**基于VR与EEG的情绪识别系统设计与实现**

摘要：

随着虚拟现实技术的不断发展，基于虚拟现实的应用也越来越多；随着脑电信号采集设备的迅速发展，脑电信号越来越多的应用于实际生活；随着社会经济的快速发展，人们面临的情绪、心理问题也越来越频繁。本文对如何通过脑电信号识别虚拟现实设备佩戴者的情绪状态进行了研究，并设计了基于虚拟现实与脑电信号的情绪识别系统。本文对现有的脑电设备进行了改进，并设计相应的虚拟现实场景对被试进行刺激，使其产生平静—兴奋的情绪状态，并采集脑电信号对这两种情绪状态进行识别。本文设计了通过脑电识别平静—兴奋情绪的有效算法，对脑电信号进行预处理去除伪迹，然后采用小波变换将脑电信号分解为五种常见频带的节律信号，分别对每种节律信号提取频带能量、时间序列统计特征、信息熵等多种特征，之后通过特征选择与降维对特征进行处理，通过多种分类算法进行模型融合的方式提高识别准确率。最终，在对10人的80次实验中，本文的情绪识别系统对情绪的识别达到平均80.2%的识别率。

关键词：虚拟现实 脑电信号 小波变换 频带能量 模型融合

目录

[第一章 绪论 5](#_Toc467942694)

[1.1研究背景 5](#_Toc467942695)

[1.2研究现状 5](#_Toc467942696)

[1.3研究目的与意义 8](#_Toc467942697)

[1.3.1 脑电检测设备对心理研究的意义 8](#_Toc467942698)

[1.3.2 脑电检测设备对心理问题治疗的意义 8](#_Toc467942699)

[1.3.3 结合VR的情绪识别的商业价值 9](#_Toc467942700)

[1.4课题来源与研究内容 9](#_Toc467942701)

[第二章 相关研究 11](#_Toc467942702)

[2.1 VR技术简介 11](#_Toc467942703)

[2.2情绪相关研究 13](#_Toc467942704)

[2.2.1 情绪的产生 13](#_Toc467942705)

[2.2.2 情绪的分类 14](#_Toc467942706)

[2.2.3 情绪的诱发 17](#_Toc467942707)

[2.3 大脑的功能结构与脑电的产生 20](#_Toc467942708)

[2.4 脑电信号的预处理 22](#_Toc467942709)

[2.5 脑电信号的特征提取 23](#_Toc467942710)

[2.5.1 脑电信号的时域特征 23](#_Toc467942711)

[2.5.2 脑电信号的频域特征 23](#_Toc467942712)

[2.5.3 脑电信号的时-频特征 25](#_Toc467942713)

[2.6 情绪识别方法 26](#_Toc467942714)

[2.7本章小结 27](#_Toc467942715)

[第三章 实验设计 28](#_Toc467942716)

[3.1 实验目的 28](#_Toc467942717)

[3.2 刺激材料 28](#_Toc467942718)

[3.3 被试选择 29](#_Toc467942719)

[3.4 实验环境 29](#_Toc467942720)

[3.5 实验流程 30](#_Toc467942721)

[3.6 本章小结 33](#_Toc467942722)

[第四章 情绪识别系统设计 34](#_Toc467942723)

[4.1 系统整体框图 34](#_Toc467942724)

[4.2 硬件设计 35](#_Toc467942725)

[4.2.1 脑电采集设备 35](#_Toc467942726)

[4.2.2 VR情绪刺激设备 36](#_Toc467942727)

[4.2.3 信号同步设计 37](#_Toc467942728)

[4.3 虚拟现实软件设计 38](#_Toc467942729)

[4.4 本章小结 40](#_Toc467942730)

[第五章 情绪识别算法设计 41](#_Toc467942731)

[5.1 情绪识别研究 41](#_Toc467942732)

[5.1.1 数据预处理 41](#_Toc467942733)

[5.1.2 特征提取 43](#_Toc467942734)

[5.1.3 特征选择与降维 51](#_Toc467942735)

[5.1.4 情绪识别算法 54](#_Toc467942736)

[5.2 情绪识别分析 57](#_Toc467942737)

[5.2.1信号分解效果分析 57](#_Toc467942738)

[5.2.2特征选择效果分析 58](#_Toc467942739)

[5.2.3模型融合效果分析 58](#_Toc467942740)

[5.3 识别软件整体效果 59](#_Toc467942741)

[5.4 本章小结 60](#_Toc467942742)

[第六章 总结与展望 61](#_Toc467942743)

[6.1 总结 61](#_Toc467942744)

[6.2 展望 61](#_Toc467942745)

# 第一章 绪论

## 1.1研究背景

情绪是日常屡见不鲜并亲身体验着的一种心理活动，它给人们带来快乐和满足，又使人不可避免地遭受苦恼和折磨。

近年来，随着社会经济飞速发展，越来越多的人们正在经受前所未有的情绪问题。在工作中，许多员工不小心进入情绪的漩涡；在生活中，太多的琐事引发情绪失控；在学习上，学生的情绪问题层出不穷，有些已酿成悲剧。长期处于消极的情绪状态会引起更加复杂的心理状态，比如抑郁症等，更有甚者会引起生理问题的出现。

对于情绪的研究已经有一个多世纪，不同学派使用的方法也不尽相同。近年来，随着感知技术和脑电信号采集设备的广泛应用，信号处理技术与机器学习技术的快速发展，计算机数据处理能力的大幅提升，虚拟现实技术的涌现，基于脑电的情绪识别研究已经成为热门课题。

在利用脑电进行情绪研究，如何获得被试在不同情绪下的脑电信号这一环节很重要，目前常用的手段包括：自发回忆、图片刺激、音乐刺激、视频刺激、想象刺激等，但目前的情绪诱导方法存在着诱发效果差、易受干扰等缺点。本文设计一种用虚拟现实技术作为诱发手段的基于脑电信号情绪识别系统，并对这种设备的有效性进行验证。

## 1.2研究现状

人们对情绪识别的研究已经有很长的历史了，使用的方法与手段也不尽相同。在心理学上情绪的研究是通过访谈、表情和情绪体验自我报告法（Self-Assessment Manikin, SAM）来进行，具有一定的客观性。随着信息处理技术的发展，越来越多的研究者通过对不同情绪状态下的生理、心理及行为参数进行研究，通过提取相应的特征使用分类识别技术以识别情绪状态。其中基于自主神经系统的情绪识别采样皮肤温度、血压、血容量、心率、呼吸、肌电、心电等自主心理信号进行情绪识别。Kim 【Emotion Recognition Based on Physiological Changes in Music Listening】等人将呼吸、心电、肌电和皮肤电等多种自主心理信号融合，实现对以音乐为诱发素材的4种情绪进行识别，达到了90%以上的分类正确率。同样基于呼吸、肌电、皮肤电、心电四种生理信号，程德福等人用双重结构粒子群和KNN相结合的算法识别了高兴、愤怒、悲伤、愉悦四种情绪状态，达到93%的高识别率【程德福，刘光远，邱玉辉.双重结构粒子群和KNN在生理信号情感识别中的应用】。

伴随着医学图像技术的发展，基于中枢神经系统信号来识别情绪称为新的研究方向。基于中枢神经系统的情绪识别研究方法，是指通过分析不同情绪状态下大脑发出的信号差异来对情绪进行识别。大脑信号相对于其他自主神经信号更容易反应出大脑的情绪状态，有着较高的识别率，因此被广泛应用于情绪识别的研究中。比较常见的基于中枢神经系统的识别方法主要包括：功能核磁共振（fMRI）、脑电波（EEG）、功能性红外光谱成像（fNIRS）。其中基于脑电的情绪识别是最简单易行也最早进行研究。

利用脑电来进行情绪识别，国内外已经有人进行了大量的相关工作与研究，从不同的情绪维度取得了比较好的识别率。

Sammle等人通过音乐诱发情绪，并对脑电信号研究分析发现，欢快的音乐能在额中区引起更多的δ波【Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music】。

赖永秀在等人2008年发现了左右额叶α波与正负情绪的相关性【音乐情绪感知的脑电研究】。

Kristina Schaaff等人通过在人脑前额区采集脑电信号来识别情绪状态时，90张国际情绪图片系统（International Affective Picture System，IAPS）图片作为刺激素材，使被试产生高兴、中性、不高兴三种情绪状态，最终结果显示识别率达到 47.11%，他的工作让人机交互下的情绪识别成为一种新的趋势【Towards an EEG-based Emotion Recognizer for Humanoid Robots】。

P.C. Petrantonakis等人多次在脑前额区（FP1、FP2、F3/F4）采集脑电信号，研究基于脑电信号的情绪识别研究，他们提出了基于混合滤波和高阶交叉的新方法，对高兴、惊讶、气愤、恐惧、厌恶、悲伤等六种情绪类别进行识别分析，识别率达到了84.72%【Eeg-based emotion recognition using hybrid filtering and higher order crossings】。

上海交通大学吕宝粮教授的研究团队是国内通过脑电信号对情绪研究情绪做的比较好的团队之一。他指导的郑伟龙等人将深度信念网络（deep belief network，DBN）应用于基于脑电的情绪分类，取得了87%的识别率【EEG-BASED EMOTION CLASSIFICATION USING DEEP BELIEF NETWORKS】。聂聃、王晓韡等人用视频作为诱发素材通过脑电识别消极与积极情绪，获得了89%的情绪识别率【EEG-based Emotion Recognition during Watching Movies】。

在用脑电做情绪研究的过程中，如何使被试产生不同的情绪状态是很重要的一环。目前常用的方法就是在实验过程中通过各种刺激素材引起被试产生不得情绪，同时记录脑电信号。

台湾大学的Lin等人利用奥斯卡影片的原声音乐片段作为刺激素材，刺激被试产生喜怒哀乐四种情绪，并利用脑电对这四种信号进行分类【EEG-based emotion recognition in music listening[J]】。

Savran 等人通过使用IAPS作为刺激素材，刺激被试产生正向、负向、中性的情绪，构建了包含脑电信号、外周生理信号、功能性近红外光谱以及脸部视频的情绪数据库“2005 emotional database”【Emotion detection in the Loop from Brain Signals and Facial Images】。

吕宝粮等人选择热门影片作为刺激素材，使被试产生积极、消极、中性情绪，并构建数据库【http://bcmi.sjtu.edu.cn/~seed/】。并在此基础上，尝试多种数据预处理算法，提取多种特征，使用不同的算法，获得了不错的情绪识别效果。

在用脑电做情绪识别的过程中，使用什么样的特征也是很重要的一步。国内外研究者在这一问题上已经做过大量的尝试，找出了很多适合情绪识别的脑电特征。

Picard 等人利用生理信号的统计学特征（均值、标准差、一阶差绝对值均值、二阶差绝对值均值、正规化的一阶差绝对值均值、正规化的二阶差绝对值均值）以及统计学特征的变换和组合作为特征，对情绪进行了分类【Toward machine emotional intelligence: Analysis of aﬀective physiological state】。

Lin 等人对脑电信号进行短时傅里叶变换，将每个电极的脑电信号在γ、β、α、θ、δ 五个频段上的能量谱作为特征，得到了不错的分类效果【EEG-based emotion recognition in music listening】。

Duan 等人提取脑电信号的微分熵特征，在情绪识别中，他们发现微分熵特征对于正向、负向情绪的识别具有较高的辨识度和准确率【Diﬀerential entropy feature for EEG-based emotion classiﬁcation】。

Huang 等人根据被试在不同的情绪状态下左右脑脑电信号的差异，提出了非对称空间滤波特征（asymmetric spatial ﬁltering，ASF)，对情绪进行分类，取得了较好的效果【Fast emotion detection from EEG using asymmetric spatial ﬁltering】。Duan在分类情绪愉快-平静-悲伤时也用到了类似的特征【基于脑电信号的视频诱发情绪识别】。

然而，目前的研究多处于实验研究阶段，面向实用的基于脑电信号的情绪识别并没有太多的实践。而本文是将基于脑电的情绪识别应用于实际工程的一次尝试，同时为接下来的实验室研究基于VR与脑电的抑郁症病人的识别打下基础。

## 1.3研究目的与意义

### 1.3.1 脑电检测设备对心理研究的意义

脑电(Electroencephalograph, EEG)是目前心理学研究中广泛采用的非入侵式记录系统。半个世纪以来，脑电被广泛应用于脑功能研究，在心理学、生理学、认知神经科学、医学临床应用等领域取得了巨大成就，被誉为“脑功能的窗口”，有很高的研究与应用价值。脑电可用于研究注意力、感知觉、记忆、语言、运动、决策和判断等信息加工过程。便携式脑电设备在数据精度要求不高情况下能为科研人员带来很多便利，同时基于VR的特定场景刺激下脑电信号也为心理研究人员带来新的实验空间。

