



(12)发明专利申请

(10)申请公布号 CN 110390272 A

(43)申请公布日 2019.10.29

(21)申请号 201910582226.X

(22)申请日 2019.06.30

(71)申请人 天津大学

地址 300072 天津市南开区卫津路92号

(72)发明人 董娜 李英杰 常建芳 高忠科

(74)专利代理机构 天津市北洋有限责任专利代
理事务所 12201

代理人 杜文茹

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

A61B 5/00(2006.01)

A61B 5/04(2006.01)

A61B 5/0476(2006.01)

A61B 5/18(2006.01)

权利要求书2页 说明书7页 附图4页

(54)发明名称

一种基于加权主成分分析的EEG信号特征降
维方法

(57)摘要

一种基于加权主成分分析的EEG信号特征降维方法:提取疲劳驾驶的m个EEG信号的样本,分为训练集和测试集进行训练,得到分类总准确率A和n个不同的分类准确率;将总准确率A分别与不同的分类准确率做差,得到n个差值;对n个差值做归一化处理得到n个权值;对n个权值构造权值对角矩阵;将m个EEG信号的样本写成一个m*n维矩阵;将m*n维矩阵与权值的对角矩阵相乘,得到加权的EEG信号特征数据;计算其协方差矩阵并分解,得到协方差矩阵的特征值和特征值所对应的单位化特征向量;选取前k个协方差矩阵的特征值对应的单位化特征向量进行组合形成映射矩阵;从而得到降维的EEG信号特征数据。本发明有效提高了分类识别精度和降低识别模型的训练时间。



1. 一种基于加权主成分分析的EEG信号特征降维方法,其特征在于,包括如下步骤:

1) 应用AR模型提取疲劳驾驶的m个EEG信号的样本,每个样本包含有n维特征数据;

2) 将m个EEG信号的样本分为训练集和测试集,使用SVM分别对训练集和测试集进行训练,并对测试集进行分类,得到分类总准确率A;

3) 去掉每个EEG信号样本的n维特征数据中的第一维特征数据,得到每个EEG信号样本的n-1维特征数据,将m个EEG信号的样本分为训练集和测试集,使用SVM分别对训练集和测试集进行训练,并对测试集进行分类,得到分类准确率 A_1 ;去掉每个EEG信号样本的n维特征数据中的第二维数据,将m个EEG信号的样本分为训练集和测试集,用SVM对剩下的n-1维数据进行训练并分类,得到分类的准确率 A_2 ,依次类推,最终得到n个分类的准确率 $A_1, A_2, \dots, A_i, \dots, A_n$;

4) 将步骤2)得到的分类总准确率A分别与步骤3)得到的分类准确率 $A_1, A_2, \dots, A_i, \dots, A_n$ 做差,得到n个差值;当差值为负数时,做差的分类准确率 A_i 所对应的去掉的第i维特征对分类有负影响,当差值为正数时,做差的分类准确率 A_i 所对应的去掉的第i维特征对分类有正影响;

5) 对n个差值分别做归一化处理得到n个权值;

6) 对n个权值 $w_i \ i=1, 2, \dots, n$,构造权值对角矩阵 $W_{n \times n}$:

$$W_{n \times n} = \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{bmatrix} \quad (2);$$

7) 将m个EEG信号的样本写成一个 $m \times n$ 维矩阵 $X_{m \times n}$:

$$\begin{Bmatrix} x_1 (x_{11}, \dots, x_{1n}) \\ \dots \dots \dots \\ x_m (x_{m1}, \dots, x_{mn}) \end{Bmatrix} \Rightarrow X_{m \times n} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \quad (3)$$

其中, $x_1 \dots x_m$ 为m个EEG信号的样本;

8) 将 $m \times n$ 维矩阵 $X_{m \times n}$ 与权值w的对角矩阵 $W_{n \times n}$ 相乘,得到加权的EEG信号特征数据 $Z_{m \times n}$:

$$Z_{m \times n} = X_{m \times n} * W_{n \times n} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} * \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{bmatrix} = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ \cdots & \cdots & \cdots \\ z_{m1} & \cdots & z_{mn} \end{bmatrix} \quad (4);$$

