

分类号: TP311.5

学校代码: 10410

密 级: 公开

学 号: 0242016001

# 江西農業大學 碩 士 學 位 論 文



## 基于脑电信号多熵特征的驾驶疲劳 检测模型研究

Driving fatigue detection model based on  
multi-entropy feature of EEG signals

申 请 人: 王海玉

指 导 教 师: 胡剑锋 教授 王映龙 教授

学 科 专 业: 计算机技术

所在院 ( 所 ): 计算机与信息工程学院

论文提交日期: 二〇一九年六月





# 目 录

摘 要.....	I
Abstract.....	II
第 1 章 绪论.....	1
1.1 课题研究背景及意义.....	1
1.2 国内外驾驶疲劳检测研究现状.....	1
1.2.1 基于车辆运行状态的监测预警.....	2
1.2.2 基于驾驶员行为特征疲劳的监测方法.....	2
1.2.3 基于驾驶员生理指标的疲劳监测方法.....	3
1.3 本文的主要工作与研究内容.....	4
1.4 本文的组织结构.....	5
第 2 章 相关技术背景及分析方法介绍.....	7
2.1 脑电伪迹去除方法.....	7
2.1.1 主成分分析.....	7
2.1.2 独立成分分析.....	8
2.1.3 小波变换.....	9
2.2 脑电特征提取常用方法.....	10
2.2.1 时域分析.....	10
2.2.2 频域分析.....	11
(1) 傅里叶变换.....	11
(2) 功率谱分析.....	11
2.2.3 时频分析.....	12
2.2.4 非线性动力学分析.....	12
2.3 脑电分类常用方法.....	13
2.3.1 Fisher 线性判别.....	13
2.3.2 随机森林.....	14
2.3.3 人工神经网络.....	15
2.3.4 集成学习.....	16
第 3 章 实验数据采集.....	18
3.1 实验范式.....	18
3.2 受试者选择与实验参数设定.....	19
3.3 脑电采集.....	20
3.4 脑电信号预处理.....	22
第 4 章 基于小波变换与多项指标的疲劳驾驶检测应用.....	25
4.1 基于小波变换的信号降噪.....	26
4.2 基于 KPCA 的特征提取.....	27

4.3 基于 LS-SVM 的脑电波数据处理.....	28
4.4 模型与总结.....	30
第 5 章 集成多特征融合与三种不同核分类器的疲劳脑电分类研究.....	31
5.1 基于 EMD 的脑电数据降噪.....	31
5.2 多特征提取.....	33
5.2.1 基于模糊熵的特征提取.....	33
5.2.2 基于样本熵的特征提取.....	34
5.3 模型选择.....	34
5.4 模型分析与评价.....	35
第 6 章 基于独立成分分析降噪与集成分类器的疲劳脑电研究.....	38
6.1 基于 ICA 数据的脑电预处理.....	39
6.2 基于信息熵的特征提取.....	40
6.3 基于 AR 系数的特征提取.....	41
6.4 基于 Adaboost 强分类器分类.....	42
第 7 章 总结与展望.....	44
7.1 本文主要工作与研究结论.....	44
7.2 未来研究展望.....	44
攻读硕士学位期间发表论文.....	53

## 摘 要

随着科技和经济的发展，交通工具种类日渐繁多，随之而来的是交通事故逐年上升，在各种恶性交通事故中，疲劳引发的交通事故的概率高出正常驾驶的 3-5 倍。因此若能够利用 EEG 实现实时检测疲劳状态，对于避免交通事故具有重要意义。为了准确的检测驾驶员疲劳状态，本文主要研究多熵值特征下的疲劳检测模型，通过对疲劳检测过程中的去噪、特征提取和分类等三个主要步骤进行分析研究，建立起三个准确率较高、特征稳定的疲劳检测分析模型：

(1)针对脑电降噪预处理，提出了一种基于小波变换与多项指标的疲劳驾驶检测应用，在小波变换降噪处理的基础上应用主成分分析（kernel principal component analysis ,KPCA）方法进行特征提取，最后应用最小二乘分类器进行分类，实验证明了降噪后的数据分类效果比未进行降噪处理的数据分类结果具有明显优势。

(2)针对融合脑电多特征，提出了一种集成多特征融合与三种不同核分类器的疲劳脑电分类研究方法。对 32 导 EEG 进行预处理，应用经验模态分解（Empirical Mode Decomposition, EMD）去掉异常值和高频噪声情况，然后对处理结果进行样本熵和模糊熵的多特征提取，将特征集合作为样本投入到集成三种不同核的自适应增强（Adaptive boostint, Adaboost）分类器中，实验结果发现在降噪处理结果与分类器选择相同的基础上，多特征提取分类结果明显高于单一特征提取效果。

(3)针对分类器的单一性以及各分类器的优势互补问题，提出了一种基于 AdaBoost 的组合型实验方法用于分析脑电检测疲劳驾驶。试验过程中针对不同受试者采用独立成分分析(Independent Component Correlation Algorithm, ICA)处理分析，继而进行样本熵、信息熵、模糊熵和 AR 系数的特征提取；最后运用 AdaBoost 将最小二乘向量机基于三种核分类器集成为一个强分类器。试验结果显示，采用 AdaBoost 分类器分类效果优于单个核分类器，对疲劳驾驶平均识别率达到 93%，五折交叉验证准确率为 91.04%，在一定程度上拓宽了基于脑电信号的安全驾驶辅助监控系统的研究道路。

**关键词：**脑电信号；小波变换；熵特征；最小二乘；Adaboost；

## Abstract

With the development of technology and economy, the variety of vehicles is increasing day by day, and the traffic accidents are increasing year by year. In various malignant traffic accidents, the probability of traffic accidents caused by fatigue is 3-5 times higher than that of normal driving. Therefore, if EEG can be used to detect fatigue state in real time, it is of great significance to avoid traffic accidents. In order to accurately detect the fatigue state of the driver, this paper mainly studies the fatigue detection model under the multi-entropy feature, and analyzes and studies the three main steps of denoising, feature extraction and classification in the fatigue detection process to establish three accurate Fatigue detection analysis model with high rate and stable characteristics:

(1) For the preprocessing of EEG noise reduction, a fatigue driving detection application based on wavelet transform and multiple indexes is proposed. Based on the wavelet transform noise reduction processing, the kernel principal component analysis (KPCA) method is applied. The feature extraction is carried out, and finally the least squares classifier is used for classification. The experiment proves that the data classification effect after noise reduction has obvious advantages compared with the data classification result without noise reduction processing.classification effect after noise reduction has obvious advantages compared with the data classification result without noise reduction processing.

(2) Aiming at the multi-features of fusion EEG, a method of fatigue EEG classification based on integrated multi-feature fusion and three different kernel classifiers is proposed. The 32-lead EEG is preprocessed, and the abnormal mode and high-frequency noise are removed by using Empirical Mode Decomposition (EMD). Then, the sample results are extracted by sample entropy and fuzzy entropy. The feature set is taken as the sample input. In the adaptive boostint classifier integrating three different cores, the experimental results show that the multi-feature extraction classification results are significantly higher than the single feature extraction results based on the same noise reduction processing results and classifier selection.

(3) Aiming at the singularity of the classifier and the complementary advantages of each classifier, a combined experimental method based on AdaBoost is proposed to analyze the EEG detection fatigue driving. During the experiment, different subjects were analyzed by Independent Component Correlation Algorithm (ICA), and then sample entropy, information entropy, fuzzy entropy and AR coefficients were extracted. Finally, AdaBoost was used to base the least squares vector machine. The three core classifiers are integrated into one strong classifier. The test results show that the classification effect of AdaBoost classifier is better than that of single nuclear classifier. The average recognition rate of fatigue driving is 93%, and the accuracy rate of five-fold

cross-validation is 91.04%, which broadens the safe driving based on EEG signals to some extent. The research path of the auxiliary monitoring system.

**Keywords: EEG signal; Wavelet transform; Entropy characteristic; LSSVM; Adaboost;**

## 第1章 绪论

### 1.1 课题研究背景及意义

疲劳驾驶：是指驾驶人在长时间的连续行车或者在本身身体技能低下的情况下操控车辆，使得客观上出现驾驶技能以及反应速度下降的一种现象。在全球人类死伤事故中，交通事故的伤亡是人类死伤的主要原因之一。2017年5月交通网资料显示，2016年，京沪高速淮安段共发生交通事故787起，其中因疲劳驾驶造成的交通事故四百起以上，约占事故总量52.6%。疲劳驾驶成为触发众多特大交通事故的敏感因素，由于疲劳驾驶引发的交通事故率占到了15-30%<sup>[1]</sup>。因此对驾驶员疲劳状态进行实时监测对降低交通事故，减少伤亡有重要意义。

脑电信号自从1920年被Hans Berger发现以后，一直是脑相关研究的重要工具，例如在心理学上，用ERP成分进行各种心理疾病的检测，例如医学上对癫痫等疾病的检测等等，随着脑电信号研究的深入，如今很多信号分析方法被引入脑电信号特征分析，例如相关<sup>[2]</sup>、AR<sup>[3]</sup>、小波分析<sup>[4]</sup>、STFT<sup>[5]</sup>、ARMA<sup>[6]</sup>、复杂度<sup>[7]</sup>、ICA<sup>[8]</sup>、HMM<sup>[9]</sup>、粒子群<sup>[10]</sup>等，这些脑电信号研究成果也促使很多研究者开始利用脑电信号进行驾驶疲劳检测研究的尝试，例如刘超<sup>[11]</sup>应用了一种经验模态分解下自适应多尺度熵的脑电特征提取算法来区分驾驶员疲劳状态与静息状态下的前额脑电信息，该方法可以为脑电疲劳驾驶检测提供了相当重要的参考价值。方毅<sup>[12]</sup>为克服单一模态驾驶疲劳检测模型的局限性提出了使用多核学习、多集典型相关分析融合多源异构特征,准确率达到89%、93%，搭建的驾驶疲劳检测模型,准确率达到95%。陈骥驰等人<sup>[13]</sup>提出了一种基于脑网络特征的疲劳检测方法，实现了利用较少电极可穿戴EEG设备进行疲劳检测。胡剑锋等人曾用熵特征<sup>[14-15]</sup>进行研究，分类识别率达到98%。

鉴于上述研究背景和研究意义，本文提出了三个基于脑电信号熵特征的疲劳检测模型，主要包括以下三个内容：(1) 针对脑电降噪预处理，提出了集成小波变换与多项指标的疲劳驾驶检测应用；(2) 针对融合脑电多特征，提出了一种集成多特征融合与三种不同核分类器的疲劳脑电分类研究方法；(3) 针对分类器的单一性以及各分类器的优势互补问题，提出了一种基于实验室模拟实验的脑电信号分析，进行疲劳检测。

### 1.2 国内外驾驶疲劳检测研究现状

疲劳驾驶会导致车辆行驶路线不定，速度变化不具有规律性，增大了发生交通事故的概率，同时，驾驶人生理与身体机能下降，反应速度下降，脑力与体力均出现明显下降，操控车辆的能力大幅度下降，发生交通事故得概率极速升高。

随着硬件技术的发展，国外如今很多便携式脑电信号采集设备的面世为车载环境下进行疲劳驾驶检测提供了硬件上的支撑，例如美泰公司的MindFlex的利用脑波



控制立体区域游戏的脑电信号采集设备、以 Nerosky 的研究为基础,通过和美国奥林匹克剑术团队的合作,开发的高尔夫游戏中的高尔夫帽中的脑电信号采集设备、东芝 MindTun 公司的观察大脑对音乐的反应的脑电采集设备、xwave 耳机的脑电采集设备等等。

基于驾驶员行为特征的监测方法主要包括外部环境检测、驾驶员头部运动和面部表情的监测等。利用机器视觉,红外线,雷达传感器和收集和融合信息的其他电子设备,可以收集反映疲劳状态和驾驶意图的特征。基于驾驶员行为特征的疲劳监测技术是最能直观提取驾驶员驾驶疲劳特征参数的方法,具有较高的实验测量准确率,但该方法算法复杂性高,检测程序繁琐并且极易收到外界光照,光线变化以及驾驶员个人生理条件变化,具有较弱的推广应用性。

### 1.2.1 基于车辆运行状态的监测预警

随着车辆日渐进入到人们的生活中,社会对车辆的监管要求越来越严格,需要随时掌握道路上行驶车辆的状态信息,为此,需要远程监控车辆发动机并实现信息管理,以提高车辆的行车安全性。车辆发动机问题的早期预警已成为汽车行业的热点。基于车辆行为特征的疲劳检测是指通过检测驾驶人操作车辆的方向盘、加速踏板、制动踏板等操作行为以及车辆的行驶速度、加速度、车身横摆角度和车道偏移量等车辆行驶相关信息来间接确定驾驶人的疲劳程度<sup>[16]</sup>。

金雪<sup>[17]</sup>收集了 25 名驾驶员的驾驶行为数据,并分析了他们在不同驾驶条件下的驾驶行为特征。提出一种基于加权投票法的融合算法和基于 GA\_BP 神经网络的融合算法,具有较好的识别效果,但前者具有较好的识别效果。

通过对现有研究成果的认真研究,我们发现不同的研究方式有自身的优势,也有不可避免的缺点,如利用车辆行为作为特征的优势在于实时性强<sup>[18]</sup>,对驾驶员无干扰,缺点是由于对道路硬件要求较高,由于车辆行为特征并不能真正反映驾驶员疲劳的生理变化,疲劳识别准确率并不高,仅在 80%上下<sup>[19]</sup>;利用驾驶员行为作为特征的疲劳检测克服了车辆行为中不能反映驾驶员疲劳生理变化的缺点,疲劳识别率达到 90%左右,缺点在于受光照变化、面部表情、复杂背景等因素的影响较大;直接利用驾驶员生理特征作为疲劳检测依据识别准确率最高,检测准确率最高能达到 98%以上,但是由于信号过于敏感,个体差异性较大,因而稳定性较差,效果不佳。

### 1.2.2 基于驾驶员行为特征疲劳的监测方法

通过驾驶人的操作行为如方向盘操作等操作推断驾驶人疲劳状态称之为基于驾驶人操作行为的驾驶人疲劳状态识别技术。该方法以实时捕捉驾驶员的行为状态驾驶员疲劳检测为基础,在此基础上准确判断疲劳状态,并且不能影响驾驶员的正常驾驶状态。

张弛<sup>[20]</sup>采用 OpenCV 中的 Haar 分类器,对驾驶员人脸、人眼进行检测,提出了“双眼检测、单眼疲劳判别”的驾驶疲劳检测方法,从视频图像处理角度深入处理疲劳数据,在一定程度上减少疲劳驾驶引发的交通事故,提高驾驶安全系数。

苏静静<sup>[21]</sup>从 KCF(Kernelized Correlation Filter,KCF)的跟踪算法出发将检测到的人脸区域进行快速跟踪;按照级联回归的方法定位面部关键点再根据关键点位置提取眼睛和嘴部区域;最后采用 CNN (Convolutional Neural Networks, CNN) 网络模型对提取出的眼睛和嘴部区域进行状态识别,计算多个疲劳参数对驾驶员的疲劳状态进行综合检测。

胡神松<sup>[22]</sup>采用平均投影和方差投影的方法构建基于人眼特征的疲劳驾驶检测系统为后续的疲劳判定奠定基础。

目前还没有有效的检测方法,因此对驾驶员疲劳检测技术的研究任务紧迫且前景十分广阔。传统的疲劳检测技术是基于驾驶员自身的感知和与交警巡逻的合作。除此之外也会有道路摄像头实时监测车辆驾驶路线,出现特殊情况进行自动报警等。但随着人均车辆持有数量的增加以及道路的不断加长,仅仅依靠人为力量以及一些基础的检测设备不足以满足日益增长的社会需求。

### 1.2.3 基于驾驶员生理指标的疲劳监测方法

由于上述两种方法的不足,驾驶疲劳监测方法的研究逐渐转向对驾驶员生理指标的研究。基于驾驶员生理指标的疲劳监测方法包括脑电图 (EEG), 眼图 (EOG) 和心电图 (ECG)。通过提取不同频段的脑电图的典型特征来比较脑电图 (EEG) 信号、眼睛信号是指不同时间闪烁的类型和变化心电图信号检测驾驶员心率的变化。口眼状况检测方法是疲劳检测的关键问题,但受光照变化和戴眼镜的影响。为解决这些问题,耿磊等人提出了一种基于面部行为分析的疲劳检测方法。基于 AdaBoost 和 Kernelized Correlation Filter (KCF) 跟踪算法,检测驾驶员的面部。通过级联回归确定特征点,并获得眼睛和嘴部区域。卷积神经网络 (CNN) 用于检测眼睛和嘴巴的状态。在此基础上,计算疲劳试验的疲劳参数。实验结果表明,该方法能够准确地检测出射眼的状态,并在许多情况下更有效地检测疲劳<sup>[23]</sup>。

Stern H S 等人<sup>[24]</sup>总结了关于研究国家科学院,工程和医学研究的商业机动车驾驶员疲劳的数据和方法问题的建议。提供了一个框架,用于识别影响驾驶员疲劳的各种因素,并将驾驶员疲劳与碰撞风险和长期驾驶员健康相关联。还描述了可用于改善对观察数据的相关关系的理解的统计方法。增强数据收集和使用现代统计方法进行因果推断的建议有可能增强我们对疲劳与公路安全和长期驾驶员健康之间关系的理解。

驾驶员疲劳越来越成为交通事故的一个促成因素,因此迫切需要一种自动检测驾驶员疲劳的有效方法。Hu J 等人<sup>[25]</sup>为了捕捉脑电信号的主要特征,计算了四种类型的熵(基于单通道的 EEG 信号)作为特征集,包括样本熵,模糊熵,近似熵和谱熵。所有特征集都被用作梯度增强决策树 (GBDT) 的输入,这是一种快速且高度

精确的增强集合方法。GBDT 的输出根据其 EEG 信号确定驾驶员是否处于疲劳状态。还采用了三种最先进的分类器，k-最近邻，支持向量机和神经网络。试验结果证明这项工作的平均识别率最高可达 94.0%，可满足日常应用的需求。