### 1.3.2 脑电检测设备对心理问题治疗的意义

通过脑电信号获取人们的情绪状态，进一步判断心理状态，这对研究心理问题有着很大的帮助。结合VR的脑电检测设备，利用VR场景的现场感、沉浸感刺激被试产生特定的反应，通过脑电信号判断被试的状态，再利用VR的现场感、沉浸感来改善用户的状态。这种思想在情绪改善、缓解压力、减少恐惧、抑郁症治疗、老年痴呆记忆力恢复等领域有着广泛的应用前景。

### 1.3.3 结合VR的情绪识别的商业价值

目前，基于VR的各项服务已经如雨后春笋般涌现。基于VR的情绪识别技术能够为佩戴VR设备的用户提供更加个性化更加贴心的服务。

VR方向的广告投放为情绪识别带来一片新的天地，头显中嵌入的传感器可以随时对使用者的情绪进行解读，从而展现合适的推荐内容。在疲倦和忧伤之余，甜品和饮品的推送会显得更有接受度。基于VR世界的零售商店，将有机会根据顾客的情绪反馈来对特定的产品进行促销，以便提供更为人性化的服务。也许阿里巴巴的“BUY+”计划在不久就会用上该项技术。

VR游戏是VR应用领域的一大热门，将情绪识别技术应用于VR头显设备，游戏制作人无需再费尽心力的揣测玩家的想法，通过数据的收集，读取玩家的情感状态，分析玩家对于游戏的哪一部分最感兴趣，从而相应的调节平衡，设置内购激励措施，或者增添非玩家控制角色（Non-Player-Controlled Character，NPC）的交互功能，最大限度的提高整体留存率。

其他服务，如银行、房地产等行业也可以有类似的应用。

## 1.4课题来源与研究内容

本课题依托广州市人体数据科学研究中心申请的科技项目《基于虚拟现实的脑功能监护设备及其在抑郁症辅助治疗中的应用》，本文设计的基于VR的脑电采集系统即是脑功能监护设备的重要组成部分，以该采集系统为基础，用VR刺激素材刺激被试产生兴奋—平静的情绪，同时记录脑电信号作为训练数据。然后设计分类算法根据脑电信号判断被试是否处于兴奋状态。修改特定的虚拟现实刺激场景与调整分类算法，该设备后续将应用于抑郁症识别与辅助治疗的研究中。

本文分为六部分，各部分内容如下：

第一章，绪论。介绍本文的研究背景和研究现状以及研究基于VR的情绪识别的意义。

第二章，相关研究。首先简单介绍了VR技术的发展历史，即VR技术发展的三次热潮。然后介绍了情绪是如何产生的，情绪的分类模型以及情绪的诱发方法，得出基于VR刺激的情绪诱发方式是一种较有效较新的情绪诱发方法。之后介绍了大脑的结构与脑电是如何产生的，脑电能反应人类各种心理情绪状态。之后介绍了脑电信号的预处理方法，脑电信号的特征提取。最后介绍情绪的识别方法，并得出基于脑电的情绪识别方法简单易行准确率高等优点。

第三章，实验设计。介绍了实验方法，包括被试的选择、刺激素材的挑选、实验环境介绍、实验步骤介绍等，一个完整的实验设计从情绪刺激到脑电采集以及被试自我评价都在这一章节呈现。

第四章，情绪识别系统设计。这一章主要从情绪系统的整体框架出发介绍系统的组成，包括硬件设计与软件设计。硬件设计中主要完成脑电采集设备和VR设备的信号同步功能，软件设计中主要介绍测试软件设计和虚拟现实场景制作。

第五章，情绪识别算法设计。作为情绪识别系统的重要组成部分，情绪识别算法包括数据预处理，特征提取，特征选择与降维，以及情绪识别算法设计。同时，在这一章对该情绪识别算法的效果做了验证分析，得到了不错的情绪识别效果。

第六章，总结与展望。针对该识别系统，对所做工作进行总结，分析工作中的不足并提出改进方案，以及对下一步研究进行展望。

# 第二章 相关研究

## 2.1 VR技术简介

虚拟现实技术（Virtual Reality，VR），指采用计算机技术为核心的现代高科技手段生成一种虚拟环境，用户借助特殊的输入/输出设备，与虚拟世界进行自然的交互，提供用户关于视觉、听觉、触觉等感官的模拟，让用户如同身历其境一般。人们对于真实的认知建立在人类感官的基础上，即便纯粹抽象理念上的推演，也无法脱离大脑这一生理结构本身的局限性。从认知科学角度讲，身体归属感（Bodily Ownership）、涉入感（Sense of Agency）以及（身体随处）态势感知（Situation Awareness）都是自我意识的重要组成部分。与其他智能终端技术相比，VR技术的优势在于自然的交互（从设备单一操作平面的束缚解脱出来，以自然的身体动作与语音进行交互），与具有临场感（Pre-sense）的虚拟环境（从智能终端窗口影像提供的旁观者感受，演变为VR环境营造置身其中的参与者体验）。

虽然VR仍在商业化应用的路途当中，但其在人们进入信息虚拟世界的过程真正从“以机器（设备）为中心”转向“以用户（人）为中心”，具有极佳的认知体验，因此我们认为VR技术将在三个层面具有巨大需求：在追求极致体验的游戏与视频领域，临在感优势将使VR迎来第一轮爆发性巨大需求；而随着多学科交叉领域研究的推进，大部分基于空间与位置稀缺性的商业模式（如运动、房地产销售、教育、医疗、演唱会等）将被VR技术颠覆，VR技术进入到生活的各个领域；在VR产品大规模生产与应用得以实现后，将如同多点触摸屏的易用性推动智能手机无所不在一样，VR技术也有潜力达到同样的普及程度，成为下一个重大通用计算平台。

VR的发展经历了三次热潮：第一次源于1960年代，确立了VR技术原理；第二次发生在1990年代，VR试图商业化但未能成功；目前正处于第三次热潮前期，技术层面比之前有质的飞越，显示设备、显卡性能都有较大幅度的提升，物理追踪技术更为精确，体感控制、手势识别、语音控制等交互方式更加丰富。

第一次热潮发生在1960年代，科学家们建立了VR的基础原理和产品光学构造。1960年，电影摄影师Morton Heilig提交了一款VR设备的专利申请文件，专利文件上的描述是“用于个人使用的立体电视设备”。1967年，Heilig又构造了一个多感知仿环境的虚拟现实系统Sensorama Simulator，这也是历史上第一套VR系统，它能够提供真实的3D体验，例如用户在观看摩托车形式的画面时，不仅能看到立体、彩色、变化的街道画面，还能听到立体声，感受到行车的颠簸、扑面而来的风还能闻到花的芳香。1968年美国计算机图形学之父Ivan Sutherlan在哈佛大学组织开发了第一个计算机图形驱动的头盔显示器HMD及头部位置跟踪系统，是VR发展史上一个重要的里程碑。进入80年代，VR相关技术在飞行、航天等领域得到比较广泛的应用。

第二次热潮发生在1990年代，这是一次如火如荼的商业化热潮，但最终没能获得成功。1989年Jaron Lanier首次提出Virtual Reality的概念，被称为“虚拟现实之父”。1991年，一款名为“Virtuality 1000CS”的设备出现在消费市场中，由于它笨重的外形、单一的功能和昂贵的价格，并未得到消费者的认可；但掀起了一个VR商业化的浪潮，众多公司推出了自己的VR产品。但这一轮商业化热潮，由于光学、计算机、图形、数据等领域技术尚处于高速发展早期、产业链也不完备，并未得到消费者的积极响应。但此后，企业的VR商业化尝试一直没有停止。

第三次热潮源于2014年Facebook 20亿美元收购Oculus，VR商业化进程在全球范围内得到加速，各种VR产品在市场上铺开。各项技术的突破，使得VR设备得以普及，开始在游戏、影视、直播、教育、医疗等领域逐步得到应用。而本文是将VR在科学研究与心理情绪研究相结合的一次尝试。

由于技术局限、用户体验、内容和应用开发以及价格问题，导致VR技术在50余年的发展与变革中一直未真正得到应用。而近两年VR行业炙手可热，VR在市场的推动下VR技术也得到了飞速的发展，越来越多的软件、硬件厂家不断加入进来。Oculus Rift、HTC VIVE、Sony PS等热门VR产品开始在业内流行开来，而本文用于情绪刺激的VR设备就是HTC VIVE，一款不错的VR设备。

## 2.2情绪相关研究

### 2.2.1 情绪的产生

情绪是如何产生的，在认知科学、神经科学领域经过漫长的研究，虽然目前并没有完全的定论，但也取得了令人比较满意的成果。

达尔文的进化观点对情绪的研究有启蒙作用，他认为情绪作为人类种族进化的证据，可能是人类行为得以延续的机制【情绪心理学】。基于达尔文进化论的影响和生物科学的发展看，美国心理学家James和丹麦生理学家Lange分别于1884年和1885年提出相同的情绪心理学说，后来被称为James-Lange情绪外周学说，它强调情绪是对身体变化的感觉【情绪心理学】。而Cannon反对James的理论，他的丘脑学说指出大脑皮层接收外界刺激并激活丘脑，并由此产生不同的情绪。Cannon的理论肯定了丘脑在情绪产生过程中的作用，但同样过于片面而全面否定了外周生理与情绪产生之间的关系。

在Cannon以后，情绪生理学的研究继续进行。精神分析学派弗洛伊德把情绪看做能够释放的过程，也承认情绪活动必须伴随有意识的体验【情绪心理学】。新精神分析学派接受弗洛伊德的情绪是能量释放、无意识和内驱力等观点，同时把情绪放在更大的心理环境中考虑。认为能量的释放，冲动，动机，知觉，认知，意识和无意识均参与情绪的形成，认为情绪本身就是一种行为动力【情绪心理学】。在这些理论的基础上，Pepez在1937年提出了情绪的“Pepez环路”理论，之后Maclean在这个环路上附加一些合团，命名为“边缘系统”，它包括皮层和皮层下的结构，扣带回、海马皮层、丘脑和下丘脑。在相当长的一段时期里，边缘系统在情绪脑机制的解释上占统治地位，然而上世纪80年代以来，这个概念由于在结构和功能上的不精确而被质疑。现在，情绪的机构定位，从下丘脑延伸到边缘系统和整个中枢神经系统各水平结构，从新皮质前额叶皮层到脊髓均包括在内。而原来的边缘系统，如海马和乳头体已被证明对认知比对情绪过程更重要。可是，杏仁核，作为边缘系统的一部分，在许多情况中被牵涉到情绪加工中。边缘系统作为情绪的机构定位存在如此之久就是由于杏仁核所起的核心作用，然而，边缘系统作为一个神经解剖概念，使用起来有缺陷，很难从解剖学的基础上来确定，而且，情绪与调节它的神经结构最终能被功能性神经行为研究证实。情绪脑的主要结构涉及杏仁核和以杏仁核为核心的广泛连接的神经环路：前额叶皮层，包括眶额回皮层；扣带回皮层，特别是前扣带回皮层；下丘脑、杏仁核；腹侧黑质、隔区和中脑边缘核团等部位【情绪心理学】。影响情绪的关键脑区示意图如图2-1所示。

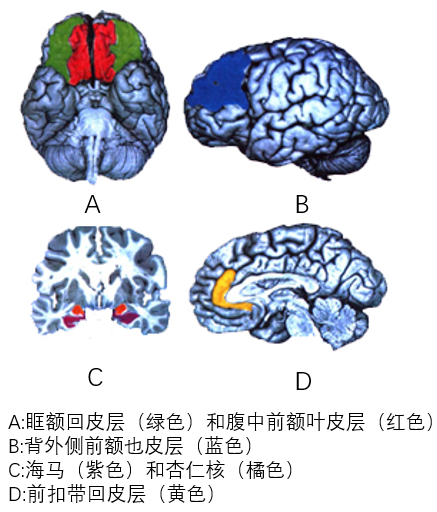


图2-1 影响情绪的关键脑区【情绪心理学】

人们对情绪的研究还在继续，但神经解剖学和认知科学的研究表明，情绪的产生与生理活动，大脑皮层，杏仁核有着密切的关系。这为研究大脑皮层的活动分析和识别人的情绪状态提供了理论依据。

### 2.2.2 情绪的分类

目前，学术界对情绪并没有统一的定义，学术界对情绪模型也存在很大的争议。但有个普遍的认同就是情绪与生理上的反应有一定的映射关系，这为利用生理信号进行情绪分类提供了理论依据。