9) 计算加权的EEG信号特征数据 $Z_{m \times n}$ 的协方差矩阵 C' ;

10) 分解协方差矩阵 C' ,得到协方差矩阵的特征值和特征值所对应的单位化特征向量;

11) 按照给定的累计贡献率选取前k个协方差矩阵的特征值对应的单位化特征向量进行组合形成映射矩阵 P' ;

12) 利用映射矩阵 P' ,通过如下公式得到降维的EEG信号特征数据 Y' :

$$Y' = Z_{m \times n} P' \quad (9)$$

式中, $Z_{m \times n}$ 为加权的EEG信号特征数据。

2. 根据权利要求1所述的一种基于加权主成分分析的EEG信号特征降维方法,其特征在于,步骤5)所述的归一化处理是采用如下公式:

$$w = \frac{d - \min}{\max - \min} \quad (1)$$

其中,d为归一化之前的数据,max是n个差值中最大的值,min是n个差值中最小的值,w为权值。

3.根据权利要求1所述的一种基于加权主成分分析的EEG信号特征降维方法,其特征在于,步骤9)所述的协方差矩阵是采用如下公式计算:

$$C' = \frac{1}{n-1} Z_{m \times n}^T Z_{m \times n} \quad (5)$$

式中, C' 为协方差矩阵;n表示每个样本特征的维数; $Z_{m \times n}$ 表示加权的EEG信号特征数据。

4.根据权利要求1所述的一种基于加权主成分分析的EEG信号特征降维方法,其特征在于,步骤10)所述的分解协方差矩阵 C' 是采用如下公式

$$\lambda_i u_i = C' u_i, i = 1, 2, 3, \dots, n \quad (6)$$

式中, λ_i 为协方差矩阵的第i个特征值, u_i 为 λ_i 对应的单位化特征向量,对得到的n个特征值按降序排列,得到 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。

5.根据权利要求1所述的一种基于加权主成分分析的EEG信号特征降维方法,其特征在于,步骤11)包括,选取前k个协方差矩阵的特征值,第i个协方差矩阵的特征值的累积贡献率计算如下:

$$\alpha_i = \frac{\sum_{j=1}^i \lambda_j}{\sum_{j=1}^k \lambda_j} \quad (7)$$

根据累计贡献率,选取前k个协方差矩阵的特征值对应的单位化特征向量进行组合形成映射矩阵 P'

$$P' = [u_1, u_2, \dots, u_k] \quad (8)$$

式中, $i \leq k$, α_i 为累计贡献率, λ_j 为协方差矩阵的第j个特征值, $u_1, u_2, \dots, u_k$ 为前k个协方差矩阵的特征值对应的单位化特征向量。

一种基于加权主成分分析的EEG信号特征降维方法

技术领域

[0001] 本发明涉及一种EEG信号特征降维方法。特别是涉及一种基于加权主成分分析的EEG信号特征降维方法。

背景技术

[0002] 近十年来,我国的汽车数量急剧增加,交通事故也随之增多[1]。据有关报道,我国已成为交通事故多发国家之一。导致交通事故的因素很多,其中,驾驶员疲劳驾驶是导致交通事故的主要原因。驾驶员在疲劳状态下会出现注意力分散、思维活动降低的现象,进而造成其反应迟钝、车辆控制力下降,增加发生交通事故的可能性[2]。因此,准确快速检测驾驶员驾驶疲劳状态显得尤为重要。基于生理信息的驾驶疲劳检测是一种可以根据生理指标变化对驾驶员疲劳状态做出检测识别的客观手段[3]。研究表明,通过体温、血压、心电图、脑电图(electroencephalogram, EEG)、肌电图等电生理信息可对驾驶员疲劳状态进行有效评价与检测。其中,脑电信号的分析是最为常用的一种判别方法。

