和 Arousal 分值评定, 按照维度模型同时针对 Valence 和 Arousal 坐标轴设立 3 和 7 两个门限值, 分为高中低三挡 (HV、NV、LV 以及 HA、NA、LA), 两两组合将 V-A 平面分为 9 个部分, 对应特征向量所属的类标签  $Y$ 。这种依据门限值进行情绪标签划分的方法在已有文献(例如文献[8, 13])中多次使用, 其目的是不被既定的情绪类型影响, 仅通过维度模型来描述情绪。本文划分情绪标签的粒度相较已有文献(如文献[13]中的四类)更细致, 因而能够在维度模型上实现更具体的情绪识别。

### 3.3 栈式自编码器深度神经网络分类

获取了特征向量与类标签以后, 情绪识别问题抽象为一个分类问题, 目的是将一个未知标签的样本分到 9 个设定的情绪类别中去。

在上文的国内外研究现状叙述中, 传统分类器如朴素贝叶斯、K 最近邻以及支持向量机(SVM)是最常用的三类, 也将作为后文实验中与栈式自编码器优化分类结果的对比。这里简单介绍这三种分类算法。朴素贝叶斯(NB)分类发源于古典数学理论, 依据贝叶斯定理, 对于给出的待分类情绪样本, 求解在已知特征向量出现的条件下各个情绪标签出现的概率, 并将该样本分入概率最大的情绪类别中。K 最近邻(KNN)分类, 主要思想是寻找与待分类样本特征向量在特征空间内距离(通常简化为欧式距离)的  $K$  个训练样本, 并将带分类样本分入这  $K$  个训练样本归属最多的类中。支持向量机(SVM)分类, 是采用核函数将原特征向量映射到新的特征空间, 以便通过优化迭代的方法寻找到使支持向量距离最远的超平面, 实现情绪样本分类。

以上三类方法在各个领域均有非常丰富的应

用。但是, 这三类分类方法都是典型的监督学习, 即需要大量的有标签训练样本。在日常情绪监测识别中, 部分样本包含类标签, 而其他大量样本并无类标签。

随着深度学习理论的提出, 栈式自编码器神经网络分类算法正适用于此类情况, 可对无标签样本特征进行预学习, 再通过有标签样本进行深度学习修正。同时, 其他数据库中无标签样本也可应用于特征集的预学习。因而, 面向日常情绪监测这一应用, 本文采用此方法优化分类过程, 并与可穿戴多模生物信息感知网络配合, 形成一套完整的日常情绪监测识别解决方案。

栈式自编码器结构如图 5 所示, 本文采用的是一个双隐层稀疏自编码器组成的神经网络。

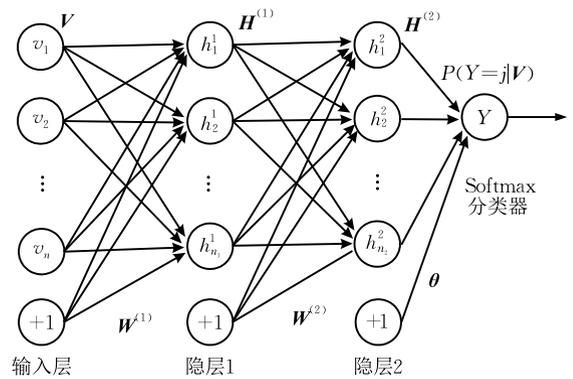


图5 栈式自编码器与 Softmax 分类器结构设计

组成栈式自编码器的基本结构是稀疏自编码器, 结构如图 6 所示, 其训练方法如下。假设稀疏自编码器输入为  $X$ (将偏置也写入输入向量), 权值矩阵为  $W$ , 选择 sigmoid 函数  $f(x) = 1/(1 + e^{-x})$  作为激活函数。自编码器编码、解码可以表示为

$$H = f(WX), \hat{X} = W^T H \quad (3)$$

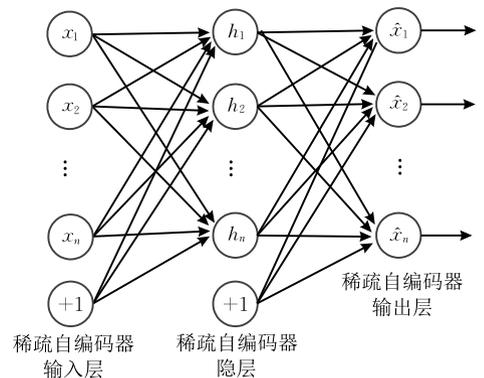


图6 稀疏自编码器结构与训练过程示意

作为一种无监督学习算法, 稀疏自编码器采用反向传播机制使目标输出值趋近输入值。自编码器尝试学习一个恒等函数, 即使

$$\hat{\mathbf{X}} = h_w(\mathbf{X}) \approx \mathbf{X} \quad (4)$$

加入稀疏限制以后,自编码器本质是在学习输入训练样本的稀疏表示,并从中挖掘到用于分类的特征结构.当神经元的输出接近于 1,即认为该神经元被激活;当神经元的输出接近于 0,即认为该神经元被抑制;使得神经元大部分时间都被抑制的限制条件称为稀疏性限制.

训练稀疏自编码器的代价  $J^{(A)}$  函数分为两项:第 1 项为输出与输入差值的二范数,用于约束输出结果使其逼近输入;第 2 项为隐藏层神经元激活输出的一范数,作为稀疏自编码器训练的稀疏性限制.  $J^{(A)}$  函数如下:

$$J^{(A)} = \|\mathbf{h}_w(\mathbf{X}) - \mathbf{X}\|^2 + \lambda \sum_j |h_j| \quad (5)$$

依据以上流程对稀疏自编码器的权值矩阵进行预训练.