疲劳会使驾驶员反应时间延长、操作失误增多等，如果能够实时检测疲劳状态，对于避免交通事故具有重要意义。

### 1.3 本文的主要工作与研究内容

本文依据脑电数据处理流程，分别对降噪、特征提取和分类三个方面进行研究，提出了三种实验模型，对比降噪与否对数据分类效果准确性的影响，从单一特征到多特征提取效果对比，从单一分类器进行分类到集成不同核分类器进行分类。另外提出了一种综合降噪、多特征提取以及集成不同核分类器优势的基于 AdaBoost 的组合型实验方法用于分析脑电检测疲劳驾驶。试验过程中针对不同受试者采用独立成分分析 (Independent Component Correlation Algorithm, ICA) 处理分析，继而进行样本熵、信息熵、模糊熵和 AR 系数的特征提取；最后运用 AdaBoost 将最小二乘向量机基于三种核分类器集成为一个强分类器。试验结果显示，采用 AdaBoost 分类器分类效果优于单个核分类器，对疲劳驾驶平均识别率达到 93%，五折交叉验证准确率为 91.04%，在一定程度上推动了基于脑电信号的安全驾驶辅助监控系统的研究<sup>[26]</sup>。主要工作如下：

1. 介绍了传统驾驶疲劳的检测方法，对比现代疲劳驾驶识别方法，简要分析其利弊，综述了国内外相关领域的研究背景，并对脑电应用领域加以总结，提出传统驾驶疲劳检测方法改善的方向以及憧憬新的脑电应用方向。

2. 针对脑电噪声预处理，提出了一种基于小波变换和多指标的疲劳驾驶检测方法。在小波变换降噪的基础上，将主成分分析 (KPCA) 应用于特征提取。最后应用最小二乘分类器进行分类，实验证明了降噪后的数据分类效果比未进行降噪处理的数据分类结果具有明显优势。

3. 针对融合脑电多功能特点，提出了一种融合多特征融合和三种不同核分类器的疲劳脑电分类方法。对 32 通道 EEG 进行预处理，通过经验模式分解 (Empirical Mode Decomposition, EMD) 消除异常值和高频噪声。然后对处理结果进行样本熵和模糊熵的多特征提取，将特征集合作为样本投入到集成三种不同核的 Adaboost 强分类器中，实验结果发现在降噪处理结果与分类器选择相同的基础上，多特征提取分类结果明显高于单一特征提取效果。

4. 针对分类器的简单性和各分类器的优势互补，提出了一种基于 AdaBoost 的联合实验方法来分析脑电疲劳驾驶。在实验过程中，采用独立分量分析 (Independent Component Correlation Algorithm, ICA) 对不同的主体进行处理和分析，然后对样本熵，信息熵，模糊熵和 AR 系数进行特征提取。最后，使用 AdaBoost 将最小二乘向量机集成到基于三种内核分类器的强分类器中。实验结果表明，AdaBoost 分类器的

分类效果优于单核分类器。疲劳驾驶的平均识别率为 93%，五次交叉验证的准确率为 91.04%。它在一定程度上促进了基于脑电信号的安全驾驶辅助监控系统的研究<sup>[26]</sup>。

5. 实验验证了文中三种实验模型的有效性，研究和分析表明了基于降噪基础上的脑电数据处理有效的增加了脑电研究疲劳与静息的分类准确率，能够有效进行分类，为更加准确的预测疲劳、预警疲劳都打下了良好的基础；除此之外，本文提出的创新性基于 AdaBoost 的组合型实验方法集成三种不同核分类器的方法在分析脑电检测疲劳驾驶具有相当高的准确性和相当大的效率优势。

## 1.4 本文的组织结构

本文的章节安排如下：

第 1 章为绪论，介绍本文所选课题的国内外研究现状及其研究意义，简要介绍了疲劳驾驶的危害，说明了疲劳检测的刻不容缓，并介绍了传统疲劳驾驶检测的利弊以及提出了改进的新方向。

第 2 章为相关技术的介绍，主要阐述了基于脑电的疲劳识别的主要降噪、特征提取以及分类的基本处理流程以及在该流程中常用的信号处理技术，主要从常用的四种脑电处理分析方法：时频分析、高阶谱分析、人工神经网络和非线性动力学分析着手，对四种处理方法进行介绍，并将国内外对应处理方法的应用结果加以展示，同时总结脑电研究的相关成果并提出脑电研究在当前社会的应用前景。并总结了论文中主要应用的 LSSVM 与 Adaboost 分类的技术。

第 3 章说明了实现实验的程序选择与数据收集，首先对目前应用比较多的程序实验软件进行介绍，国际脑电软件处理平台主要有包括 BESA、ASA、EMSE、CURRY 等各具特色，在一定程度上推动了国际脑电的研究，但其引进成本较高，适应性不强，不能融合创新方法，国内也有基于 Neuroscan、matlab、Python 等开发的程序对脑电信号进行处理，本文通过 matlab 编程实现，通过引入不同特征对已有算法进行改进，并通过实验验证。

第 4 章为针对脑电降噪预处理，提出了一种基于小波变换与多项指标的疲劳驾驶检测应用，在小波变换降噪处理的基础上应用主成分分析（KPCA）方法进行特征提取，最后应用最小二乘分类器进行分类<sup>[27]</sup>，实验证明了降噪后的数据分类效果比未进行降噪处理的数据分类结果具有明显优势。

第 5 章针对融合脑电多特征，提出了一种集成多特征融合与三种不同核分类器的疲劳脑电分类研究方法。对 32 导 EEG 进行预处理，应用经验模态分解（Empirical Mode Decomposition, EMD）去掉异常值和高频噪声情况，然后对处理结果进行样本熵和模糊熵的多特征提取，将特征集合作为样本投入到集成三种不同核的 Adaboost 强分类器中，实验结果发现在降噪处理结果与分类器选择相同的基础上，多特征提取分类结果明显高于单一特征提取效果。

第 6 章为针对分类器的单一性以及各分类器的优势互补问题，提出了一种基于 AdaBoost 的组合型实验方法用于分析脑电检测疲劳驾驶。试验过程中针对不同受试

者采用独立成分分析(Independent Component Correlation Algorithm, ICA)处理分析, 继而进行样本熵、信息熵、模糊熵和 AR 系数的特征提取; 最后运用 AdaBoost 将最小二乘向量机基于三种核分类器集成为一个强分类器。试验结果显示, 采用 AdaBoost 分类器分类效果优于单个核分类器, 对疲劳驾驶平均识别率达到 93%, 五折交叉验证准确率为 91.04%, 在一定程度上推动了基于脑电信号的安全驾驶辅助监控系统的研究<sup>[26]</sup>。

第 7 章为总结与展望, 将论文的主要工作和研究成果进行了归纳总结, 并针对系统设计实现中发现的不足对提出了改进的方向和未来进一步深入研究方向。



## 第2章 相关技术背景及分析方法介绍

大脑的复杂结构决定着其具有复杂的网络,所产生得脑电信号(EEG)是大脑皮层神经元细胞群自发性节律性的电生理活动,其频率变动范围在每秒1—30次之间的,可划分为四个波段,即 $\delta$  (1—3Hz)、 $\theta$  (4—7Hz)、 $\alpha$  (8—13Hz)、 $\beta$  (14—30Hz),它们分别从不同方面反映了大脑的生理状态和活动规律,是揭示大脑功能本质和信息传递处理方式的主要手段,基于此,目前脑电的研究也越来越丰富。

脑电研究学科日渐丰富,从单一集中于医学、神经方面等生理的研究扩展到涉及神经病学、外科学、生物医学、心理学、信号处理、计算机软件与计算机应用技术甚至少数应用到到驾驶疲劳检测等领域;脑电的研究方式也从单一的波形描述、脑电地形图到多形态分析并取得了卓越的成就。脑电信号分析方法主要有时域分析、频域分析、双谱分析、时频分析、非线性分析等。常用的特征提取方法有:自回归模型(AR模型)、功率谱密度估计、小波变换,混沌法、共空间模式、新型描述,多维统计分析,提取特征主要有时域特征,频域特征和非线性动力学特征。由于 EEG 信号的非线性非平稳及频域特征较为突出的特点,分析比较适用非线性、时域分析方法,近年来,波形特征描述、自回归 AR 模型、傅里叶变换、功率谱密度、双谱分析、小波变换、熵、人工神经网络、非线性动力学分析等脑电分析处理方法得到了深入研究<sup>[27]</sup>。

### 2.1 脑电伪迹去除方法

#### 2.1.1 主成分分析

主成分分析(Principal Component Analysis, PCA)通过将高维空间的问题转化到低维空间处理,从而使问题简易化、明确化,并且这些综合指标之间不具有互相关性,同时又可以提供原有指标包含的大部分信息<sup>[27]</sup>。作为一种由线性到非线性之间的桥梁, KPCA 核方法是通过引入核函数把非线性变换后的特征空间内积运算转换为原始空间的核函数计算,从而缩减计算量。

实验脑电特征参数样本集合为:

$$X = \{\mathbf{x}_1, \dots, \mathbf{x}_m\} \quad (2-1)$$

其中  $m$  为维数,这里  $m=8N$ 。KPCA 方法将数据隐式映射到高维线性可分空间。样本在特征空间中的内积可以用一个核函数来计算:

$$k(\mathbf{x}, \mathbf{y}) = (\phi(\mathbf{x}))^t \phi(\mathbf{y}) \quad (2-2)$$

核函数通过映射  $\Phi$  把样本集合中的数据  $x$  映射到  $F$ , 计算特征空间中的协方差矩阵为公式(2-3):

$$\bar{C} = \frac{1}{m} \sum_{i=1}^m \phi(\mathbf{x}_i) \phi'(\mathbf{x}_i) \quad (2-3)$$

在协方差公式中令  $\lambda$  为  $\bar{C}$  的特征值， $\mathbf{V}$  为对应的特征矢量，使得

$$\mathbf{v} = \sum_{i=1}^M \alpha_i \Phi(x_i) \quad (2-4)$$

上式左乘  $\phi'(\mathbf{x}_j)$  即有

$$\lambda [\phi'(\mathbf{x}_j) \cdot \mathbf{V}] = \phi'(\mathbf{x}_j) \cdot \bar{C} \mathbf{V}, \quad j = 1, \dots, m \quad (2-5)$$

则特征样本在特征矢量上的投影即为主元

$$\lambda \sum_{\mu=1}^M (\Phi(X_\nu) \bullet \Phi(X_\mu)) \quad (2-6)$$

### 2.1.2 独立成分分析

假设每一个数据源  $S_i$  由密度函数  $P_s$  给出，数据源  $S$  的联合分布：

$$p(s) = \prod_{i=1}^n p_s(S_i) \quad (2-7)$$

需要注意的是，上面式子之所以成立是是基于数据源是彼此独立这一假设基础上的。

根据上面讨论的密度函数和线性变换的关系以及根据公式(2-8)

$$\mathbf{x} = A\mathbf{s} = W^{-1}\mathbf{s} : \quad (2-8)$$

得到公式(2-9)

$$p(x) = \prod_{i=1}^n p_s(w_i^T x) \cdot |W| \quad (2-9)$$

对于一个连续实数的随机变量  $Z$ ，它的分布函数  $F$  的定义：

$$F(z_0) = P(z \leq z_0) = \int_{-\infty}^{z_0} p_z(z) dz \quad (2-10)$$

所以变量  $Z$  的密度函数就是它的分布函数的导数：

$$p_z(z) = F'(z) \quad (2-11)$$

为了找到一个具体的  $S$  的密度函数  $p_s(s)$ ，只要能找到它的分布函数即可。分布函数是由 0 到 1 的单调递增函数，由前面的讨论知道这个分布函数不能取高斯分布函数。

通常取公式(2-10)的 sigomid 函数：

$$g(s) = 1/(1 + e^{-s}) \quad (2-10)$$

作为默认的概率分布函数，计算得到公式 2-11

$$p_s(s) = g'(s) \quad (2-11)$$

方阵  $W$  是模型的参数，给定训练集  $\{x^{(i)}; i=1, \dots, m\}$ ，那么参数的似然函数：

$$\ell(W) = \sum_{i=1}^m \left[ \sum_{j=1}^n \log g'(w_j^T x^{(i)}) + \log |W| \right] \quad (2-12)$$

找到最大化这个参数  $W$  的似然函数，利用公式 2-13 得到

$$\nabla_w |W| = |W| (W^{-1})^T \quad (2-13)$$

推导出参数在随机梯度下降算法中的更新规则，对于训练样本  $x^{(i)}$ ：

将  $W_i$  写成行向量形式。那么得到：

$$W := W + \alpha \begin{pmatrix} \begin{bmatrix} 1 - 2g(w_1^T x^{(i)}) \\ 1 - 2g(w_2^T x^{(i)}) \\ \dots \\ 1 - 2g(w_n^T x^{(i)}) \end{bmatrix} x^{(i)T} + (W^T)^{-1} \end{pmatrix} \quad (2-14)$$

其中  $\alpha$  是梯度上升速率。当迭代求出  $W$  后，便可得到公式(2-15),并通过该公式还原出原始信号。

$$S^{(i)} = Wx^{(i)} \quad (2-15)$$

### 2.1.3 小波变换

小波变换升级了 Gabor 的窗口傅里叶变换，优化了傅立叶窗口的局部化思想，改善了傅里叶变换窗口大小不随频率变化的缺点，采用了快速衰减和振荡的函数<sup>[28]</sup>，然后进行缩放和将其移位以获得小波基函数，以便信号可以基于其小波基函数。通过时频分解，基函数在时频相平面上具有可变的时频窗口，其可以适应不同的分辨率。在实验中，基于小波变换的数据积分变换定义为：

$$[W_\psi f](a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \overline{\psi\left(\frac{x-b}{a}\right)} f(x) dx \quad (2-16)$$

通过公式(2-17)计算其小波系数

$$C_{jk} = [W_\psi f](2^{-j}, k2^{-j}) \quad (2-17)$$

设平方可积的实数空间  $\psi(t)$ , 当  $\psi(t)$  满足公式(2-18)时：

$$C_{\Psi} = \int_R \frac{|\psi(t)|^2}{|w|} dw < \infty \quad (2-18)$$

称  $\psi(t)$  为一个基本小波或母小波，将母小波函数  $\psi(t)$  经伸缩和平移后，则得到一个小波序列<sup>[29]</sup>：

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad a, b \in R, a \neq 0 \quad (2-19)$$

其中  $a$  为伸缩因子， $b$  为平移因子。

对于任意的函数的连续小波变换为：

$$W_f(a, b) = \langle f, \psi_{a,b} \rangle = \frac{1}{\sqrt{|a|}} \int_R f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (2-20)$$

其逆变换如公式(2-21)：

$$f(t) = \frac{1}{C_{\Psi}} \int_{R^+} \int_R \frac{1}{a^2} W_f(a, b) \psi\left(\frac{t-b}{a}\right) da db \quad (2-21)$$

当通过小波变换处理信号时，应选择适当的小波函数来分解信号。其次，应该通过阈值处理分解的参数，并选择适当的阈值进行分析。最后，使用处理后的参数通过逆小波变换重建信号<sup>[30]</sup>，以获得变换后的新数据。

## 2.2 脑电特征提取常用方法

### 2.2.1 时域分析

较早应用的 EEG 分析方法是从时域中直接提取出有用波形特征的信息，其中 AR 模型、方差分析 (ANOVA)、波形参数分析和波辨识、直方图分析 (Histogram)、相关分析 (CA)、峰值检测等都是使用较多的时域分析方法。

由于脑电信号在频域往往比在时域更加简单直观，所以大多数研究是在频域下进行的。频谱分析、倒频谱分析、包络分析、阶比谱分析以及全息谱分析等方法是使用较多的频域信号分析处理方法。近年来，吴玉鹏<sup>[31]</sup>用功率谱中 AR 谱技术和 FFT 技术进行正常人的脑电图和癫痫病人痫性发作脑电图的比较，说明 AR 谱性能优于 FFT 功率谱，为临床诊疗癫痫提供技术上的支持。郝冬梅<sup>[32]</sup>等人将 6 阶 AR 模型拟合 EEG 信号，以 LVQ 神经网络作分类器，通过比较网络选择不同参数对分类正确率的影响得到竞争层神经元数目直接影响了正确率的这一结论。冯春辉<sup>[33]</sup>发现在不进行主成分分析的基础上支持向量机 (SVM) 与 CSSD 的组合识别正确率最高达 74%。还有众多研究者对其进行深入研究并在一定程度上取得了相应的成果。

然而时域和频域分析方法，主要是用于平稳信号的分析。对于非平稳、非线性信号的分析，应用时域和频域时都难以提取出有效的特征信息。时频分析方法，弥

补了时域和频域分析方法在非线性信号分析的缺陷，它可以准确的提取到非线性信号随时间变化的特征信息。目前，在信号处理的工程领域和众多学科当中，时频分析的技术已被大范围应用，比如语音分析、图像识别<sup>[34]</sup>、机械设备故障诊断<sup>[35]</sup>、生物工程等，在处理脑电信号过程里，脑电信号特征的多分辨率分析、波形辨识、诱发电位特征提取、信号的预处理等应用是时频分析的主要涉及领域。

随着信息科技与计算机技术的飞跃式发展，信号处理与运用发展突飞猛进。时频分析是现代信号处理中最重要的处理方法之一，提供频域和时域的联合分布信息是其作为一种新兴的信号处理方法的一大特征，清晰地阐明信号频率因时间变化而变动的关系，近年来受到越来越多的重视。

### 2.2.2 频域分析

EEG 信号是大脑皮质中大脑神经元电活动的一般效应。它很容易受到主观因素的影响,如心理活动和声光刺激等客观因素,因此,EEG 具有高度随机性和不规则波形,这特征增加了 EEG 时域分析的难度。

相对稳定的脑电功率谱可以揭示脑电图中隐含的一些病理信息。因此,频域分析得方法成为脑电图临床应用相当重要的方法。频域信号分析有两种:首先,傅立叶变换分析确定性信号,分析频谱信息;其次,随机信号的功率谱和自相关函数是傅立叶变换对(Wiener-Sinichin 定理)。

#### (1) 傅里叶变换

傅里叶变换是一种特殊的积分变换,它将时域信号分解为不同频率的正弦或余弦函数之和其分解为幅度和频率分量,广泛应用于物理,信号处理,密码学,声学,光学等领域,其四种形式及特征如表 2-1。

表 2-1 傅里叶变换的四种形式及其特征  
Table2-1 Four Forms of Fourier Transform and Their Characteristics

变换	时间	频率
连续傅里叶变换	连续, 非周期性	连续, 非周期性
傅里叶级数	连续, 周期性	离散, 非周期性
离散时间傅里叶变换	离散, 非周期性	连续, 周期性
离散傅里叶变换	离散, 周期性	离散, 周期性