情绪模型是在情绪识别前需要搞清楚的。一般的情绪分类主要有两大基本观点：离散模式和连续模式。离散模式认为情绪具有完全不同的结构，连续模式认为不同情绪之间有着过渡阶段。而离散模型在研究中被广泛使用。早在中国的《黄帝内经》中将情绪分为“喜怒忧思悲恐惊”，Ekman受达尔文进化论启发而形成的离散情绪模型将情绪分为愉悦、惊讶、愤怒、厌恶、悲伤和恐惧六种情绪状态。

很多时候，情绪是几种基本情绪混合而成的综合感受，这就产生了概率化、多标签、模糊化的情绪模型，即多维连续模型【视频诱发情绪】。最常见的维度分类方法是James-Lange维度情绪分类模型，如图2-2所示，该模型以James-Lange情绪外周学说为基础指导，其影响广泛并获得了普遍的认可。James-Lange理论将人类的情绪划分在维度空间中的矢量点上，不同情绪的差别由矢量空间中对应点之间的距离度量，相似的情绪状态在空间中的距离也相近，不同情绪的转变可以视为一种渐变的转换。如图，横坐标表示心情的愉悦程度，从悲伤到愉快过渡；纵坐标用来表示兴奋程度，从无聊逐渐过渡到兴奋状态。这样，不同的情绪就分解到二维坐标系中。本文以James-Lange二维情绪模型为基础，针对情绪的平静—兴奋维度进行情绪的识别，期望通过脑电信号能对情绪有很好的识别效果。



图2-2 James-Lange二维情绪模型

Mehrabian和Russell【Albert M. Basic Dimensions for a General Psychological Theory: Implications for Personality, Social, Environmental, and Developmental Studies [M]. Cambridge: Oelgeschlager, Gunn & Hain, 1980】提出以PDA情感模型来描述和测量情绪状态，P代表愉悦度（Pleasure），表示个体情绪状态的正负特性；Ａ代表兴奋度（Arousal），表示个体的生理激活水平；Ｄ代表支配度（Dominance）,表示个体对情景和他人的控制状态。

PAD将情绪通过三个维度进行表达：首先是两个基本的维度：愉悦度用以描述情绪从愉快到不愉快的感觉，分类为正、负两极，位于正极的称积极情绪，通常带来愉悦感受（如欢乐、幸福等）；位于负极的称消极情绪，通常产生不愉悦感受（如愤怒、害怕等）。唤醒度是指激活与情感状态相联系的机体能量的程度，描述情绪从兴奋到平静的感觉，唤醒度越大，所产生的情绪就越强烈。再加上与强度相关的维度“支配度”，代表情绪的控制和主导性，例如愤怒和害怕都属于不愉悦的情绪，其中愤怒是支配性的情绪，而害怕是服从性的情绪。

SAM是佛罗里达大学情绪和注意研究中心教授Bradley和Lang设计的一种测量顾客情绪反应的情绪自我评价等级系统，以Mehrabian和Russel的情绪PAD 维度模型（愉悦度、唤醒度、支配度）为基础【An Approach to Environmental Psychology】。SAM通过卡通人物的抽象绘图的形式阐明愉悦度、唤醒度、支配度三个维度，如图 2-3所示。其中，微笑的图像到皱眉图像代表愉悦度；从兴奋地、睁大眼睛的图像到放松的、欲睡的图像表示唤醒度；支配度通过 SAM 尺寸的大小表示控制方面的改变（大的图像表示对现状最大限度的控制），被试需要表明哪个人物形象更能代表他们的情绪状态。



图2-3 SAM情绪自我评价等级系统

本文针对脑电信号在平静—兴奋维度上进行情绪分类，以唤醒度作为分类标准。

### 2.2.3 情绪的诱发

情绪研究的重要前提条件之一就是诱发人的不同情感，情绪材料诱发即向被试呈现具有情绪色彩的材料, 从而诱发被试相应情绪的方法。根据材料呈现感觉通道的不同, 可以将其分为视觉刺激材料, 听觉刺激材料和嗅觉刺激材料【情绪诱发方法述评】。

视觉刺激：视觉刺激是比较常用的情绪诱发方法, 即给被试呈现具有情绪色彩的文字、图片等刺激材料,以此来诱发被试的目标情绪。目前, 视觉刺激已

经形成了较为完善的标准刺激材料库, 在文字方面, 美国国立精神卫生研究所(National Instituteof Mental Health, NIMH)推出的英语情感词系统(Affective Norms for English Words, ANEW, 1999a)和英文情感短文系统(Affective Norms for English Text, ANET, 2007)都是得到广泛认可的文字情绪刺激材料库(Kousta, Vigliocco, Vinson, Andrew, &Del Campo, 2011; Kousta, Vinson, & Vigliocco,2009; Lang, 2010); 而在图片方面, NIMH 建立了国际情绪图片系统(International Affective Picture System, IAPS, 2008), 为情绪诱发研究提供了更多选择(Frantzidis et al., 2010)。

由于文字和图片刺激往往会受到文化背景的影响, 国内研究者在相关研究的基础上, 对国外的刺激材料进行了本土化修订和完善, 推出了汉语情感词系统(Chinese Affective Words System, CAWS)【汉语情感词系统的初步编制及评定】和中国情绪图片系统(Chinese Affective Picture System, CAPS)【中国情绪图片系统的编制——在46名中国大学生中的试用】。这些系统的开发, 为国内研究者提供了一系列情绪诱发的重要工具【情绪调节内隐和外显态度在青少年阶段的发展特点】 【负性情绪干扰行为抑制控制:一项事件相关电位研究】。

听觉刺激：研究发现, 自然界的声音录音、非言语音节以及音乐都可以作为情绪诱发的材料。例如，NIMH通过采集鸟叫、婴儿哭泣、炸弹爆炸、下雨等一系列声音, 对其愉悦度和唤醒度进行评定，建立了国际情感数码声音系统 (International Affective Digital Sounds, IADS, 1999b)，2007年又对其进行了修订, 推出了IADS2。这两个系统为研究听觉刺激对认知、情绪、行为的影响提供了标准化的工具, 被应用于大量实验研究中【Musical experience and neural efficiency–effects of training on subcortical processing of vocal expressions of emotion】 【When room size matters: acoustic influences on emotional responses to sounds.】。而国内研究者同样在大量收集各种声音的基础上建立了中国情感数码声音系统(Chinese Affective Digital Sounds, CADS)【本土化情绪声音库的编制和评定】。

随着音乐在消费者情绪行为控制、情绪紊乱的心理治疗、个体自我情绪调节【The influence of in-store music on wine selections.】等众多领域得到了越来越广泛的应用, 音乐情绪诱发也开始受到了心理学家的重视。经过十多年的积累, 一些音乐和情绪的对应关系逐渐达成共识, 例如, 巴赫的“勃兰登堡协奏曲”或贝多芬的“第六交响乐”通常能够诱发愉快情绪; 霍尔斯特的“火星：战争使者”能够诱发恐惧情绪; 而使用巴伯的“弦乐柔板”能够诱发出悲伤情绪【An exploratory study of musical emotions and psychophysiology.】【Music and emotion: perceptual determinants, immediacy, and isolation after brain damage】 【From emotion perception to emotion experience: Emotions evoked by pictures and classical music】等等。

虽然音乐作为优秀的情绪诱发方式，但目前还缺乏标准化的情绪诱发材料库。

嗅觉刺激：嗅觉是人类的另一种重要的感觉器官, 在嗅觉诱发情绪的研究中, 主试通常让被试有意或无意识地嗅闻某种气味, 以此达到情绪诱发的目的。研究发现, 嗅觉刺激和其他感觉通道的刺激一样, 能够诱发被试积极或消极的情绪, 进而对个体的认知、行为产生【Sex differences in human olfaction: between evidence and enigma】【Mall atmospherics: the interaction effects of the mall environment on shopping behavior】【Common and specific effects of fine fragrances on the mood of women】;而进一步研究表明, 阈下的嗅觉刺激也能够起到相同的作用(Walla, 2008)。此外, 研究发现, 气味还存在着联结诱发作用, 即被试往往会将特定的气味与闻到该气味时的情绪体验之 间产生联结(【Olfactory Comfort: Smelling a Partner's Clothing During Periods of Separation】【What Color is that Smell? Cross-Cultural Color-Odor Associations】【Garlic ingestion by pregnant women alters the odor of amniotic fluid】, 再次向其呈现该气味就能够诱发出相应的情绪。一些研究者利用这一特点, 设计了一系列情绪诱发实验, 收到了很好的诱发效果。例如, 在Herz, Schankler和Beland (2004)的研究中, 主试在被试(无论是成人还是儿童)遭遇挫折的同时让其闻到一种特别的气味, 在随后的无关任务中, 同样气味的再次出现成功诱发了被试的相应情绪, 并降低了其完成任务的动机。而类似的情况在大量研究中都有所报告【Basic emotions evoked by odorants: comparison between autonomic responses and self-evaluation】【Ambient odors associated to failure influence cognitive performance in children】【Sex differences in human olfaction: between evidence and enigma】。

与音乐诱发类似，嗅觉诱发目前尚未有标准化材料库，刺激材料往往很难得到。

多通道刺激：多通道刺激是指组合使用视觉、听觉、嗅觉等诱发材料的刺激，以达到更佳的情绪诱发效果。其中视频刺激就是融合视觉听觉的刺激。而本文提到的基于虚拟现实技术的情绪刺激则是融合视觉、听觉、触觉等多种感觉的刺激手段，同时虚拟现实技术的沉浸式体验让被试处于不被打扰的刺激环境中，情绪诱发效果更加明显。

本论文采用的就是具有多通道刺激的VR头显设备进行情绪诱发的，具有听觉视觉同时具有身体感觉的全方位沉浸式刺激。

## 2.3 大脑的功能结构与脑电的产生

大脑是中枢神经系统的最高级部位，包括左、右两个半球并由胼胝体相连。大脑半球被覆灰质，称大脑皮层，其下方为白质，称为髓质。髓质内的灰质核团为基底神经节，在大脑两半球间有巨束纤维相连。



图2-4 大脑结构示意图



图2-5 大脑皮层结构分布示意图【http://baike.baidu.com/view/179125.htm】

大脑皮层结构如图2-5所示，大脑各叶的位置、结构和主要功能如下：

额叶：位于中央沟以前，在中央沟和中央前沟之间为中央前回。在其前方有额上沟和额下沟，被两沟相间的是额上回、额中回和额下回。额下回的后部有外侧裂的升支和水平分支。主要负责高级的认知功能，学习功能、语言功能、抽象思维、情绪等。

顶叶：位于中央沟之后，中央后回位于中央沟与中央后沟之间，横行的顶间沟将顶叶分为顶上小叶和顶下小叶，顶下小叶又包括缘上回和角回。顶叶主要负责躯体感觉功能，空间信息处理功能，以及视觉信息和体感信息的整合。

颞叶：颞叶位于外侧裂下方，由颞上、中、下三条沟分为颞上回、颞中回、颞下回。颞叶主要负责听觉的处理，与长期记忆有关。

枕叶：枕叶位于枕顶沟和枕前切线之后，主要负责视觉信息处理。

脑电信号源于大脑内多个神经元的共同放电活动，通过大脑皮层传递到头皮表面。神经元包括细胞体、轴突和树突。其中细胞体负责整个细胞的营养代谢，是细胞存活的重要条件，其结构包括细胞膜、细胞质以及细胞核三部分。树突用于接收临近神经元发出的的信息，一般情况下树突具有多个较短的树突，每个树突都具有多个分支。轴突的主要作用是传递神经活动，每个神经元只有一个轴突。轴突末梢的突触与其他神经元树突通过神经递质传递信息，形成轴突-树突突触。

脑电实际上是一种生物放电现象，就是神经元在进行活动时随之产生的各种电位动作，如图2-6。大量的实验研究表明，神经元突触后电位是脑电信号的主要来源，构成了脑电波的大部分信号。神经元之间的信号传递是通过突触的连接来完成的，其传递的过程可以描述为：当神经元产生刺激冲动时，将通过突触从前一神经元传递给后一神经元。这种刺激的传递是通过突触内部释放的神经递质来实现的，神经递质的释放改变了突触后膜两端电位差，从而引起细胞间的局部电流，当这种电流累积到一定程度时便实现了信号的传递。根据神经元反馈回路学原理，脑电信号实际上是一种具有节律性的电信号。大量的神经元在互相连接时是一种闭合的来回形式，并且存在有正、负反馈机制。当回路中的某一个或多个神经元兴奋时，这种刺激将根据具体的连接状态形成抑制或者激励，从而完成大脑的特定功能。而在这个活动中，神经元的不断放电便形成了脑电信号，并且能够通过设备采集。