以提取的特征  $\mathbf{V}$  作为隐层 1 的输入,训练隐层 1.接着,将预训练完成的隐层 1 稀疏自编码器的输出作为隐层 2 自编码器的输入.采用同样的方法,完成对隐层 2 稀疏自编码器的预训练.完成对特征进行无监督自学习以后,再使用有标签数据与 Softmax 分类器实现分类,并对各层的权值进行修正(Fine Tunning).

Softmax 分类器是 Logistic 回归模型在多元分类问题上的推广.情绪识别涉及到的正是多元分类问题,类标签个数  $k=9$ .

依据预训练完成后的栈式自编码器隐层 2 的输出  $\mathbf{H}^{(2)}$ ,Softmax 分类器能够计算出对应类标签的概率  $P(Y=j|\mathbf{H}^{(2)})$ ,对于整个分类器来说,即为求取  $P(Y=j|\mathbf{V})$ ,其输出形式为  $k$  维向量,分别表征着样本属于四类情绪标签的概率.具体形式如下:

$$h_{\theta}(\mathbf{V}^{(i)}) = \begin{bmatrix} P(Y=1|\mathbf{V}^{(i)}) \\ P(Y=2|\mathbf{V}^{(i)}) \\ \vdots \\ P(Y=k|\mathbf{V}^{(i)}) \end{bmatrix} = \frac{1}{\sum_{l=1}^k e^{\theta_l^T \mathbf{H}^{(2)(i)}}} \begin{bmatrix} e^{\theta_1^T \mathbf{H}^{(2)(i)}} \\ e^{\theta_2^T \mathbf{H}^{(2)(i)}} \\ \vdots \\ e^{\theta_k^T \mathbf{H}^{(2)(i)}} \end{bmatrix} \quad (6)$$

模型参数  $\theta = [\theta_1, \theta_2, \dots, \theta_k]^T = \{\theta_{ij}\}_{k \times (n_{\theta}+1)}$ ,记为一个尺寸  $k \times (n_{\theta}+1)$  的矩阵.概率分布在上式中已经进行归一化,所有概率之和为 1.假设有  $n_T$  个训练样本,Softmax 分类通过最小化代价函数  $J^{(S)}$  实现:

$$J^{(S)} = -\frac{1}{n_T} \sum_{i=1}^{n_T} \sum_{j=1}^k 1\{Y^{(i)}=j\} \log \frac{e^{\theta_j^T \mathbf{H}^{(2)(i)}}}{\sum_{l=1}^k e^{\theta_l^T \mathbf{H}^{(2)(i)}}} \quad (7)$$

其中,  $1\{eq\}$  是示性函数,当表达式  $eq$  为真时,函数值为 1;当表达式  $eq$  为假时,函数值为 0.采用梯度下降法求解优化问题.在实际运算中,需要添加一个权重衰减项使代价函数转化为严格的凸函数:

$$J^{(S)} = -\frac{1}{n_T} \sum_{i=1}^{n_T} \sum_{j=1}^k 1\{Y^{(i)}=j\} \log \frac{e^{\theta_j^T \mathbf{H}^{(2)(i)}}}{\sum_{l=1}^k e^{\theta_l^T \mathbf{H}^{(2)(i)}}} + \lambda' \sum_{i=1}^k \sum_{j=0}^{n_{\theta}} \theta_{ij}^2 \quad (8)$$

在 Softmax 分类器中,将  $\mathbf{V}^{(i)}$  待分类样本分类为  $j$  的概率为

$$P(Y^{(i)}=j|\mathbf{V}^{(i)}) = \frac{e^{\theta_j^T \mathbf{H}^{(2)(i)}}}{\sum_{l=1}^k e^{\theta_l^T \mathbf{H}^{(2)(i)}}} \quad (9)$$

微调可以大幅提升栈式自编码神经网络的性能表现.本质上,微调是将整个网络进行整体迭代优化.这里采用反向传播算法,从 Softmax 输出层到隐层 1 逐层对代价函数求偏导,计算调整向量,并迭代实现权重矩阵的微调.具体过程如下:

$$\delta^{(S)} = -\nabla J^{(S)} f'(\theta^T \mathbf{H}^{(2)}) \quad (10)$$

$$\nabla J^S = \theta^T (\mathbf{I} - h_{\theta}(\mathbf{V})) \quad (11)$$

其中  $h_{\theta}(\mathbf{V})$  为条件概率向量,  $\mathbf{I}$  为输入数据对应的类别标签.逐层向前计算各隐层的微量量:

$$\delta^{(2)} = (\mathbf{W}^{(2)})^T \delta^{(S)} f'(\mathbf{W}^{(2)} \mathbf{H}^{(1)}) \quad (12)$$

$$\delta^{(1)} = (\mathbf{W}^{(1)})^T \delta^{(2)} f'(\mathbf{W}^{(1)} \mathbf{V}^{(1)}) \quad (13)$$

综上所述,提取出的特征向量输入训练完成后的双隐层栈式神经网络,输出即为类标签(选择概率最大的类标签),从而实现情绪状态的分类识别.选择此方法优化情绪识别分类的原因如下:

(1) 情绪识别数据包含大量无标签数据,采用面向无标签数据的自编码器特征预学习可以充分使用实验数据;

(2) 情绪特征的提取存在不确定性,目前并无标准的特征选取范围与流程,栈式自编码器可以自行挖掘离散特征间的关联特征,提高情绪识别特征提取的效果;

(3) Softxmxax 分类器适用于各类别互斥的多元分类场景,本文构建的基于情绪维度模型的多类分类任务是具有互斥性的多元分类任务.

基于可穿戴多模传感网络的栈式自编码器优化情绪识别实验结果也将印证以上结论.

