

基于虚拟现实和动态加权决策融合的恐高情绪识别^①



俞洋钊¹, 何炳蔚¹, 白丽英², 张彧³, 俞广杰¹, 钟发忻¹

¹(福州大学 机械工程及自动化学院, 福州 350116)

²(福州大学 人文社会科学院, 福州 350116)

³(福州大学 马克思主义学院, 福州 350116)

通信作者: 何炳蔚, E-mail: mebwhe@fzu.edu.cn

摘要: 目前恐高情绪分类中的生理信号主要涉及脑电、心电、皮电等, 考虑到脑电在采集和处理上的局限性以及多模态信号间的融合问题, 提出一种基于 6 种外周生理信号的动态加权决策融合算法。首先, 通过虚拟现实技术诱发被试不同程度的恐高情绪, 同步记录心电、脉搏、肌电、皮电、皮温和呼吸这 6 种外周生理信号; 其次, 提取信号的统计特征和事件相关特征构建恐高情感数据集; 再次, 根据分类性能、模态和跨模态信息提出一种动态加权决策融合算法, 从而对多模态信号进行有效整合以提高识别精度。最后, 将实验结果与先前相关研究进行对比, 同时在开源的 WESAD 情感数据集进行验证。结论表明, 多模态外周生理信号有助于恐高情绪分类性能的提升, 提出的动态加权决策融合算法显著提升了分类性能和模型鲁棒性。

关键词: 恐高症; 虚拟现实; 决策融合; 多模态生理信号; 情绪分类

引用格式: 俞洋钊, 何炳蔚, 白丽英, 张彧, 俞广杰, 钟发忻. 基于虚拟现实和动态加权决策融合的恐高情绪识别. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/9193.html>

Fear of Heights Emotion Recognition Based on Virtual Reality and Dynamic Weighted Decision Fusion

YU Yang-Zhao¹, HE Bing-Wei¹, BAI Li-Ying², ZHANG Yu³, YU Guang-Jie¹, ZHONG Fa-Xin¹

¹(School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350116, China)

²(School of Humanities and Social Sciences, Fuzhou University, Fuzhou 350116, China)

³(School of Marxism, Fuzhou University, Fuzhou 350116, China)

Abstract: Currently, the physiological signals in the classification of acrophobia emotions mainly involve electroencephalogram (EEG), electrocardiogram (ECG), and skin electromyography (EMG). However, due to the limitations of EEG acquisition and processing as well as the fusion between multimodal signals, a dynamic weighted decision fusion algorithm based on six peripheral physiological signals is proposed. Firstly, the different levels of acrophobia are induced in the subjects through the virtual reality technology, while six peripheral physiological signals (ECG, BVP, EMG, EDA, SKT, and RESP) are recorded. Secondly, the statistical and event-related features of the signals are extracted to construct a dataset of acrophobia emotions. Thirdly, a dynamic weighted decision fusion algorithm is proposed according to the classification performance, modal, and cross-modal information, so as to effectively integrate multi-modal signals to improve the recognition precision. Finally, the experimental results are compared with previous relevant research, and then verified on the open-source WESAD emotion dataset. The conclusions show that multi-modal peripheral physiological signals are conducive to enhancing the classification performance of acrophobia emotions, and

① 基金项目: 国家科技特区创新资助项目 (18-163-15-ZT-001-007-46); 2020 年福建省本科高校教育教学改革研究一般项目 (FBJG20200199)

收稿时间: 2022-12-24; 修改时间: 2023-02-03; 采用时间: 2023-03-14; csa 在线出版时间: 2023-05-24

the proposed dynamic weighted decision fusion algorithm significantly improves both the classification performance and model robustness.

Key words: acrophobia; virtual reality (VR); decision fusion; multimodal physiological signals; emotion classification

恐高症是一种常见的心理疾病,其特征表现为当患者处于一定高度时会产生过度焦虑和恐慌情绪,严重程度的恐高症甚至伴有眩晕、恶心、食欲不振等症状。统计显示,约91%的人都曾出现过恐高症状,约10%的人患有轻度或重度恐高症^[1]。

众所周知,恐高症可以通过暴露疗法进行缓解,该疗法主要分为想象暴露疗法和实景暴露疗法^[2]。前者需要患者有足够的想象力,能够逐步创建令自己产生恐高情绪的场景;后者则需要人工来搭建真实场景,往往实施起来难度较大,并具有一定的危险性。特别是对于恐高症患者的治疗而言,若对其实施实景暴露可能会更加发生危险。虚拟现实(virtual reality, VR)是近年来迅速发展起来的一项技术,利用计算机技术来生成一个视听触嗅等感觉逼真的虚拟时空世界,具有可控性、安全性、可重复性等优势,在精神治疗和康复领域已被广泛应用^[3]。虚拟现实暴露疗法(virtual reality exposure therapy, VRET)则是将VR提供的特定应激源与暴露疗法相结合,以计算机生成的虚拟环境辅助执行暴露疗法的治疗,适用于大规模患者的治疗,具有操作简单、可重复性好等优点,已被逐渐应用于心理治疗领域^[4]。因此,这项工作利用VR技术来实现更加有效、真实的恐高情绪诱发。

在先前工作中,研究人员主要通过访谈、问卷量表等方式来进行恐高症的症状识别。例如,使用恐高症问卷、行为逃避测试、焦虑敏感指数表等问卷量表形式来识别被试的恐高水平^[5-8]。然而,该识别方法存在强主观性的缺陷,许多人难以真实地表达自己的精神状态或是缺乏相关专业知识,导致其提供的信息在有效性和确凿性上仍有待商榷。

随着信息处理技术的发展,国外学者通过VR诱发个体的恐高情绪,并将收集到的生理指标应用于恐高情绪的识别。一些研究将从虚拟场景中获得的部分外周生理信号(肌电、心率、皮电、眼电)和唾液皮质醇的统计变化作为辅助识别手段^[9-12],但仍以问卷量表作为被试恐高状态的主要识别方法。为弥补问卷量表方式强主观性、对生理信号的统计分析并不能有效量

化恐高情绪等问题,一些学者使用支持向量机(support vector machine, SVM)、随机森林、深度神经网络等模型对脑电或将脑电与心电、皮电融合进行分析从而识别恐高情绪,部分研究的准确率已达到90%以上^[13-16]。但脑电信号的特征维数较高且具有高度的非平稳性和信噪比低的特点,对人体运动产生的生理伪影具有高度敏感性。因此仅适用于实验室环境下,无法用于实际环境下的可靠分析^[17,18]。而在先前恐高情绪识别的相关研究中并未涉及的肌电、脉搏、呼吸和皮温这4种外周生理信号通常会随着恐惧情绪的诱发而发生变化^[19,20],同时具有对个体情绪状态波动敏感、采集装置更为便携以及数据噪声低等优点。受此启发,这项工作利用心电、脉搏、肌电、皮电、皮温和呼吸6种外周生理信号实现恐高情绪的识别。

对多模态信号进行有效的融合是情绪识别的关键挑战。若仅对单模态生理信号分析获得的结果可靠性不足、泛化性低,而多模态生理信号利用不同类别的信号进行优势互补,极大地提升了情绪识别模型的鲁棒性^[21]。因此在设计多模态情绪识别方法时,更为有效的模态融合方式可以显著提升识别精度。例如,Bălan等^[13]搭建了一个虚拟高空场景来收集被试的脑电、心率和皮电,随后使用特征级融合对所有特征进行处理,最后通过多种机器学习和深度学习模型进行恐高情绪的识别,测试准确率从42.5%到89.5%不等。Zheng等^[15]同样使用VR诱发个体恐高情绪,从而进行了基于心电和眼电信号的对照实验,通过对特征级融合和基于多数投票规则的决策级融合的实验结果,得出恐高症和多模态生理信号之间的关系。然而,上述的融合方式存在信息冗余、增加运算成本、无法有效发挥每个生理信号的独特优势等问题。因此在情绪识别领域,近年的相关研究主要使用决策级融合来整合不同模态的信息。Albraikan等^[22]根据4种生理信号的识别准确率设计权重进行情绪的三分类,Li等^[23]则基于生理信号自身的稳定性设计权重以提高识别精度。这种通过单权重实现决策级融合在效果上尽管优于特征级融合,但在融合过程中,极易丢失模态之间的相关性且