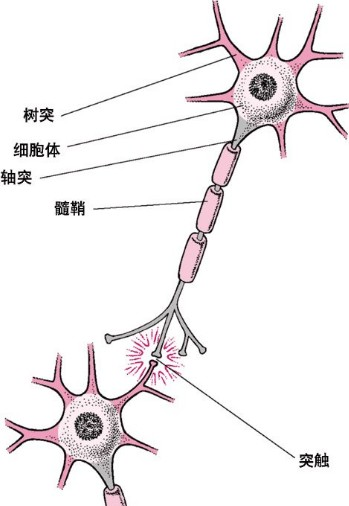


图2-6 神经元结构图【http://baike.sogou.com/v656225.htm】

根据脑电信号产生的条件可以分为自发脑电信号、诱发脑电信号两种。其中

自发脑电信号是大脑自身发生的一种有规律的电信号，其主要代表了大脑对身体

各个器官功能的控制用以维持自身的生存。而诱发脑电信号则是经由外界刺激产

生的，是大脑对外界环境信息进行感知、加工处理的一种表现。从信号特征来看

通常来讲自发脑电信号较诱发脑电信号波幅较大，不需要刺激，信号随时可以测

得，方便简单，但信号的处理相对较难【驾驶员对交通标志的视觉信息】。

## 2.4 脑电信号的预处理

脑电活动最早是在1875年被英国利物谱的内科医生Richard发现的，而其首次被记录则是1926年由奥地利精神病学家Hans Berger完成了对人脑EEG信号的采集【Dvorak. Taken versus multichannel reconstruction in EEG correlation exponent estimates [J]. Physics Letters A. 1990, 151(5):225-233】。

由于脑电信号十分微弱，因此在采集过程中很容易受到其他噪声信号的干 扰。脑电信号的预处理主要是指去除采集到的脑电信号中所掺杂的伪迹。在情绪识别研究中，所要去除的伪迹主要包括眼电、肌电、心电、工频干扰、电磁干扰和任务不相关的脑电等。目前比较常用的伪迹去除方法主要包括滤波和独立成分分析等。由于工频干扰和电磁干扰往往发生在高频段，因此可以通过带通滤波或低通滤波的方式将容易产生干扰的频段过滤掉，只保留有效的频段的脑 电信号。对于不易通过滤波去除的伪迹，通常采用独立成分分析、主成分分析等方法，找出干扰信号并将其与脑电信号分离开。Bartels等将盲信号分离、独立成分分析和支持向量机相结合，提出了一种有效的去伪迹处理方法。通过分析每个算法的特点，得出了用盲信号分离中的Amuse算法去除眼电，用独立成分分析中的Infomax算法去除肌电的预处理方法，并得到了较好的去伪迹效果【基于脑电的情绪识别综述】。【Automatic Artifact Removal from EEG - A Mixed Approach Based on Double Blind Source Separation and Support Vector Machine】

## 2.5 脑电信号的特征提取

基于脑电的情绪识别研究中，情绪提取环节至关重要，只有找出了与任务相关的代表性特征才能为后续的情绪识别提供保证。常见的脑电特征主要分为3类：时域特征、频域特征和时-频特征。

### 2.5.1 脑电信号的时域特征

时域特征往往把去除伪迹后的脑电信号在时域上的信息或将时域上的信号统计量作为特征。时域信号的均值、标准差、偏斜度、峰值、原始信号首次差异的平均绝对值、归一化后信号的首次差异的平均绝对值、振幅差等，可以作为时域特征。时域特征在情绪识别中的典型应用就是事件相关电位分析法。事件相关电位是指当外加一种特定刺激，作用于感觉系统或脑的某一部分，在给予或撤销某种刺激时，或当某种心理因素出现时，在脑区产生的电位变化【基于脑电的情绪识别综述】。在利用事件相关电位进行情绪识别研究中，通常是诱发情绪然后观察相关脑区的诱发电位变化。

### 2.5.2 脑电信号的频域特征

脑电信号的频率范围主要在0-100Hz，不同的文献对脑电波的频率范围规定不同，但共识是把脑电波分为五个频段的节律波，各节律波形如图2-7所示，每个频段的主要频率范围相差不大，按频率可以分为以下几段：



图2-7 脑电节律波形图

γ波(gamma)：30-60Hz，参与更高任务处理以及认知功能。γ波在学习、记忆和信息处理有重要作用，40 Hz的γ波被认为参与感知与学习新东西。人们已经发现，精神失常或者学习障碍时，γ波活动往往低于平均水平；焦虑、兴奋、压力时γ波活动处于高水平；比较合适的状态是在认知、信息处理、学习、感知、快速眼动睡眠时。

β波(beta)：12-30Hz，β波被称为高频低振幅的脑电波，当我们清醒的时候容易观察到。参与意识、逻辑思维、应急反应；适量的β波有助于集中精力、轻松完成实践任务。β波活动太强可能导致过度的压力和焦虑，β波频率越高表明兴奋性越高。当你喝咖啡因或兴奋剂，β波活动自然会增加。β波是比较容易出现的脑电波，大多数人每天的活动都会出现，比如：批判性思维、写作、阅读、和社会化。肾上腺素分泌、焦虑、高兴奋、压力等会导致β波活动增加。

α波(alpha)：8-12Hz，α波表征的是意识思维和潜意识之间的桥梁，换句话说，α波是β和θ之间的频率范围。它帮助我们在必要时冷静下来并促进深度放松的感觉。如果我们感到压力， 可能发生“阿尔法阻塞”的现象，这涉及到β波和少量α波，本质原因是在我们过于兴奋时，β波“块”的生产而引起的α波。在白日梦、放松是α波活动较高。

θ波(theta)：4-8Hz，做梦和睡眠时θ波比较强烈，θ波也关系到沉浸、专注的状态。太多θ活动可能使人们抑郁，也可能使他们注意力高度集中。θ有它的好处,帮助改善我们的直觉、创造力，而且可以使我们感觉更自然放松。它也参与恢复性睡眠。只要θ不产生超过在我们醒着的时间，这是一个非常有用的脑波范围。抑郁、冲动时θ波活动比较强烈。

δ波(delta)：0.5-4Hz，正常成人在清醒状态下，几乎是没有δ波的，但在睡眠期间可出现δ波。在婴儿时期，脑电频率比幼儿更慢，常可见到δ波。一般认为，高幅度的慢波（δ或θ波）可能是大脑皮层处于抑制状态时电活动的主要表现。

常见的频域特征有功率谱、功率谱密度、能量等。这些特征的提取一般都建立在功率谱估计的基础上。经典谱估计是以一定时间段数据的傅里叶变换作为基础的估计，通常通过直接法和间接法两种途径实现。直接法是将功率谱看作是幅频特性平方的总体均值与持续时间之间的比值来计算；间接法是先计算出相关函数，再进行傅里叶变换，从而得到所要求的功率谱估计。在此基础上就可以得到功率谱密度和能量等频域特征。

### 2.5.3 脑电信号的时-频特征

由于脑电信号的不稳定性，单纯考虑时域特征或频域特征都是不全面的，因此越来越多的研究开始将时域和频域联系起来，找出能够同时反映时域和频域的脑电特征，这里称其为时频特征。提取时频特征的常见方法，主要有短时傅里叶变换和小波变换等。

短时傅里叶变换是在传统傅里叶变换的基础上加入窗函数，通过窗函数的不断移动来决定时变信号局部弦波成分的频率和相位。Lin等以音乐作为刺激材料， 通过32导的电极帽采集情绪实验中被试的脑电信号，再利用短时傅里叶变换将脑电信号映射到上述5个常用频段上，之后分别计算出每个频段所对应的功率谱密度，最后利用电极间的对称关系组合出4组特征，分别是：对称电极间的差，对称电极间的商，除去中间电极后的电极功率谱密度，以及所有电极各自的功率谱密度。通过这4组特征进行后续的情绪识别。

小波变换是指用某些特殊函数作为基函数，并以此对数据进行变换操作，从而发现其类频谱特征的变换过程。与短时傅里叶变换一样，小波变换也将时域和频域联系起来，与短时傅里叶变换不同的是小波变换的窗函数可以随频率的改变而改变，从而能够给出更好的特征表示。Murugappan等用视频片段作为情绪实验的刺激材料，在经过预处理之后对脑电信号进行小波变换，从而得到一系列的小波系数，通过小波系数得到频段的能量。Murugappan等将子频段的能量在总体 能量中的比率，alpha频段小波系数的均方根及能量作为脑电特征。

## 2.6 情绪识别方法

通过计算机识别情绪的研究方法主要分为两大类：基于非生理信号的识别和基于生理信号的识别。

非生理信号的识别方法主要包括对面部表情、语音语调、动作和瞳孔来识别。面部表情的识别多通过图像识别的方法来实现【A Real-Time Automated System for The Recognition of Human Facial Expressions】。面部表情的识别方法是根据表情与情绪的对应关系来识别不同的情感，在特定情绪状态下人们会产生特定的面部肌肉运动和表情模式，比如高扬的眉毛、微笑的嘴巴、上扬的嘴角表示快乐，眼睛圆睁、嘴唇无意识张开表明惊讶，眉头紧蹙下沉、瞪眼、双唇就紧闭代表生气等，多种情绪状态都可以通过面部表情识别出来。用面部表情识别情绪的研究有很多，像MindReader、Affdex等情绪识别软件就是通过图像识别面部表情得到人们情绪状态的。通过语音识别情绪也已经得到广泛的研究，人们说话的声音里面隐藏着许多有价值的情绪信息，一个悲伤的人说话声音嘶哑低沉、字句停顿时间长、语调单一、讲话速度慢，一个狂躁的人说话声音大、语速快、压迫感强。密歇根大学的精神病学家Melvin McInnis和计算机专家Emily Mower Provost、Zahi Karam于2013年启动一个叫做PRIORI的项目，通过分析语音波形里的音调、音量和节奏等特征，在情绪真正开始波动之前，预测情绪的变化。人的肢体动作可以反映出情绪状态，高兴的时候手舞足蹈、愉快的笑容，不高兴的时候垂头丧气，这就是身心互动的原理。但是，通过机器从肢体动作中提取情绪状态存在较大的难度，因为情绪与动作是多对多的关系，很多动作状态并没有情绪表达。韩智攀基于3D运动数据与识别视频动作数据识别情绪做了尝试，但效果并不理想【基于动作识别的情绪提取方法研究】。瞳孔的大小与情绪状态也是密不可分的，不同情绪状态下被观察者瞳孔直径及直径变化情况差异有统计学意义，愉快组瞳孔直径及变化程度与平静组、 焦虑组差异均有统计学意义【正常人群的情绪状态与瞳孔大小的关系】。但是检测瞳孔大小来识别情绪存在很低的可操作性，检测瞳孔的装置本身就会影响被试情绪，这为通过瞳孔识别情绪状态这种方法的实际使用造成了障碍。通过非生理信号识别情绪的方法优点是直观，但缺点在于如肢体的不可靠性、瞳孔的不可操作性，面部表情和语音语调可以被伪装为情绪识别带来不可靠因素。并且，对于残疾人、未成年人等有些方法并不能适用。

生理信号识别方法主要包括脑电信号、外围生理体征数据、脑血流数据等。外围生理体征数据包括心率、皮肤阻抗、呼吸、肌电、眼动等生理信号来识别对应的情绪。这些外围信号虽然无法伪装也能够得到真实的数据，但是与情绪并没有特别强的关联性，识别准确率低且缺乏合理评价指标，因此也不适合应用于实际。

脑电信号识别方法是通过分析在不同情绪状态下检测大脑发出的不同脑电信号来识别相应的情绪。这种方法不易被伪装，且容易测量，相对去其他生理信号识别情绪有着较高的识别率，同时检测到的脑电信号可以用于脑机接口研究中，因此这种方法被越来越多的应用于情绪识别的研究中。常见的基于脑电信号识别情绪的方法包括核磁共振(NMR)和脑电图(EEG)。由于核磁共振设备体积庞大、价格昂贵，不适合实验室研究。因此目前人们更多的采用脑电信号来进行情绪识别。

本文设计一种基于VR刺激与多种脑电信号特征相结合并运用机器学习的方法来识别情绪的一个有效系统。

## 2.7本章小结

本章首先介绍了VR技术的发展与特点；其次阐述了情绪的基础知识，包括情绪产生的理论学说，情绪的分类与情绪的诱发方法，指出基于VR技术从多通道诱发的情绪相比现有的刺激方法效果要更好；再次探讨了脑电产生的生理基础；然后梳理了脑电信号的预处理以及从脑电信号提取特征的方法；最后列举了一些情绪识别的研究方法，各种识别方法做了对比分析优缺点，指出基于脑电识别情绪的可行性与便利性。

# 第三章 实验设计

## 3.1 实验目的

采集不同人在虚拟现实刺激下不同情绪状态的脑电信号，通过对脑电信号的分析处理，选择合适的数据处理算法，提取有效的脑电信号特征，设计鲁棒的情绪识别算法，达到通过脑电信号识别情绪状态的目的。同时，验证这套基于VR与脑电的情绪识别系统是能够正常有效地工作的。为今后这套系统应用于其他方面(比如，抑郁症识别、阿尔茨海默症早期诊断)提供基础。