### 3.4 多模传感网络情绪识别结果评价

可穿戴传感网络提取的情绪特征质量、栈式自编码器对特征的学习效果以及情绪状态的识别率是三项需要评价和比对的指标。依据 US. Berkeley 在文献[26]中提出的面向可穿戴生物信号分类特征的评价方法, 定义任意两个样本特征向量  $V_m$  和  $V_n$  间的相似度:

$$\text{Similarity}(V_m, V_n) = \frac{V_m \cdot V_n}{\|V_m\| \cdot \|V_n\|} \quad (14)$$

假设多数特征向量是可靠的。某类标签下, 所有特征向量之间的相似度均值称为该类标签的自相似度 (Self-similarity), 所有特征向量与他类标签下特征向量的相似度均值称为该类标签的互相似度 (Cross-similarity)。自相似度与互相似度的差值称为相似度差值 (Diff-similarity)。

自相似度表征了该类样本特征的集中性 (Centrality)。同类标签样本应该具有更高的相似度, 因而自相似度越高, 则意味着同类标签样本的集中性、一致性越好。相似度差值表征了该类样本特征的可区分性 (Distinguishability)。同类标签样本之间的相似程度越高, 异类标签样本之间的相似程度越低, 两者的差值越大, 则样本的可分性就越好。可穿戴多模生物信息感知网络提取的特征以及栈式自编码器对特征的改进效果即用该类指标评估<sup>[26]</sup>。

最终某类情绪状态的识别率使用正样本检出率 (Positive Predictive Value) 来评估:

$$\text{RecognitionRate} = TP / (TP + FP) \quad (15)$$

其中,  $TP$  表示被分类为正样本, 实际也为正样本的样本个数;  $FP$  表示被分类为正样本, 实际是负样本的样本个数。该指标表征了当被测个体处于某种情绪状态时, 被正确识别出的概率。

## 4 实验结果与分析

### 4.1 实验设置

实验被测对象共 8 名, 均为男性, 右利手, 在 5 个月的跨度中选择了 71 天, 记录 8 个被测个体的脑电、脉搏、血压信息用于情绪识别任务。记录过程中, 被测个体穿戴上文所述的多模传感网络, 同时完成日常生活任务 (如阅读图书、观看影像、玩手机游戏等), 并依据量表对所处情绪状态进行自评, 确定类标签。无确切类标签或无法评价的样本, 可用于栈式自编码器的特征学习。

数据段共 570 段, 进一步使用时间窗口平移分成 10 段, 获得 5700 段数据样本。其中, 脑电采样率 500 Hz, 测量误差  $0.5 \mu V$ , 十米内蓝牙串口通讯,

12 位 AD 转换, 由于每秒产生 6000 bit 数据量, 直接使用 Arduino IDE 调用串口读取语句读写速度不够, 不能满足采样率的需要, 因而需要在 Arduino UNO 3 编程中设立缓存读取机制, 具体方法是采用在程序体通讯语句之前添加如下语句: `while (Serial.available() < Buff), Buff` 为预先设立的数据长度, 这样当 `Buff` 达到设立的长度以后再进行数据操作, 可以数据读写速度过快的造成的问题。

脉搏传感器测量脉搏频域范围为  $0.05 \sim 200$  Hz, 放大倍数为 331, 脉搏传感器原始波形为模拟输出信号, 接入身体主站的 AI 口, 频段符合 AI 口的输入范围, 经由主站的解析程序每 20 ms 产生 32 位脉搏原始波形数据, 数据格式: 字符串“标识符 S + 幅值数据”, 如 S600, 则表示脉搏波形数据输出值为 600 mV。心率 (Heartbeats Per Minute) 的数据格式为: 字符串“标识符 B + 心率数据”, 如 B60, 则表示当前心率为 60 次/min。心跳间隔时间 (Time Between Heartbeats) 的数据格式为: 字符串“标识符 I + 心跳间隔数据”, 如 I1000, 则表示当前心跳间隔时间为 1000 ms。心率与心跳间隔时间从原始信号中计算得到, 每产生一个 R 波由传感器给出, 采样率不固定。脉搏传感器接入身体主站同样需要设立缓存读取机制。所以, 脉搏信号输出数据量约为每秒 1600 bit 的数据量。

血压传感器基于 Hanback HBE-Zigbex II Platform 开发, 人工启停血压测量过程, 产生结果后更新输出数据, 如果没有开始新的测量过程, 则一直输出原测量结果至身体主站。由于血压不存在瞬态变化测量, 因而这里人为设置每 2 秒输出一次结果, 每次输出数据量为两个 8 位整数, 平均 8 bit/秒。各个传感器与主站的通信方式并不相同, 因而信道是不同的, 采取 Arduino UNO 3 及其通讯、供电板块叠插的方式搭建主站, 并实现通讯。WiFi 模块将主站设置为网络环境中的客户端 Client, 向远程监控端 (固定 IP 与端口) 输出数据。远程数据服务器为 NI-PXI 1065。远程监控端为 iPhone 5C。

### 4.2 多模生物信息数据样本分析

数据样本的类标签分布如图 7 所示。V-A 平面被分割成 9 类不同的情绪标签 LVHA、NVHA、HVHA、LVNA、NVNA、HVNA、LVLA、NVLA、HVLA。由于本文是面向日常状况的情绪监测识别, 各类情绪状态出现的概率并不相同, 因而类标签分布不均匀。除去标出的带标签样本数据, 还有无标签样本数据, 这些数据可用于特征预学习过程, 提高多模生物信号情绪特征的一致性与可分性。