#### (2) 功率谱分析

功率谱密度可分为经典谱分析和现代谱分析。经典方法基于傅里叶变换,主要包括相关图法和周期图法,以及在此基础上的改进方法。在静电分析脑电图的平均光谱特性时,传统的 FFT 周期图估计方法得到了广泛的应用。但是这种方法是一种低分辨率的谱估计方法,估计的功率谱很难与信号的实际功率谱相匹配,为了快速分析脑电信号并找到并准确描述神经系统损伤,人们一直在努力获得对脑电信号及



其变化的进一步定量描述和估计。

功率谱研究信号的能量（或功率）的频率分布，并突出信号频谱中的主频率，应用领域在驾驶员疲劳状态脑电特征、事件、活动对青年女子的脑电影响、环境致使脑电变化等方面应用较为广泛。常用的频域信号处理方法包括：频谱分析、倒频谱分析、包络分析、阶比谱分析、以及全息谱分析。功率谱和傅里叶变换等方式吸引较多学者研究，应用也较为集中，而谱分析正处于逐步完善阶段。

### 2.2.3 时频分析

时域和频域分析方法主要用于静态信号分析。然而，当需要进行对于非平稳和非线性信号进行分析时，时域和频域中分析难以提供有效的特征信息，成为了实际操作分析过程得阻碍。联合时频分析（JTFA）弥补了时域和频域分析方法的缺点。目前，时频分析技术已广泛应用于信号处理和工程领域的各个领域，如语音分析，图像识别，生物工程，机械设备故障诊断等<sup>[36]</sup>。

### 2.2.4 非线性动力学分析

随着科学技术的腾飞发展,非线性问题在许多学科之中显现身影,非线性时间序列是非线性系统中呈现的和随机信号相像的一种确定性的信号。

李冬梅<sup>[37]</sup>用 EMD 的方法分解癫痫脑电信号，并运用 ICA(Independent Component Correlation Algorithm, ICA)降维 Hurst 指数、小波熵、Lyapunov 指数、排列熵、样本熵这五个非线性动力学指标，对发作时间做出预测，为治疗癫痫患者提供了的理论基础,推进了癫痫患者治愈的进程。李树春<sup>[38]</sup>等人将非线性动力学理论通过特征值用以对青年人和老年人脑电的差异进行分辨，实验发现青年人脑功能与老年人在能量、发放量等方面的差异。除此以外，相关维数、最大李雅普诺夫指数、复杂度、近似熵等，提取 EEG 脑电信号的相应特征也在文献中有较多讨论，印证了非线性动力学混沌算子分析 EEG 信号的应用具有一定科学基础。

学者们试图通过统计叠加，频谱分析和相关分析来分析和处理这些信号，并在一定程度上取得了相应的结果。然而，从另一个角度来看，这些方法基于线性，并且适用于稳定性的平衡线性系统分析方法将导致非线性信号分析的精度损失。随着非线性学科的迅速发展，许多学者对传统的线性方法不满足于处理这种时间序列，并致力于非线性分析方法。非线性动力学的发展为学者解密大脑和脑电图提供了一种新的有效途径。Lyapunov 指数，复杂性度量和分形维数是用于研究脑电图的主要非线性方法。

## 2.3 脑电分类常用方法

### 2.3.1 Fisher 线性判别

Fisher 线性判别就是把样本都投影到一个方向，然后在这个一维空间中确定一个分类阈值，而 Fisher 要找的就是这个投影方向。衡量投影方向的标准：选择投影方向，使投影后两类相邻尽可能远，同时每一类内部的样本又尽可能近。

针对二分类问题，我们将数据投影到一维，通过调整权向量，使类别之间分开最大。

训练样本每一个样本是一个  $d$  维向量，即  $x_i \in R^d$ ，我们将寻找一个投影方向  $\omega$  使投影后的样本为

$$y_i = \omega^T x_i \quad (2-22)$$

在原来的样本空间中，类别均值为

$$m_i = \frac{1}{N_i} \sum_{x_j \in x(i)} x_j \quad (2-23)$$

式中  $i=1,2$  表示两个类别， $x(i)$  表示第  $i$  类的集合。定义各类内离散度矩阵为

$$S_i = \sum_{x_j \in x(i)} (x_i - m_i)(x_j - m_i)^T \quad (2-24)$$

总的类内离散度矩阵为  $S_w = S_1 + S_2$ ，类间复杂度矩阵定义为：

$$S_b = (m_1 - m_2)(m_1 - m_2)^T \quad (2-25)$$

投影以后的一维空间里，两类的均值中将  $x_i$  用  $y_i$  替换，得到  $m_i' = \omega^T m_i$  而类内复杂度此时是一个值

$$S_i'^2 = \sum_{y_i \in y(i)} (y_i - m_i')^2 \quad (2-26)$$

类内总复杂度为  $S_w' = S_1' + S_2'$ ，而类间复杂度就变成了两类均值差的平方

$$S_b' = (m_1' - m_2')^2 \quad (2-27)$$

为了让类内尽可能聚集，而类间尽可能离散，将这一目标表述为如下 Fisher 准则函数

$$\max J_F(w) = \frac{S_b'}{S_w'} \quad (2-28)$$

$$\begin{aligned} S_b' &= (m_1' - m_2')^2 = (w^T m_1 - w^T m_2)^2 \\ &= w^T (m_1 - m_2)(m_1 - m_2)^T w = w^T S_b w \end{aligned} \quad (2-29)$$

以及同理  $S_w' = w^T S_w w$

因此 Fisher 判别准则为

$$\max_w J_F(w) = \frac{w^T S_b w}{w^T S_w w} \quad (2-30)$$

由于  $w$  的幅值不会影响  $w$  的大小, 不会影响  $JF$  函数的值, 因此可以将分母常数化而优化分子最大, 把问题建立约束条件, 通过 Lagrange 乘子法优化为无约束极值问题, 引入拉格朗日乘子  $\lambda$

$$L(\omega, \lambda) = \omega^T S_b \omega - \lambda(\omega^T S_w \omega - c) \quad (2-31)$$

上式极值满足

$$\frac{\partial L(\omega, \lambda)}{\partial \omega} = 0 \quad (2-32)$$

求解上式得到

$$\lambda \omega^* = S_w^{-1} (m_1 - m_2) (m_1 - m_2)^T \omega^* \quad (2-33)$$

式中  $\omega^*$  为极值解, 其中  $S_w$  假定了其为非奇异矩阵, 当样本数量大于特征维数是, 一般容易满足此限制。由于公式 2-30 为标量, 不影响其方向, 而我们要找的权值向量只考虑其方向

$$(m_1 - m_2)^T \omega^* \quad (2-34)$$

所以取公式(2-35)为 Fisher 线性判别的最佳投影方向:

$$\omega^* = S_w^{-1} (m_1 - m_2) \quad (2-35)$$

### 2.3.2 随机森林

设原始样本集为下集合

$$T = \{(x_{i1}, x_{i2}, \dots, x_{iM}, y_i)\}_{i=1}^N \quad (2-36)$$

原始样本集的样本数为  $N$ , 其中  $x_{i1}, x_{i2}, \dots, x_{iM}$  为样本集的特征属性共  $M$  个,  $y_i$  为分类标签,  $y_i$  取值为  $c \geq 2$  个, 代表  $c$  分类。图 2-1 为随机森林生成森林和决策过程示意图。

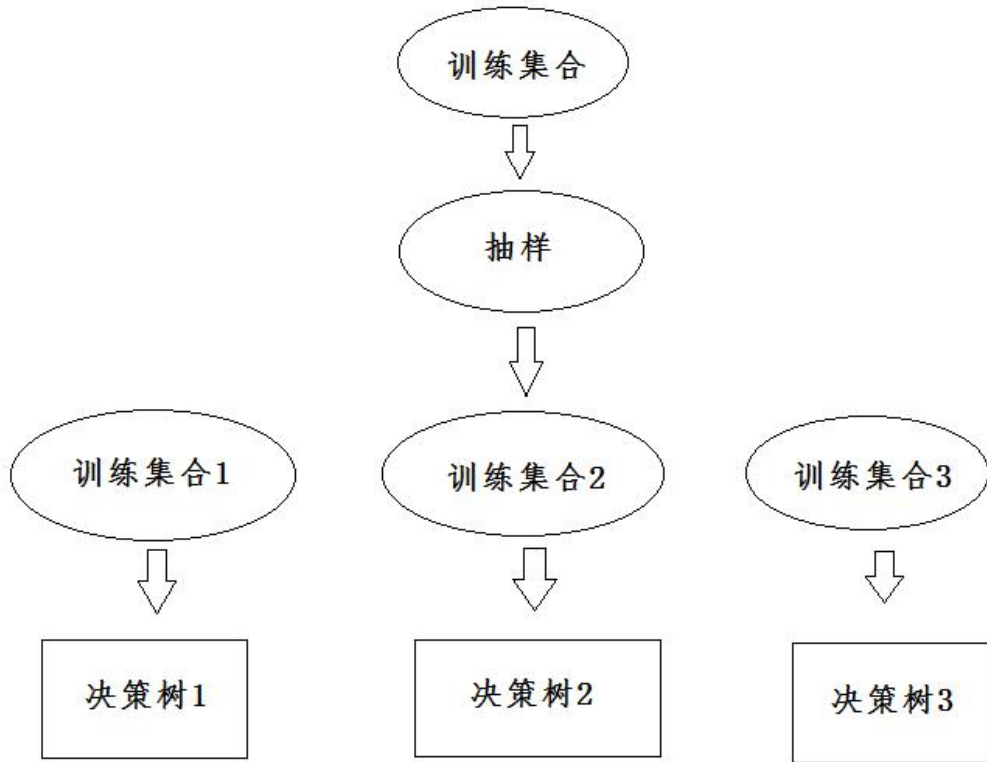


图 2-1 随机森林生成森林

Fig 2-1 Random forest-generated forest

随机森林实际上是一种特殊的 bagging 方法，首先，bootstrap 方法用于生成  $m$  个训练集。然后，对于每个训练集，构建决策树。当节点找到要拆分的特征时，不是要找到可以使索引最大化的所有特征（例如信息增益），而是从特征中随机提取某些特征并在提取的特征中找到最优解，这可以应用于节点拆分。由于整合的想法，随机森林法相当于采样样本和特征，因此可以避免过度拟合。

### 2.3.3 人工神经网络

人工神经网络（ANN，Artificial Neural Network）是实现非线性分析处理，能进行复杂逻辑操作的一个理论化数学模型。其具有快速寻找优化的能力，但化学能力不够强，很难找到通用的模型，而且容易进入局部是其典型的不足处理，但其应用简单，分类结果也比较准确的优势，许许多多的学者仍将其应用于电学的众多研究当中，但在众多研究问题中，由于一些假象信息或者信息内容不完整，决策规则时而相互冲突、互相矛盾，时而无则可依，都是传统的信息处理方式的巨大挑战。而人工神经网络具有自动诊断、问题求解，具有模拟或替代与人的思维有关的优势功能，能给出合理的识别与判断以解决传统方法难以实现的难题。

高群霞<sup>[39]</sup>等人对比了 Hilbert-Huang 变换和 WT 这两种常用的脑电特征提取方式，

以及 SVM、ANN 两类模式识别方法在睡眠分期中的应用，归纳了近几年这一领域的发展趋势走向和研究近况。初孟<sup>[40]</sup>等人运用 Jensen 函数与时频分析相结合的方法进行棘波的波形特征提取，结合 ANN 进一步降低棘波检测的错误率,这一方法仿真实验应用中取得了较有成效的结果。

### 2.3.4 集成学习

通过构建和组合多个学习者来完成学习任务的方法是集成学习，也称为多分类器系统。常见的集成策略包括 boosting、bagging、stacking 等。

#### (1) Boosting 策略

Boosting 是一种可以将弱学习器提升为强学习器的算法，其实现方法是通过一个弱学习器学习训练数据，通过对训练数据分布的调整，提高错误学习的样本的权值比重，使得新的学习器实现更好的学习效果，如此重复训练多个学习器，调整各个分类器权值比重最后将所有学习器的结果结合。其中最为代表性的算法是 Adaboost。

Adaboost 可以表示为“加性模型”，通过多个基学习器的线性组合来最小化指数损失函数。

$$H(x) = \sum_{t=1}^T \alpha_t h_t(x) \quad (2-37)$$

$$\ell_{\text{exp}}(H | D) = IE_{x \sim D}[e^{-f(x)H(x)}] \quad (2-38)$$

#### (2)Bagging 策略

在集合算法中，Bagging 方法形成一类算法，其在原始训练集的随机子集上构建黑盒估计器的若干实例，然后聚合它们的各个预测以形成最终预测。这些方法用作减少基本估计器的方差的方法，方法是将随机化引入其构造过程，然后从中构建整体。由于它们提供了减少过度拟合的方法，因此对于通常最适合弱模型（例如，浅决策树）的增强方法，该方法对于强而复杂的模型最有效。

Bagging 方法有很多种，但是它们通过绘制训练集的随机子集的方式大多不同：

当数据集的随机子集被绘制为样本的随机子集时，该算法被称为 Pasting。当用替换绘制样本时，该方法称为 Bagging。

#### (3)Stacking 策略

Boosting 和 Bagging 策略中所训练的子学习器都为相同类型学习器，学习器差异越大，越可以增加集成学习器的泛化能力，较为简单的结合方法是平均法和投票法，在相同的训练集上训练多个不同类型的学习器，并在此基础上将所有学习器的结果进行加权平均或者投票，其中加权平均的权值由人为设置，难以达到很好的结合效果。

Stacking 是一种能力强大的策略，它使用一个学习器将其他所有学习器结合在一起。Stacking 先从数据集中学习初级学习器，然后将各个初级学习器的输出作为



输入再训练一个学习器，从而得到一个最终的结果。

### 第3章 实验数据采集

基于脑电的疲劳检测其基本流程首先采集数据，脑电信号幅度较为微弱，仅在 $1\mu\text{V}$ - $200\mu\text{V}$ 之间，这些干扰主要源于受试者自身其他身体部位的生理活动和脑电收集装置运行产生的噪声，因此在数据采集中设备以及环境要求较为严格，本文通过受试者在单调环境下驾驶诱发驾驶疲劳，从而进行脑电信号收集，将收集的脑电信号进行预处理，值得提及的是，本文针对改进脑电降噪、特征提取和分类三个方面分别提出了分析模型，并取得了不错的效果，图3-1为主要流程图。

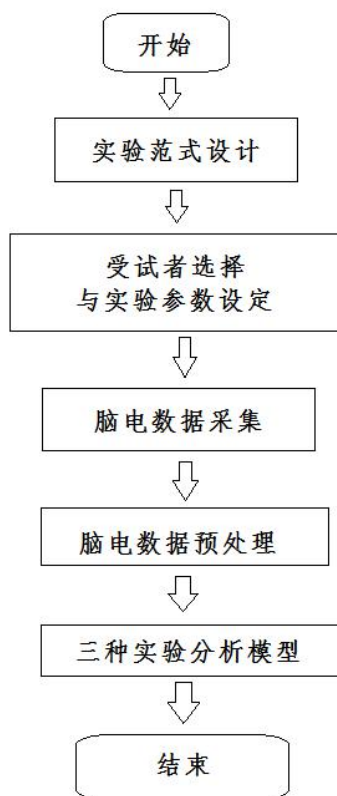


图3-1 实验流程图

Fig 3-1 flow chart of experiments

#### 3.1 实验范式

本实验属于疲劳驾驶实验，在数据采集方面选取状态良好的高校学生，参照国内车辆持有男女比例对受试者进行进行男女比例筛选，确定男士多于女士并确定各位受试者在实验前对各位受试者进行装置使用培训以确保数据收集高效保证质量。

关于设备和场景的选择，选用北京中宇有限公司生产的静态驾驶模拟器进行驾驶疲劳仿真实验，该设备结合计算机软件数据，虚拟如晴天，雾天或雪山，高速公路，农村等不同的驾驶环境，还原最真实的驾驶环境。本次实验选择的驾驶环境为

交通密度低的公路，较容易诱发单调驾驶。研究表明，单调的驾驶环境中大脑更容易陷入疲劳状态，静息与重度疲劳信号特征区分明显，且脑电信号更加稳定，更有利于下一次的数据记录。

关于受试者的选择，22 名大学生（14 名男性，19~24 岁）参加了这个实验。所有受试者被要求在实验前和实验过程中禁用任何类型的刺激，如酒精，药物或茶。实验前，受试者练习几分钟的驾驶任务，熟悉实验的程序和目的。这项工作得到了学校的认可。此外，所有受试者根据本研究中的人类研究方案提供书面知情同意书。

在实验过程中，首先根据受试者头部大小挑选好适应受试者的 32 导电极帽子并给受试者佩戴好电极帽。在完成导电膏的注射任务时，先进行受试者面部去油去角质，然后接触电极部分注射接地电极位置，其他按从前到后的电极顺序依次注射完成，在将受试者全部电极以及参考电极注射完毕后连接实验设备让受试者通过虚拟驾驶舱模拟真实的公路驾驶环境。由于受试者百分之八十以上具有车辆驾驶培训或者驾车经验，所以能够在较短的时间内对虚拟驾驶的操作达到一个较熟悉的状态，这里我们希望受试者按照交通规则进行形势，并且保持行车速度在 40km/h，图 3-2 为实验预备

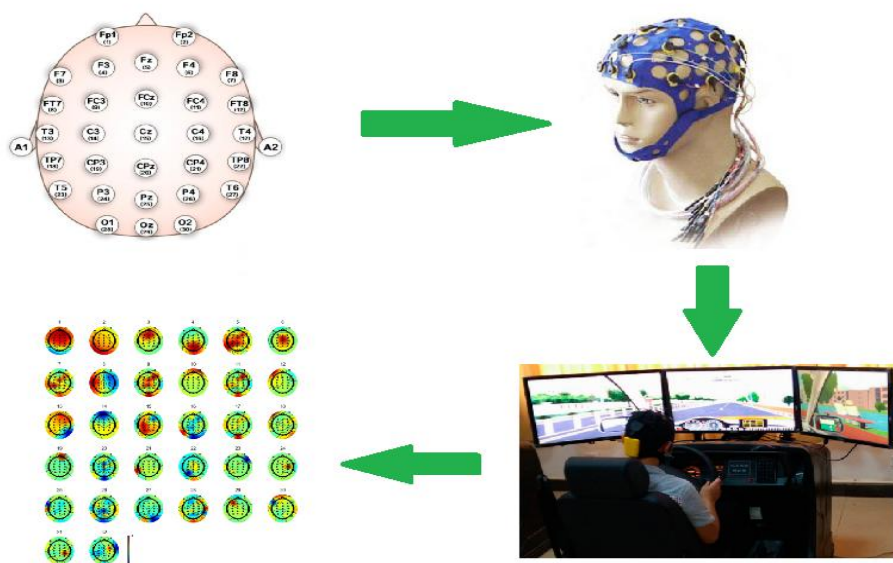


图 3-2 实验预备

Fig 3-2 Experimental preparation

### 3.2 受试者选择与实验参数设定

本文在数据选取方面参照国内车辆持有男女比例对受试者进行进行男女比例筛选，确定男士多于女士选取状态良好的 10 名在校学生，年龄在 19-25，并确定各位受试者在实验前保证休息以及无献血、嗜酒、熬夜等其他影响数据准确性的行为，在静态驾驶模拟器（北京中宇有限公司生产的 ZY-31D 型汽车驾驶模拟器）上进行驾驶疲劳模拟实验。同时本次试验得到了学校的许可并与各位受试者签订书面同意