无法充分体现每种模态的优势。因此,一些研究将上述的融合方式进行整合,提出包含多权重方式的融合方法以实现稳健的情绪分类^[21,24]。然而,上述的融合方式未充分考虑类不平衡和信号在融合后的相对重要性问题。因此本研究基于所使用的6种外周生理信号,提出一种包括3种权重方式的动态加权决策融合算法,从而实现稳定、高效的恐高情绪识别。

综上,针对恐高情绪识别相关研究在多模态生理信号的选择和融合方面的局限性,使用VR技术诱发被试的恐高情绪并收集6种外周生理信号,最后提出一种动态加权决策融合算法整合不同模态的生理信号,实现稳健的恐高情绪识别。本文主要贡献如下。

(1) 使用VR技术构建虚拟高空场景以诱发恐高情绪,并收集场景中被试的心电、脉搏、肌电、皮电、皮温和呼吸6种外周生理信号,为后续实验分析奠定基础。

(2) 针对现有融合方式未充分考虑类不平衡问题和生理信号对于分类性能的相对重要性,根据子分类器的分类性能、模态间的相关性和模态失真性设计3种权重,提出一种动态加权决策融合算法,从而实现不同生理信号的优势互补、极大地提升模型的鲁棒性。

(3) 将实验结果与先前相关研究进行对比分析,并使用开源的WESAD情感数据集进一步验证所提出方法的可行性。

本文其余部分安排如下:第1节中,概述本文中使用的方法;第2节阐述了本文进行的实验,包括被试信息、生理信号采集以及实验流程;第3节介绍了数据处理的细节,包括预处理、特征工程以及所提出的动态加权决策融合算法;第4节展示了对实验结果进行了分析讨论,并使用开源的WESAD数据集进一步验证所提方法的可靠性;第5节阐述了结论和未来的工作方向。

1 方法

1.1 基于VR的恐高情绪诱发实验设计

VR是一种基于计算机和各种高级技术的环境模拟系统,综合了视听触嗅觉等感觉器官的优势,从而创造一个沉浸感十足的多维信息空间^[3]。依据其独有的沉浸性、交互性、想象性等特性,VR可以提高临床治疗和情感计算中生态有效性方面的潜力。然而,这些研究通常使用简单的静态刺激,缺乏类似现实世界活动和

交互等许多潜在重要因素。因此,VR技术可以在未来神经经济学中扮演更为重要的角色。它提供了一个更加生态化的框架,使得研究者可以在这个框架内开展具有隐性测量的实验研究。

本文通过VR替代传统情绪诱发方式,搭建一个以现代化都市为背景的虚拟高空场景,使用3ds MAX对场景内的部分模型进行建模,在Unity 2017.3.0f1游戏引擎中利用MonoBehavior类库以及C#语言应用程序接口为场景添加行人、猫、直升机、薄雾、电梯上升以及空中坠落等动画事件。场景中包括4种类型的刺激,分别为3D听觉(风声、人声、车辆噪音)、静态视觉(摩天大楼、城区建设)、动态视觉(行人、车辆、云朵、薄雾)和3D动态视听(直升机、鸟群)。

由于绝大多数的恐高患者不敢尝试乘坐透明电梯,因此本实验安排被试通过透明电梯抵达115层的高空,如图1所示为电梯内视角。为模拟现实世界中电梯上升的感觉,从平地上升到高空的过程中使用3D音效,平地噪声(车辆噪音、人声)逐渐变小,空中噪声(直升机、鸟群、风声)逐渐变大。此外,从高处俯视与平时看到的景象是大不相同的,此时视野中的景象大幅度缩小且更为模糊,视觉信息大大减小,进而产生眩晕问题,导致恐高程度的加剧。因此场景通过添加似隐似现的薄雾用于进一步模糊被试的视线,实现恐高情绪的进一步诱发。



图1 电梯内视角

VR场景搭建完毕后,招募合适的被试进行实验数据采集。考虑到脑电在采集和处理方面的局限性,实验选取心电、脉搏、肌电、皮电、皮温和呼吸6种外周生理信号作为后续动态加权决策融合模型的输入。

1.2 基于动态加权决策融合的情绪识别

在情绪识别领域,基于单模态信号的识别早已被证实由于易受其余信号的干扰,其分类准确率和可靠性显著低于多模态情绪识别精度^[25]。为了解决该问题,近年来的大量工作都是基于多模态的信号模式进行研

究。如今多模态信息融合最常用的方法是特征级融合和决策级融合,前者虽然较为简单、计算复杂度低,但它创建的特征集通常含有冗余信息,并且无法考虑到不同信号针对不同情绪分类时的差异性。决策级融合则不存在上述缺点,它允许使用不同的权重来调整不同模态对于最终分类的贡献,从而利用每个信号的模态优势^[20]。

最常用的决策级融合方法是多数投票法^[25],但这种等权重的决策方式在某种程度上与特征级融合类似,无法体现不同信号的特点。因此一些研究根据每个生理信号的识别准确率进行权重定义,最终通过线性融合权重得出最优结果^[26]。Khezri 等^[24]则考虑了不同子分类器在生成分类结果时的一致性程度。此外,还有研究从信号可靠性方面来制定权重的分配^[21]。综上,现有研究未充分考虑数据类不平衡问题和模态对于分类时的相对重要性问题,提出开发一种动态加权决策融合算法 (dynamic weighted decision fusion algorithm, DWDFA)。该算法包含 3 种权重分配方式,分别为每个模态的分类性能系数权重、体现跨模态信息的最大互信息系数权重以及体现模态稳定的失真性系数权重。

先前已有大量研究表明, SVM 和 K 近邻 (K-nearest neighbors, KNN) 在基于生理信号的情绪识别领域被证明具有显著效果^[13,20]。因此,本文首先将收集到的数据进行预处理和特征工程,随后通过 DWDFA 将 6 种多

模态生理信号进行有效结合,最后使用 KNN 和 SVM 两种分类器识别不同程度的恐高情绪。

1.3 方法概述

提出的恐高情绪识别方法包括一个带有手柄控制器的 VR 头戴显示设备 HTC Vive Pro Eye、一个 Ergo-LAB 人机环境同步云平台和一台用于信息处理的台式电脑,系统流程图如图 2 所示。首先使用 VR 技术搭建“高空救猫”虚拟高空场景诱发被试的恐高情绪,同步记录被试在虚拟高空场景中执行任务时的外周生理信号(心电、脉搏、肌电、皮电、皮温和呼吸)和任务结束后的主观痛苦感觉单位(subjective units of distress scale, SUDS)量表记录,从而构建生理信号-初始情绪关联数据集。随后对初始数据集进行异常值处理和滤波方式以增加数据的可靠性,并提取多模态生理信号的统计特征和反应信号上下文相关信息的事件相关特征,引入过滤法、包装法和嵌入法筛选进行特征选择,并使用 SVM 和 KNN 进行分类。为提高模型的精度和稳定性,考虑到 SVM 更优异的表现和 KNN 的效率问题,选用 6 种生理信号对应的 SVM 子分类器,并根据 DWDFA 分配权重实现动态加权融合,其中权重参数包括子分类器分类性能系数权重、子分类器最大互信息系数权重和生理信号受外界干扰的失真性系数权重,最后根据测试集和 WESAD 数据上的性能指标评测该方法的可行性。