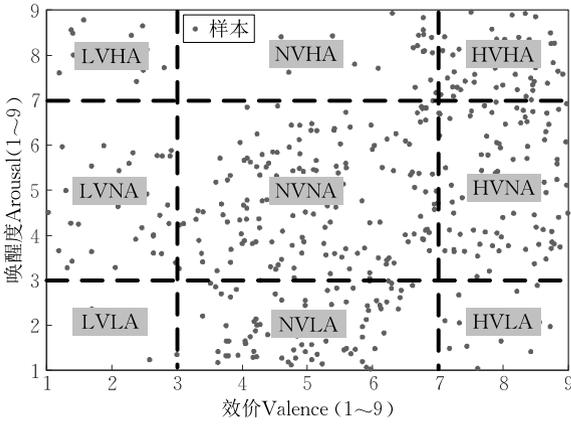


图 7 样本类标签在情绪 V-A 平面上的分布

数据样本原始波形如图 8 所示,生物信号原始波形长度共 240 s,包含脑电、脉搏、心率、血压.图 8 中对应的被测个体正在观看包含恐怖情节的视频片段,部分信号时域上能够体现变化,脑电信号等则需要进一步提取特征.采用上文所述方法对数据进行预处理,经试验优化选取时间窗口长度为 24 s,从 0~24 s 为第一段、24 s~48 s 为第二段,以此类推,该数据段被分成 10 段分别提取特征.从原始数据波形可以看出,无线稀疏可穿戴测量结构导致测量数据均存在一定程度的不稳定,例如血压信号出现的短时间漂移,选择 24 s 的时间跨度进行平均化处理,可以满足上文列出的情绪特征提取要求(相关特征瞬态特性不强),同时可以减少测量段突发状况对测量结果的影响.

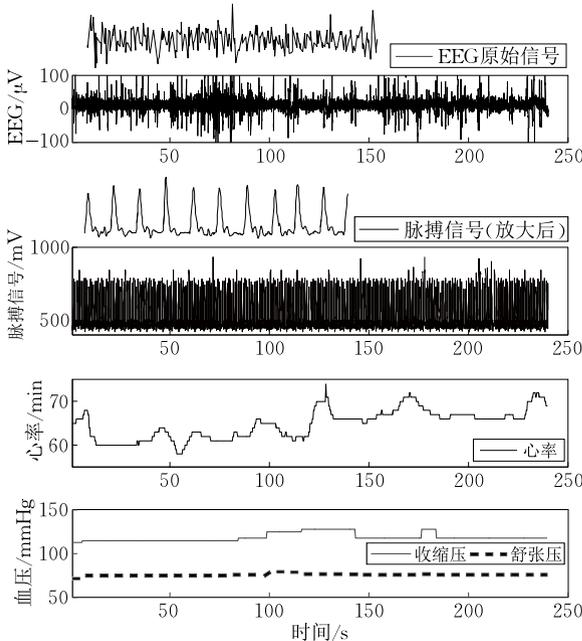
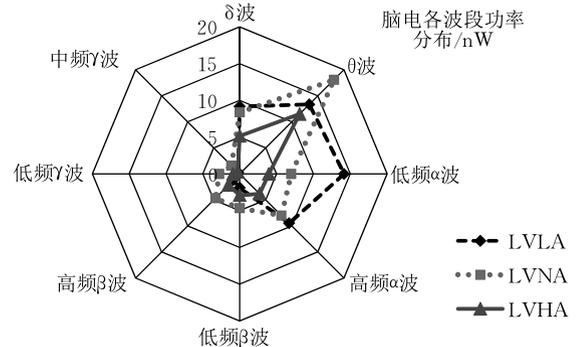


图 8 多模情绪相关生物信号原始波形图

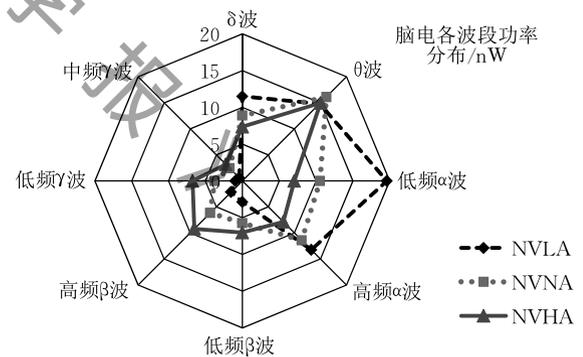
提取特征进一步细化为脑电低频  $\alpha$  波(8 Hz~10 Hz)、高频  $\alpha$  波(10 Hz~13 Hz)、低频  $\beta$  波(13 Hz~

20 Hz)、高频  $\beta$  波(20 Hz~31 Hz)、低频  $\gamma$  波(31 Hz~50 Hz)、中频  $\gamma$  波(50 Hz~100 Hz)、 $\theta$  波(4 Hz~8 Hz)、 $\delta$  波( $<4$  Hz)的功率谱分布、功率谱峰值点对应的频率值、基于 TGAM 脑电模组及其配套 eSense 脑波测量技术对被测对象专注程度和冥想深度的测量平均值和标准差、以及时域标准差、均值;心率时域均值和标准差;心跳间隔时间;HRV 心率变异性分析,包括全部窦性心搏 RR 间期的标准差、RR 间期平均值标准差、相邻 RR 间期差值的均方根;血压时域均值和方差.以上特征依据已有医学结论或文献给出,这里结合具体数据段,做具体分析说明.

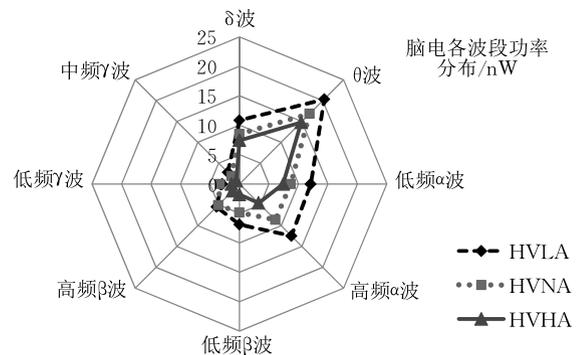
各情绪状态下脑电信号各波段的平均功率分布特征如图 9 所示.