书以及填写主观疲劳调查表辅助实验数据的准确性，由于脑电信号幅度较为微弱，仅在 1 $\mu$ V-200 $\mu$ V 之间，在收集分析过程中极易受到自身身体部位生理活动或者脑电收集装置运行的干扰降低数据真实性，因此在实验正式开始前对受试者进行数据收集装置的使用培训，以确保数据的真实性，提高数据的质量和代表性。

实验的总时长是 40-130 分钟。第一步是熟悉模拟软件，然后持续单调驾驶，直到驾驶员疲劳被确定并且实验终止。

当驱动持续 10 分钟时，最后 5 分钟的脑电信号记录为正常状态。当连续驾驶持续 30~120 min（直到自我报告的疲劳问卷结果显示受试者处于驾驶疲劳状态），服从 Borg 疲劳量表和 Lee 的主观疲劳量表，最后 5 分钟的脑电信号被标记为疲劳。EOG 也被用来分析眨眼模式作为验证疲劳状态的客观部分。关于疲劳状态验证是基于受试者填写的自我报告的疲劳调查问卷斯坦福嗜睡量表。

表 3-1 斯坦福嗜睡量表

Table 3-1 Stanford Sleepiness Scale( SSS)

等级	瞌睡程度
1	感觉非常振奋、充满生机与活力
2	身体机能处于较高水平，不过不在峰值，可以集中注意力很清醒，不过身体及思维都比较放松
3	能及时反映但不够灵敏
4	有些倦意、松懈
5	充满倦意，不再想保持清醒，非常松懈，开始打瞌睡，头晕眼花，不再与睡意作斗争
6	只想躺下来休息
7	睡眠初期，开始出现梦境 <sup>[41]</sup>

采用问卷调查，以确定疲劳状态的这种方法不仅已经在我们的研究中使用，而且在许多其他研究。所有通道数据均参考 A1 和 A2 处的两个电连接的乳突，以国际 10-20 系统为基础，以 32 通道电极帽（包括 30 个有效通道和两个参考通道）以 1000Hz 进行数字化，并存储在一台电脑上进行离线分析。通过记录水平和垂直 EOG 来监视眼球运动和眨眼。

采集 EEG 信号后，采用 Neuroscan（澳大利亚 Compumedics 公司）的 Scan 4.3 软件进行数据预处理的主要步骤。原始信号首先由 50Hz 陷波滤波器滤波，并使用 0.15-45Hz 带通滤波器。继而将来自 30 个通道的 5 分钟 EEG 信号分成 1 个时期，分成 300 个时期。

### 3.3 脑电采集

本文采用客观监测方法采集脑电信号，针对不同驾驶员进行针对性分析评价，并且辅助主观调查表对所采集数据进行校核，首先，使用 Neuroscan 公司生产的 32 导联采集设备对脑电信号进行采集，并且采用了 10-20 国际标准进行实验。

### (1) 电极帽准备

实验中电极帽选择 Neuroscan 的 Quik-Cap 37 导联电极帽简便、准确、耐用和舒适，采用标准 EEG 电极定位系统，其电极安放位置分别如下图 3.3 所示（其中 A1 为参考电极，A2 为接地电极）：

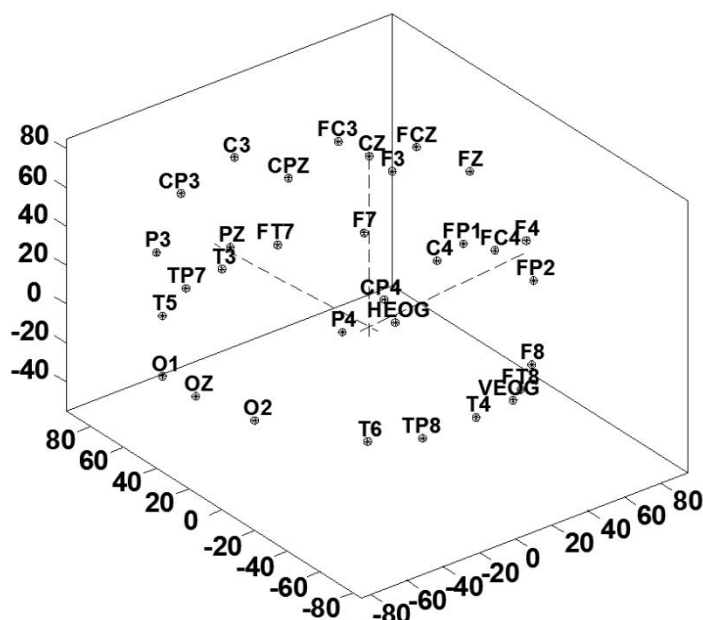


图 3-3 脑电电极 3D 分布图

Fig 3-3 EEG electrode 3D distribution

### (2) 佩戴电极帽

电极帽的佩戴工作首先需要按照受试者头部尺寸进行电极帽尺码选择，尺码包含：新生儿（34-40cm）、婴儿（42-48cm）、儿童或小号（48-54cm）、中号（54-62cm）、大号（62-68cm）。尽可能选择与受试者头部尺寸相匹配的电极帽，以尽量减少受试者在实验中的不适症状。

### (3) 注射导电膏

导电膏是一种类白色粘稠凝胶，根据受试者头围取出适量导电膏放入搅拌杯中，加入适量的水，然后用一次性注射器搅拌均匀并把稀释后的导电膏依次注入已经安置好在受试者头上电极帽的 32 个极点中，从接地电极开始。HEOL 是水平左眼点，贴在受试者左眼角往后一指宽左右的位置；HEOR 是水平右眼点，贴在受试者右眼角往后一指宽左右的位置；VEOL 是垂直下眼电，贴在受试者左眼眶下一指左右的位置；VEOU 是垂直上眼电，贴在受试者眉毛往上一指左右的位置。

### (4) 待阻抗稳定、记录数据

在注射完导电膏之后，让受试者于虚拟驾驶舱内 1-3 分钟熟悉环境以及操作。实验员对受试者进行简要说明，使受试者初步练习遵守交通规则驾驶操作，使受试者全程保持车辆驾驶速度保持在 40km/h，与此同时，实验员在室外观察电极帽电



极与受试者头皮之间阻抗强度的变化情况，针对导电不良的电极进行导电膏补注射以及相应搅动，直到各电极导电良好的时候，实验员指示受试者开始进行模拟驾驶实验，记录所需数据并记录实时时间、天气受试者年纪等信息。

### 3.4 脑电信号预处理

#### 1. Linear Derivation 单极导联合并为双极导联

软件默认的去除眼电的算法为双极导联模式，单极导联合并为双极导联，先载入一个名为 EOG.ldr 的设置文件（软件会附带），设置输出路径，点击 OK 即可。

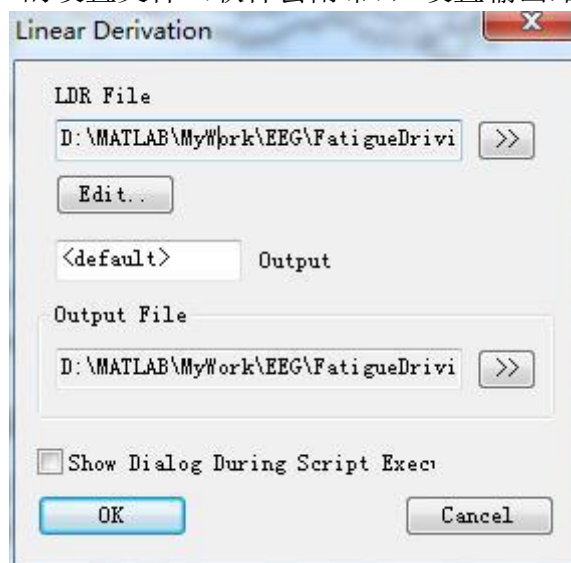


图 3-4 单极导联合并为双极导联

Fig 3-4 Unipolar leads merged into bipolar leads

#### 2. DC offset correction DC 校正

DC 校正，去除直流漂移，默认参数，直接确认。

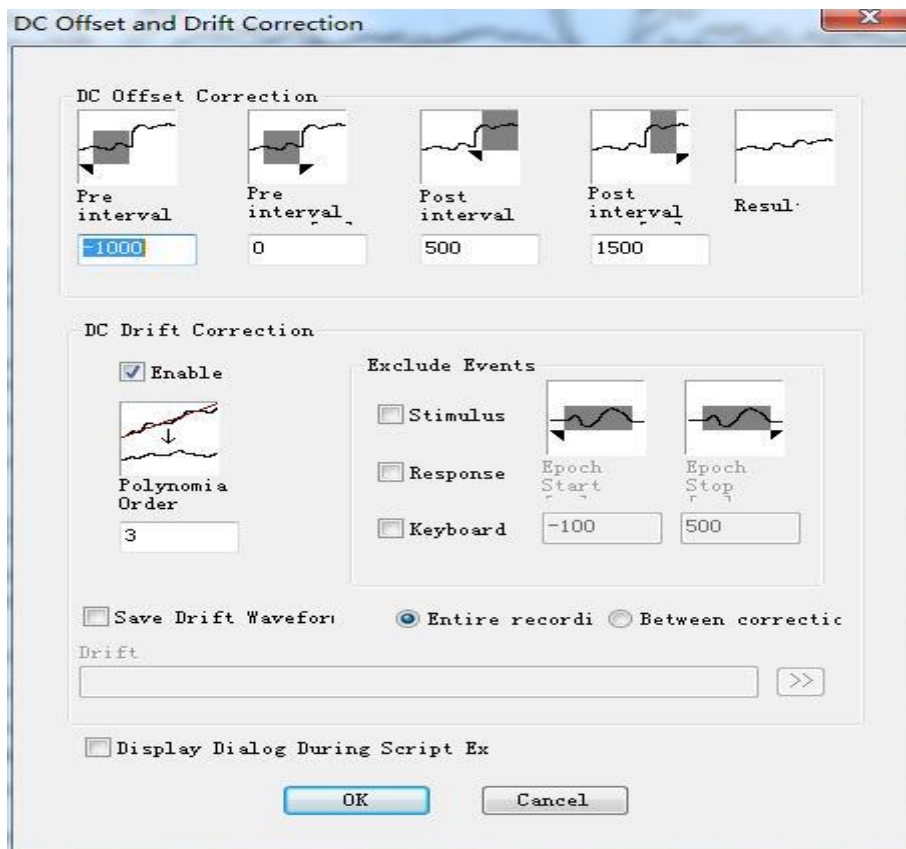


图 3-5 DC 校正

Fig 3-5 DC correction

### 3. Ocular Artifact Reduction 去除眼电

首先点击工具栏 (Mark a block) 命令，分别在波形漂移比较明显的波动起始和结束位置单击鼠标左键，选择 Reject Block；去除眼电需参数设置如下图所示，依据眼电波形选择 Positive (正波) 或 Negative (负波) 并点击 Channel 选择所对应的导联进行操作。

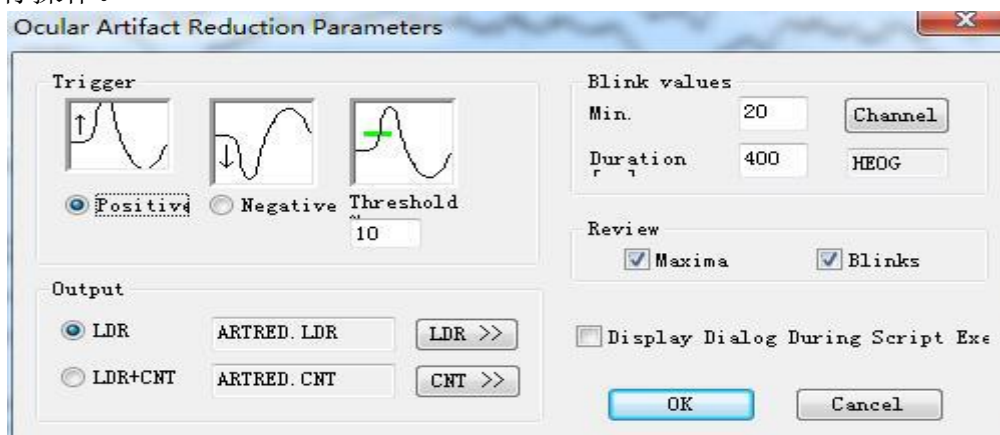


图 3-6 去除眼电

Fig 3-6 Remove eye electricity

#### 4.Artifact rejection 去除伪迹

根据接收数目（接收数占总数的 80%以上）规定最大值最小值（范围为 $\pm 50$ - $\pm 100$  之间），点击 Select，一般选择除眼电、心电、肌电等其他的所有导联，点击 OK 确认，完成不明显伪迹的去除。

## 第4章 基于小波变换与多项指标的疲劳驾驶检测应用

疲劳驾驶是引发交通事故的一大因素，脑电作为携带众多生理信息的生物电成为研究疲劳驾驶的可靠生理指标，本章针对脑电降噪预处理，提出了一种基于小波变换与多项指标的疲劳驾驶检测应用，在小波变换降噪处理的基础上应用主成分分析（KPCA）方法进行特征提取，最后应用最小二乘分类器进行分类，实验证明了降噪后的数据分类效果比未进行降噪处理的数据分类结果具有明显优势。

依据受试者体征以及脑电装置收集状况进行预处理，对实验脑电数据利用小波变换提取实验中  $\alpha$  波、 $\beta$  波、 $\theta$  波和  $\delta$  波四种频段的均幅值  $(\alpha+\beta)/\beta, \alpha/\beta, (\delta+\alpha)/(\alpha+\beta), (\alpha+\beta)/\theta$  共八项合成指标集成为脑电特征参数。通过 KPCA 提取贡献率较大的主元特征信息形成特征集合,并将特征信息输入最小二乘支持向量机（LSSVM），建立 KPCA-LSSVM 预测模型并对比其他四种模型试验，最终求得该模型平均正确率达到 89.47%，实验证明了该实验的有效性及在数据处理速度上的优势。

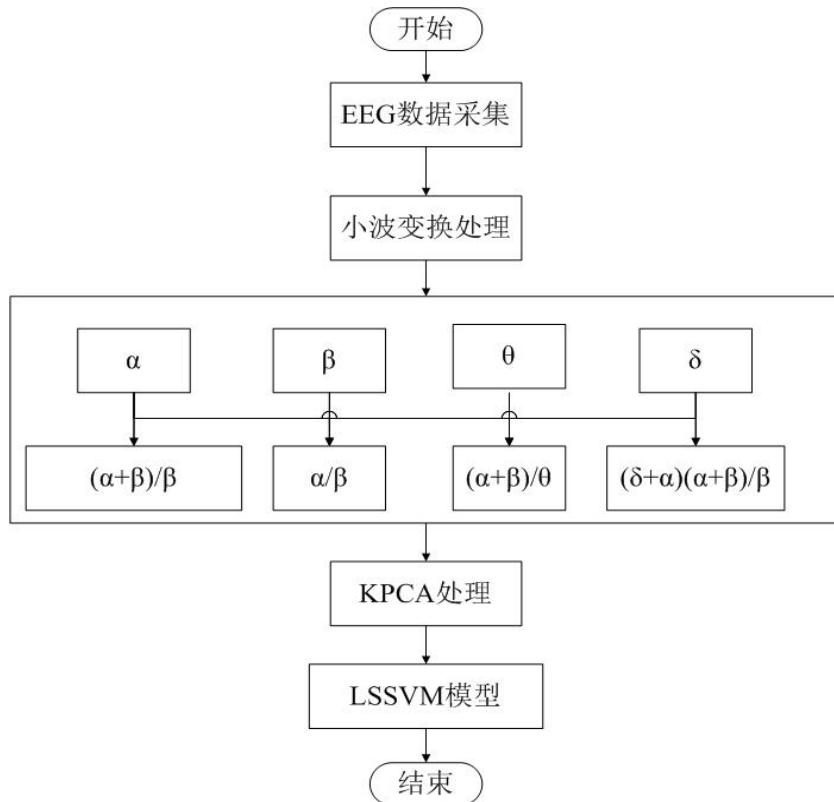


图 4-1 基于小波变换与多项指标的疲劳驾驶检测实验模型

Fig 4-1 Experimental model of fatigue driving detection based on wavelet transform and multiple indicators

## 4.1 基于小波变换的信号降噪

将采样频率为 1000Hz 的脑电信号进行两倍提取后，其采样频率降低为 500Hz，依据采样定理，250Hz 则成为待分析的焦点脑电数据，由于脑电信号自身特征以及受到工频干扰的频带范围，实验中运用小波变换将 EEG 分解为  $\alpha$  波、 $\beta$  波、 $\theta$  波和  $\delta$  波四种频段，应用 Daubechies D10 小波函数进行四层分解，d1 为原始信号，d2 到 d5 分别为  $\alpha$  波、 $\beta$  波、 $\theta$  波和  $\delta$  波四种频段重构图，图 4-2 为小波分解的过程。