## 3.2 刺激材料

在本试验中，尽量挑选在VR中有沉浸感能引起被试情绪共鸣的VR格式的视频。由于实验时间有限同时目前的VR设备长时间佩戴会引起眩晕等不适，所以视频时长尽量短测试时间尽量短，同时使情绪能在一段时间内维持在一定的水平之上，以获得在有限的时间内获得足够多的有效数据。刺激材料一般都在两分钟左右。

本实验，要求刺激材料能够诱发被试产生兴奋、平静两种情绪状态。

由于被试都是以中文为母语，都是研究生学历具有一定的英语视听能力，接受能力强，眼界较开阔，所有没有特别筛选中文影片。对于正面兴奋情绪我们选择了像高空跳伞、过山车等刺激能引起被试兴奋感觉的材料，对于负面平静情绪我们选择了以自然风光为主的刺激材料。



图3-1 刺激素材

## 3.3 被试选择

参与实验的志愿者系华南理工大学22-26岁之间的学生10名，6名男生4名女生，右利手、身体健康、听力正常、无脑神经损伤和精神病史，睡眠状态良好。实验之前告知被试实验目的并讲解实验的操作步骤，同时告知脑电采集设备无创无害。被试均有不定期观看影片的习惯，被试均有观看VR视频的经历以至于在测试中不会产生紧张，被试英语水平都在四级以上对视频中的英文字幕、语言无障碍理解。

## 3.4 实验环境

为了保证实验不受外界打扰，实验在实验室特定房间进行，安静舒适，光线柔和，温度适宜。

试验中，我们采用德力凯NSD-7101神经监护仪作为脑电采集设备，HTC VIVE作为VR内容刺激设备。NSD-7101神经监护仪电脑屏幕实时显示被试脑电信号，VR显示器实时显示被试看到的虚拟现实画面。在VR刺激材料开始播放与结束播放时刻，VR主机会通过TCP/IP协议传递时间戳到NSD系统，NSD记录时间节点以备后续脑电导出。在刺激材料播放间隙，有SAM情绪自我评价表以供被试填写。在准备工作做好之后，整个测试过程无需其他人员参与。实验条件如图3-1所示。



图3-1 实验环境

## 3.5 实验流程

实验程序的严谨保证数据的准确性与有效性。在本试验中，挑选大量候选刺激视频供筛选者（不参与测试）打分，挑选唤醒度平均值最大的四个视频和唤醒度平均值最小的四个视频共八段视频作为测试刺激素材。

整个脑电采集过程如下：

Step1：告知被试实验目的和注意事项，解释SAM情绪自我评价表的含义，指导被试操作设备，SAM自我评价量表如图3-2所示；



图3-2 SAM自我评价表(觉醒度-支配度)

Step2：由实验员佩戴好脑电帽，找准电极位置，注入导电胶使得电极阻抗在5K左右但不超过10K，佩戴VR头显设备，导电膏注射与阻抗检测图如图3-3所示；



图3-3 电极阻抗示意图

电极的位置分布采用国际统一标准的扩展10-20系统，本实验采用的16导各电极的位置如图3-4所示。



图3-4 16导的国际10-20系统电极分布示意图

Step3：被试平复心态，调整合适的姿势，实验员开启脑电测试VR程序。脑电测试VR程序流程如图3-5所示。



图3-5 脑电采集程序流程

## 3.6 本章小结

本章主要介绍了被试在VR刺激素材下脑电信号采集的相关事宜。简单介绍实验目的、实验刺激素材、被试选择以及实验过程等。我们选择十名被试参与实验，被试接受能诱发兴奋、平静情绪的实验素材刺激，并根据刺激做自我评价，评估自己的情绪状态。根据实验记录下特定刺激下的脑电数据以及对应的情绪状态，以进行后续情绪识别的研究。

# 第四章 情绪识别系统设计

## 4.1 系统整体框图

如图4-1所示，本文设计的情绪识别系统包括硬件、软件以及情绪识别算法实现。其中算法设计会在第五章详细介绍，本节介绍情绪识别系统的硬件组成，NSD主机(脑电采集设备型号NSD-7101)与VR主机的信号同步设计以及虚拟现实场景设计。

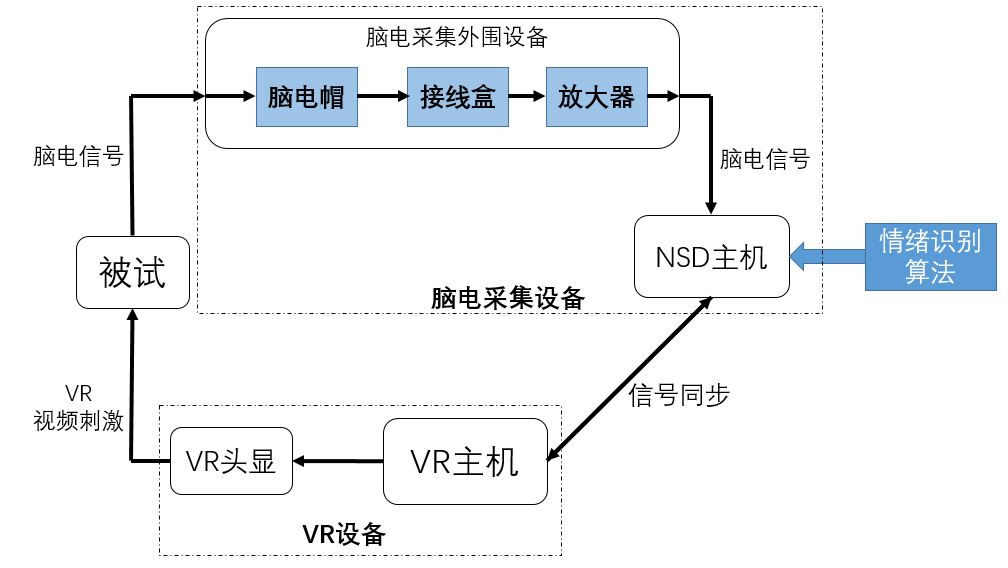


图4-1 情绪识别系统示意图

如图所示，VR设备为被试提供特特定的VR视频刺激，为被试提供多通道刺激；脑电采集设备将脑电帽采集的脑电模拟信号经过接线盒转接到放大器采样量化放大后转化为数字信号传输给NSD主机，情绪识别算法根据接受到的脑电信号识别出情绪状态并显示出来；NSD主机和VR主机需要保持数据同步，即开始采集脑电波的时刻与开始播放VR视频的时刻一致，停止时刻也一致，同时两主机可以进行数据传输。

下面分别介绍该情绪识别系统的硬件设计和软件设计。

## 4.2 硬件设计

### 4.2.1 脑电采集设备

本实验采用德力凯公司NSD-7101神经监护仪作为脑电采集主体设备，但结合VR头显需要做一定的改进。

NSD-7101神经监护仪是深圳市德力凯公司研发的脑电采集设备，其能同时监测脑电、心电、呼吸、血压、心率等信息，其中脑电输入为16导联，电压测量误差不超过±10%，共模抑制比大于110dB,噪声电平小于5uV。该神经监护仪具有高性能的前置放大器，灵敏度高、抗干扰能力强，能够如实的记录各种信号。其采用全电极同时采样保持的方式，是电极间的数据不会发生时间延迟现象，采样数据在放大器内部进行数字转换后再传送主机，使得传输过程中不受打扰。

|  |  |  |
| --- | --- | --- |
| C:\Users\VR_team\AppData\Local\Microsoft\Windows\INetCacheContent.Word\IMG_20161023_205708.jpg | C:\Users\VR_team\AppData\Local\Microsoft\Windows\INetCacheContent.Word\IMG_20161023_205719.jpg | C:\Users\VR_team\AppData\Local\Microsoft\Windows\INetCacheContent.Word\IMG_20161023_205745.jpg |
| 脑电帽 | 接线盒 | 放大器 |

图4-2 脑电帽、接线盒、放大器实物图

NSD-7101主要由脑电帽、接线盒、放大器和电脑主机四部分组成，脑电帽、接线盒、放大器实物如图所示。其中脑电帽负责脑电采集，在接近头皮的电极区域注入导电胶来降低阻抗；采集到的信号通过接线盒传输到EEG放大器，在EEG放大器内进行前置放大、采样保持、滤波等操作后将数据传入电脑主机；电脑主机安装有NSD系统，数据在这里统一处理、交互并在显示屏显示。NSD系统具有采集参数设定、设备调试、数据处理滤波、显示实时脑电信号、按键反馈值信息等功能，为我们检测脑电提供了极大帮助。

由于该神经监护仪监有测心电、心率、脑血流等功能，十六导联中只有十导联有效数据，所以这为我们后面的数据采集与情绪分类增加了难度。同时，由于该设备主要用于医院脑电波检测等检测项目，脑电电极阻抗也只能最低降低到5K左右，相对于Neuroscan等脑电设备其数据精度并不理想，这也为我们通过脑电信号识别情绪提供了挑战。但是，由于该设备的开放性以及与德力凯公司的合作，使得我们有机会对它进行改进。当然，价格相对于Neuroscan也要便宜几倍。为了与VR设备进行配合使用，以及在该设备上开发识别程序，我们增加了该设备的通信功能，相关细节在4.2.3描述。

### 4.2.2 VR情绪刺激设备

本实验采用的VR设备是HTC VIVE，HTC Vive是当今市场上性能体验较好的VR产品之一。其通过以下三个部分致力于给使用者提供沉浸式体验：一个头戴式显示器、两个单手持控制器、一个能于空间内同时追踪显示器与控制器的定位系统。头戴显示器采用4K屏幕，用户基本不会感觉到纱门效应。在显示技术上采用菲涅尔透镜，其能让图像亮度保持一致，成像画质均匀，避免了边角变暗、模糊等问题。即使不佩戴眼镜，400度左右近视也能看清画面的细节。控制器定位系统Lighthouse不需要借助摄像头，而是靠激光和光敏传感器来确定头戴显示器和手持控制器的位置，允许使用者在一定范围内走动。

|  |  |
| --- | --- |
| C:\Users\VR_team\AppData\Local\Microsoft\Windows\INetCacheContent.Word\kf0hUdc.jpg | http://www.vr186.com/d/file/vr_news/vr_product_information/2016-06-27/cda90b2bb5908b143b0a93c2d07b2fda.jpg |
| 图4-3 HTC VIVE设备组成 | 图4-4 菲涅尔透镜 |

HTC VIVE需要配合PC机一起使用，PC机负责控制与信息处理功能，在PC机安装SteamVR平台以驱动和控制VIVE，使用Unreal Engine 4等虚拟引擎工具制作虚拟场景内容，本文的测试系统以及部分情绪诱发素材就是用该引擎制作的。

### 4.2.3 信号同步设计

由于脑电信号采集设备的主机是经过特殊改造的机器，而HTC VIVE需要 的PC主机配置要求比较高，所以本系统暂时使用两台主机来完成整个工作。 如图，主机A是NSD-7101神经监护仪的专用主机设备，安装有EEG采集盒驱动和NSD脑电检测系统以及其他应用程序；主机B是VIVE配置的主机设备，其显卡为Nvidia GTX 1070，CPU为英特尔i5，HDMI接口，安装有SteamVR虚拟现实平台和VIVE应用软件。这两台主机都是各自部分的主要设备，负责数据处理与系统控制功能。



图4-5 主机A与主机B

为了精确、自动记录不同情绪诱发材料下的脑电波以及两设备间的交互，需要对主机A和主机B进行信号同步与数据通信。本文采用的是TCP协议通信以保证两主机信号同步与数据通信，其中主机B作为服务器端，主机A作为客户端，构成C/S系统结构，以降低通讯开销。TCP（Transmission Control Protocol，传输控制协议）是一种面向连接的、可靠的、基于字节流的传输层全双工通信协议，由IETF的RFC 793定义。在简化的计算机网络OSI模型中，它与用户数据报协议（User Datagram Protocol ,UDP）共同完成第四层传输层所指定的功能。

主机A与主机B的TCP数据协议帧结构如表4-1所示。

表4-1 数据协议帧结构

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **包头** | **包长** | **命令** | **数据头** | **数据** |
| **4 Byte** | **4 Byte** | 1. **Byte** | **…** | **…** |

其中命令的数据结构如表4-2所示。

表4-2 命令的数据结构

|  |  |
| --- | --- |
| **命令** | **说明** |
| 0x01 | NSD-7101上传数据 |
| 0x02 | 开始记录 |
| 0x03 | 停止记录 |
| 0x04 | 开始播放 |
| 0x05 | 停止播放 |
| 0x06 | 事件 |
| 0x07 | 时间戳 |