(a) 脑电功率分布特征(LVLA、LVNA、LVHA)



(b) 脑电功率分布特征(NVLA、NVNA、NVHA)



(c) 脑电功率分布特征(HVLA、HVNA、HVHA)

图 9 多模情绪相关生物信号原始波形图

图 9 采用雷达图的形式,较直观地表达出不同情绪状态下脑电各波段平均功率的分布状况.例如目前较明确的、已经被医学验证的单一特征是  $\alpha$  波的功率与放松状态有关.从图 9(a)、(b)、(c)中均可以看出,包含“LA”(低唤醒度,个体更为放松)的情绪状态,明显具有较高的  $\alpha$  波功率,且  $\alpha$  波功率随着唤醒度提高(放松程度下降)而降低.包含“LA”“NA”与“HA”的情绪状态,其对应样本的  $\alpha$  波平均功率依次为 27.03nW、18.10nW 与 11.57nW.低唤醒度情绪比高唤醒度情绪的  $\alpha$  波平均功率高出 134.3%.

除去单一指标,各波段功率分布的特征将作为分类器的输入,借由栈式自编码器对其联合特征进行预学习与分类优化.不同情绪状态下脑电专注指数与冥想指数(由 eSense TGAM 模块提供,范围 0~100)的平均值如表 1、表 2 所示.

表 1 不同情绪状态下脑电专注指数特征

情绪标签	专注指数		
	LV	NV	HV
HA	56.30	56.94	57.80
NA	46.89	49.31	48.60
LA	44.08	41.84	39.59

表 2 不同情绪状态下脑电冥想指数特征

情绪标签	冥想指数		
	LV	NV	HV
HA	43.18	45.74	46.69
NA	46.89	50.97	51.14
LA	48.18	55.13	62.18

随着唤醒度的提升(从 LA 变化到 HA),专注指数有较为明显的上升.高唤醒度情绪比低唤醒度情绪的平均专注指数高出 45.9%.而冥想指数则在第四象限,对应标签(HVLA)取得最高值 62.18,与冥想状态所在 V-A 平面的位置一致.冥想指数随着距离 V-A 第四象限位置的距离增大二逐渐降低.其他模态信号提取的特征基本是较为常见的时域指标或统计指标,这里不做详细论述.这些特征都将作为分类器的输入,由栈式自编码器神经网络实现分类的优化,优化后的特征与分类结果将在下文进行具体分析与讨论.

#### 4.3 栈式自编码器情绪优化特征提取与识别

将原始(Raw)特征向量与训练完成的栈式自编码器预学习(Pre-learned)后的特征向量分别进行相似度评价.

表 3、表 4 为自相似度和相似度差值评价结果,表格中斜杠左侧为栈式自编码器预学习优化后的特征相似度评价,右侧为原始特征相似度评价.相似度差值均为正数,即表示相同情绪标签下的样本相似度要高于不同情绪标签下的样本,表明可穿戴传感网络获取的多模情绪相关信息具有可用性.具体而言,包含“NV”“NA”标签的样本自相似度值较低,表明越中性(Neutral)的情绪,其样本集中性越差.对于“NVNA”标签的情绪样本,其自相似度值取得最小值 0.859/0.846(预学习样本/原始样本).栈式自编码器优化后,样本的集中性明显增强,自相似度平均值提高了 1.9%.相似度差值评价中,同样中性的情绪样本可分性较差,栈式自编码器能较为明显地提高样本的可分性,相似度差值提高了 18.4%.

表 3 情绪识别特征向量自相似度(集中性)评价

情绪标签	自相似度(预学习/原始)			
	LV	NV	HV	平均
HA	0.977/0.951	0.920/0.894	0.944/0.940	<b>0.947/0.928</b>
NA	0.924/0.920	0.859/0.846	0.925/0.914	<b>0.903/0.893</b>
LA	0.982/0.956	0.893/0.871	0.964/0.939	<b>0.946/0.922</b>
平均	<b>0.961/0.942</b>	<b>0.891/0.870</b>	<b>0.944/0.931</b>	<b>0.932/0.914</b>

表 4 情绪识别特征向量相似度差值(可分性)评价

情绪标签	相似度差值(预学习/原始)			
	LV	NV	HV	平均
HA	0.207/0.177	0.116/0.082	0.127/0.117	<b>0.150/0.125</b>
NA	0.118/0.084	0.074/0.050	0.118/0.098	<b>0.103/0.077</b>
LA	0.160/0.155	0.112/0.108	0.188/0.152	<b>0.153/0.139</b>
平均	<b>0.162/0.139</b>	<b>0.101/0.080</b>	<b>0.144/0.122</b>	<b>0.135/0.114</b>

特征向量自相似度与相似度差值在 V-A 平面的分布如图 10 所示.灰度值越高则样本一致性或可分性越好,由图可见,中性情绪灰度值较低,样本质量较差,强烈的情绪状态对应的灰度值较高,样本质量也较好.

样本相似度评价过程体现了可穿戴传感网络获取多模情绪相关生物信号样本的可用性,也论证了栈式自编码器通过对特征向量(包括无标签样本)的预学习与训练微调,优化了分类训练样本的集中性和可分性.

图 11 所示为相似度差值评价与情绪识别率的关系.本文提出的方法对应 9 类情绪标签的识别率依次(按照坐标轴顺序)为 62%、61%、73%、55%、52%、56%、66%、57%、64%.相似度差值与情绪识别率变化规律一致.从最终识别率的角度验证了相似度评价方法评价特征质量的有效性.