[0003] 脑电反映了神经系统的电生理信号,通过脑电信号可以对驾驶员的疲劳状态进行检测与分析。由于脑电信号所提特征通常是高维特征信号,这给我们的处理带来了一定的困难,为了解决这个问题,我们需要对高维特征信号进行降维处理。主成分分析(principal component analysis, PCA)是一种常用的降维方法,它可以从多元事物中解析出主要影响因素,揭示事物的本质,简化复杂的问题。

[0004] Liu J等对于情绪分类问题提出了一种混合维度特征减少方案[4],从EEG中提取14种不同的特征,最小冗余最大相关性(mRMR)用于最大化特征与分类变量之间的相关性和最小化特征与特征之间的相关性,通过PCA进一步减少生成的特征,以提取主成分。Bousseta R等通过连续小波变换(CWT)和经验模态分解(EMD)两种方法提取EEG特征[5],并使用PCA进行特征降维,通过线性和径向基函数(RBF)核函数SVM分类器进行左右手运动想象分类。孙颖等提出了基于PCA的非线性全局特征和功率谱熵的融合算法[6],提取EEG的非线性几何特征作为新的情感脑电特征,并结合EEG的功率谱熵以及Hurst指数等非线性属性特征,通过PCA技术进行降维以及特征融合,以SVM为分类器进行情感识别。Neshov NN等提出了使用6通道EEG数据识别五个心理任务的算法[7],主要思想是将原始EEG信号分成几帧并计算其频谱,应用高斯的二阶导数来提取特征,并使用PCA降低特征维度。Xin L I等提取8种正负情绪代表各个脑区的14个通道脑电数据[8],基于小波分解重构delta、theta、alpha、beta四种节律波,使用主成分分析(PCA)融合小波特征、近似熵和Hurst指数的脑电特征并降低特征维数。Zarei R等提出了一种特征提取方法[9],结合PCA和交叉协方差技术(CCOV),在脑机接口技术应用中基于EEG信号从心理状态中提取判别信息,应用基于相关性的变量选择方法,对提取的特征进行最佳的第一次搜索,以识别用于表征精神状态信号分布的最佳特征集。

[0005] 前人在对PCA的应用中,都是基于数据方差降维的,平均的对待每一维特征。然而不同的特征在识别过程中所起的作用是不同的[10],因此不能同等的对待每一维特征。由

以上分析可见,目前对于PCA降维方法还有待进一步的完善和更新。

[0006] 参考文献

[0007] [1]王福旺,王宏.长途客车驾驶员疲劳状态脑电特征分析[J].仪器仪表学报,2013,34(5):1146-1152.

[0008] [2]于旭蕾,李相泽.基于便携式脑电数据的实时疲劳驾驶检测系统[J].现代信息科技,2018,v.2(04):45-47.

[0009] [3]叶建芳,刘强,李雪莹.基于随机森林的疲劳驾驶检测识别模型的优化研究[J].汽车实用技术,2018,No.268(13):46-50.

[0010] [4]Liu J,Meng H,Li M,et al.Emotion detection from EEG recordings based on supervised and unsupervised dimension reduction[J].Concurrency and Computation:Practice and Experience,2018:e4446.

[0011] [5]Bousseta R,Tayeb S,Ouakouak I E,et al.EEG efficient classification of imagined right and left hand movement using RBF kernel SVM and the joint CWT_PCA[J].2018.

[0012] [6]Ying S,Jianghe M A,Xueying Z.EEG emotion recognition based on nonlinear global features and spectral feature[J].Computer Engineering and Applications,2018.

[0013] [7]Neshov NN,Manolova AH,Draganov IR,et al.Classification of Mental Tasks from EEG Signals Using Spectral Analysis,PCA and SVM[J].Cybernetics and Information Technologies,2018.

[0014] [8]李昕,蔡二娟,田彦秀,et al.一种改进脑电特征提取算法及其在情感识别中的应用[J].生物医学工程学杂志,2017(04):32-39+50.

[0015] [9]Zarei R,He J,Siuly S,et al.A PCA Aided Cross-Covariance Scheme for Discriminative Feature Extraction from EEG Signals[J].Computer methods and programs in biomedicine,2017,146:47.