主机B：C/S结构的服务器端，提供VR视频刺激内容，开放IP地址和端口号。在VR视频开始播放或者在播放的某些时刻向客户端主机B发送数据包，通知主机B开始或者结束记录脑电波数据。同时主机B可以接收来自客户端主机A的命令开始播放或者暂停。主机B的虚拟现实场景协议代码截图如4-6所示。



图4-6 TCP协议的枚举类型

主机A：C/S结构的客户端，对主机A的NSD软件新增交互指令，接收来自主机B的指令开始与停止记录数据，同时可以发送开始与停止命令给主机B。这项功能原本的NSD信号采集系统程序是没有的，我们利用其授权的接口新增了TCP协议与控制功能，根据接受到的命令可以进行相应的任务。

同时，主机A与主机B兼有数据传输功能，主机A采集的脑电数据可以通过TCP协议传输到主机B，作为预留功能本文未传输脑电数据。

## 4.3 虚拟现实软件设计

不同于一般的桌面应用开发，VR虚拟现实应用开发需要用到Unity、UE4等引擎工具，在设计实验程序和情绪检测程序以及虚拟现实场景时均会用到。实验程序需要播放不同的刺激素材并在播放刺激视频时候展示SAM自我调查量表，记录被试选择的自我调查量表的数据，提示被试进入休息、缓解情绪、准备下一段测试，整个测试过程尽量不需要他人参与即可完成。与试验程序类似，情绪检测程序为被试呈现特定虚拟现实场景，以诱发被试的情绪。为了更加充分地诱发被试的情绪，需要我们自己设计特定的虚拟现实场景，某些场景需要被试亲自参与其中。下面以UE4为例简要介绍虚拟现实场景开发流程。

虚幻4引擎（Unreal Engine 4，UE4）是由游戏开发者制作并供游戏开发者使用的一整套游戏开发工具，是Epic Games公司发布的业内顶级游戏引擎。SteamVR是VIVE的虚拟现实平台，UE4开发的虚拟现实项目需要搭载在SteamVR上才能在VIVE上提供虚拟沉浸式体验。

在虚幻引擎中，最基础的建造单元叫做Object，UE4中几乎所有的东西都继承于Object。虚幻引擎场景的内容包括：蓝图（blueprint）、角色（Actor）、组件（component）、关卡（level）。蓝图是UE4提供的一个可视化脚本系统，提供一种直观的、基于节点的可视化脚本设计界面，通过使用各种不同功能的节点把事件、函数和变量互相连接，构成图表以实现各种复杂的功能。引擎中拥有多种不同类型的Actor，如静态网格物体、摄像机、玩家起始点等，Actor是放置在关卡中的任意对象，支持三维变化如平移，旋转和缩放。组件是一种特殊类型的对象，组件一般用于需要简单地切换部件的地方，以便改变具有该组件的Actor的某个特定方面的行为或功能，例如飞机和轮船的发动机是不同的，而飞机和汽车也是不同的。关卡是定义的游戏场景，也被称为地图，主要通过Actor的属性来创建、查看及修改关卡。

如图4-6所示，本设计中用到的虚拟现实刺激场景示例。



图4-7 虚拟现实场景过山车制作示例

## 4.4 本章小结

本章主要介绍了基于脑电的情绪识别系统整体结构，分别介绍了脑电采集设备和VR虚拟现实设备以及这两种设备的同步问题；然后介绍了虚拟现实场景设计。硬件系统与软件系统的设计为基于VR虚拟现实刺激与脑电信号的情绪识别系统提供基础，对脑电采集设备的改造使得本情绪识别系统能顺利工作。

# 第五章 情绪识别算法设计

## 5.1 情绪识别研究

本章的内容是为本文的情绪识别系统提供可靠的情绪识别算法，主要探究在两种情绪下对应的脑电信号的不同特征模式，从线性动力学和非线性动力学的角度进行特征参数的提取，通过机器学习的分类方法进行情绪的正确判别。本文对脑电信号的处理流程如图5-1所示。



图5-1 脑电信号处理流程

我们从脑电采集设备NSD-7101得到的是经过放大滤波后的原始信号，并不是纯粹的脑电信号。首先需要经过预处理去除掉掺杂脑电信号中的眼动伪迹、肌电伪迹等噪声成分；其次通过小波变换得到不同频段的脑电波节律信号，对每种节律信号以及原始脑电波信号取小波能量、近似熵、样本熵、fisher information等多种特征数据作为脑电信号的特征；然后使用特征选择方法以及降维技术对提取到的脑电特征进行降维；最后选择合适的分类算法模型对情绪进行识别并对识别结果做分析。

下面对在脑电信号处理流程中使用到的算法进行详细说明。

### 5.1.1 数据预处理

脑电帽采集到的信号经过放大器放大抽样后得到的并不是真实的脑电信号，脑电信号非常微弱，比较容易受到外界的干扰；同时，由于被试有意或无意的动作比如眨眼、吞咽，头动等都不可避免的引入噪声。这些干扰信号就是脑电的伪迹（artifact），要得到真正的脑电信号需要对噪音和伪迹进行去除，减少噪音对信号的干扰。

本文采用小波变换(wavelet)与独立成分分析（ICA）相结合的方法(wICA)，有效去除眼电伪迹的同时尽可能多的保留脑电信号【Recovering EEG brain signals- Artifact suppression with wavelet enhanced independent component analysis】。由于伪迹成分的能量集中在一定的频率范围内，且能量幅度较大；而脑电成分的能量分布于整个频段，且能量幅度较小。对伪迹分量进行离散表博变换，那么伪迹成分的小波系数幅度比脑电成分的小波系数大，同时伪迹成分的能量集中于一些较大的小波系数而脑电成分的能量分布于整个小波域【Recovering EEG brain signals- Artifact suppression with wavelet enhanced independent component analysis】。于是，去除小波系数较大的伪迹而保留小波系数较小的脑电信号。小波变换的原理在下一节进行介绍。小波变换阈值去燥的效果主要取决于小波函数的选择和小波分解的层数。采用wICA滤波的基本步骤如下：

Step1:对采集到的多通道脑电信号进行ICA分解（本文使用FastICA算法），得到混合矩阵M和N个独立成分{s1(t), s2(t),…,sN(t)}；

Step2:对各个独立成分进行小波变换得到对应的小波逼近系数矩阵{W(j,k)}；去除伪迹独立成分中的对应小波系数集中且较大的伪迹而保留脑电成分；

Step3:对每层分支采用硬阈值的方法对每层小波系数进行阈值处理：

其中，N为信号长度，，W(d,b)是小波逼近系数，median(A)是A的中位数。每层的小波系数与对应的阈值进行比较，高于阈值的小波系数置零，其他小波系数保持不变；

Step4:将阈值处理后的小波系数进行逆小波变换，得到只包含脑电成分的信号{}；

Step5:进行ICA逆变换，得到不含伪迹较纯净的脑电信号

。

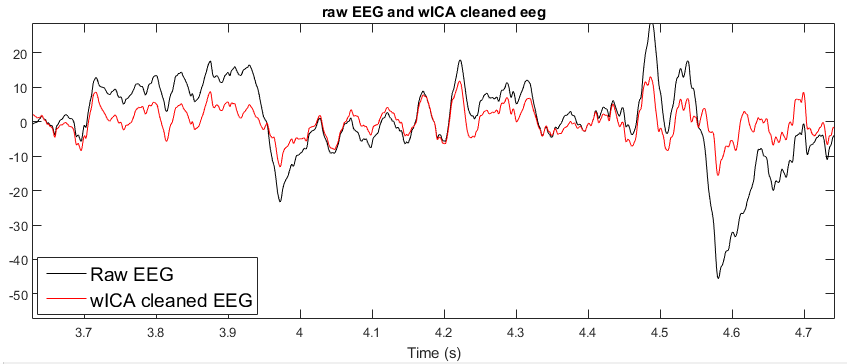
图5-2是脑电信号经过是wICA算法滤波前后的波形对比图，可以看到黑色的脑电波有些凹陷很深的地方，这些正是眼动或者头动引起的，滤波后这种情况有所缓解。

图5-2 原始脑电通过wICA滤除伪迹对比图

### 5.1.2 特征提取

特征工程是机器学习、模式识别工作中比较重要的一环，而提取到有用的特征对最终的识别准确率起着至关重要的作用。本文以非动力学分析方法为基础提取脑电信号的非线性特征，在已有的研究基础上，提取各节律信号功率谱、熵函数、时间序列相关性指数、分形维数等多种特征。这些特征在不同的文献中已经应用于脑电信号的研究分析，并将其作为脑电信号的重要属性。

前文已经介绍过，原始脑电信号可以通过小波变换分解为Delta波(0-4Hz)、Theta波(4-8Hz)、Alpha波(8-12Hz)、Beta波(12-30Hz)、Gamma波(30-60Hz)五种节律波，这五种节律波形在神经科学、心理学等领域都被认为与人类大脑活动有着密切的联系。针对每种波形以及原始信号计算频带能量与信息熵等多种特征。然后结合原始脑电信号的相关性指数与复杂性指数等一起构成脑电信号的特征。

接下来详细介绍特征提取过程用到的各种方法。

#### 5.1.2.1 小波变换

脑电信号是一种典型的短时非平稳随机信号，并且具有一定的非高斯性和非线性【结合小波包和ICA的脑电信号特征波提取方法】。传统的分析处理方法将脑电看成近似线性、准平稳的随机信号，这样的处理结果并不令人满意。小波变换是一种时频信号处理方法，与快速傅里叶变换（FFT）和短时傅里叶变换（STFT）相比，小波变换是在时域和频域的局部变换，对原始信号进行多尺度细化，在高频处对时间进行细分，在低频出对频率进行细分。小波变换能够自动适应时频信号分析的要求，克服了傅里叶变换的不足。小波变换包括：连续小波变换(continuous wavelet transform, CWT)和离散小波变换（discrete wavelet transform，DWT）。主要区别在于连续变换在所有可能的缩放和平移上操作，这使得连续小波变换复杂度高冗余量大；而离散变换采用缩放和平移值得特定子集。本文针对脑电波处理只需要提取五段信号，采用离散小波变换。

离散小波变换主要建立在二进制小波变换的基础上，以位移为2的幂次基础进行离散化。小波变换的实质就是把信号按照不同的频带成分分别提取出来，不同尺度的小波函数相当于不同频带的带通滤波器。如图5-3，采用小波提取框图表示小波分解过程，其中每层都含有近似分量A和细节分量D。输入信号x[n]长度为N，g[n]是低通滤波器去掉信号的高频部分保留的低频成分；h[n]是高通滤波器滤掉低频部分保留的高频成分；↓Q表示欠采样滤波器，如果出入x[n]则输出y[n]=x[Qn]，此处取Q=2。



图5-3 离散小波分解示意图

本文针对脑电信号采用db8小波函数作为小波基函数，五层小波分解，抽样频率为Fs=256Hz，根据奈奎斯特采样定理，原始信号的最大有效频率为128Hz。其分解频率树图如图5-4所示。



图5-4 小波分解频率树图

细节分量D1、D2、D3、D4、D5和近似分量A5分别对应着噪声、Gamma、Beta、Alpha、Theta、Delta节律波。至此，我们可以针对每种节律波提取其时域、频域特征了。图5-5即某一导联上小波分解得到的五种节律波形图。



图5-5 小波分解的各层节律波形

#### 5.1.2.2 频带功率谱密度与相对强度比

小波变换已经把原始脑电信号分解为五种节律信号，由帕森瓦尔定理，对于每个频带的节律信号的能量可以由时间序列的幅度值的平方和表示，这样每个电极的在五个频带上的频带能量即可用如下公式计算：

其中，i=1,2,…,N，k为节律波序号，为不同节律波的能量,为k频带的功率谱密度(Power Spectral Intensity ,PSI），为第k层的第i个幅值。

信号总能量，K为节律波总层数。相对强度比(Relative Intensity Ratio,RIR)为：

,k = 1,2,…,K

#### 5.1.2.3分形维数

脑电独立成分分析方法将脑电分解为几个独立的子系统，维数分析方法认为在大脑皮层测得的脑电信号是这几个独立子系统电活动的时空叠加。运用维数分析将脑电信号从一维时间序列提取出这些独立子空间的信息。本文采用Petrosian Fractal Dimension (PFD)【A Comparison of Waveform Fractal Dimension Algorithms】和Higuchi Fractal Dimension (HFD)【Parameters analyzed of Higuchi’s fractal dimension for EEG brain signals】作为脑电信号的特征：

其中，N是时域信号的长度，是信号符号改变的次数。

Higuchi 【1988，Approach to an irregular time series on the basis of the fractal theory】在他的文章中解释了HFD的计算方法。通过原始信号[]构造k个新的序列