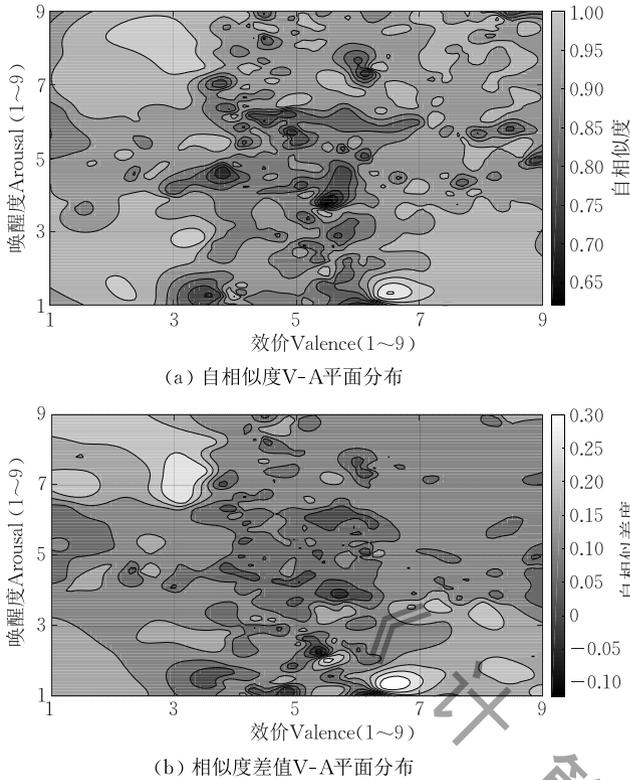


图 10 V-A 平面情绪特征相似度评价

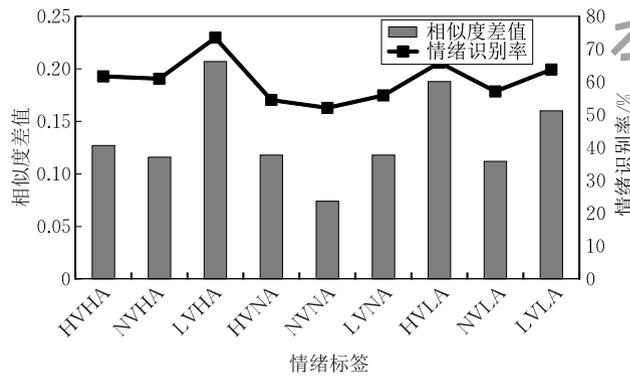


图 11 相似度差值评价与情绪识别率的关系

进一步比较传统方法朴素贝叶斯(Naive Bayes, NB)、K 最近邻(K-Nearest Neighbor, KNN,  $K=5$ , 使用欧拉距离)、支持向量机(Support Vector Machine, SVM, RBF 核函数)与本文提出的面向多模可穿戴传感网络的栈式自编码器(SAE, 双隐层栈式神经网络, 稀疏自编码器预训练方法, 隐层节点数 50/25)优化情绪分类的识别率, 结果如图 12、图 13 所示。

NB、KNN、SVM 以及 SAE 的平均识别率依次为 50.43%、54.35%、55.89%和 60.61%。对于不同的被测个体, 本文提出的面向多模传感网络的栈式自编码器优化情绪识别率达到了 53.77%~73.48%, 较传统方法提高了约 5%。该识别率在可穿戴传感

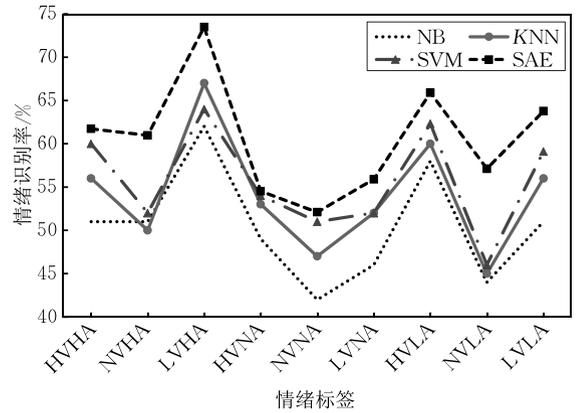


图 12 不同方法对不同情绪标签的识别率结果

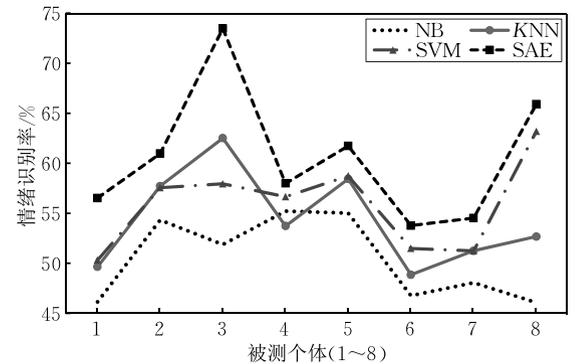


图 13 不同方法对不同被测个体的识别率结果

网络的测量条件下, 达到或超过了部分医学级情绪识别研究成果<sup>[1,8,10,13]</sup>的水平。

进一步讨论方法实现的能耗与时间代价。对于本文采用的结构, 数据存储、分类与计算均在服务器上实现, 因而传感器端的能耗仅为数据测量与通信的基本能耗, 锂电池模块供电可以满足基本使用, 可进一步引入数据压缩<sup>[27]</sup>、WiFi 自动启停机制<sup>[28]</sup>、全局优化策略<sup>[29]</sup>等各类无线传感网络能效优化方法以降低能耗、延长电池使用寿命。另外, 本文采用的栈式自编码器优化算法, 在权值矩阵训练过程中耗时较长, 权值矩阵训练完成后, 识别过程相较于 NB 与 KNN 中的概率与距离运算具有优势。由于本文假设是在计算资源充沛的服务器上运行算法, 重点是实现面向日常情绪监测识别的架构, 目前尚未在身体主站或智能手机终端直接实现分类, 因而算法的实时性将作为未来优化的重要方向。