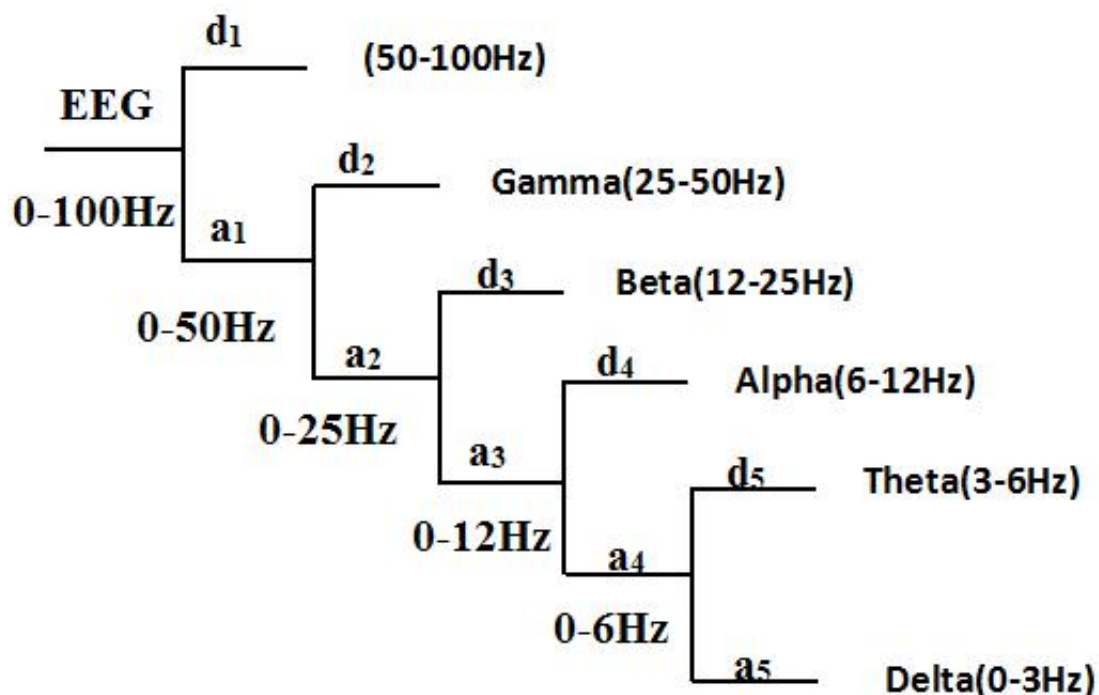


图 4-2 小波变换分解示意图

Fig 4-2 Schematic diagram of wavelet transform decomposition

脑电信号降噪处理分解出的脑电信号因选择不同的小波基函数而变化，去除噪声的效果也各有不同，因此选择适当高效的小波基函数成为影响信号降噪效果的关键性因素。支撑宽度、对称性、正交性、正则性是描述小波特性的几项基本指标。实验中对对比选用不同小波基的实验效果，通过对比发现  $\text{coifN}$  与  $\text{dbN}$  具有较大优势，其中  $N$  为小波的阶数。小波函数  $\psi(t)$  和尺度函数  $\varphi(t)$  中的支撑区为  $2N-1$ ， $\psi(t)$  的消失矩为  $N$ 。 $\text{dbN}$  小波具有较好的正则性，即将其作为稀疏基所引入的光滑误差不明显更具有和易性，图 4-3 为光滑化信号重构过程。

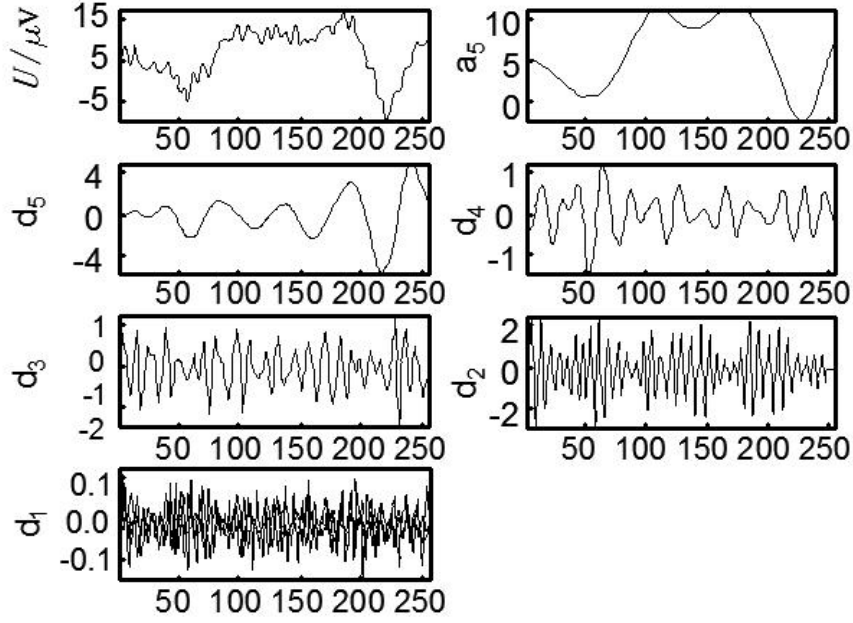


图 4-3 疲劳脑电信号分解图

Fig. 4-3 breakdown of fatigue EEG signals

实验中提取  $\alpha$  波、 $\beta$  波、 $\theta$  波和  $\delta$  波四种频段的幅值序列并求均值，参照文献 51<sup>[42]</sup>同时计算  $(\alpha + \beta)/\beta$ ,  $\alpha/\beta$ ,  $(\delta + \alpha)/(\alpha + \beta)$ ,  $(\alpha + \beta)/\theta$  这四项合成指标，对  $N$  个电极的脑电信号进行处理得到  $8N$  个脑电特征参数形成脑电参数特征集合  $x$ 。

## 4.2 基于 KPCA 的特征提取

降维方法包括属性选择和映射方法。其中映射又分为线性映射方法（PCA、FDA 等）和非线性映射方法：核方法（KPCA、KFDA 等）、二维化、流形学习（ISOMap、LLE、LPP 等），还包括神经网络和聚类等其他方法<sup>[43]</sup>。脑电波是非线性非平稳信号，用线性关系去刻画会损失精度且导致低效，因此本文主要选用 KPCA 主成分分析。

主成分分析(Principal Component Analysis, PCA)通过将高维空间的问题转化到低维空间处理，从而使问题简易化、明确化，并且这些综合指标之间不具有互相关性，同时又可以提供原有指标包含的大部分信息<sup>[44]</sup>。作为一种由线性到非线性之间的桥梁，KPCA 核方法是通过引入核函数把非线性变换后的特征空间内积运算转换为原始空间的核函数计算<sup>[45]</sup>，从而缩减计算量。

依据上文方法计算特征空间中的协方差矩阵为<sup>[47]</sup>：

$$\bar{C} = \frac{1}{m} \sum_{i=1}^m \phi(\mathbf{x}_i) \phi'(\mathbf{x}_i) \quad (4-1)$$

本文应用 KPCA 对所得  $8N$  个脑电参数进行特征降维<sup>[48]</sup>，将主元按照贡献率降

序排列，取前  $k$  ( $k < m$ ) 个主元使构成的疲劳识别脑电指标累计贡献率达到 90% 以上形成新的集合  $T$ 。以受试者 1 脑电数据进行分析，其静息及重度疲劳状态下脑电信号各贡献率分布情况如图 4-4:

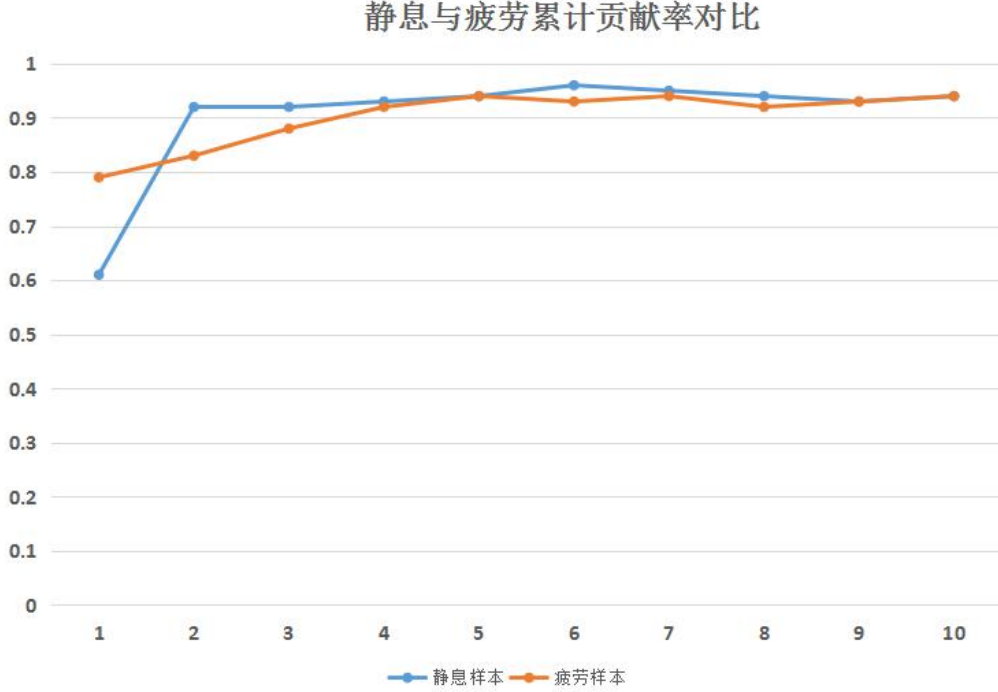


图 4-4 静息与疲劳累计贡献率对比

Fig 4-4 Comparison of cumulative contribution rate of rest and fatigue

### 4.3 基于 LS-SVM 的脑电波数据处理

支持向量机 (support vector machine, SVM) 区别于其他机器学习算法的经验风险最小原则，是建立在结构风险最小化原则基础上的一种基于统计学习理论而来的分类方法，具有训练样本量小，且局部优化和劣化能力强的优势。

上文得到的样本

$$\mathbf{T} = (\mathbf{x}_i, \mathbf{y}_i), i = 1, 2, \dots, m \quad (4-2)$$

其中， $\mathbf{x}_i$  为样本里的第  $i$  个  $n$  维输入向量， $\mathbf{y}_i$  为对应的输出向量，则基于支持向量机的学习可以理解为寻找输入量和输出量之间的映射关系  $f(\cdot)$ ，使得  $f(\mathbf{x}_i) = \mathbf{y}_i$  成立<sup>[50]</sup>。一般  $f(\cdot)$  表示为：

$$f(\mathbf{x}) = \mathbf{w}\Phi(\mathbf{x}) + b \quad (4-3)$$

式中  $\mathbf{w}$  称为权向量， $b$  称为阈值，把  $n$  维输入向量  $\mathbf{x}$  映射到特征空间。若使预测模型精度尽可能高，须调试适当的参数，最大程度降低风险，构造如下式的经验风险函数：

$$R_{emp} = \frac{1}{m} \sum_{i=1}^m L(y_i - f(x_i)) \quad (4-4)$$

其中  $L$  称为损失函数，其表达式为：

$$L(\cdot) = \begin{cases} 0 & \text{if } |y_i - f(x_i)| < \varepsilon \\ |y_i - f(x_i)| - \varepsilon & \text{else} \end{cases} \quad (4-5)$$

在上式中  $\varepsilon$  是大于 0 的小数，根据实验情况进行调整设定。根据 SRM 准则可选取非线性估计函数为：

$$f(x) = \sum_{i=1}^m w_i \Phi_i(x) + b \quad (4-6)$$

构造如下优化目标函数：

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) \quad (4-7)$$

$$s.t. \begin{cases} y_k - f(x_k) \leq \varepsilon + \xi_i \\ f(x_k) - y_k \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0, \xi_i^* \geq 0 \end{cases} \quad (4-8)$$

上式中  $\xi, \xi^*$  为两个松弛变量， $C$  称之为容量因子。

将上式转化为拉格朗日对偶问题求解，可得：

$$f(\mathbf{x}) = \sum (\alpha_i^* - \alpha_i) K(\mathbf{x}_i, \mathbf{x}) + b \quad (4-9)$$

$$b = \sum_{SVS} (\alpha_i - \alpha_i^*) (K(x_j, x_i) + K(x_k, x_i)) \quad (4-10)$$

其中  $\alpha_i$  为拉格朗日乘子， $K(\cdot)$  称为核函数，要求满足：

$$K(x_i, x_j) = (\Phi(x_i) \cdot \Phi(x_j)) \quad (4-11)$$

引入核函数可以实现实验在运算维数和运算复杂程度上的优化。本实验选用径向基核函数：

$$K(x, x_i) = \exp(-\|x - x_i\|^2 / 2\delta^2) \quad (4-12)$$

实验需要确定惩罚因子  $C$  和径向基函数中  $\delta$ ，不同参数值得分类效果相差较大，本文通过对训练数据采用模拟退火算法并结合“留一法”进行交叉验证得到<sup>[51]</sup>，并由训练样本求出 Lagrange 乘子  $a$ 、偏置值  $b$ ，再计算函数  $f(x)$ ，得到  $\text{sgn } f(x)$  为 +1，则为重度疲劳状态， $\text{sgn } f(x)$  为 -1，则为静息状态。表 4-1 为特征集合相同情况下两种分类器分类效果对比。



表 4-1 两种分类器分类效果对比

Table 4-1 Comparison of classification effects of two classifiers

分类器		分类效果		训练分类器时间/s
		ROG 曲面下面积	分类准确率/%	
SVM	$C=3.1492$ $\delta=77.23129$	0.9641	89.71	0.129705
LS-SVM	$C=1.209771$ $\delta=0.3718$	0.9718	89.47	0.059381

LS-SVM 方法有较好的学习能力及泛化性能,能够在一定程度上提高建模水平,训练和泛化误差也可以较好的被控制在一定范围内。同时该模型简化了数据计算量,只需调整正规化参数和核参数,提高了计算效率且能够达到较高的计算精度<sup>[52]</sup>。

#### 4.4 模型与总结

本实验对比了单独应用小波变换、KPCA 以及同时应用这两种方法的实验,结果如下表:

表 4-2 四种方法的分类效果对比

Table 4-2 Comparison of classification effects of four methods

方法编号	数据处理	特征提取	分类器	正确率/%
A	小波变换	/	LS-SVM	76.47
B	/	KPCA	LS-SVM	72.12
C	小波变换	KPCA	SVM	85.56
本章	小波变换	KPCA	LS-SVM	89.47

从上表中可以看出在控制同等特征集合时 LS-SVM 在分类效果、训练分类器耗时上都占有一定的优势;但应用到实际驾驶疲劳预警系统上由于客观条件的限制,目前研究仍然存在一些局限性。例如没有考虑模拟驾驶与实际驾驶的环境差异,模拟驾驶时间比较短,以及“单通道视觉”实验影响实验效果等因素对疲劳监测的影响。因此,使用 KPCA-LSSVM 的分析方法还有待进一步完善。

针对驾驶疲劳的检测,尽可能准确预测疲劳发生,并发出语音、震动等提示,提高安全驾驶系数,减少交通事故的发生,通过强设备的改进以及算法的优化进一步推进脑电在疲劳驾驶智能检测系统中的应用。

## 第5章 集成多特征融合与三种不同核分类器的疲劳脑电分类研究

在我们研究过程中发现，疲劳驾驶特征带有个体差异性的<sup>[53]</sup>，例如，我们发现男女受试者的差异性，同时我们发现，相同的信号特征随着分类器的不同带有差异性<sup>[54,55]</sup>，但是从对疲劳驾驶研究现状分析，但是现有研究中，对疲劳检测少有加入个体差异性的研究。

针对融合脑电多特征，提出了一种集成多特征融合与三种不同核分类器的疲劳脑电分类研究方法。实验结果发现在降噪处理结果与分类器选择相同的基础上，多特征提取分类结果明显高于单一特征提取效果。

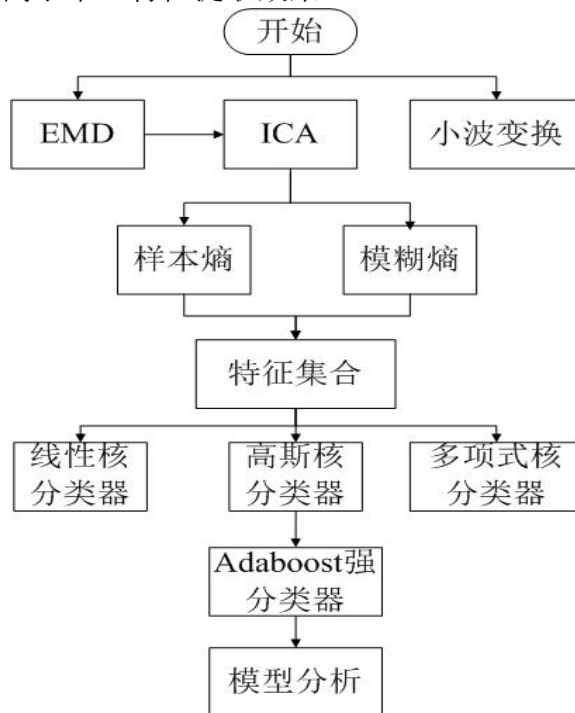


图 5-1 集成多特征融合与三种不同核分类器的疲劳脑电分类模型流程图

Fig 5-1 Flow chart of fatigue EEG classification model with integrated multi-feature fusion and three different kernel classifiers

### 5.1 基于 EMD 的脑电数据降噪

由于脑电信号幅度较为微弱，仅在 1uV-200uV 之间，在收集分析过程中极易受到自身身体部位生理活动或者脑电收集装置运行的干扰降低数据真实性，因此在实验正式开始前对受试者进行数据收集装置的使用培训，以确保数据的真实性，提高数据的质量和代表性。

经验模态分解<sup>[56]</sup>（EMD）能根据脑电信号自身的特征时间尺度将信号分解为若干 IMF 以及一个余项的线性和。其中 IMF 反映了信号的自身特征，余项则反映

了信号的变化趋势。

找到信号  $s(n)$  的极值点，用曲线连接所有极大值点和极小值点，得到上、下包络  $l_{\max}$  和  $l_{\min}$ ；

求平均值  $a = (l_{\max} + l_{\min}) / 2$ ；

将数据去掉均值得到  $h_1 = s - a$ ；

继而检查  $h_1$  是否满足 IMF 定义，若不满足要求，则标记为新数据并重复上述步骤，若满足则分离为 IMF 分量，记录  $c_1 = h_1$ ；

记  $r_1 = s - c_1$ ，并将  $r_1$  标记为新数据重复上述步骤直到  $r_n$  变为单调函数、常量或者仅有一个极值点的函数时分解阶数，此时  $r_n$  称为余量，则原信号可表示为<sup>[57]</sup>

$$s = \sum_{i=1}^m c_i + r_m \quad (5-1)$$

实验最后得到  $m$  个 IMF 和一个余量，标号从低到高的 IMF 分别对应原始信号时间尺度从细到粗的特征信息。以受试者 1 疲劳信号为例，原始信号经过 EMD 分解为 7 个 IMF 分量，各分量所属频率依次降低，舍弃残余值，IMF 分量图如 5-2 所示：