[0016] [10]王永欣,张化祥,王爽.基于属性加权的主成分分析算法[J].济南大学学报(自然科学版),2015,29(6):438-443.

发明内容

[0017] 本发明所要解决的技术问题是,提供一种能够提高疲劳驾驶状态分类准确率的基于加权主成分分析的EEG信号特征降维方法。

[0018] 本发明所采用的技术方案是:一种基于加权主成分分析的EEG信号特征降维方法,包括如下步骤:

[0019] 1)应用AR模型提取疲劳驾驶的m个EEG信号的样本,每个样本包含有n维特征数据;

[0020] 2)将m个EEG信号的样本分为训练集和测试集,使用SVM分别对训练集和测试集进行训练,并对测试集进行分类,得到分类总准确率A;

[0021] 3)去掉每个EEG信号样本的n维特征数据中的第一维特征数据,得到每个EEG信号样本的n-1维特征数据,将m个EEG信号的样本分为训练集和测试集,使用SVM分别对训练集和测试集进行训练,并对测试集进行分类,得到分类准确率A₁;去掉每个EEG信号样本的n维

特征数据中的第二维数据,将m个EEG信号的样本分为训练集和测试集,用SVM对剩下的n-1维数据进行训练并分类,得到分类的准确率 A_2 ,依次类推,最终得到n个分类的准确率 $A_1, A_2, \dots, A_i, \dots, A_n$;

[0022] 4) 将步骤2)得到的分类总准确率A分别与步骤3)得到的分类准确率 $A_1, A_2, \dots, A_i, \dots, A_n$ 做差,得到n个差值;当差值为负数时,做差的分类准确率 A_i 所对应的去掉的第i维特征对分类有负影响,当差值为正数时,做差的分类准确率 A_i 所对应的去掉的第i维特征对分类有正影响;

[0023] 5) 对n个差值分别做归一化处理得到n个权值;

[0024] 6) 对n个权值 $w_i \ i=1, 2, \dots, n$,构造权值对角矩阵 $W_{n \times n}$:

$$[0025] \quad W_{n \times n} = \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{bmatrix} \quad (2);$$

[0026] 7) 将m个EEG信号的样本写成一个 $m \times n$ 维矩阵 $X_{m \times n}$:

$$[0027] \quad \begin{cases} x_1 (x_{11}, \dots, x_{1n}) \\ \dots \dots \dots \\ x_m (x_{m1}, \dots, x_{mn}) \end{cases} \Rightarrow X_{m \times n} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \quad (3)$$

[0028] 其中, $x_1 \dots x_m$ 为m个EEG信号的样本;

[0029] 8) 将 $m \times n$ 维矩阵 $X_{m \times n}$ 与权值w的对角矩阵 $W_{n \times n}$ 相乘,得到加权的EEG信号特征数据 $Z_{m \times n}$:

$$[0030] \quad Z_{m \times n} = X_{m \times n} * W_{n \times n} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} * \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{bmatrix} = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ \cdots & \cdots & \cdots \\ z_{m1} & \cdots & z_{mn} \end{bmatrix} \quad (4);$$

[0031] 9) 计算加权的EEG信号特征数据 $Z_{m \times n}$ 的协方差矩阵 C' ;

[0032] 10) 分解协方差矩阵 C' ,得到协方差矩阵的特征值和特征值所对应的单位化特征向量;

[0033] 11) 按照给定的累计贡献率选取前k个协方差矩阵的特征值对应的单位化特征向量进行组合形成映射矩阵 P' ;