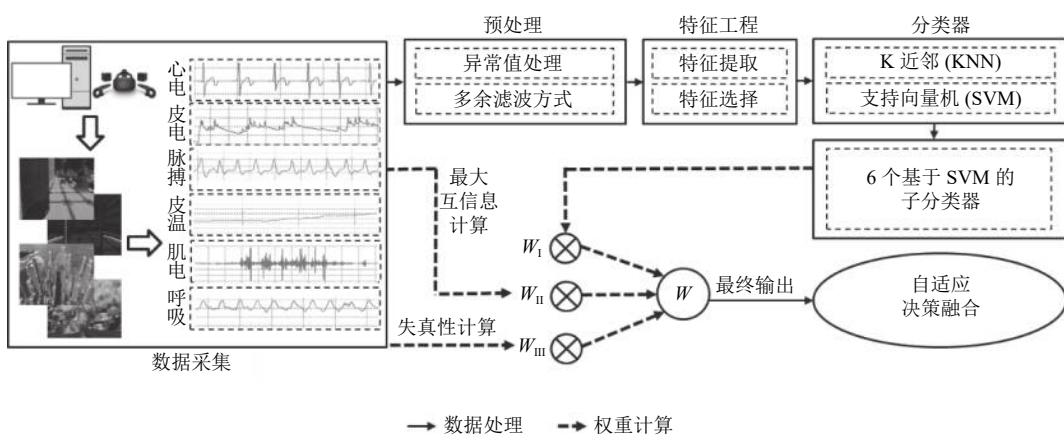


图 2 恐高情绪识别流程图

2 实验

2.1 被试信息

实验被试来自当地某大学的 55 名硕士研究生(6 名女生和 49 名男生,平均年龄为 23.5 ± 3.5 岁),所有被试均无任何脑部或精神疾病,视听均正常或者矫正

正常,惯用手均为右手。实验前,每位被试均被告知实验的具体内容和程序,并签署一份实验知情同意书。

2.2 实验流程

被试如有不适可随时提出终止实验,实验结束后工作人员根据 SUDS 记录为被试打上相应的情绪标签。

具体实验流程如下.

(1) 工作人员为被试佩戴 VR 头盔和生理采集设备, 并提前告知被试具体实验流程.

(2) 设备佩戴完毕后, 开始执行传感器校准. 被试在实验开始前需要进行两轮眼动信号校准: 第 1 轮的校准需要被试在保证头部不动的情况下, 使用瞳孔追踪 VR 校准场景中的移动点; 第 2 轮在此基础上, 通过调整 VR 头盔侧边的对焦按钮完成校准. 两轮校准的目的是将被试的视野调整到最佳范围, 保证他们在最舒适的状态下进行沉浸式的 VR 体验.

(3) 校准完毕后, 被试进入 VR 场景. 由于大多数被试先前没有体验 VR 的相关经验, 因此如图 3(a) 所示, 首先使其在街道上对 VR 场景熟悉一段时间, 缓解由初次接触 VR 所产生的情绪波动.

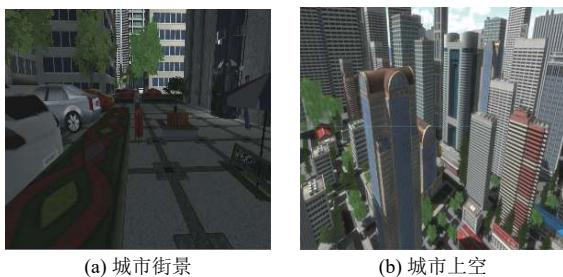


图 3 虚拟场景部分截图

(4) 场景熟悉完毕后, 被试进入电梯. 使用手柄选择楼层开始上升, 电梯上升过程速度模拟真实电梯上升速度, 并可以实时看到电梯上升时的外景, 如图 3(b) 所示. 到达指定楼层后, 被试需要走出电梯, 踏上面前的木板. 木板的边缘有一只小猫, 被试需要使用手柄触发抱起小猫的事件, 随后安全回到电梯内部, 即任务成功. 反之则自动触发木板断裂、高空坠落等相应事件, 即任务失败. 实验过程中存在部分患有严重恐高症的被试无法走到木板的边缘, 他们被允许站在可以到达的极端位置环视场景.

(5) 任务结束后, 被试填写 SUDS 和模拟器不适量表. SUDS 被划分分为平静、低恐高和高恐高 3 个类别, 作为初始数据集的标签; 模拟器不适量表记录则用于评估 VR 暴露可能产生的副作用.

3 数据处理

3.1 预处理

数据采集时可能会发生各种情况导致传感器信号

劣化、噪声和外部干扰, 即用户移动、电极断开、环境温湿度不稳定、用户相关的生理功能障碍和静电伪影, 造成采集到的数据存在数据缺失、异常点等问题. 为提高情绪识别的准确率, 需要对原始数据进行预处理操作^[27].

由于不同模态的生理信号具有不同的数据特征, 因此它们的预处理方法也不同. 但对于具有相同特点的信号, 可以对其进行一致的预处理操作. 如皮电在低频范围内的数据特征才具有生理意义, 皮温经常含有高频噪声, 所以采用低通滤波器对皮电和皮温进行处理. 对于心电和脉搏, 它们通常受到电力线干扰 (50–60 Hz)、电极接触噪声、基线漂移、电子设备产生的仪器噪声的污染, 所以使用具有 0.002 Hz 和 100 Hz 截止频率的三阶巴特沃斯滤波器来消除心电和脉搏信号中的噪声影响. 肌电的幅度从 50 μ V 到 30 mV 不等, 其频谱在 10–500 Hz 范围内, 因此通过截止频率高于 100 Hz 的高通滤波器剔除诸如心电伪影、运动伪影等噪声. 对于呼吸信号, 采用带通滤波器减少噪声干扰, 随后通过平滑信号消除滤波后的信号趋势^[20]. 最后将预处理后的数据进行标准化处理, 并按照 7:3 的预设比例分割为训练集和测试集.

3.2 特征工程

特征工程包含众多内容, 是成功应用机器学习的一个关键环节, 而如何充分利用数据进行预测建模就是特征工程要解决的首要问题.

原始数据经过预处理和分割后, 便可以提取描述自身信息的特征指标. 它以紧凑的方式来表征原始数据, 并允许在变换的维度中比较不同模态, 从而增强信息信号的特征. 表 1 展示了所提取的特征, 这些特征既可用于表达生理信号变化趋势的统计特征, 也有反映生理信号上下文相关信息的事件相关特征^[17].

由于不同特征变量之间的依赖关系可能会降低分类器的性能, 不是所有特征都适合作为分类器的输入. 为了找到表示恐高情绪的最佳特征子集、提升分类器的性能, 采用几种常见的特征选择技术, 包括过滤法、嵌入法和包装法^[19]. 过滤法和嵌入法仅需一次训练即可完成特征选择, 其不同之处在于过滤法主要根据发散性或相关性的评分来选择特征, 嵌入法则需要根据选择的算法来确定特征的优劣. 相比过滤法和嵌入法, 包装法则根据目标函数进行多次训练达到特征选择的目的. 尽管它的速度更为缓慢、计算成本更高, 但是却

更加精确^[20]. 考虑到本文的样本数据集不算太大, 最终选用包装法作为 SVM 分类器的特征选择方法.