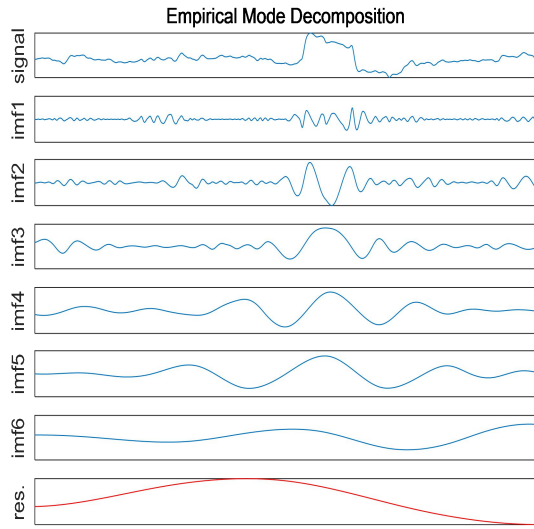


图 5-2 IMF 分量图

Fig. 5-2 IMF component diagram

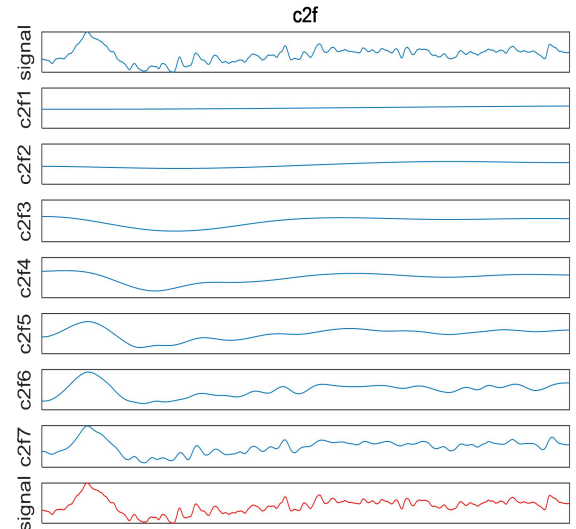


图 5-3 cf2 图

Fig 5-3 cf2 Diagram

$f2c$  和  $c2f$  分别从 2 个相反的方向重构原始信号，这里以  $c2f$  为例， $f2c1$  即为 IMF1，而  $f2c2$  则为 IMF1 与 IMF2 合成的， $f2c3$  由 IMF1、IMF2、IMF3 合成的；并且  $c2f1$  就是 IMF4，以此递推。

小波变换<sup>[58]</sup>具有低熵性、多分辨率特性等特点具有良好的非平稳性信号的描述，并且基函数的选择具有相当大的灵活性。脑电信号降噪处理，需去掉异常值和高频噪声，去除脑电的冗余信号信息，保证光滑性和与原信号的一致性，应用了基于中值滤波的自适应小波阈值去噪法<sup>[59]</sup>进行降噪实验，并与 EMD 降噪方法处理结果对比。

将去噪信号作为标准信号  $f(n)$ ，原始信号为  $s(n)$ ，信号长度为  $l$ ，依据公式定

义统计两种方法的实验参数如下表

表 5-1 去噪效果评价指标比较

Table 5-1 Denoising effect evaluation index comparison			
方法选择	SNR	RMSE	相关系数R
EMD去噪	4.9494	0.848725	0.8630
DB5去噪	2.0798	1.897257	0.6243

表 5-1 中 SNR 值反映了降噪后信号中含有噪声成分的多少，表中 EMD 降噪的信噪比比 DB5 降噪的信噪比大，说明 EMD 降噪保留的噪声成分更小；RMSE 值反映了降噪后信号与原信号之间的离散程度，EMD 的均方误差值较小，值越小越好，R 反映了降噪后信号与原信号之间的相关程度，综上，EMD 在三指标中均优于 DB5，因此本实验在 EMD 降噪基础上进一步进行特征提取。

## 5.2 多特征提取

对原始脑电数据进行合理的降噪，尽可能保留其真实有效数据，将降噪后的数据进行特征提取，使原始数据转化为具有明显物理意义或者统计意义或核的特征，能够有效的减少冗余数据、减少数据存储，更加高效准确的对脑电信号进行探索，本文在经验模态分析后得到的数据集  $x_i$  提取模糊熵、信息熵作为特征。

### 5.2.1 基于模糊熵的特征提取

熵<sup>[60]</sup>度量了一个系统或一段信息的不确定性，模糊熵描述了一个模糊集的模糊性程度，上文实验得到的模糊集

$$\{X_i\} = \{X_1, X_2 \cdots X_n\} \quad (5-2)$$

转化向量表示为

$$X = (Ux(X_1), Ux(X_2), Ux(X_3), \cdots Ux(X_n),) \quad (5-3)$$

根据模糊熵定义

$$\pi_A(X_i) = \frac{U_A(X_1)}{\sum_{i=1}^n U_A(X_i)} \quad (5-4)$$

则公式 5-5 即为模糊熵，

$$\begin{aligned} H(\pi_A(X_1), \pi_A(X_2), \pi_A(X_3), \cdots \pi_A(X_n),) \\ = -\frac{1}{\ln n} \sum \pi_A(X_i) \ln(\pi_A(X_n)) \end{aligned} \quad (5-5)$$

其中， $i = 1, 2, 3, \cdots n$ 。

### 5.2.2 基于样本熵的特征提取

原始数据为集合  $X$ ，总长度记为  $N$ 。设嵌入维数  $m$  和相似容限  $r$ ，依据原始信号重构一个相似容限  $r$  和  $m$  维向量

$$X_i = [X_i, X_{i+1}, \dots, X_{i+m-1}] \quad (5-6)$$

定义  $x_i$  与  $x_j$  间的距离  $d_{ij}$  为两者对应元素差值绝对值的最大值，即

$$d_{ij} = d[x_i, x_j] = \max[|x_{i+k} - x_{j+k}|] \quad (5-7)$$

对每个  $i$ ，计算  $X_i$  与其余向量的距离  $d_{ij}$ ，统计  $d_{ij}$  小于  $r$  的数目及此数目与距离总数  $N-m-1$  的比值，记作  $B_i^m(r)$ ，再求  $B_i^m(r)$  的平均值  $B^m(r)$ ，即

$$B^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^m(r) \quad (5-8)$$

再对维数  $m+1$ ，重复上述步骤，得到，计算平均值。原始序列的样本熵定义为：

$$SampEn(m, r) = \lim_{N \rightarrow \infty} [-\ln \frac{B^{m+1}(r)}{B^m(r)}] \quad (5-9)$$

当  $N$  为有限数时，上式表示成：

$$SampEn(m, r, N) = \ln B^m(r) - \ln B^{m+1}(r) \quad (5-10)$$

通常  $m$  取值 1 或 2，李立等人<sup>[61]</sup>研究得到，样本熵比近似熵更适合作为脑电特征，在本实验中引用样本熵作为特征之外增加了对参数的讨论。

表 5-2 不同的  $m$  和  $r$  组合时的  $\delta$  值

Table 5-2 The value of $\delta$ for different combinations of $m$ and $r$			
	$m=1$	$m=2$	$m=3$
0.15xSD	42.37%	45.21%	49.21%
0.20xSD	41.25%	47.39%	47.83%
0.25xSD	44.78%	52.71%	45.26%
0.30xSD	42.51%	46.32%	44.81%

由表 5-2 知在本研究中取  $m = 2$ ，而  $r = 0.25 \times SD$  时差别极大更容易进行区分，其中  $SD$  表示时间序列的标准差。

### 5.3 模型选择

实验对比了小波变换和经验模态分析降噪，在降噪效果较好的 EMD 基础上进行样本熵、模糊熵的提取。采用 AdaBoost<sup>[62]</sup>方法来形成一个强分类器。对全部受试者的实验结果进行记录，以全面了解实验的效果和模型的价值，计算其准确率、灵

敏感度、查准确、查全率等，正确的对模型进行评估，将分类效果差或错分的数据进行记录，同时对时耗做出相应记录，将列为日后改进的方向。

本文分别提取同一受试者的静息特征集和重度疲劳数据集，将数据进行归一化之后，将样本随机分成训练样本和测试样本，这里抽 600 个样本为测试样本，其中，正负样本各一半。

Boosting 算法通过与其他分类算法组合使用，以集成的方式实现对基本分类算法高精度的估计<sup>[63]</sup>。设每个样本初始为均匀分布  $D$ ，每个样本加权。

$$D_1(i) = (w_1, w_2, \dots, w_n) = \left(\frac{1}{N}, \dots, \frac{1}{N}\right) \quad (5-11)$$

将式子进行迭代，更新样本权重。所有被错分类的样本点的权值之和即为误差率，这里我们设线性核、高斯核和多项式核三个分类器分别为  $f_1$ 、 $f_2$ 、 $f_3$  以及最终的强分类器为  $f$ ，并根据误差率  $e_1$  计算分类器  $f_1$ 、 $f_2$ 、 $f_3$  的系数。

$$e_1 = P(H_1(x_i) \neq y_i) = \sum_{i=1}^N w_i I(H_1(x_i) \neq y_i) \quad (5-12)$$

计算出的该次迭代  $H$  在强分类器  $f$  中的比重。

$$a_t = \frac{1}{2} \ln \left( \frac{1 - e_t}{e_t} \right) \quad (5-13)$$

$a$  为以错分最小的分类器为该次迭代的基本分类器，若某个样本点已经被准确分类，就降低其权重；反之若样本点分类错误，那么就将它的权重提高<sup>[64]</sup>。

$$f(x) = \sum_{t=1}^T a_t H_t(x) \quad (5-14)$$

根据分类器的分类效果，将被错分的样本点的权重，继而用下一个分类器  $C_{k+1}$  训练将权值更新过的样本集，然后将每次训练得到的分类器按照公式 5-11 进行融合作为最后的决策分类器。当强分类器错误率小于等于 1% 时停止迭代<sup>[65]</sup>。以受试者 1 数据为例，线性分类器错误率比较低，高斯核分类器核多项式分类器略高，但彼此优势互补，高斯核 LSSVM 分类器为 89.71%；线性核 LSSVM 分类器分类效果较其他两种分类器效果明显，达到 90.2%；多项式核 LSSVM 分类器为 91.4%；最终 Adaboost 最终正确率为 95%。

$$H_{final} = \text{sign}(f(x)) = \text{sign} \left( \sum_{t=1}^T a_t H_t(x) \right) \quad (5-15)$$

## 5.4 模型分析与评价

从实际出发，在受试者性别选取上参考了目前中国车辆持有人的男女比例，以最大限度提升研究的价值；分类正确率是衡量实验效果的标杆，除此之外，本文结

合简单交叉验证（simple cross validation）进一步对结果进行评估。

表 5-3 受试者实验数据统计

Table 5-3 Subject experimental data statistics

受试者序号	性别	高斯核分类器	线性分类器	多项式分类器	综合分类器
1	男	0.94	0.96	0.92	0.97
2	女	0.89	0.92	0.96	0.99
3	女	0.91	0.93	0.87	0.91
4	男	0.87	0.92	0.94	0.91
5	男	0.93	0.89	0.95	0.96
6	男	0.92	0.94	0.89	0.95
7	女	0.88	0.91	0.95	0.96
8	女	0.90	0.93	0.89	0.92
9	男	0.91	0.96	0.93	0.95
10	男	0.93	0.96	0.94	0.98
综合	-	0.908	0.932	0.924	0.95

表 5-3 以受试者 1 为例，随机将  $S_1$  分为训练数据集  $S_{train}$ （70%的样本）和简单交叉验证集  $S_{test}$ （剩下的 30%样本），通过训练数据集训练所得出的最强分类器，并依次对所有受试者数据进行验证，结果如图 5-4 所示。

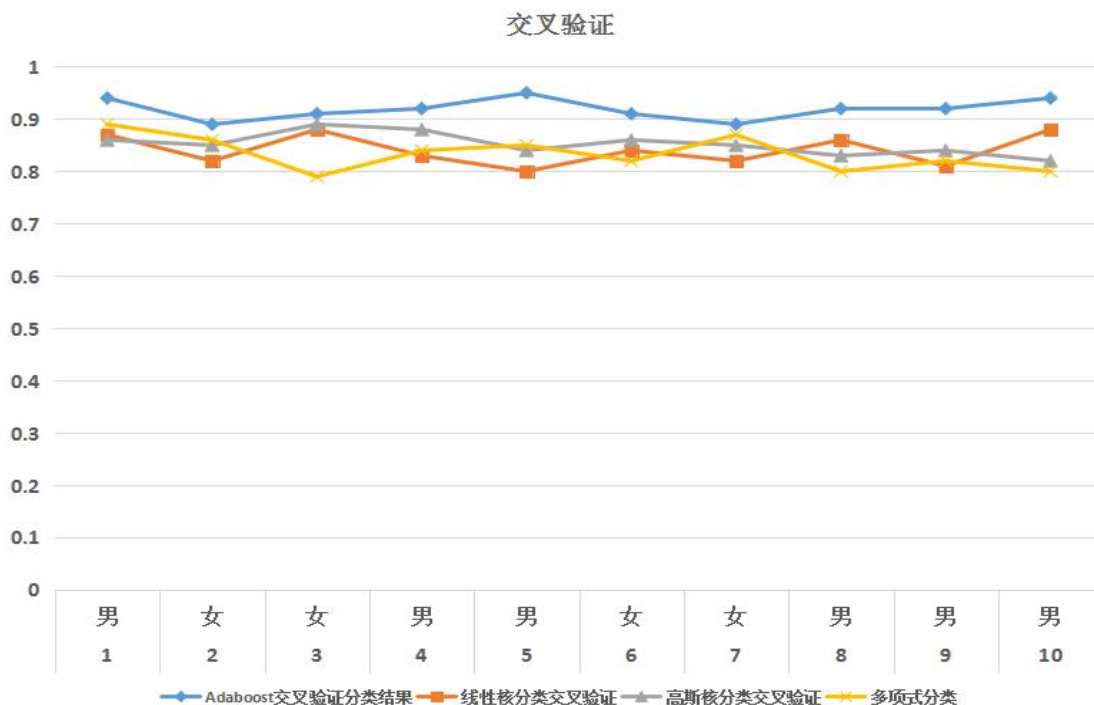


图 5-4 交叉验证分类结果

Fig 5-4 cross validation classification results

该方法能够根据不同脑电数据特征综合运用三种分类器选择最优集成分类器进

行分类,更具有灵活性和通用性。同时交叉验证 10 位受试者脑电数据平均准确率为 92%,对比线性核、高斯核核多项式核分类器方法的交叉验证结果明显准确度较高,说明该方法具有一定的鲁棒性,但简单交叉验证的不足之处在于未能充分利用验证集合中的数据,在脑电数据有限的情况下需进一步改善。



## 第 6 章 基于独立成分分析降噪与集成分类器的疲劳脑电研究

随着交通的日益便利，交通事故发生的频率也愈加频繁。本章提出了一种基于 AdaBoost 的组合型实验方法用于分析脑电检测疲劳驾驶。试验过程中针对不同受试者采用独立成分分析(Independent Component Correlation Algorithm, ICA)处理分析，继而进行样本熵、信息熵、模糊熵和 AR 系数的特征提取；最后运用 AdaBoost 将最小二乘向量机基于三种核分类器集成为一个强分类器。试验结果显示，采用 AdaBoost 分类器分类效果优于单个核分类器，对疲劳驾驶平均识别率达到 93%，五折交叉验证准确率为 91.04%，在一定程度上推动了基于脑电信号的安全驾驶辅助监控系统的研究。除此之外，本章优化选择了特征提取阶段的 AR 阶数；对比分析了三种不同核分类器的分类效果，同时选取受试者 1 数据进行自适应实验模式选择简要进行分析。

针对分类器的单一性以及各分类器的优势互补问题，提出了一种基于 AdaBoost 的组合型实验方法用于分析脑电检测疲劳驾驶。试验过程中针对不同受试者采用独立成分分析(Independent Component Correlation Algorithm, ICA)处理分析，继而进行样本熵、信息熵、模糊熵和 AR 系数的特征提取；最后运用 AdaBoost 将最小二乘向量机基于三种核分类器集成为一个强分类器。试验结果显示，采用 AdaBoost 分类器分类效果优于单个核分类器，对疲劳驾驶平均识别率达到 93%，五折交叉验证准确率为 91.04%，在一定程度上推动了基于脑电信号的安全驾驶辅助监控系统的研究，图 6-1 为本章实验分析模型。

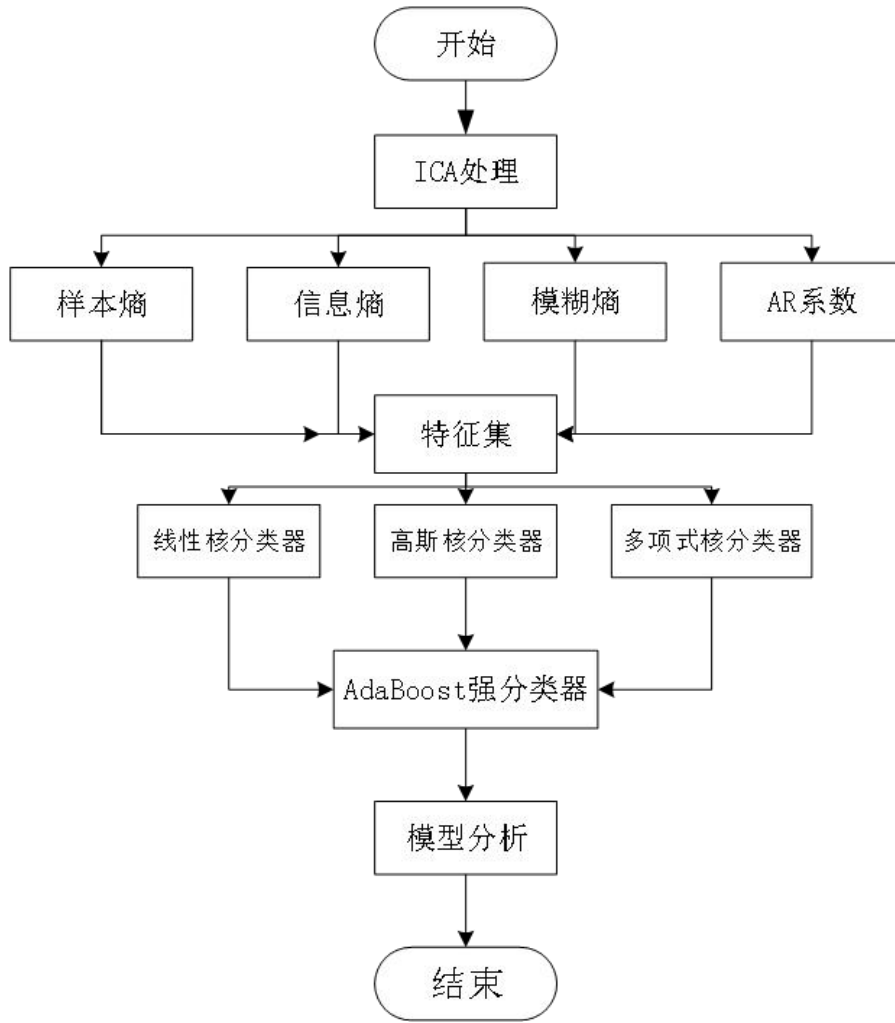


图 6-1 基于独立成分分析降噪与集成分类器的疲劳脑电研究模型

Fig 6-1 Fatigue EEG Research Model Based on Independent Component Analysis Noise Reduction and Integrated Classifier