其中m=1,2,…,k

对上式中的每个序列计算长度：

平均长度定义为：

,

k=1,2,…,Kmax，则该公式需要计算Kmax次，Kmax是需要调节的参数。而HFD系数是重对数(ln(L(k))—ln(1/k))图的线性回归的最佳拟合直线的斜率。如图5-6所示。



图5-6 重对数ln(L(k))—ln(1/k)的散点图

#### 5.1.2.4 Hjorth参数

Hjorth【EEG analysis based on time domain properties Analyse EEG basee sur les series temporelles】分析了脑电信号的时域特征对我们分析脑电有较大的参考价值。它统计的时间信号的斜率均方根与振幅均方根的比率称之为移动性（mobility），用于估算均数频率；信号的斜率变化与理想的曲线比率的均方根称之为复杂度（complexity），被用于估算信号的带宽。对于时间序列[], Hjorth移动性（Hjorth mobility）和Hjorth复杂度（Hjorth complexity）分别为：

其中，，，，。

#### 5.1.2.5 熵函数

随着非线性动力学技术的应用于发展，研究人员发现了许多代表脑电信号特征的属性，其中熵函数对情绪的识别有着重要的作用。苏建新等人验证了近似熵(Approximate Entropy,ApEn)、样本熵（Sample Entropy,SE）、排列熵（Permutation Entropy,PE）在识别平静与压力态两种情绪状态的有效性【基于脑电信号的情绪识别研究】。本文采用谱熵(spectral entropy)、近似熵、SVD熵（SVD Entropy）作为特征。其中谱熵【Quantification of EEG irregularity by use of the entropy of the power spectrum】定义为：

其中RIR在5.1.2.2已经定义过为频谱相关强度比。

SVD熵【Temporal and spatial complexity measures for electroencephalogram based brain-computer interfacing】因使用了奇异值分解(SVD)得名。对于输入原始脑电信号[]构造延时序列

，其中τ是时延，是嵌入维数(embedding dimension)。本文中采用，τ=2。嵌入空间的构造方法为：

对嵌入矩阵计算SVD分解得到奇异值向量，M为向量的长度。最后SVD熵定义为：

其中，标准化的奇异值。

相似熵是用来衡量时间序列复杂度的一种参数，该方法描述的是若干相似向量从m维空间增加到m+1维向量时任然保持较高相似性的概率。近似熵算法对确定信号和随机信号都合适，同时适合由随机信号和确定信号相混合的信号。近似熵的计算步骤如下：

Step1：对于原始输入信号构造新的子序列，，其中m是子序列的长度，取1,2或3；

Step2：r定义为信号噪声等级，r=k·SD，SD是信号x[n]的标准差，k=0,0.1,0.2,…,0.9;

Step3：构造子空间矩阵，对于矩阵中的每个元素计算

Step4：计算近似熵

本文选择长度为N=1000的数据计算近似熵，嵌入维数选择m=2，有效阈值取r=0.25·SD。由于输入数据程度较长，所以分多段每段N=1000计算ApEn值然后计算平均作为原始脑电序列的近似熵ApEn。

#### 5.1.2.6 Fisher Information

Fisher信息作为脑电信号的复杂度特征被Christopher【Extracting multisource brain activity from a single electromagnetic channel】用于分析多源脑电信号。Fisher信息计算：

其中，为5.1.2.5计算SVD熵用到的标准化奇异值。

#### 5.1.2.7 Hurst指数

Hurst指数是判断时间序列数据遵从随机游走还是有偏的随机游走过程的指标，能够很好的描述脑电信号的变化趋势及稳定性，是基于重标极差（Rescaled Range statistics,R/S)）分析方法得到的参数。对于原始时间信号先计算累积偏差

其中。然后计算

Hurst指数即重对数(ln(R(n)/S(n))—ln(n))图的线性回归的最佳拟合直线的斜率。如图5-7所示。



图5-7 ln(R(n)/S(n))—ln(n)的重对数坐标散点图

#### 5.1.2.8 去趋势波动分析

去趋势波动分析（detrended fluctuation analysis，DFA）是随机信号、混沌系统和时间序列分析中统计自仿射性的一个有效参数，用于分析时间序列的长程相关性，它可以有效的滤除序列中的各阶趋势成分，能检测含有噪声并叠加有多项式趋势信号的长程相关，适合非平稳序列的长程幂律相关分析。

对于原始时间信号其平局值为，计算新的序列信号，其中。然后将序列y[n]切片成长度为M的多段子区间，以最小二乘法对每一个子区间的m个数据拟合一条直线，Y坐标轴截距用表示，序列的均方波动函数，其中被称为趋势。

#### 5.1.2.9 其他参数

M. Murugappan【Discrete Wavelet Transform Based Selection of Salient EEG Frequency Band for Assessing Human Emotions】将各电极的小波系数的标准差、方差、功率谱以及各小波层的能量熵作为脑电信号的特征来分析情绪。功率谱与熵前面介绍过，标准差表征不同情绪状态下电极信号的潜在偏差，公式为

其中是小波系数的均值，是小波系数的的细节分量，k为小波系数序列元素的序号，j为小波分解层数。

方差表征不同情绪状态下电极信号的变化范围，计算公式为

### 5.1.3 特征选择与降维

在上一节，我们提取了多种与脑电信号有关的特征，每一种特征都有文献应用于相应研究与应用，其中不乏有几种特征应用于基于脑电的情绪分类，但是我们并不确定脑电信号的哪些特征与平静-兴奋这个维度的情绪状态相关性较大，所以需要做特征选择以去除不相关的冗余特征，降低情绪识别的难度，提升准确度。采用机器学习、数据挖掘的方法从众多特征中选取有用的特征帮助情绪识别是本文的创新点之一。

由于降维技术也有可以达到特征选择的目的，而特征选择也可以达到降维的目的。于是本节将特征选择与降维技术一起阐述。

#### 5.1.3.1 特征选择算法

常见的特征选择方法大致可分为三类：过滤式(filter)、包裹式(wrapper)和嵌入式(embedding)。本文采用的是过滤式选择中的信息增益（information gain）方法和嵌入式选择中的正则化方法。

信息增益在下一节情绪识别算法中有较为详细的描述，它关注的是以某种特征作为类别划分依据则样本纯度提升的性能，纯度提升越大的特征在决策树中越接近根结点。

正则化方法是融合在学习器训练过程中的，特征选择过程和学习器训练过程在同一个优化过程中完成，逻辑回归分类算法、支持向量机均有使用。给定数据集，以线性回归模型为例，以平方误差为损失函数，则优化目标为

当样本特征很多而样本数相对较少，式（）很容易陷入过拟合。为了缓解过拟合问题，对式（）引入正则化项，若使用L2范数正则化，则有

其中正则化参数λ>0。上式称为“岭回归”(ridge regression)。若使用L1范数正则化，则有

其中正则化参数λ>0。上式称为LASSO(Least Absolute Shrinkage and Selection Operator)。

L1范数和L2范数是正则化中使用最多的正则化项，它使得不那么重要的特征的系数尽可能小，最终得到的w会是一个稀疏的向量。这就意味着初始的d维特征中仅有对应着w的非零分量的特征才会出现在最终的模型中，于是达到了特征选择的目的。当然，L1范数和L2范数正则化的另一个重要作用就是防止模型过拟合。

#### 5.1.3.2 降维算法

本文采用的降维技术是主成分分析(Principal Component Analysis,PCA),这是一种最常用的降维方法，其主要思想就是将特征从高维空间线性投影到低维空间。它的主要思想就是将高维空间数据投影到超平面上，使得样本点到这个超平面的距离足够近同时样本点在这个超平面上的投影能尽可能分开，即最近重构性和最大可分性。实际上，最近重构性和最大可分性是等价的，可以分别从最近重构和最大可分推导出主成分分析的优化目标函数。

从最近重构性来推导。数据样本进行中心化，即；投影变换后得到的新坐标系为，其中是标准正交基向量，，。若丢弃新坐标系中的部分坐标，即将维度降低到，则样本点在低维坐标系中的投影是，其中是在低维坐标系下第维的坐标。若基于来重构，则会得到。

考虑整个训练集，原样本点与基于投影重构的样本点之间的距离为

其中，const为常数，tr()表示求矩阵的迹。根据最近重构性，式（）应被最小化，考虑到是标准正交基，是协方差矩阵，有

这是主成分分析的优化目标。

从最大可分性来推导。样本点在新空间中超平面上的投影是，若所有样本点的投影能尽可能分开，则应该使投影后样本点的方差最大化。投影后样本点的方差，于是优化目标可写为

显然式（）与式（）等价。

对式（）或式（）价可以使用拉格朗日乘子算法得

于是，只需对协方差矩阵进行特征值分解，将求得的特征值排序：，再取前个特征值对应的特征向量构成得到主成分分析的解。PCA的算法描述如图5-8所示【机器学习，周志华】。



图5-8 PCA算法

### 5.1.4 情绪识别算法

情绪识别系统算法结构的最后一步就是设计分类算法，在分类算法中我们提取脑电信号的特征，做特征选择和降维之后，利用分类器对所对应的情绪状态进行识别。这里着重介绍决策树算法和逻辑回归分类算法，然后将多种不同的分类算法模型融合，形成最终的识别结果。

#### 5.1.4.1 决策树分类算法

决策树（Decision Tree，DT）是一类常见的机器学习方法。顾名思义，决策树是基于树结构来进行决策的，这恰是人类面临决策问题时一种自然的处理机制。决策树由结点和有向边组成，结点分为内部结点和叶子结点，内部结点表示一个特征或属性，叶子结点表示一个类。

用决策树分类，从根结点开始对实例的某一特征进行测试，根据测试的结果将实例分配到其叶子结点；这时，每一个叶子结点对应着该特征的一个取值，递归地对实例进行测试并分配，直至达到叶子结点。最后将实例分配到叶子结点的类中。如图5-9，模拟了是否打羽毛球的决策树直观过程。



图5-9 是否打羽毛球的决策树

那么，如何选择最优划分属性。一个准则就是随着划分过程的不断进行，我们希望决策树的分支结点所包含的样本尽可能属于同一类别，即结点的“不纯度”降低。常用的ID3决策树分类算法【Quinlan,1986】是以信息增益为准则来选择划分属性的。

离散属性a有V个可能的取值，若使用a来对样本集D进行划分，则会产生V个分支结点，其中第v个分支结点包含了D中所有在属性a上取值为的样本，记为。属性a对样本集D进行划分的信息增益

其中，Ent(D)为样本集合D的信息熵，

为样本集合D中第k类样本所占的比例。Ent(D)的值越小，则D的不纯度越高。

通过比较每个属性的信息增益值，选择信息增益最大的作为最优特征作为当前根结点。

实际上，信息增益准则对可取数目较多的属性有所偏好，为减少这种偏好可能带来的不利影响，C4.5决策树算法【Quinlan,1993】使用信息增益比来选择最优划分属性。信息增益比定义为

其中

称为属性a的“固有值”，属性a的可能取值数目越多，IV(a)的值就会越大。信息增益比则对可取数值数目较少的属性有所偏好。

信息增益和信息增益比是决策树算法中最常用的特征选择准则，各有优缺点。决策树算法中，为了尽可能正确分类样本，结点划分过程将不断重复，有时会造成分类器过拟合现象，则需要用到剪枝策略。

5.1.4.1逻辑回归分类算法

逻辑回归（Logistic Regression，LR）虽然它的名字是“回归”，但它实际上是一种分类算法。对于二分类任务，用对数几率函数作为判别函数

其中为分类面，x为样本训练集合。对数几率函数是一种“Sigmoid函数”，其输出值在z=0附近变化狠陡，输出值为0到1的值，可以作为样本属于类0或者类1的概率。则只需要找到这样的分类面即可将原始样本点分开，当z>0时预测值为类1；当z<0时预测值为类0；当z=0时随机取类标记。



图5-10 Sigmoid函数示意图

#### 5.1.4.2 支持向量机

支持向量机（Support Vector Machine，SVM）是应用广泛效果较好的分类器，其可作为线性分类器也可应用于非线性可分情况。其核心思想就是找到一个划分超平面将空间中的点划分开来，对于线性不可分的情况，将数据集映射到高维空间中找出划分超平面，从而达到鲁棒的分类效果。算法具体描述如下。

首先从线性可分的二分类问题来讨论。给定训练样本集为。在样本空间中，划分超平面可通过如下线性方程来描述：

其中为法向量，决定超平面的方向；b为位移项，决定了超平面与原先的距离。超平面可由法向量w和位移b确定，样本空间中任意一点x到超平面(w,b)的距离可以写为

假设超平面(w,b)能将训练样本正确分类，即对于，若，则有；若，则有。令

距离超平面最近的是上式成立的训练样本称为 “支持向量”（Support Vector），两个异类支持向量到超平面的距离之和为，被称为间隔。欲找到具有最大间隔的划分超平面则需要满足约束条件：