表 1 事件相关特征和统计相关特征

特征	信号	特征向量
	心电	相邻R-R间期的连续差异的标准偏差,
	脉搏	平均心率, 心动间隔均值等
	皮电	皮肤电导水平, 皮肤电导反应个数,
		平均延迟时间, 平均反应强度等
事件相关特征	肌电	超过阈值的信号积分, 信号激活次数,
		积分肌电值等
	呼吸	呼吸速率, 吸气和呼气的持续时间, 吸入/呼出空
		气量, I/E之间的比率等
	皮温	最小和最大动态范围和斜率的标准差等
	心电	
	脉搏	平均值, 最大值, 最小值,
统计相关特征	皮电	标准差, 均方根, 方差频率标准差,
	肌电	频率均方根, 小波能量比,
	呼吸	小波能量熵, 小波尺度熵, 心率变异性等
	皮温	

包装法中最典型的目标函数是递归特征消除法, 它是一种贪婪优化算法, 通过每次迭代时创建模型、剪除最差特征, 直到所有的特征都被计算过. 然后根据剩余特征的顺序来对特征进行排名, 最终选出一个最佳子集. 它的参数 estimator 是所选择的实例化评估器, n_features_to_select 是最终要选择的特征个数, step 代表每次迭代要移除的特征个数, 使用学习曲线选择最优的特征数量.

由于 KNN 分类器没有公开 coef 和 feature_importances_ 属性, 所以无法对其应用包装法. 于是首先采用过滤法作为 KNN 分类器的特征选择方法, 但是仅过滤了 2 个特征, 没有显著改善 KNN 分类器的分类性能. 随后计算这些特征之间的 Pearson 相关性, 如图 4 所示, 这些特征之间大部分是高度不相关的, 可以作为 KNN 分类器的输入特征^[28], 所以最终将基于 SVM 分类器的包装法筛选后的特征作为 KNN 分类器的输入.

由于恐高情绪的诱发往往伴随着个体压力的产生, 而先前研究表明当个体产生压力时, 其呼吸和心率会随着交感神经和副交感神经活动做出变化. 图 5 显示经过特征选择后得到的 20 个特征的特征重要性权重, 权重越大说明该特征对于情绪的准确识别贡献也越大, 其中特征的索引和描述如表 2 所示. 结果显示从呼吸和心电中提取的特征对于恐高情绪识别的贡献度较高, 符合先前研究结论.

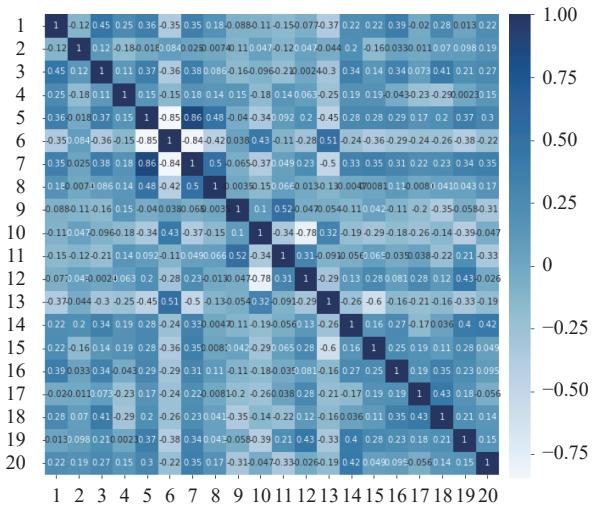


图 4 Pearson 热力图

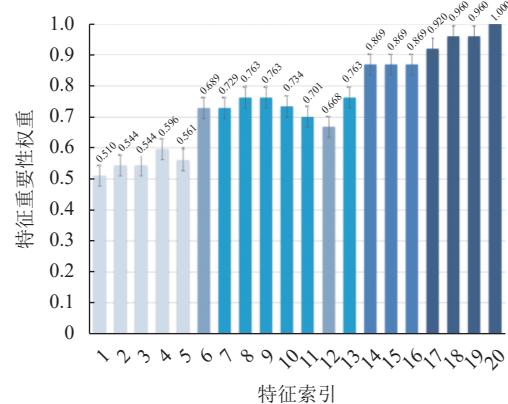


图 5 特征重要性权重

表 2 特征选择后的特征

索引	生理信号	特征
1~5	脉搏	大于 20 ms 的心率个数占心跳总数的百分比, 低频和高频能量占其综合的百分比, 高频功率占总功率的百分比, 低频带和高频带功率的比值, 差值散点图中第一象限点的个数.
6~9	皮电	皮肤电导平均值, 皮肤电导最小值, 渐变信号最小值, 突变信号最小值.
10~13	肌电	最大值, 最小值, 平均绝对值, 积分肌电均值.
14~18	心电	低频和高频能量占其综合的百分比, 平均心率, 高频的功率, 大于 50 ms 的心率个数占心跳总数的百分比, 高频功率占总功率的百分比.
19, 20	呼吸	平均值, 标准差.

特征级融合虽然能最大程度地保留原始信号的特征信息, 但它并没有考虑到不同模态之间的差异性, 因此提出一种动态加权决策融合的技术以解决上述问题. 对于先前的决策融合相关研究^[21~24,26], 研究者主要聚焦

于心电、肌电、皮电、脉搏这4种常用的生理信号，所以本文添加了已被证明对识别具有显著效果的呼吸信号。此外，由于皮电和皮温在某种程度上的关联，尽管其对于识别的贡献较低，但可以通过提出的DWDFA减少其权重占比，从而保证模型的性能。

3.3 基于DWDFA的分类

SVM和KNN是两种经典的监督学习机器学习分类器，已被广泛用于各种领域的识别任务^[13]。因此在特征工程结束后，将二者作为恐高情绪识别的分类器。

SVM是一种广义的线性分类器，其基本模型是定义在特征空间上的间隔最大的线性分类器，并且它还包括不同的核函数，这使它成为实质上的非线性分类器。它的目标是求解能够正确划分训练数据集并具有最大几何间隔的分离超平面。而对于非线性问题，SVM可以将非线性特征空间映射到线性空间，随后使用线性SVM方法对新的特征空间进行分类^[29]。本文采用以径向基函数为核函数的SVM，并采用学习曲线对其参数进行优化。

与SVM不同，KNN是一种基于实例的非参数分类器，没有训练过程。它的核心思想是：如果一个样本在特征空间中最相邻的K个样本中的大部分属于某个类别，那么该样本也属于该类别。本文通过网格搜索对KNN进行调参，最终选择K值为5，距离的权重为distance。

考虑到KNN的分类精度和效率显著低于SVM（具体结果见第4节），因此本文将SVM作为最终的分类器和DWDFA权重计算过程的子分类器。所提出的3种权重分配方式如下所示：(1) W_I 为子分类器分类性能系数权重；(2) W_{II} 为子分类器最大互信息系数权重；(3) W_{III} 为生理信号受外界干扰的失真性系数权重。

3.3.1 分类性能系数权重 W_I

分类准确率就是分类正确的比例，在分类领域中被看作是最常用的评估指标，尤其在数据集类别平衡时，其数值越高，代表该分类器的性能越好。然而大多数的数据集常存在类别不平衡问题，即数据偏斜。在这种情况下，依据准确率所得出的结论往往是极具迷惑性且意义不大。因此对于存在数据偏斜的数据集，大量研究采用的是对于不平衡数据更为敏感的F1分数，它被定义为精确率和召回率的调和平均值。

考虑到本文的数据集结构并不均衡，本文根据不同子分类器的分类准确率和F1分数的加权融合结果来计算权重 W_I 如式(1)所示， W_{II} 是一个N阶对角矩阵，

其中 N 表示恐高情绪类别数， β 是准确率和 F_1 分数的加权系数。 a_{NN} 表示在第 i 种生理信号对于第 N 种情绪的分类准确率， E 是 N 阶单位矩阵。

$$W_{II} = (1 - \beta) \begin{pmatrix} a_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{NN} \end{pmatrix} + \beta \cdot E \cdot F_1 \quad (1)$$