## 6.1 基于 ICA 数据的脑电预处理

ICA(Independent Component Correlation Algorithm, ICA)是用以分析多维数据的线性分析方法, 试图将一组随机变量表示成统计独立变量的线性组合<sup>[66]</sup>。

设(6-1)是一组由  $m$  个隐含变量产生的  $n$  维观测信号, 其中  $T$  表示数据样本的个数;  $t$  为观测数据长度。

$$X(t)=[x_1(t), x_2(t), x_3(t), \dots, x_n(t)] \quad (6-1)$$

设(6-2)为隐含在  $X(t)$  中的  $m$  维的一组随机变量, 其中  $m \leq n$ 。

$$S(t)=[s_1(t), s_2(t), \dots, s_m(t)] \quad (6-2)$$

转化为线性问题处理, 即

$$x_i(t) = \sum_{j=1}^n a_{ij} s_j(t), (i=1,2,3,\dots,n; j=1,2,\dots,m) \quad (6-3)$$

用矩阵形式定义推导  $X = AS$ ，其中  $s_i$  为独立分量，

$$A = [a_1, a_2, \dots, a_m] \in R^{n \times m} \quad (6-4)$$

式子(6-4)为一满秩矩阵； $a_i (i=1,2,3,\dots,m)$  是混合矩阵的基向量，即观测信号  $x_i$  是由独立源  $s_i$  通过不同的权值  $a_{ij}$  线性加权构成。在处理 ICA(Independent Component Correlation Algorithm, ICA)算法过程中，系统目标通过匹配混合信号  $X$  的一个线性变换矩阵  $W$ ，输出  $y_i$ ，即

$$y(t) = Wx(t) = WAs(t) \quad (6-5)$$

图 6-2 为受试者 1 的疲劳数据的 ICA(Independent Component Correlation Algorithm, ICA)独立源成分，经试验每次经过 ICA 分析得到的独立成分的次序会发生变化，但所得独立分量在头皮电极的能量分布情况具有极高的相似度，即从多导脑电信号提取出来的独立源成分拥有较多的共同特征<sup>[67]</sup>。

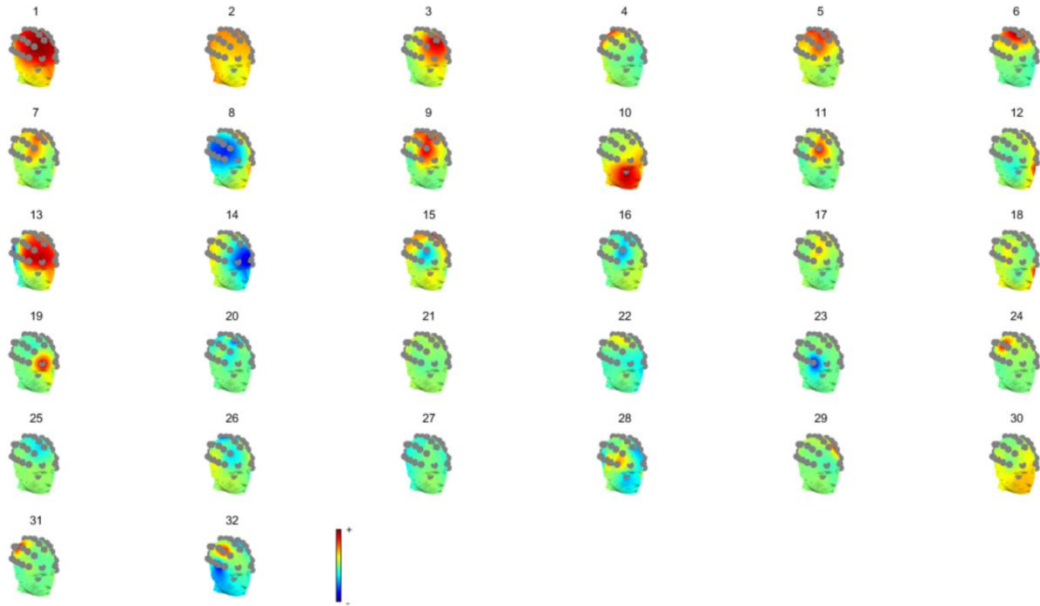


图 6-2 疲劳脑电 ICA 地形图

Fig 6-2 fatigue EEG ICA topographic map

## 6.2 基于信息熵的特征提取

张爱华等人<sup>[68]</sup>证明了模糊熵方法在表征脑电睡眠各分期的复杂度的优势,因此将信息熵作为特征集合之前。假设  $X$  是一个离散随即变量，即它的取值范围  $R=\{x_1, x_2, \dots\}$  是有限可数的。

设  $p_i=P\{X=x_i\}$ ， $X$  的熵定义为：

$$H(X) = \sum_{i=1}^n p_i \log \frac{1}{p_i} = - \sum_{i=1}^n p_i \log_a p_i \quad (6-6)$$

其中信息熵的单位依自信息的单位决定的,由对数换底公式得到  $r$  进制的信息熵  $H_r(X)$  与二进制信息熵  $H(X)$  之间的换算关系

$$\begin{aligned} H_r(X) &= - \sum_{i=1}^n p_i \log_r p_i \\ &= - \sum_{i=1}^n p_i \log_a p_i \cdot \frac{1}{\log r} = \frac{H(X)}{\log r} \end{aligned} \quad (6-7)$$

不同的信源其统计特性的不同,信息熵也具有差异。

除此之外,本实验还提取模糊熵、样本熵作为特征集合的一部分,提取流程同本文 5.5.1 和 5.2.2 章节。

### 6.3 基于 AR 系数的特征提取

EEG 具有典型的非线性特征,将 AR 模型应用到 EEG 这一随机过程中,参照实际 EEG 选取合适的阶数。AR 模型即自回归模型,用差分方程可表示为:

$$x(n) = - \sum_{i=1}^p a_p(i) x(n-i) + \varepsilon(n) \quad (6-8)$$

其中  $\varepsilon(n)$  是方差为  $\delta^2$ 、均值为 0 的白噪声序列,  $p$  为 AR 模型阶数,  $a_p(i), i=1,2,\dots$ 。构建 AR 参数模型的阶数选择,阶数过低则会使分辨率不够,阶数过高会出现谱估计出现谱分裂现象。

对于时间序列  $\{x_t\}, t=1,2,\dots,N$ , 其中  $AR(n)$  为

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_n x_{t-n} + a_t, a_t \sim NID(0, \delta^2_a) \quad (6-9)$$

参数估计就是按照一定的方法估计出参数,依据下式

$$a_t = x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \dots - \phi_n x_{t-n} \quad (6-10)$$

由分析得, AR (n) 模型的参数估计本质上就是对式中  $\phi_1, \phi_2, \dots, \phi_n (t=1,2,3,\dots,n)$  这些参数的估计。参数估计方法包括直接和间接两种,本章引用文献 69<sup>[69]</sup>的最小二乘法进行参数估计对参数进行确定,表示为  $Y = X\phi + a$ , 其中

$$\begin{aligned} Y &= [x_{n+1}, x_{n+2} \dots x_N]^T, \\ \phi &= [\phi_1, \phi_2 \dots \phi_n]^T \\ a &= [a_{n+1}, a_{n+2} \dots a_N]^T \end{aligned} \quad (6-11)$$

则  $\phi$  的最小二乘估计  $\hat{\phi} = (X^T X)^{-1} X^T Y$ , 多种模型的适应性检验准则本实验中以最小二乘估计和 Akaike 信息检验准则的 FPE (Final Prediction Error)、AIC (An Information Criterion)、BIC 准则 进行 AR 阶数讨论。

FPE 准则函数:

$$\begin{aligned} FPE(n) &= \frac{N+n}{N-n} \delta_a^2 \\ AIC(n) &= N \ln \delta_a^2 + 2n \\ BIC(n) &= N \ln \delta_a^2 + n \ln N \end{aligned} \quad (6-9)$$

实验通过 matlab 程序执行,  $n \in [1, 30]$  迭代计算, 匹配适应的 AR 阶数, 在 AIC 和 BIC 准则共同作用下获得最能反映本质特征的最佳阶次和参数。

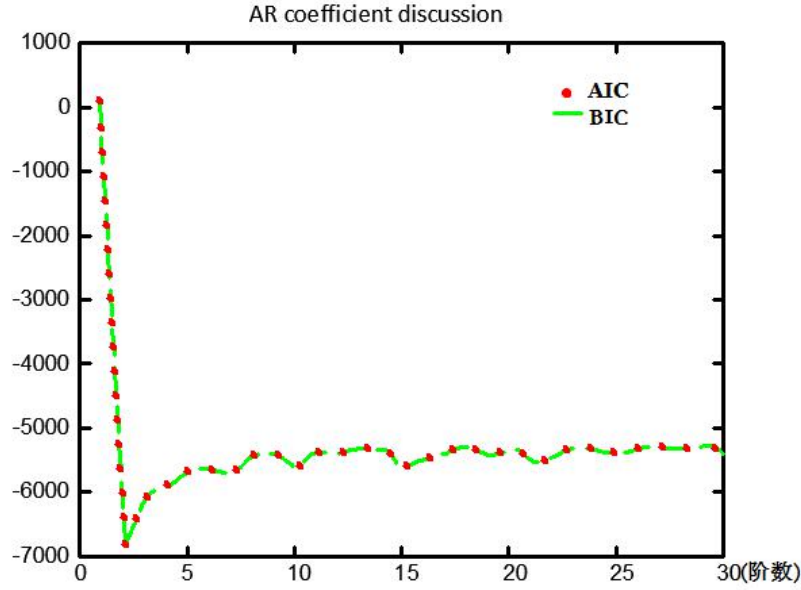


图 6-3 AR 阶数讨论

Fig. 6-3 discussion on AR order

由图 6-3 知, 检验准则函数的变换趋势基本一致, 当  $n=2$  时取得最佳值, 即  $P=2$ 。

## 6.4 基于 Adaboost 强分类器分类

最小二乘支持向量机将最小二乘估计引入支持向量机中, 是基于结构风险最小化原则的数据分类和回归工具 [70]。

该模型给定训练样本集合(6-10), 其中训练样本  $x_i \in R$ , 输出数据  $y_i \in R$ 。依据结构风险最小化原则构造最小化目标函数如(6-11)

$$\{x_i, y_i\} (i=1, 2, \dots, M), \quad (6-10)$$

$$\min \frac{1}{2} w^T w + \frac{C}{2} \sum_{i=1}^M \varepsilon_i^2 \quad (6-11)$$

$$y_i = w^T \phi(x_i) + b + \varepsilon_i, (i=1, 2, \dots, M)$$

其中  $w \in R^n$  为权矢量,  $\phi(x)$  将  $x$  从原始空间非线性地映射到高维特征空间,  $\varepsilon \in R$  为误差项,  $b$  为偏差量,  $C$  为正则化参数。将上式转换为无约束目标函数, 得到最小

二乘支持向量机的回归函数为

$$y(x) = \sum_{i=1}^M \alpha_i K(x, x_i) + b \quad (6-12)$$

其中，常用的核函数有线性核函数<sup>[71]</sup>、多项式核函数、径向基核函数等形式，这里将提取的特征组成特征集合，分别输入三种核函数的最小二乘分类器<sup>[72]</sup>，具体实现算法同第五章。实验中方法综合利用了不同核函数优势。多项式核函数具有较好的全局性，并且其阶数越低，推广能力越强<sup>[73]</sup>；RBF 可以将样本映射到一个更高维空间，区别于线性核函数，可以处理非线性类别标签和特征之间的关系的样本<sup>[74]</sup>。脑电随受试者主体、身体状况、设备配置等不同而变化，实验结果如图 6-3，三种核函数各自发挥优势，以 AdaBoost 方法迭代给三种分类函数分类器赋值权重，最终选择形成强分类器以最大程度提高分类的准确率。

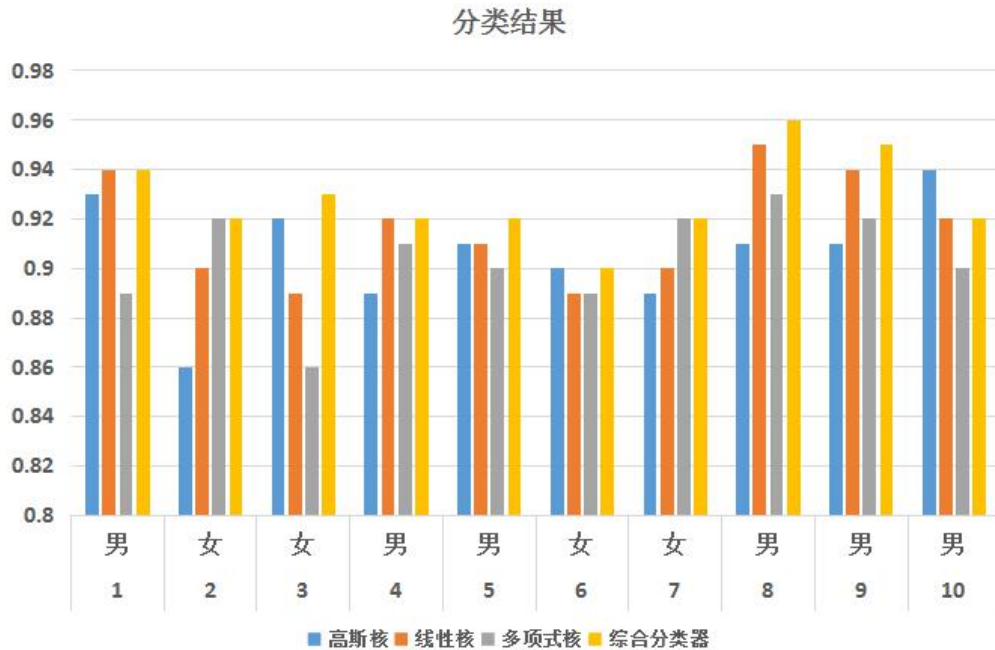


图 6-4 分类结果

Fig 6-4 Classification results

分析图 6-4 可知，综合分类器分类准确率总体高于单独分类效果；同一受试者数据应用不同分类器分类准确率具有不确定性；针对单独分类器分析，不同受试者适用于不同核的分类器；就本实验数据而言，女性分类效果平均准确率达到 0.935，要高于平均准确率约为 0.928 的男性数据分类效果，这可能由于女性偏细心脑电数据收集装置使用更为标准，所以数据收集较为准确。

## 第7章 总结与展望

### 7.1 本文主要工作与研究结论

本文主要从时域分析、频域分析、时频分析、人工神经网络和非线性动力学分析等几方面对脑电研究涉及到的关键科学问题、实验目的、脑电信号分析处理方法、实验结果等内容进行分析总结。

(1) 从涉及到的关键学科和实验目的方面看,脑电的研究不仅限于对生理、心理等方面的研究,更可以渗透到各种症状的预测、状态监测以及意念控制等,近年来脑电波已被应用到解锁、控制飞行器等领域当中,脑电波作为独特性较强的数据信息,可以结合移动设备等更多设备,使脑电信号收集更便捷,更加广泛应用于消费类电子、医疗设备、汽车行业、玩具、公司考勤以及教具等未来科技发展领域,多学科融合,拓宽脑电的应用领域;

(2) 从数据来源看,多数来源于竞赛数据,缺乏针对性,且多为离线数据,不利于实时监测,未提供数据采集中的特殊因素,会降低数据分析结果的准确率以及预测的误差性,若实现针对性数据实时监测并分析处理,结合多种采集设备综合记录,将大大提高脑电数据分析的准确率,大大推动脑电的应用与发展;

(3) 从脑电处理方式看,脑电数据的分析大多基于统计分析,可尝试结合优势互补的不同脑电处理方式综合处理数据,着重定量深入分析,更深层次的挖掘脑电信息,为更广泛充分的探索脑电的奥秘奠定基础。

与此同时,依据脑电数据处理流程,分别对降噪、特征提取和分类三个方面进行研究,提出了三种分析模型,取得了不错的分类效果。

EEG 是自然界中最为复杂的网络,反映了大脑皮层神经元细胞群自发性节律性的电生理活动,含有丰富信息,不断探索研究脑信号,科学分析将为脑电造福人类开辟新的天地。脑电信号分析发展将造福人类未来生活的诸多方面,相信随着脑信号处理的发展和完善,人类总将会揭示大脑的奥秘!