[0034] 12) 利用映射矩阵 P' ,通过如下公式得到降维的EEG信号特征数据 Y' :

$$[0035] \quad Y' = Z_{m \times n} P' \quad (9)$$

[0036] 式中, $Z_{m \times n}$ 为加权的EEG信号特征数据。

[0037] 步骤5)所述的归一化处理是采用如下公式:

$$[0038] \quad w = \frac{d - \min}{\max - \min} \quad (1)$$

[0039] 其中,d为归一化之前的数据,max是n个差值中最大的值,min是n个差值中最小的值,w为权值。

[0040] 步骤9)所述的协方差矩阵是采用如下公式计算:

$$[0041] \quad C' = \frac{1}{n-1} Z_{m \times n}^T Z_{m \times n} \quad (5)$$

[0042] 式中, C' 为协方差矩阵;n表示每个样本特征的特征维数; $Z_{m \times n}$ 表示加权的EEG信号特征数据。

[0043] 步骤10)所述的分解协方差矩阵 C' 是采用如下公式

$$[0044] \quad \lambda_i u_i = C' u_i, i=1,2,3,\dots,n \quad (6)$$

[0045] 式中, λ_i 为协方差矩阵的第 i 个特征值, u_i 为 λ_i 对应的单位化特征向量,对得到的 n 个特征值按降序排列,得到 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。

[0046] 步骤11)包括,选取前 k 个协方差矩阵的特征值,第 i 个协方差矩阵的特征值的累积贡献率计算如下:

$$[0047] \quad \alpha_i = \frac{\sum_{j=1}^i \lambda_j}{\sum_{j=1}^k \lambda_j} \quad (7)$$

[0048] 根据累计贡献率,选取前 k 个协方差矩阵的特征值对应的单位化特征向量进行组合形成映射矩阵 P'

$$[0049] \quad P' = [u_1, u_2, \dots, u_k] \quad (8)$$

[0050] 式中, $i \leq k$, α_i 为累计贡献率, λ_j 为协方差矩阵的第 j 个特征值, $u_1, u_2, \dots, u_k$ 为前 k 个协方差矩阵的特征值对应的单位化特征向量。

[0051] 本发明的一种基于加权主成分分析的EEG信号特征降维方法,有益效果是:

[0052] 第一,主成分分析法是基于数据方差降维的,平均的对待每一维特征,然而不同的特征在识别过程中所起的作用是不同的。本发明基于特征数据的属性进行降维,加强某些对识别关键的特征,同时减弱某些与识别关系不大的特征。

[0053] 第二,应用本发明对脑电特征数据进行降维,可以有效的提高分类识别精度和降低识别模型的训练时间。

附图说明

[0054] 图1是本发明一种基于加权主成分分析的EEG信号特征降维方法的流程图;

[0055] 图2a是受试者1的实验结果图;

[0056] 图2b是受试者2的实验结果图;

[0057] 图2c是受试者3的实验结果图;

[0058] 图2d是受试者4的实验结果图;

[0059] 图2e是受试者5的实验结果图。

具体实施方式

[0060] 下面结合实施例和附图对本发明的一种基于加权主成分分析的EEG信号特征降维方法做出详细说明。

[0061] 发明的一种基于加权主成分分析的EEG信号特征降维方法,先依次去除一个特征,并统计不同特征对疲劳状态分类性能的影响,然后将不同特征的准确率下降值归一化处理作为该特征的权值,最后引入主成分分析的方法建立加权主成分分析的方法对特征进行降维。

[0062] 发明的一种基于加权主成分分析的EEG信号特征降维方法,包括如下步骤:

[0063] 1) 应用AR模型提取疲劳驾驶的 m 个EEG信号的样本,每个样本包含有 n 维特征数据;

[0064] 2) 将 m 个EEG信号的样本分为训练集和测试集,使用SVM分别对训练集和测试集进行训练,并对测试集进行分类,得到分类总准确率 A ;

[0065] 3) 去掉每个EEG信号样本的n维特征数据中的第一维特征数据,得到每个EEG信号样本的n-1维特征数据,将m个EEG信号的样本分为训练集和测试集,使用SVM分别对训练集和测试集进行训练,并对测试集进行分类,得到分类准确率 A_1 ;去掉每个EEG信号样本的n维特征数据中的第二维数据,将m个EEG信号的样本分为训练集和测试集,用SVM对剩下的n-1维数据进行训练并分类,得到分类的准确率 A_2 ,依次类推,最终得到n个分类的准确率 $A_1, A_2 \cdots A_i, \cdots, A_n$;