3.3.2 最大互信息系数权重 W_{II}

在大多数情况下，不同模态的信息不是完全没有关联的，而决策融合的过程往往会丢失它们之间的相关性，进而导致准确率的下降^[25]。且大量的研究是根据两个变量的互信息来度量二者的相关性，其定义如式(2)所示： m 和 n 代表两个独立的子分类器， $p(m,n)$ 是二者的联合概率分布， $p(m)$ 和 $p(n)$ 是边缘概率分布。

$$I(m;n) = \int p(m,n) \log_2 \frac{p(m,n)}{p(m)p(n)} dx dy \quad (2)$$

然而当数据集存在数据偏斜时，互信息更容易倾向少数类，缺乏公平性；此外，互信息仅限于特定的函数类型，而通过单一函数建模得出变量之间的复杂关系并不精准^[30]。本文将最大互信息(maximal information coefficient, MIC)系数作为不同生理信号之间相关性，它相较于互信息具有更多的普适性和公平性，因此拥有更高的准确率。 W_{II} 的计算过程如图6所示。

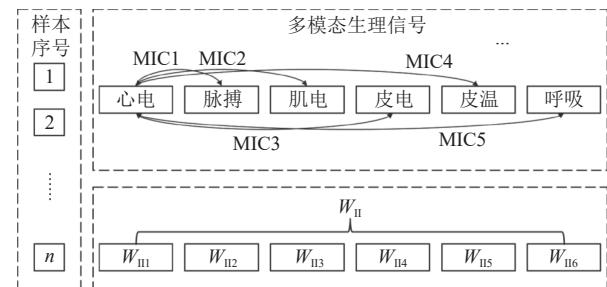


图6 权重 W_{II} 的计算流程图

式(3)是两种信号之间的MIC具体计算方式。它将两个变量的关系离散在二维空间，并沿着各自的方向划分区间数，其中 a 和 b 是划分个数， c 为情绪类别。根据文献[30]的设定， B 一般设置为数据量的 0.6^2 左右。

$$mic_{m,n}^c = \max_{a \cdot b < B} \frac{I(m;n)}{\log_2 \min(a,b)} \quad (3)$$

所有样本的MIC系数权重 W_{II} 如式(4)所示，其中 N 为情绪类别， M 为生理信号总数。

$$W_{IIi} = \frac{1}{N(M-1)} \sum_{j=1}^N \sum_{\substack{m=1 \\ m \neq i}}^M mic_{i,m}^j \quad (4)$$

3.3.3 生理信号失真性系数权重 W_{III}

众所周知, 生理信号的失真主要源于被测个体的身体运动, 它会对信号的质量产生不利影响. 第*i*个生理信号的第*k*种失真点判断指标的失真指数 λ_i^k 定义如式(5)所示:

$$\lambda_i^k = \begin{cases} 1, & n \geq \alpha \\ \frac{n}{\alpha}, & n < \alpha \end{cases} \quad (5)$$

其中, n 为失真数据点的总数, α 是判断该信号是否不失真的阈值. 当 n 大于等于阈值, 失真指数设置为 1; 反之则计算二者的比值. 由于不同信号的结构不同, 因此判定方式也不尽相同. 如图 7 所示, 心电信号的缺失会导致 R 波波峰的消失, 影响识别精度, 因此将 R 波波峰的异常变化情况作为心电信号失真的判定指标之一. 此外, 由于测量过程的基线漂移等情况, 呼吸信号的表面凹凸不平, 这就导致如图 8 所示在检测时, 其波峰波谷容易出现较大的误差.

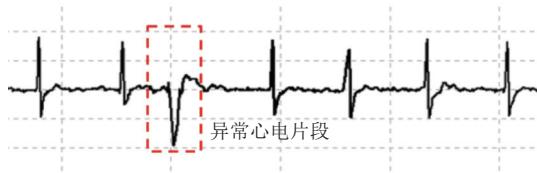


图 7 心电信号失真点判断

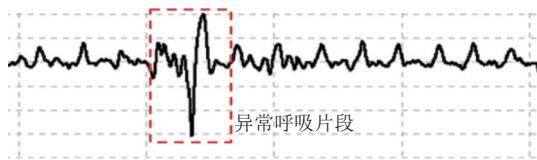


图 8 呼吸信号失真点判断

当每种模态的失真点判定完毕后, 便可以计算每种模态的失真性系数权重 W_{IIIi} , 其定义如式(6)所示:

$$W_{IIIi} = \frac{\left(1/K \sum_{k=1}^K (1 - \lambda_i^k)\right)^\theta}{\sum_{i=1}^M \left(1/K \sum_{k=1}^K (1 - \lambda_i^k)\right)^\theta} \quad (6)$$

其中, K 表示每个模态的信号有 K 种失真点判定指标, θ 代表每种信号的惩罚参数, 它与加权参数 W_{IIIi} 成反比, 对越容易失真的信号给予越高的惩罚, 使 W_{IIIi} 越低.

三权重计算完毕后, 将其进行加权融合组成最终权重 W_i , 由于分类性能系数权重是一个 N 阶对角矩阵, 而最大互信息系数权重和生理信号失真性系数权重是

一个向量, 因此计算时需将其乘以 N 阶的单位矩阵 E . 权重融合的计算方式如式(7)所示:

$$W_i = W_{Ii} + E \cdot (W_{IIi} + W_{IIIi}) \quad (7)$$

W_i 权重计算完毕后, 需要将其与子分类器判断的结果联合考虑, 进而得出最终的分类结果. 第*i*个子分类器的识别向量为 C_i , 因此经过加权融合后的分类器结果如式(8)所示.

$$C = \sum_{i=1}^M W_i \cdot C_i \quad (8)$$

最后根据最大值规则的判断策略得到最终的识别结果, 若 C 的最大值序号为 n ($1 \leq n \leq N$), 则识别结果为该序号对应的恐高情绪标签. DWDFA 算法如算法 1.

算法 1. DWDFA 算法

输入. 心电、脉搏、肌电、皮电、皮温和呼吸信号
输出. 恐高情绪类别

- 1) 训练 6 个基于 SVM 的子分类器.
 - 2) 根据子分类器的准确率和 F1 分数计算分类性能系数权重 W_1 (式(1)).
 - 3) 计算不同生理信号之间的最大互信息系数权重 W_{II} (式(4)).
 - 4) 计算生理信号的失真性系数权重 W_{III} (式(6)).
 - 5) 根据 W_1 、 W_{II} 和 W_{III} 计算权重 W (式(7)).
 - 6) 动态加权融合子分类器, 得出最后的恐高情绪类别 (式(8)).
-

4 实验结果与分析

本文算法实现和实验结果处理平台为 Intel Core i7-6700 CPU @ 3.40 GHz, 16 GB 内存, 使用 Matlab 和 Python 进行数据分析与处理.

4.1 自评量表结果

实验结果来自当地某大学 40 名健康的硕士研究生 (4 名女生, 36 名男生), 其余被试由于测试过程中传感器信号丢失、身体不适等缘故被排除在研究之外. 实验后填写的 SUDS 显示标签为平静的有 14 人, 标签为低恐高的有 11 人, 标签为高恐高的有 15 人, 不存在样本严重不均衡问题, 并且自评结果在统计学方面存在显著性差异 ($p < 0.05$).

4.2 性能指标

选择 70% 的样本作为训练集的样本, 剩余的数据作为评估分类器性能的测试集, 并以几何平均准确率、平均准确率和 F1 分数作为性能指标. 准确率代表所有样本中被正确分类实例的数量, F1 分数则是衡量机器学习模型精确度的一种综合指标, 它是精确率和召回率的调和平均值^[21].