这样就把问题转化为凸二次规划问题，利用拉格朗日乘子法可解决这个问题。构造拉格朗日函数：

其中。令对w和b的偏导为零可得

将式（）带入式（），即可将中的w和b消去，在考虑式的约束，就可得到式（）的对偶问题

解出α后，求出w与b即可得到模型

。

这样，原问题就转化为了最大化对偶问题，而解决最大化对偶问题可以通过SMO(Sequential Minimal Optimization)等标准方法得到解决。从而原始问题求解最大间隔问题也得到了解决。

前面的讨论中，我们假设训练样本都是线性可分的，即存在一个划分超平面能将训练样本正确分类。然而现实任务中，原始样本空间内也许不存在一个能正确划分两个样本的超平面。对于这样的问题，可将样本从原始空间映射到一个更高维的特征空间，使得样本在这个特征空间内线性可分。我们用φ(x)表示将x映射后的特征向量，原始样本的内积运算就变成了，直接计算是很困难的，我们定义核函数来表示内积运算

即与在特征空间内的内积等于它们在原始样本空间内通过函数计算的结果。将式带入到式中，就得到了拉格朗日对偶极子函数

这样问题如前文所示，同样可以得到解决。核函数可以由很多种，最常用的为线性核、多项式核和高斯核（RBF）。不同的核函数对意味着原始样本被映射到了不同的特征空间，不同的核函数导致分类效果不同。

前面的讨论一直假设训练样本在样本空间或特征空间中是线性可分的，即存在一个超平面能将不同类的样本完全划分开。然而在实际任务中往往很难找到这样的超平面将所有样本完全划分开来。缓解该问题的办法就是允许支持向量机在某些样本上出错，我们用ξ来表示这些误差，，这样式（）就变成

其中，是惩罚系数。这仍是一个二次规划问题，解决方案同样采用拉格朗日乘子法。

对于SVM参数的最优选择问题不同的问题有不同的最优参数，本文采用网格搜索法来选择最佳参数，即惩罚系数C取值范围内选取多个取值和多个核函数构成取值网格，对这些参数进行遍历搜索分类准确率最高的惩罚系数C和核函数作为参数。

#### 5.1.4.3 集成学习与模型融合

集成学习(ensemble learning)通过构建并结合多个好而不同的学习器来完成学习任务，先产生一组个体学习器，再用某种策略将他们融合起来，获得比单一学习器显著优越的泛化性能。个体学习器通常由现有的学习算法从训练数据产生，本文采用的就是C4.5决策树算法、逻辑回归、支持向量机为个体学习器的集成。



图，集成学习示意图

融合策略，本文采用的是绝对多数软投票法。即不同的个体学习器产生不同的所属类别概率，将预测为某类的概率相加超过所有预测结果一半的则预测为该类。

## 5.2 情绪识别分析

实验采样10名被试的脑电数据，每名被试测试8段VR视频，共采集到80个样本，取觉醒度值高于5的为正样本，低于等于5的负样本，则正负样本为43:37。运用基于改进的网格搜索法对各模型进行参数调优，采用5折交叉验证，取平均值作为该次输入特征参数的最终分类识别率。

本文采用多种方法来提高基于脑电信号的情绪准确率，其中以分解原始脑电信号为γ、θ、α、β、δ五种节律波再提取特征，对特征进行特征选择与降维以及通过模型融合来提高识别率这三种方法效果最为突出。以下分别针对这三种提升进行分析。

### 5.2.1信号分解效果分析

在提取特征之前，我们已经将原始脑电信号滤波后通过小波变换分解为γ、β、α、θ、δ这五个常用频带上，然后对原始信号和各节律脑电波提取特征，根据前面的介绍每个电极提取到25维特征。为了探究那种节律波对我们识别平静—兴奋情绪更重要，我们首先分别取原始脑电波和五个独立节律波作为特征对样本进行学习，之后再将所有特征结合起来对样本进行学习，采用单模型SVM作为分类器，核函数取RBF核。分类结果如表5-1所示。

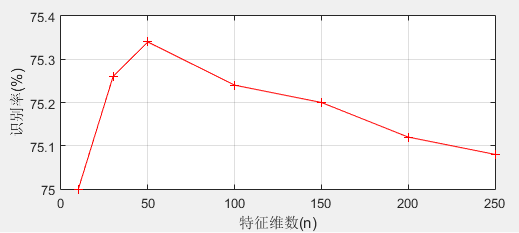
表5-1 不同节律波分类效果对比

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 原始信号 | Gamma | Beta | Alpha | Theta | Delta | ALL |
| 0.604 | 0.742 | 0.664 | 0.686 | 0.682 | 0.607 | 0.776 |

由表5-1可以看出，单纯用未做分解的原始信号提取特征作为训练数据则识别率很低，而分解后的高频信号γ波的识别率较高，δ波对兴奋的识别率较低。而结合多个节律波一起作为训练数据这识别率得到大幅度提升（表中ALL这一列表示融合多种节律信号作为特征训练数据）。这在一定程度上说明高频信号相对于低频信号与情绪的关联更加紧密，多种频率的脑电信号一起反应情绪状态。后续的分析采用多种节律信号融合的方式对作为训练数据。

### 5.2.2特征选择效果分析

前面已经介绍特征提取的过程，每个电极提取25维特征，共10导有效数据，则总共有25×10=250维特征。本文采用PCA算法将原始特征投影到低维空间中，如图，降维效果与低维空间维数的关系，坐标轴横轴为低维空间维数，纵坐标为情绪识别准确率。其中分类算法采用单模型SVM，核函数取RBF核。

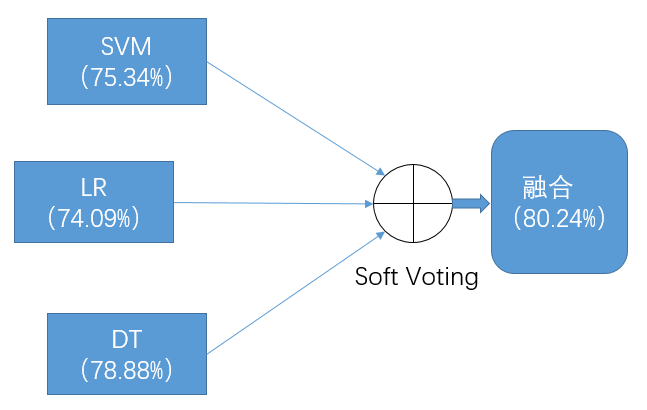


图，识别率与特征维数关系图

高维空间与低维空间的识别效果确有不同，在高维时随着维度的降低识别率有所提高，而降低到50维以下时随着维度的降低识别率急剧下降，这是由于降维的效果舍弃了部分特征向量。舍弃的这一部分信息使得样本的采样密度更大，同时舍弃的这一部分特征特征值较小而噪声较大，所以降维有去燥的效果。但是，如果维度过低则损失过多的信息导致识别率降低，于是找到合适的低维空间维度很重要。后面的分析采用特征维数为50维作为降维参数。

### 5.2.3模型融合效果分析

为了提高模型对情绪的识别率，我们设计多种“好而不同”的模型采用“软投票”方式结合起来以达到提高准确率的目的。这里的“好而不同”指的是各个子模型识别效果好但各个模型之间相差较大。如图，将SVM、LR、DT子模型融合，最终的识别率提高到80.24%。这里的决策树DT模型采用了离散化，即将脑电信号特征分桶离散化后采用决策树算法进行分类，为最终的识别率提高做了很大的贡献。



图，多模型融合示意图

## 5.3 识别软件整体效果

经过上一节5.2的分析与改进，最终该基于VR与脑电的情绪识别系统对平静—兴奋的识别准确率率达到了80%以上，单一模型的识别率也在74%以上。这样的识别效果可以与其他研究者对于二分类情绪相媲美，如苏建新用音乐视频刺激被试产生特定情绪然后基于脑电识别压力与平静状态，其最高识别率为77.8%【基于脑电信号的情绪识别研究，苏建新】；聂聃以视频作为刺激素材基于脑电信号识别积极情绪与消极情绪达到最高85%的识别率【聂聃】；张弛以中国情绪图片系统为刺激素材通过脑电信号识别正性情绪与负性情绪，识别率未达到80%【张弛】；段若男通过脑电信号识别以视频为刺激素材引起的不同情绪，其识别正向情绪—中性情绪—负向情绪的平均识别率为81.57%【段若男】。

以上研究，采集脑电的设备都是价格昂贵的进口设备，在数据精度上要比我们采用的NSD系统要高，噪声消除与滤波做的更好；我们采用的16导联有效数据10导的脑电帽，而其他研究人员采用的是64或者32导的脑电数据；刺激被试产生特定情绪的素材为图片、音乐或者视频，被试在测试时可以保持身体不动，避免头动等伪迹掺杂在脑电波数据中。由于现有VR技术未能解决的原因，长期佩戴VR设备会引起眩晕等状况，所以测试时间不能太长，造成本实验样本数据较少。精度不高、噪声大、脑电通道数低、头动影响、样本少等因素为我们识别情绪增加了难度，这在我们的试验中一定程度上得到了克服。通过wICA算法去除伪迹大大降低了噪声成分，通过小波分解得到不同节律波信号、提取多维特征、降维以及模型融合整体提高了识别准确率。

虽然通过算法部分解决了试验中存在的问题，但是提高脑电数据的精度与准确性、降低噪声、增加训练样本，势必会提高最终的识别率。

## 5.4 本章小结

本章介绍了情绪识别系统的算法设计，包括数据预处理、特征提取、特征选择与降维、算法设计等步骤，然后分析了该系统的情绪识别效果，跟其他研究结果做了比较，并找出自己的不足与提高方向。

# 第六章 总结与展望

## 6.1 总结

本论文主要设计了一个基于VR刺激被试产生不同情绪状态，通过脑电信号识别情绪状态的系统，根据情绪模型理论，主要识别的情绪维度为平静—兴奋，最终的识别率达到80%。在实验数据的基础上，我们提出并搭建了基于脑电的情绪识别系统，能够根据脑电波数据判断被试被VR内容刺激时所处的情绪状态。

该情绪识别系统主要分为两个部分，其一是系统设计，其二是算法设计。系统设计包括硬件设计和软件设计。硬件设计主要是结合脑电采集设备和VR设备的功能改造与创新，由于两台设备均需要特定配置的主机而无法共用一台主机。所以需要对这两台设备进行信号同步与数据传输，本文采用TCP协议进行两机交互。软件设计主要刺激素材虚拟现实场景设计与测试及识别软件设计。虚拟现实场景和测试识别程序均采用虚幻4引擎设计，被试与VR设备的监护界面均在VR头显完成。

识别算法构成了该情绪识别系统的核心。对于采集的脑电数据，我们先通过wICA算法去除眼动、头动等伪迹成分；然后通过小波变换将每个电极的原始脑电信号分解为五种常见的节律波，对每种节律波提取能量特征、信号统计特征、熵特征等作为脑电信号的特征；之后对提取到的特征进行特征选择与降维，以降低计算量和去除噪声；最后通过多种分类算法模型对情绪状态进行识别，然后通过软投票的方法将不同的模型融合以提高识别率。

本文的主要贡献在于，设计了基于VR刺激与脑电的情绪识别系统，设计了一套情绪识别算法。

## 6.2 展望

本文虽然设计了相对可靠的通过脑电识别VR佩戴者情绪状态的系统，但由于条件限制，仍然存在一些不足：

首先，由于脑电采集设备采用湿电极需要涂导电膏才能降低阻抗，脑电实验费时费力，VR头显长期佩戴会引起眩晕，以至于测试样本较少。实验室另外的小组对干性电极的研究可能会解决湿电极存在的问题，干电极的使用为本系统商用提供基础。接下来增加被试样本数量有助于我们形成更一般性的结论，对识别率提高也有很大的帮助。

其次，本文只研究了平静—兴奋维度上的情绪识别，在后续的工作中可以增加情绪维度，比如高兴、紧张、恐惧等，对更多种情绪进行识别。

再次，由于我们的刺激素材使虚拟现实场景，主要为被试提供视觉、听觉上的沉浸式体验，后续工作中可以增加更多维度的刺激，比如触觉、嗅觉等多维虚拟现实刺激。

然后，在虚拟现实场景中，被试不可避免的要进行头部转动，以及其他脑部信号对脑电信号有着很大的影响，在后续的研究中可以着重对VR刺激下的脑电信号组成做进一步的研究，这对滤除伪迹会有很大的帮助。

最后，在情绪识别算法的设计中，还可以增加其他特征参数，出了本文用到的常见分类器还可以尝试更多的分类算法模型，对最终的识别结果也可以进行强度等级划分，以使识别结果更加准确。