[0066] 4) 将步骤2)得到的分类总准确率A分别与步骤3)得到的分类准确率 $A_1, A_2 \cdots A_i, \cdots, A_n$ 做差,得到n个差值;当差值为负数时,做差的分类准确率 A_i 所对应的去掉的第i维特征对分类有负影响,当差值为正数时,做差的分类准确率 A_i 所对应的去掉的第i维特征对分类有正影响;

[0067] 5) 对n个差值分别做归一化处理得到n个权值;所述的归一化处理是采用如下公式:

$$[0068] \quad w = \frac{d - \min}{\max - \min} \quad (1)$$

[0069] 其中,d为归一化之前的数据,max是n个差值中最大的值,min是n个差值中最小的值,w为权值。

[0070] 6) 对n个权值 $w_i \quad i=1, 2, \cdots, n$,构造权值对角矩阵 $W_{n \times n}$:

$$[0071] \quad W_{n \times n} = \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{bmatrix} \quad (2);$$

[0072] 7) 将m个EEG信号的样本写成一个 $m \times n$ 维矩阵 $X_{m \times n}$:

$$[0073] \quad \begin{cases} x_1 (x_{11}, \cdots, x_{1n}) \\ \cdots \cdots \cdots \\ x_m (x_{m1}, \cdots, x_{mn}) \end{cases} \Rightarrow X_{m \times n} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \quad (3)$$

[0074] 其中, $x_1 \dots x_m$ 为m个EEG信号的样本;

[0075] 8) 将 $m \times n$ 维矩阵 $X_{m \times n}$ 与权值w的对角矩阵 $W_{n \times n}$ 相乘,得到加权的EEG信号特征数据 $Z_{m \times n}$:

$$[0076] \quad Z_{m \times n} = X_{m \times n} * W_{n \times n} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} * \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{bmatrix} = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ \cdots & \cdots & \cdots \\ z_{m1} & \cdots & z_{mn} \end{bmatrix} \quad (4)$$

[0077] 9) 计算加权的EEG信号特征数据 $Z_{m \times n}$ 的协方差矩阵 C' ;所述的协方差矩阵是采用如下公式计算:

$$[0078] \quad C' = \frac{1}{n-1} Z_{m \times n}^T Z_{m \times n} \quad (5)$$

[0079] 式中, C' 为协方差矩阵;n表示每个样本特征的维数; $Z_{m \times n}$ 表示加权的EEG信号特征数据。

[0080] 10) 分解协方差矩阵 C' ,得到协方差矩阵的特征值和特征值所对应的单位化特征向量;所述的分解协方差矩阵 C' 是采用如下公式

$$[0081] \quad \lambda_i u_i = C' u_i, i=1, 2, 3, \cdots, n \quad (6)$$

[0082] 式中, λ_i 为协方差矩阵的第i个特征值, u_i 为 λ_i 对应的单位化特征向量,对得到的n

个特征值按降序排列,得到 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。

[0083] 11) 按照给定的累计贡献率选取前k个协方差矩阵的特征值对应的单位化特征向量进行组合形成映射矩阵 P' ;包括,选取前k个协方差矩阵的特征值,第i个协方差矩阵的特征值的累积贡献率计算如下:

$$[0084] \quad \alpha_i = \frac{\sum_{j=1}^i \lambda_j}{\sum_{j=1}^k \lambda_j} \quad (7)$$

[0085] 根据累计贡献率,选取前k个协方差矩阵的特征值对应的单位化特征向量进行组合形成映射矩阵 P'

$$[0086] \quad P' = [u_1, u_2, \dots, u_k] \quad (8)$$

[0087] 式中, $i \leq k$, α_i 为累计贡献率, λ_j 为协方差矩阵的第j个特征值, $u_1, u_2, \dots, u_k$ 为前k个协方差矩阵的特征值对应的单位化特征向量。