以上实验结果表明, 可穿戴多模生物信息传感网络在简化的可用于日常情绪健康监测的测量架构下, 实现了对情绪相关信号的测量感知和有效特征提取。栈式自编码器优化的深度特征比原始特征具有更好的分类效果, 最终情绪识别率达到了 53.77%~73.48%, 较传统方法提升了 5%, 在可穿戴传感网

络测量条件下达到或超过了部分医学级情绪研究识别研究成果, 论证了该测量机制的可靠性。

## 5 结 论

本文针对日常情绪健康状况的监测识别, 构建了多模可穿戴生物信息传感网络测量被测个体的多模情绪相关信号(脑电、脉搏以及血压), 经由身体主站将信号传输至远程网络数据中心, 并将情绪识别的结果进行网络发布, 简化了测量结构, 实现了网络环境下的可穿戴多模情绪识别。

本文提出了双隐层栈式自编码器神经网络(基于深度学习理论)优化的情绪识别算法。两层稀疏自编码器对初始特征向量进行预学习, 再经由 Softmax 分类器对权值矩阵进行微调。整个 71 天长时间跨度的实验结果表明, 栈式自编码器预学习后的特征向量具有更高的集中性与可分性, 自相似度值提高了 1.9%, 相似度差值提高了 18.4%。优化分类算法提高了情绪识别率, 最终情绪识别率达到了 53.77%~73.48%, 较传统方法提升了 5%。未来的工作还需重点积累大量的实验数据, 以优化栈式自编码器对无标签特征预训练的效果。同时优化前端传感器的能效性与后端算法的实时性, 探讨在可穿戴智能终端实现情绪识别算法的可能性。

本文提出的面向可穿戴多模生物信息传感网络的栈式自编码器神经网络优化情绪识别方法与实现架构, 可为人类个体情绪健康监测和情绪相关精神心理疾病初筛提供依据。

## 参 考 文 献

- [1] Chanel G, Kronegg J, Grandjean D, et al. Emotion assessment: Arousal evaluation using EEG's and peripheral physiological signals//Proceedings of the International Workshop on Multimedia Content Representation, Classification and Security (MRCS 2006). Istanbul, Turkey, 2006: 530-537
- [2] Iacoviello D, Petracca A, Spezialetti M, et al. A real-time classification algorithm for EEG-based BCI driven by self-induced emotions. *Computer Methods and Programs in Biomedicine*, 2015, 122(3): 293-303
- [3] Tao J, Tan T. Affective computing: A review//Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII 2005). Beijing, China, 2005: 981-995
- [4] Jatupaiboon N, Pan-Ngum S, Iprasena P. Real-time EEG-based happiness detection system. *The Scientific World Journal*, 2013, 2013: 1-12
- [5] Khezri M, Firoozabadi M, Sharafat A R. Reliable emotion recognition system based on dynamic adaptive fusion of forehead biopotentials and physiological signals. *Computer Methods and Programs in Biomedicine*, 2015, 122(2): 149-164
- [6] Li Yue-Long, Jin Yan, Wang Jian-Ming, et al. Face feature points extraction: A review. *Chinese Journal of Computers*, 2016, 39(7): 1356-1374(in Chinese)  
(李月龙, 靳彦, 汪剑鸣, 等. 人脸特征点提取方法综述. *计算机学报*, 2016, 39(7): 1356-1374)
- [7] Jiang Teng-Jiao, Wan Chang-Xuan, Liu De-Xi, et al. Extracting target-opinion pairs based on semantic analysis. *Chinese Journal of Computers*, 2017, 40(3): 617-633(in Chinese)  
(江腾蛟, 万常选, 刘德喜等. 基于语义分析的评价对象-情感词对抽取. *计算机学报*, 2017, 40(3): 617-633)
- [8] Soleymani M, Lichtenauer J, Pun T, et al. A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, 2012, 3(1): 42-55
- [9] Chanel G, Kierkels J J M, Soleymani M, et al. Short-term emotion assessment in a recall paradigm. *International Journal of Human-Computer Studies*, 2009, 67(8): 607-627
- [10] Chanel G, Rebetz C, Bétrancourt M, et al. Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 2011, 41(6): 1052-1063
- [11] Lin Y, Wang C, Jung T, et al. EEG-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, 2010, 57(7): 1798-1806
- [12] Murugappan M, Nagarajan R, Yaacob S. Combining spatial filtering and wavelet transform for classifying human emotions using EEG signals. *Journal of Medical and Biological Engineering*, 2011, 31(1): 45-51
- [13] Picard R W, 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
- [14] Healey J A, Picard R W. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 2005, 6(2): 156-166
- [15] Picard R W. Automating the recognition of stress and emotion: From lab to real-world impact. *IEEE Multimedia*, 2016, 23(3): 3-7
- [16] Lisetti C L A T, Nasoz F. Using noninvasive wearable computers to recognize human emotions from physiological signals. *EURASIP Journal on Advances in Signal Processing*, 2004, 2004(11): 929414
- [17] Koelstra S, Mühl C, Soleymani M, et al. Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 2012, 3(1): 18-31