### 7.2 未来研究展望

为了取得较好的识别效果,模糊系统、SVM、BP 神经网络等多种方法被用来进行疲劳检测,也取得了不错的成果,但是单一的分类器由于通常需要借助梯度下降、统计分析、高纬度分析的方法,因此需要大量的训练时间,甚至需要加入人工标记以提高识别率,因此单一的分类器在多生理特征的分类研究中有一定的局限性,由此产生了集成学习的方法,把多个分类器进行组合,利用不同的决策结果把分类结果有机的结合起来作为分类结果。由于集成学习可以组合多个弱监督模型以期得到一个更好更全面的强监督模型,已有研究者把集成学习引入生理信号分析,特别

是脑电信号相关分析。

通过认真学习现有研究成果,我们发现不同的研究方式有自身的优势,也有不可避免的缺点,如利用车辆行为作为特征的优势在于实时性强,对驾驶员无干扰,缺点是由于对道路硬件要求较高,由于车辆行为特征并不能真正反映驾驶员疲劳的生理变化,疲劳识别准确率并不高,在80%上下;利用驾驶员行为作为特征的疲劳检测克服了车辆行为中不能反映驾驶员疲劳生理变化的缺点,疲劳识别率达到90%左右,缺点在于受光照变化、面部表情、复杂背景等因素的影响较大;直接利用驾驶员生理特征作为疲劳检测依据识别准确率最高,检测准确率最高能达到98%以上,但是由于信号过于敏感,个体差异性较大,因而稳定性较差,项目组曾以脑电信号为工具进行驾驶疲劳研究,在实验室环境下采集的脑电信号,识别率达最高可达98%,但是移植到真实环境下,受到天气和驾驶员自身状态影响较大,识别率很不稳定。

基于此,将未来的研究思路定位为融合脑电特征和眼动特征进行驾驶疲劳特征分析,以脑电特征为主,以眼动特征为辅进行驾驶疲劳脑区分析,以混合集成学习为工具进行驾驶疲劳检测研究。首先,探索驾驶疲劳相关的功能团。确定哪些大脑区域在驾驶疲劳状态中起关键作用,探究驾驶疲劳产生的相关神经机理,建立相对应的关系模型。其次,运用FBN分析驾驶疲劳相关EEG信号全局网络特性。合理构建FBN,对比分析驾驶疲劳相关的FBN网络拓扑特性、动力学特性、大脑区域和网络指标的变化特性,以期发现驾驶疲劳产生的内在大脑网络变化规律。然后,构建多特征融合模型,运用集成学习对驾驶疲劳进行检测。对比分析不同疲劳状态下局部特征、全局特征以及两者的融合的检测效果,组建疲劳状态多特征融合脑网络模型,对驾驶疲劳状态进行识别,提高系统的识别率和稳定性、减少个体差异带来的影响、增强系统的抗干扰能力。



## 参考文献

- [1] Klauer SG, Dingus TA, et al. The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. U.S. Department of Transportation Washington D.C, 2006.
- [2] Charbonnier S, Roy RN, et al. EEG index for control operators' mental fatigue monitoring using interactions between brain regions. *Expert Systems with Applications*, 2016, 52, 91-98
- [3] Chai R, Naik G, et al. Driver Fatigue Classification with Independent Component by Entropy Rate Bound Minimization Analysis in an EEG-based System. *IEEE Journal of Biomedical & Health Informatics*, 2016:1. DOI: 10.1109 / JBHI. 2016. 2532354
- [4] Khushaba RN, Kodagoda S, et al. Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm. *IEEE Trans Biomed Eng*, 2011, 58:121-31.
- [5] Shi L, Lu B. EEG-based vigilance estimation using extreme learning machines. *Neurocomputing*, 2013, 102, 135-143.
- [6] Zhao C, Zheng C et al. Multivariate autoregressive models and kernel learning algorithms for classifying driving mental fatigue based on electroencephalographic. *Expert Systems with Applications*, 2011, 38: 1859 – 1865.
- [7] Zhang C, Wang H, Fu R. Automated detection of driver fatigue based on entropy and complexity measures. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15: 168-177.
- [8] Chuang CH, Huang CS, et al. An EEG-based perceptual function integration network for application to drowsy driving. *Knowledge-Based Systems*, 2015, 80:143-152.
- [9] Fu R, Wang H, Zhao W. Dynamic driver fatigue detection using hidden Markov model in real driving condition. *Expert Systems with Applications*, 2016, 63:397 – 411.
- [10] Wu Q, Mao JF, et al. Hybrid BF – PSO and fuzzy support vector machine for diagnosis of fatigue status using EMG signal features. *Neurocomputing*, 2015, 173: 483-500.
- [11] 刘超. 基于经验模态分解的多尺度熵在疲劳驾驶前额脑电特征抽取方法上的研究[D].南昌: 南昌大学, 2018.
- [12] 方毅. 基于信息融合的驾驶员疲劳状态检测方法研究[D].哈尔滨: 哈尔滨工业大学, 2018.
- [13] 陈骥驰, 王宏, 王翹秀, 化成城, 刘冲. 基于脑电信号的疲劳驾驶状态研究[J]. *汽车工程*, 2018, 40(05): 515-520.
- [14] Hu, J. (2017). Comparison of different features and classifiers for driver fatigue detection based on a single eeg channel. *Comput Math Methods Med*, 2017(3), 5109530.
- [15] Mu, Z., Hu, J., & Min, J. (2017). Driver fatigue detection system using electroencephalography signals based on combined entropy features. *Applied Sciences*, 7(2), 150.

- [16] 陈昕,朱国华,张驰,高令顺.基于眼睛闭合状态的驾驶疲劳检测方法[J].辽宁工业大学学报(自然科学版),2018,38(03):182-186+191.
- [17] 金雪. 基于驾驶行为的疲劳驾驶检测方法研究[D].北京: 北京工业大学,2015.
- [18] 郑培,宋正河,周一鸣.机动车驾驶员驾驶疲劳测评方法的研究状况及发展趋势[J].中国农业大学学报,2001,6(6):101-105.
- [19] 彭军强,吴平东,殷罡.疲劳驾驶的脑电特性探索[J].北京理工大学学报,2007,27(7):585-588.
- [20] 张驰. 高危车辆驾驶员疲劳驾驶监测及预警系统研究[D].辽宁: 辽宁工业大学,2018.
- [21] 苏静静. 基于面部行为分析的驾驶员疲劳检测方法研究[D].天津: 天津工业大学,2017.
- [22] 胡神松. 基于人眼特征疲劳驾驶检测系统的设计与实现[D].甘肃: 西北师范大学,2016.
- [23] 耿磊,袁菲,肖志涛,张芳,吴骏,李月龙.基于面部行为分析的驾驶员疲劳检测方法[J].计算机工程,2018,44(01):274-279.
- [24] Stern H S , Blower D , Cohen M L , et al. Data and methods for studying commercial motor vehicle driver fatigue, highway safety and long-term driver health[J]. Accident Analysis & Prevention, 2018:S0001457518300836.
- [25] Hu J, Min J. Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model[J]. Cognitive Neurodynamics, 2018(12):1-10.
- [26] 韦明,杨辉,孙晓,牟锴钰,郭万鑫.两种去除脑电信号中眼动伪迹的方法比较及其在脑力疲劳测量中的应用[J].航天医学与医学工程,2012,25(01):50-53.
- [27] 王海玉,王映龙,闵建亮,胡剑锋.基于独立成分分析降噪与集成分类器的疲劳脑电分析[J].科学技术与工程,2018,18(32):200-205.
- [28] 王海玉,胡剑锋,王映龙.脑电信号处理方法的研究综述[J].计算机时代,2018(01):13-15+19.
- [29] 康利平. 光伏三电平逆变器的故障建模及其诊断方法研究[D].南昌: 南昌大学,2015.
- [30] 金纯, 曾伟. 基于脑电的实时疲劳监测算法的研究[J]. 科学技术与工程,2015,15(34):231-234+241.
- [31] 吴玉鹏.AR 谱在皮层痫样脑电信号分析应用[D].河北: 河北医科大学, 2015.
- [32] 郝冬梅, 阮晓钢.基于 LVQ 的自发脑电信号的分类研究[J].医疗设备信息, 2003, (10):7-9.
- [33] 冯春辉.基于脑电节律的模式识别方法研究[D].秦皇岛: 燕山大学, 2011.
- [34] 陈珊珊.时域分析技术在机械设备故障诊断中的应用[J].机械传动,2007,(03):79-83+4.
- [35] 刘海红,杜英举,姚伟华.不同意识脑电的双谱切片特分析与分类[J].电子设计工程,2015,23(08):42-45+48.
- [36] 赵卫峰. 基于时频分析的特征提取与模式分类方法研究[D].重庆: 重庆大学,2016.
- [37] 李冬梅.基于非线性动力学特征指标的癫痫脑电信号分类、定位与预测方法研究[D].新疆: 新疆医科大学, 2017.
- [38] 李树春, 李殿奎, 朱晓峰, 刘晓敏, 刘阳.基于关联维数和 Lyapunov 指数高频脑电非线性

- 动力学研究[J].现代生物, 2017, 17(03):561-566.[2017-09-30].DOI10.13241/j.cnki.pmb.2017.03.043
- [39] 高群霞, 周静, 吴效明.基于脑电信号的自动睡眠分期研究进展[J].生物医学工程学杂志, 2015, 32(05):1155-1159.
- [40] 初孟,邱天爽,鲍海平.基于人工神经网络的癫痫棘波检测方法[J].中国生物医学工程学报,2006,(04):421-425+429.
- [41] 卢章平,尹传斌,李瑞,何仁.基于生理信号的驾驶疲劳分级检测研究[J].重庆交通大学学报(自然科学版),2017,36(01):77-81.
- [42] 郭孜政,牛琳博,吴志敏,肖琼,史磊.基于 EEG 的驾驶疲劳识别算法及其有效性验证[J].北京工业大学学报,2017,43(06):929-934.
- [43] 谈玲.基于小波变换的信号滤波和去噪研究[J].科技信息(科学教研),2007(19):83+100.
- [44] 张壹.一种基于小波变换的脑电信号特征分析方法[A].中国仪器仪表学会.2004 全国测控、计量与仪器仪表学术年会论文集(下册)[C].中国仪器仪表学会:,2004:5.
- [45] 甘俊英.核主元分析特征提取法的研究[A].中国电子学会信号处理分会、中国仪器仪表学会信号处理分会.第十一届全国信号处理学术年会(CCS-2003)论文集[C].中国电子学会信号处理分会、中国仪器仪表学会信号处理分会:,2003:4.
- [46] 徐宝国,宋爱国,费树岷.在线脑机接口中脑电信号的特提取与分类方法[J].电子学报,2011,39(05):1025-1030.
- [47] 赵春临,郑崇勋,赵敏.基于核学习算法的驾驶精神疲劳分级研究[J].数据采集与处理,2009,24(03):335-339.
- [48] 韦振中.基于核主成分分析的特征提取方法[J].广西工学院学报,2006(04):27-31.
- [49] 高群霞.基于 EEG 的睡眠分期与睡眠评估方法研究[D].广东:华南理工大学,2015.
- [50] 周展鹏,孔万增,王奕直,周凌霄.基于心电和脑电的驾驶疲劳检测研究[J].杭州电子科技大学学报,2014,34(03):25-28.
- [51] 刘冲,于清文,陆志国,王宏.基于最小二乘支持向量机的脑电信号分类[J].东北大学学报(自然科学版),2016,37(05):634-637.
- [52] 王强,田学民.基于 KPCA-LSSVM 的软测量建模方法[J].化工学报,2011,62(10):2813-2817.
- [53] Hu JF. An Approach to EEG-Based Gender Recognition using Entropy Measurement Methods, Knowledge-Based Systems, 2017, 140, doi:10.1016/j.knosys.2017.10.032.
- [54] Hu JF. Comparison of Different Features and Classifiers for Driver Fatigue Detection Based on a Single EEG Channel, Computational and Mathematical Methods in Medicine, 2017, doi:10.1155/2017/5109530.
- [55] Min JL, Wang P, Hu JF. Driver fatigue detection through multiple entropy fusion analysis in an EEG-based system, PLoS ONE, 2017, 12, (12), e0188756. <https://doi.org/10.1371/journal.pone.0188756>.

- [56] 刘岩,李幼军,陈萌.基于固有模态分解和深度学习的抑郁症脑电信号分类分析[J].中国医学物理学杂志,2017,34(09):963-967.
- [57] 王银辉. 基于 EMD 和混合特征的癫痫脑电信号识别研究[D].大连: 大连理工大学,2016:17-24.
- [58] 刘超,林晓焕,廖文等.基于小波变换与 ICA 结合的 EP 信号提取研究[J].国外电子测量技术,2017,36(09):35-39.
- [59] 李昕,田彦秀,侯永捷等.小波变换结合经验模态分解在音乐干预脑电分析中的应用[J].生物医学工程学杂志,2016,33(04):762-769.
- [60] Lian Siyao.Blind source separation application in EEG artifact removal[D].Harbin: Harbin Engineering University, 2014.(in Chinese)廉思瑶. 盲信号分离在脑电信号伪迹去除中的应用[D].哈尔滨: 哈尔滨工程大学,2014.
- [61] Hu J, Wang P. Noise Robustness Analysis of Performance for EEG-Based Driver Fatigue Detection Using Different Entropy Feature Sets[J]. Entropy, 2017, 19(8).
- [62] 李立,曹锐,相洁.脑电数据近似熵与样本熵特征对比研究[J].计算机工程与设计,2014,35(03):1021-1026.
- [63] 田银,李沛洋,徐鹏.基于 AdaBoost 的脑机接口分类算法研究[J].电子科技大学学报,2013,42(05):791-793.
- [64] 雷雨,赵丹宁.最小二乘支持向量机回归的卫星钟差非线性组合预报[J].测绘科学,2015,40(05):33-36.
- [65] 迟恩楠,李春祥,郑晓芬.基于小波和乘法混合核函数 LSSVM 的顺风向非高斯空间风压预测[J].振动与冲击,2017,36(09):116-121.
- [66] Mu, ZD., Hu, JF., Yin, JH.: 'Driving Fatigue Detecting Based on EEG Signals of Forehead Area', International Journal of Pattern Recognition and Artificial Intelligence, 2016, 31, (05), pp 40-44.
- [67] 公昱文.独立分量分析及其在事件相关电位中的应用研究[D].天津: 天津师范大学,2010.
- [68] 张爱华,王琦,黄玲,等.意识任务识别中事件相关脑电的时频域信息熵表达[J].中国生物医学工程学报,2008,(04):515-520+542
- [69] 陈国强,赵俊伟,黄俊杰,等.基于 Matlab 的 AR 模型参数估计[J].工具技术,2005,(04):39-40.
- [70] 朱华,黄张裕,赵仲荣,等.基于不同核函数的 LSSVM 在 GPS 高程拟合中的应用[J].测绘工程,2014,23(05):18-20.
- [71] Zhu Hua, Huang Zhang-yu, Zhao Zhong-rong, et al. Application of LSSVM based on different Kernel function in GPS height fitting [J]. Surveying and mapping engineering. 2014,23(05):18-20.
- [72] 席茜,张凤琴,李小青,等.基于对偶随机投影的线性核支持向量机[J].计算机应用,2017,37(06):1680-1685.
- [73] Ram Bilas Pachori,Rajeev Sharma,Shivnarayan Patidar. Classification of Normal and Epileptic

Seizure EEG Signals Based on Empirical Mode Decomposition[M]. Springer International Publishing: 2015-06-15.

[74] Dehuri S, Jagadev A K, Satapathy S K. SMOTE and ABC optimised RBF network for coping with imbalanced class in EEG signal classification[J]. International Journal of Medical Engineering & Informatics, 2018, 10(3):215.

## 致谢

三年的研究生生活一晃而过，往昔岁月历历在目，心中倍感充实，在即将毕业之际，感慨良多。

首先诚挚的感谢我的导师王映龙教授，硕士期间，王映龙教授给我们提供了优越的学习环境，创造良好的学术氛围，提供了很多难得的锻炼机会，让我们能够在有限的研究生生涯体验更多的知识、开阔我们的学术眼界，为提高我们的学术水平创造良好的条件。王映龙教授还对我的研究方向提出了导向性意见，从研究方向的确定、论文的修改以及定稿过程中都倾注着王映龙教授的心血。王映龙教授作为学院的院长以严谨的治学之道、积极乐观的生活态度，为我树立了终身学习的典范，他的教诲将激励我在科学和教育的道路上励精图治，开拓创新。

我还要以诚挚的心意感谢胡剑锋博士，为我指明了研究方向，从实验的思路、经费、人员到实验仪器，感谢他的团队——协同创新中心信息技术研究所的全体同仁们。感谢重点实验室穆振东、闵建亮老师和王平老师，带我走进脑电研究的领域；感谢闵建亮老师，教会我编程，学会脑电信号的处理；感谢穆振东老师，在我迷茫时期的各种帮忙，鼓励我、为我分担各种困难，指点我实验思路、实验方法，穆老师的认真博学为我解答知识难题，在我的研究方向上起到了很大得助力作用，从我论文的初始思路到实验的实施再到实验数据的分析，每一步都包含着穆老师的认真负责；感谢王平老师，是他教会了我实验仪器的使用，带我完成完整的实验，积累实验经验、处理实验中的各种状况，更深层次的接触到脑电实验的领域，让我成长了很多。

感谢协同创新中心信息技术研究所的伙伴们，感谢他们一年多来给我带来的帮忙和关怀，与他们在一起的日子开心而充实，给我单调枯燥的科研生活增添了绚丽的色彩，在我实验陷入瓶颈时期给予我鼓励温暖，总是尽自己全力帮助我，在我迷茫时为我指引方向帮我分析利弊，从生活的点点滴滴中给我帮忙、给我鼓励温暖，成为我在校园最美好的回忆。

感谢我的辅导员刘津，学习和工作上都是我的人生导师，给予了我很多的帮助和开导，让我在人生道路上能够越走越平稳；感谢陈清华师兄等师兄师姐，在我的人生规划中给出了相当重要的指导性建议；感谢我的舍友，黄琴、彭莉莎、邹诗雨，在我情绪低落时为我排忧解难，快乐而温馨的宿舍生活总能带给我一份好情绪；感谢施海林同学，无论是工作还是学习上都给予了我很多帮助，以幽默的方式带给大家很多欢乐，带给我们很多正能量，带我们走上积极向上的生活方式；感谢刘芝鹏同学，踏实认真的工作态度成为了我们学习的榜样，感染着身边人向着更优秀的工作生活模式进步；还要感谢班级的同学们，陪伴我渡过美好的校园生活。

## 致谢

---

感谢我的父母和家人，他们在背后的默默支持是我前进的动力，为了论文和实验，在读期间距离家远，能够陪伴父母的时间少之又少，心中甚感愧疚。

最后，真诚的感恩所有在研究生阶段给予我帮助的老师，同学，朋友们！

攻读硕士学位期间发表论文

[1]王海玉,王映龙,闵建亮,胡剑锋.基于独立成分分析降噪与集成分类器的疲劳脑电分析[J].科学技术与工程,2018,18(32):200-205.

[2]王海玉,王映龙,闵建亮,胡剑锋.基于小波变换与多项指标的疲劳驾驶检测应用[J].计算机与现代化,2018(10):32-35+43.

[3]王海玉,胡剑锋,王映龙.脑电信号处理方法的研究综述[J].计算机时代,2018(01):13-15+19.

[4]王海玉,王映龙,闵建亮,胡剑锋.集成三种不同核分类器的疲劳脑电分类研究.已被《计算机工程与设计》录用。