4.3 结果评测

合理的特征选择将有助于情绪分类,本节首先评测了特征选择对于识别的精度是否具有显著影响。第3.2节已经阐述了本次工作选用包装法作为特征选择的理由,图9显示使用6种单模态生理信号和多模态生理信号(基于DWDFA的识别方法)在有无特征选择情况下的分类平均准确率。其中单模态信号的分类选取各子分类器中的最高准确率,除了皮电信号(无特征选择)和皮温信号(特征选择)是基于KNN分类器得出的最优结果,其余的最高平均准确率均来自SVM分类器。对于单模态生理信号,心电和呼吸信号获得了最高平均准确率的前两名,分别为76.92%和74%,符合第3.2节特征重要性权重的计算结果,并证实呼吸和心电与恐高程度具有极大的关联。对于多模态生理信号,基于DWDFA的识别方法获得了89.75%的平均准确率,显著优于单模态生理信号。此外,虽然特征选择技术无法大幅提高某些模态的分类准确率,但它可以降低特征向量维数,从而减少分类器的运算时间。

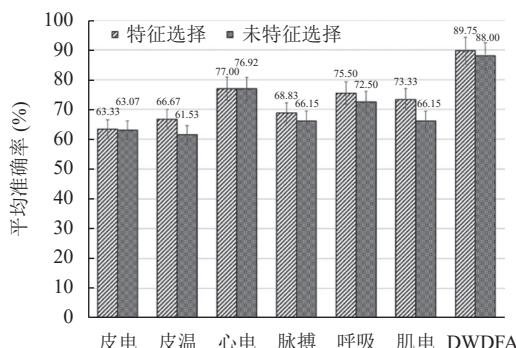


图9 有无特征选择的分类准确率

表3显示了不同生理信号对于恐高情绪分类的识别准确率,明显可以发现平静情绪类别更容易被区分。对于综合的分类性能,心电信号取得了最高的平均准确率和F1分数,分别为77%和79.23%,同时也是最为稳定的单模态信号。然而,呼吸信号在对于平静类别和低恐高类别的识别准确率却高于综合性能最优的心电信号。根据上述分析结果,不同的信号模态在对特定情绪类别进行分类时会表现出不同的效果。因此,为充分发挥不同信号模态在识别不同情绪方面的独有优势,引入呼吸信号且赋予每个信号模态不同的权重是至关重要的。

如第3.3节提出的DWDFA,它为6种模态的生理信号设计了不同的权重分配,下面进行了对照实验来评估不同方法的识别性能。从表4可以看出,使用特征

级融合获得的平均准确率仅高于基于权重,仅为84.72%。此外,如果只使用单一权重的方法进行分类,基于权重的方法获得了最高的平均准确率(87.22%)和最低的标准差($\pm 8.11\%$),是对分类结果贡献最大的权重。当三权重融合后,即基于DWDFA的识别方法达到了89.75%的最高平均准确率和最高的稳定性,同时该方法也在几何平均准确率和F1分数方面取得了最优结果。同时从图10的混淆矩阵可以得出,所提出的方法可以很好地进行恐高情绪三分类的识别,其性能超过任何的子分类器和其余融合方式,表明所提出的决策融合方式显著提高了识别的准确率和稳定性。

表3 不同生理信号的识别准确率(%)

分类	平静	低恐高	高恐高	平均准确率	F1分数
心电	81.6 \pm 15.24	72.22 \pm 12.03	74.70 \pm 13.19	77.00 \pm 8.93	79.23 \pm 9.81
皮电	69.5 \pm 17.21	53.08 \pm 16.78	63.46 \pm 17.04	63.33 \pm 12.47	61.75 \pm 12.95
肌电	80.0 \pm 17.12	70.77 \pm 15.33	69.23 \pm 16.70	73.33 \pm 11.70	72.61 \pm 12.88
脉搏	74.2 \pm 16.94	62.31 \pm 14.42	67.69 \pm 15.81	68.83 \pm 10.48	67.69 \pm 13.02
呼吸	83.8 \pm 14.78	73.8 \pm 13.10	69.6 \pm 13.25	75.5 \pm 9.05	74.92 \pm 10.86
皮温	71.43 \pm 16.33	61.54 \pm 18.02	65.38 \pm 17.16	66.67 \pm 12.21	65.76 \pm 13.07

表4 不同生理信号的识别准确率(%)

融合方式	几何平均准确率	平均准确率	F1分数
特征融合	83.06 \pm 9.74	84.72 \pm 8.78	82.78 \pm 9.51
W_I	86.35 \pm 8.92	87.22 \pm 8.11	85.32 \pm 9.26
W_{II}	83.53 \pm 9.57	85.28 \pm 9.05	82.99 \pm 9.89
W_{III}	80.23 \pm 10.79	81.39 \pm 10.24	78.96 \pm 11.02
DWDFA	88.92 \pm 8.57	89.75 \pm 7.42	88.86 \pm 8.05

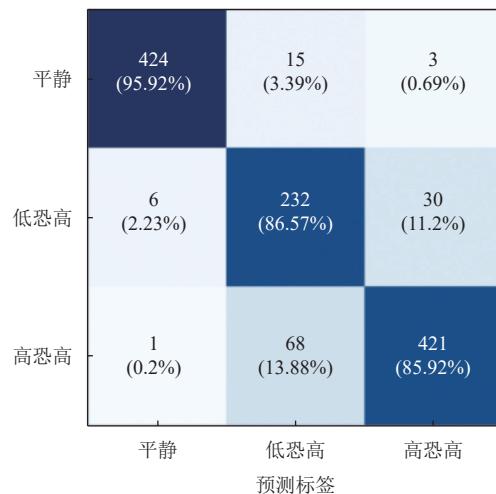


图10 基于DWDFA的恐高情绪识别混淆矩阵

表5比较了现有恐高或焦虑情绪识别的相关研究,尽管二者的压力源并不相同,但考虑到恐高和焦虑情绪的识别均属于精神压力检测领域,因此也将焦虑情绪识别的相关研究与本研究进行对比分析。在文献[16]

中, Hu 等人利用 Steam 中的 VR 游戏诱发被试的恐高情绪, 采集了 60 名被试的脑电信号, 通过卷积神经网络获得了 88.77% 的准确率. Kaur 等^[14]仅采集 5 名被试的脑电信号, 同样也使用卷积神经网络获取了 89.20% 的二分类结果. Bălan 等^[13]开发了一款虚拟高空场景并采集了 8 名被试的脑电、心电和皮电信号, 采用 SVM 获得了二分类 89.5%、四分类 42.5% 的最高准确率. Zheng 等^[15]通过分析心电和眼电信号, 获得了最高 93.93% 的恐高情绪二分类结果. 在焦虑情绪识别中, Salkevicius 等^[27]采集了 30 名被试的皮电、皮温和脉搏, 获得了 86.30% 的四分类结果, Zhang 等^[31]则采集了 22 名被试的皮电信号, 其中二分类的准确率为 86.70%. 显然, 从表 5 得出脑电信号在恐高情绪分类中的应用极为广泛且在同样的分类类别中准确率也相对较高, 但考虑到脑电信号具有噪声极大、侵入性强、特征维度高等缺点, 其在非实验室条件下的可穿戴设备应用中具有一定的局限性^[32]. 而本研究采用基于 DWDFA 的识别方法获得的准确率在表 5 中仅次于中的结果, 但其仅涉及 17 名被试, 并只进行了二分类的情绪识别. 综上, 对比分析表明本研究提出的决策融合方法 (DWDFA) 在具有较大规模样本量的基础上成功地融合了 6 种模态的信号, 从而提升了情绪识别的性能.