- [18] Choi J, Bang J, Heo H, et al. Evaluation of fear using noninvasive measurement of multimodal sensors. *Sensors*, 2015, 15(7): 17507-17533
- [19] Jirayucharoensak S, Pan-Ngum S, Israsena P. EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. *The Scientific World Journal*, 2014, 2014: 1-10
- [20] Khosrowabadi R, Quek C, Ang K K, et al. ERNN: A biologically inspired feedforward neural network to discriminate emotion from EEG signal. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, 25(3): 609-620
- [21] Dai Y, Wang X, Li X, et al. Reputation-driven multimodal emotion recognition in wearable biosensor network//Proceedings of the IEEE International Instrumentation and Measurement Technology Conference (I2MTC 2015). Pisa, Italy, 2015: 1747-1752
- [22] Dai Yi-Xiang, Wang Xue, Li Xuan-Ping, et al. Fuzzy rough emotion recognition based on sparse EEG sensing in biosensor network. *Chinese Journal of Scientific Instrument*, 2014, 35(8): 1693-1699(in Chinese)  
(戴逸翔, 王雪, 李宣平等. 面向生物信息感知网络稀疏脑电测量的模糊粗糙情绪识别. *仪器仪表学报*, 2014, 35(8): 1693-1699)
- [23] Sene A, Kamsu-Foguem B, Rumeau P. Telemedicine framework using case-based reasoning with evidences. *Computer Methods and Programs in Biomedicine*, 2015, 121(1): 21-35
- [24] Wu Xin-Dong, Ye Ming-Quan, Hu Dong-Hui, et al. Pervasive medical information management and services: Key techniques and challenges. *Chinese Journal of Computers*, 2012, 35(5): 827-845(in Chinese)  
(吴信东, 叶明全, 胡东辉等. 普适医疗信息管理与服务的关键技术与挑战. *计算机学报*, 2012, 35(5): 827-845)
- [25] Dai Y, Wang X, Li X, et al. Sparse EEG compressive sensing for web-enabled person identification. *Measurement*, 2015, 74: 11-20
- [26] Chuang J, Nguyen H, Wang C, et al. I think, therefore i am: Usability and security of authentication using brainwaves //Proceedings of the International Conference on Financial Cryptography and Data Security (FC 2013). Okinawa, Japan, 2013: 1-16
- [27] Wang Ju, Chen Xiao-Jiang, Chang Li-Qiong, et al. Compressive sensing based device-free moving target trajectory depiction. *Chinese Journal of Computers*, 2015, 38(12): 2361-2374(in Chinese)  
(王举, 陈晓江, 常俪琼等. 基于压缩感知的被动式移动目标轨迹测绘. *计算机学报*, 2015, 38(12): 2361-2374)
- [28] Peng Ying, Wang Gao-Cai, Huang Shu-Qiang, et al. An energy consumption optimization strategy for data transmission based on optimal stopping theory in mobile networks. *Chinese Journal of Computers*, 2016, 39(6): 1162-1175(in Chinese)  
(彭颖, 王高才, 黄书强等. 移动网络中基于最优停止理论的数据传输能耗优化策略. *计算机学报*, 2016, 39(6): 1162-1175)
- [29] Zhang Fa, Wang Lin, Hou Chen-Ying, et al. Network energy consumption models and energy efficient algorithms. *Chinese Journal of Computers*, 2012, 35(3): 603-615(in Chinese)  
(张法, 王林, 侯晨颖等. 网络能耗系统模型及能效算法. *计算机学报*, 2012, 35(3): 603-615)



**DAI Yi-Xiang**, born in 1990, Ph. D. candidate. His current research interests include biosensor network and EEG emotion recognition.

**WANG Xue**, born in 1963, Ph. D., professor, Ph. D. supervisor. His research interests include precision measurement and sensor technology and wireless sensor networks

## Background

With the enhancement of public health consciousness, emotional health draws great concern. Emotional health is closely related to the quality of personal life. Even for some special group of people, like pilots, soldiers, etc., their emotional states will have impacts on the stability of communities.

measurement.

**DAI Peng**, born in 1990, Ph. D candidate. His current research interests include visual sensor networks and computer vision.

**ZHANG Wei-Hang**, born in 1992, Ph. D. His current research interests include medical image processing and bio signal sensing.

**ZHANG Peng-Bo**, born in 1992, Ph. D. candidate. His research interests are physiological signal measurement and analysis.

Traditionally, to evaluate emotional states of human beings relies on the doctors or psychologists to communicate with subjects and give scores based on various questionnaires. This approach is not scientific enough and leads to the difficulties in the emotional health monitoring in daily-life. Emotion recognition enables lifeless sensors and computers to

measure and interpret human emotions. It is a procedure consisting of emotion-related bio signals recording, features extraction and emotional states classification, providing scientific evidence for emotional health monitoring and primary diagnosis of potential mental diseases. In the related works concerning emotion recognition, the application scenarios are usually restricted in the hospitals or labs and the common-used classifiers are not suitable for the daily emotion recognition data set.

This paper develops a multimodal biosensor network to simplify the sensing framework so that it can finish emotion recognition tasks when subjects are participating in daily tasks. Several wearable biosensor nodes record multimodal bio signals (EEG, pulse and blood pressure) and transmit them to a body station. The body station connects to a web data center on the Internet by wireless routers or personal hotspots. The recognition algorithm is carried out in the data center and the results are displayed for authorized web terminals. Additionally, due to the uncertainties in signal acquisition and feature extraction, a stacked auto-encoder (based on the deep learning theory) optimized emotion recognition method is proposed to improve the recognition process. The experiment convinces that the feature vectors

pre-learned by stacked auto-encoder are of higher quality while the final recognition rate is also improved approximately 5% compared to related works. The main contributions of this paper are the wearable network-based sensing structure, the stacked auto-encoder optimized multimodal emotion recognition algorithm and the quantitative analysis on 71-day experimental data. This is a novel system for daily emotional health monitoring and can provide scientific suggestions for doctors or guardians. However, in the future, large scale of data should be accumulated. Moreover, the dynamic performance and energy efficiency also need improving.

This paper is supported by the National Natural Science Foundation of China under Grant No. 61472216, and by the PhD Programs Foundation of Ministry of Education of China under Grant No. 20120002110067. And these projects mainly focus on the wireless sensor networks and multimodal wearable bio signal sensing. The projects aim to provide solutions to a new model for healthcare in the next decade named active health, which means the human beings can have an access to their health information and intelligent suggestions in daily-life. The research group has made breakthrough in sparse EEG sensing, web-enabled compressive bio signal sensing and intelligent computation.