表 5 与其余相似文献的比较

文献	生理信号	刺激类型	被试数量	类别数量	平均准确率 (%)
Hu 等 ^[16]	脑电	恐高	60	4	88.77
Kaur 等 ^[14]	脑电	恐高	5	2	89.20
Bălan 等 ^[13]	脑电、心电、皮电	恐高	8	2	89.5
Zheng 等 ^[15]	心电、眼电	恐高	17	2	93.93
Salkevicius 等 ^[27]	皮电、皮温、脉搏	焦虑	30	4	86.30
Zhang 等 ^[31]	皮电	焦虑	22	2	86.70
DWDFA	心电、脉搏、肌电、皮电、皮温、呼吸	恐高	40	3	89.75

本研究进一步对几种常用的融合方式进行了对比, 即基于最小值规则的决策融合、基于中值的决策融合、多数投票法和基于 4 种生理信号 (心电、脉搏、肌电和皮电) 的 DWDFA 方法. 图 11 显示了分类的对比结果, 与其余方法相比, 本文提出的基于 6 种生理信号的 DWDFA 方法分类准确率最高, 比排名第 2 的基于 4 种生理信号 (心电、脉搏、肌电和皮电) 的 DWDFA

方法高出 2.67%, 并具有更高的稳定性. 这表明相比于先前提出的决策融合方法, 本研究所提出的 DWDFA 方法有效地融合了更多模态的外周生理信号.

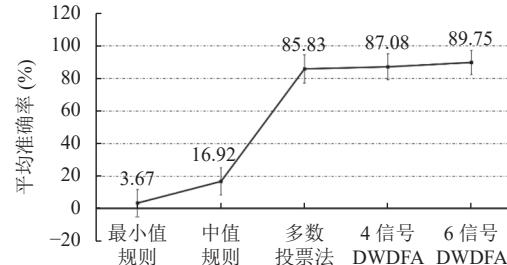


图 11 与其他融合方式的对比结果

考虑到目前暂无可公开的 VR 情感数据集, 本研究将所提出的 DWDFA 方法应用于 WESAD 数据集以验证该方法的可行性. WESAD^[33]是由 Schmidt 等人发布的多模态情感数据集, 通过视频、中性杂志和数学题目诱发 15 名被试开心、中性和压力 3 种情绪, 包括心电、脉搏、肌电、皮电、皮温和呼吸 6 种生理信号. 对该数据集进行 7:3 的数据集分割, 预处理和特征工程与第 3 节中的操作一样, 最终获得 90.93% 的准确率. 表 6 展示了使用 WESAD 数据集的相关文献, 所有文献都使用多模态生理信号进行二或三分类的模型构建. 其中, 文献 [34,35] 采用岭回归分类器和线性判别分类器进行有无压力的识别, 最高获得了 87.4% 的平均准确率. 对于三类的研究, Schmidt 等^[33]选取的 6 种生理信号与本研究一致, 通过自适应增强算法获得了最高 79.86% 的平均准确率. Yan 等^[21]则采用基于 SVM 的自适应决策融合技术对心电、脉搏、皮电和肌电进行分析, 获得了 87.41% 的平均准确率. 与其余研究相比, 基于 DWDFA 的识别方法获得了最高的平均准确率, 比排名第 2 的研究提高了 3.52%, 这些结果证实了该技术的有效性, 其优异的分类性能源于动态加权的调节不同模态的权值.

5 结语

考虑到脑电信号在恐高情绪识别研究中应用的局限性以及多生理信号的融合问题, 通过 VR 技术模拟真实高空环境并在安全的实验环境下收集被试的心电、脉搏、肌电、皮电、皮温和呼吸 6 种外周生理信号. 随后提出一种包括分类性能系数、最大互信息系数和生理信号失真性系数 3 种权重的动态加权融合算法 (DWDFA) 以整合不同模态之间的信息, 实现在多模态决策融合过程中权重的动态融合, 从而提升识别

准确率和稳定性。

将实验结果与先前相关研究进行对比分析,表明提出的 DWDFA 方法相比其余研究的分类模型具有更高的分类性能。此外,使用一个名为 WESAD 的开源情感数据集进一步地验证所提出方法的优越性,在与其余研究中的对比分析获得了最高的平均准确率。

表 6 与其余相似文献的比较

来源	生理信号	类别	
		数量	平均准确率 (%)
Siirtola等 ^[34]	皮温、脉搏、心电	2	87.4
Alshamrani等 ^[35]	心电、脉搏、肌电、皮电、皮温、呼吸	2	87
Schmidt等 ^[33]	心电、脉搏、肌电、皮电、皮温、呼吸	3	79.86
Huynh等 ^[36]	皮电、皮温、脉搏	3	81.78
Lin等 ^[37]	心电、肌电、皮电、皮温、呼吸	3	83
Yan等 ^[21]	心电、脉搏、肌电、皮电	3	87.41
DWDFA	心电、脉搏、肌电、皮电、皮温、呼吸	3	90.93

综上所述,该方法可以安全地应用于识别在虚拟环境中人的恐高水平,不仅为恐高症的辅助诊断提供科学客观依据,同时也为后期恐高症的治疗提供可靠参考。未来的工作拟从以下 3 方面着手:首先,考虑更多的生理信号失真性判定指标,以进一步改进生理信号失真性系数权重,并将其与能自动提取特征的深度学习技术相结合;其次,利用外周生理信号的便携性、数据易处理性等优势,开发融合多信号的可穿戴设备,进一步提升在 VR 场景中活动的自由度;最后,所提出的 DWDFA 方法不应仅局限于恐高情绪的识别,未来将涉及包括社交恐惧症、特定恐惧症、焦虑症等症状程度的识别,并与虚拟现实暴露疗法结合,从而提供相应的治疗手段。

参考文献

- Zeng N, Pope Z, Lee JE, et al. Virtual reality exercise for anxiety and depression: A preliminary review of current research in an emerging field. *Journal of Clinical Medicine*, 2018, 7(3): 42. [doi: [10.3390/jcm7030042](https://doi.org/10.3390/jcm7030042)]
- Wang QX, Wang H, Hu F, et al. Using convolutional neural networks to decode EEG-based functional brain network with different severity of acrophobia. *Journal of Neural Engineering*, 2021, 18(1): 016007. [doi: [10.1088/1741-2552/abfdb](https://doi.org/10.1088/1741-2552/abfdb)]
- 刘崇进, 吴应良, 贺佐成, 等. 沉浸式虚拟现实的发展概况及发展趋势. *计算机系统应用*, 2019, 28(3): 18–27. [doi: [10.15888/j.cnki.csa.006732](https://doi.org/10.15888/j.cnki.csa.006732)]
- Meyerbröker K, Emmelkamp PMG. Virtual reality exposure therapy in anxiety disorders: A systematic review of process-and-outcome studies. *Depression and Anxiety*, 2010, 27(10): 933–944. [doi: [10.1002/da.20734](https://doi.org/10.1002/da.20734)]
- Coelho CM, Silva CF, Santos JA, et al. Contrasting the effectiveness and efficiency of virtual reality and real environments in the treatment of acrophobia. *Psychology Journal*, 2008, 6(2): 203–216.
- Carl E, Stein AT, Levihn-Coon A, et al. Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders*, 2019, 61: 27–36. [doi: [10.1016/j.janxdis.2018.08.003](https://doi.org/10.1016/j.janxdis.2018.08.003)]
- Emmelkamp PMG, Bruynzeel M, Drost L, et al. Virtual reality treatment in acrophobia: A comparison with exposure in vivo. *CyberPsychology & Behavior*, 2001, 4(3): 335–339.
- Coelho CM, Silva J, Pereira AF, et al. Visual-vestibular and postural analysis of motion sickness, panic, and Acrophobia. *Acta Neuropsychologica*, 2017, 15(1): 21–33. [doi: [10.5604/1231966.1237325](https://doi.org/10.5604/1231966.1237325)]
- Krupić D, Žuro B, Corr PJ. Anxiety and threat magnification in subjective and physiological responses of fear of heights induced by virtual reality. *Personality and Individual Differences*, 2021, 169: 109720. [doi: [10.1016/j.paid.2019.109720](https://doi.org/10.1016/j.paid.2019.109720)]
- Wuehr M, Breitkopf K, Decker J, et al. Fear of heights in virtual reality saturates 20 to 40 m above ground. *Journal of Neurology*, 2019, 266(S1): 80–87. [doi: [10.1007/s00415-019-09370-5](https://doi.org/10.1007/s00415-019-09370-5)]
- Cleworth TW, Horslen BC, Carpenter MG. Influence of real and virtual heights on standing balance. *Gait & Posture*, 2012, 36(2): 172–176. [doi: [10.1016/j.gaitpost.2012.02.010](https://doi.org/10.1016/j.gaitpost.2012.02.010)]
- Bălan O, Moldoveanu A, Petrescu L, et al. Sensors system methodology for artefacts identification in virtual reality games. *Proceedings of the 2019 International Symposium on Advanced Electrical and Communication Technologies*. Rome: IEEE, 2019. 1–6.
- Bălan O, Moise G, Moldoveanu A, et al. An investigation of various machine and deep learning techniques applied in automatic fear level detection and acrophobia virtual therapy. *Sensors*, 2020, 20(2): 496. [doi: [10.3390/s20020496](https://doi.org/10.3390/s20020496)]
- Kaur R, Korolkov M, Hernandez ME, et al. Automatic identification of brain independent components in electroencephalography data collected while standing in a virtually immersive environment—A deep learning-based approach. *Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Paris: IEEE, 2020. 1–4.

- Conference of the IEEE Engineering in Medicine & Biology Society. Montreal: IEEE, 2020. 95–98.
- 15 Zheng RZ, Wang TL, Cao JW, et al. Multi-modal physiological signals based fear of heights analysis in virtual reality scenes. *Biomedical Signal Processing and Control*, 2021, 70: 102988. [doi: [10.1016/j.bspc.2021.102988](https://doi.org/10.1016/j.bspc.2021.102988)]
- 16 Hu F, Wang H, Chen JC, et al. Research on the characteristics of acrophobia in virtual altitude environment. *Proceedings of the 2018 IEEE International Conference on Intelligence and Safety for Robotics*. Shenyang: IEEE, 2018. 238–243. [doi: [10.1109/IISR.2018.8535774](https://doi.org/10.1109/IISR.2018.8535774)]
- 17 Němcová A, Svozilová V, Bucsu házy K, et al. Multimodal features for detection of driver stress and fatigue: Review. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(6): 3214–3233. [doi: [10.1109/TITS.2020.2977762](https://doi.org/10.1109/TITS.2020.2977762)]
- 18 权学良, 曾志刚, 蒋建华, 等. 基于生理信号的情感计算研究综述. *自动化学报*, 2021, 47(8): 1769–1784. [doi: [10.16383/j.aas.c200783](https://doi.org/10.16383/j.aas.c200783)]
- 19 Shu L, Xie JY, Yang MY, et al. A review of emotion recognition using physiological signals. *Sensors*, 2018, 18(7): 2074. [doi: [10.3390/s18072074](https://doi.org/10.3390/s18072074)]
- 20 Bota PJ, Wang C, Fred ALN, et al. A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals. *IEEE Access*, 2019, 7: 140990–141020. [doi: [10.1109/ACCESS.2019.2944001](https://doi.org/10.1109/ACCESS.2019.2944001)]
- 21 Yan MS, Deng Z, He BW, et al. Emotion classification with multichannel physiological signals using hybrid feature and adaptive decision fusion. *Biomedical Signal Processing and Control*, 2022, 71: 103235. [doi: [10.1016/j.bspc.2021.103235](https://doi.org/10.1016/j.bspc.2021.103235)]
- 22 Albraikan A, Tobón DP, El Saddik A. Toward user-independent emotion recognition using physiological signals. *IEEE Sensors Journal*, 2019, 19(19): 8402–8412. [doi: [10.1109/JSEN.2018.2867221](https://doi.org/10.1109/JSEN.2018.2867221)]
- 23 Li X, Yan JZ, Chen JH. Channel division based multiple classifiers fusion for emotion recognition using EEG signals. *ITM Web of Conferences*, 2017, 11: 07006. [doi: [10.1051/itmconf/20171107006](https://doi.org/10.1051/itmconf/20171107006)]
- 24 Khezri M, Firoozabadi M, Sharafat AR. 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. [doi: [10.1016/j.cmpb.2015.07.006](https://doi.org/10.1016/j.cmpb.2015.07.006)]
- 25 潘家辉, 何志鹏, 李自娜, 等. 多模态情绪识别研究综述. *智能系统学报*, 2020, 15(4): 633–645.
- 26 Wei W, Jia QX, Feng YL, et al. Emotion recognition based on weighted fusion strategy of multichannel physiological signals. *Computational Intelligence and Neuroscience*, 2018, 2018: 5296523. [doi: [10.1155/2018/5296523](https://doi.org/10.1155/2018/5296523)]
- 27 Šalkevičius J, Damaševičius R, Maskeliunas R, et al. Anxiety level recognition for virtual reality therapy system using physiological signals. *Electronics*, 2019, 8(9): 1039. [doi: [10.3390/electronics8091039](https://doi.org/10.3390/electronics8091039)]
- 28 Cui Y, Luo SH, Tian Q, et al. Mutual information-based emotion recognition. New York: Springer, 2013. 471–479.
- 29 Cortes C, Vapnik V. Support-vector networks. *Machine Learning*, 1995, 20(3): 273–297.
- 30 Reshef DN, Reshef YA, Finucane HK. Detecting novel associations in large data sets. *Science*, 2011, 334(6062): 1518–1524. [doi: [10.1126/science.1205438](https://doi.org/10.1126/science.1205438)]
- 31 Zhang X, Wen WH, Liu CY, et al. Recognition of public speaking anxiety on the recurrence quantification analysis of GSR signals. *Proceedings of the 6th International Conference on Information Science and Technology*. Dalian: IEEE, 2016. 533–538.
- 32 Sepúlveda A, Castillo F, Palma C, et al. Emotion recognition from ECG signals using wavelet scattering and machine learning. *Applied Sciences*, 2021, 11(11): 4945. [doi: [10.3390/app11114945](https://doi.org/10.3390/app11114945)]
- 33 Schmidt P, Reiss A, Duerichen R, et al. Introducing WESAD, a multimodal dataset for wearable stress and affect detection. *Proceedings of the 20th ACM International Conference on Multimodal Interaction*. Boulder: ACM, 2018. 400–408.
- 34 Siirtola P. Continuous stress detection using the sensors of commercial smartwatch. *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. London: ACM, 2019. 1198–1201. [doi: [10.1145/3341162.3344831](https://doi.org/10.1145/3341162.3344831)]
- 35 Alshamrani M. Remote intelligent healthcare system based on the ROCKET technique. *Arabian Journal for Science and Engineering*, 2021, 46(9): 9263–9277. [doi: [10.1007/s13369-021-05805-2](https://doi.org/10.1007/s13369-021-05805-2)]
- 36 Huynh L, Nguyen T, Nguyen T, et al. StressNAS: Affect state and stress detection using neural architecture search. *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. ACM, 2021. 121–125.
- 37 Lin JH, Pan SR, LEE CS, et al. An explainable deep fusion network for affect recognition using physiological signals. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. Beijing: ACM, 2019. 2069–2072.

(校对责编: 牛欣悦)