[0088] 12) 利用映射矩阵 P' ,通过如下公式得到降维的EEG信号特征数据 Y' :

$$[0089] \quad Y' = Z_m * P' \quad (9)$$

[0090] 式中, $Z_m * n$ 为加权的EEG信号特征数据。

[0091] 本发明实施例的EEG信号获取,是在模拟驾驶实验中,5名年龄从19岁到26岁(3名男性和2名女性;平均年龄:22.73)的右利手大学生自愿参与实验,他们都没有与精神病相关的疾病。在实验前两天,要求受试者避免服用抗疲劳饮料或引起嗜睡的药物。同时,他们需要保持合理的休息,每晚睡眠时间超过7小时。由于所有受试者都没有接触驾驶模拟器,所以在实验之前,他们要练习驾驶直到可以熟练操作。

[0092] 实验使用PGFD001驾驶模拟器并配备了踏板、方向盘和离合器。在虚拟驾驶软件3DInstructor2中,使用普通汽车Phaeton2.0L,默认为自动换档。实验道路为一条几乎没有弯道的高速公路。此外,增加了一个网络摄像头360D618,一台投影仪和一个立体声音箱,以提高感知能力。

[0093] 为了提高受试者的疲劳感,实验时间为14:00-15:30,有研究表明这是一个容易陷入疲劳的时期,一次实验时间大约90分钟。在隔离和安静的房间收集受试者的全程头皮EEG信号。此外,我们还通过前置摄像头监测受试者的面部状态,以验证疲劳程度。在实验开始之前,有10分钟的场景设置和大约20分钟的驾驶练习。实验开始后,受试者将连续驾驶直到报告他的轻度疲劳,这通常持续约30分钟作为警觉状态。然后,在连续驾驶过渡10分钟之后,受试者经历了另外30分钟的驾驶,其被用作疲劳状态。在此阶段之后,要求每个受试者报告一些问题和实验后评论。每个受试者的记录时间接近90分钟,这随个体差异而略有变化。随着实验的进行,受试者的疲劳程度变得越来越强烈。作为疲劳状态的重要生理测量,他们的主观评价和实际行为显示出越来越多的疲劳。

[0094] 数据采集与数据预处理:

[0095] 作为驾驶员疲劳状态的测量,EEG记录装置配备有ESI Neuroscan系统,具有40个电极并根据国际标准10/20系统排列,采样频率为200Hz。在采集之前,通过注入导电凝胶将EEG电极的皮肤阻抗调节到5k Ω 以下。在实验期间,所有受试者尽可能限制不必要的身体动作,尽量保持恒定速度并避免发生汽车碰撞。在这40个电极中,除了作为内部结构的四个电极之外,还有两个被定义为参考电极,并且四个(横跨水平和垂直方向放置)用于监视眼睛运动。原始EEG信号会受到基于高频噪声和基于眼电的低频噪声的干扰,因此需要使用

EEGLAB工具箱进行预处理。在消除了噪声和眼电伪影的干扰后获得了30通道EEG信号。

[0096] 选取警觉状态的前10min采集的数据作为非疲劳数据,选取疲劳状态的后10min采集的数据作为疲劳数据。使用具有固定长度的滑动窗口进行数据分割,窗口长度为1秒,没有重叠。数据分割完成后,每个受试者都获得1200组样本。

[0097] 本发明实施例应用AR模型对脑电数据提取特征,AR建模原理如下:

$$[0098] \quad x(t) = \sum_{k=1}^p a(k)x(t-k) + e(t) \quad (10)$$

[0099] 其中 $x(t)$ 表示时间 t 的EEG数据, p 表示AR模型阶数, $e(t)$ 是白噪声序列, $a(k)$ 表示AR模型系数。

[0100] AR建模需要选择模型阶数。阶数分别为3阶、4阶和5阶。AR特征的大小等于AR阶数乘以30个EEG通道单位,因此AR 3阶获得90个特征单位,AR 4阶获得120个特征单位,AR 5阶获得150个特征单位。

[0101] 为了比较,本发明实施例也采用了功率谱密度作为特征提取方式。脑电图包括四个波段: δ (1-3Hz)、 θ (4-7Hz)、 α (8-13Hz) 和 β (14-30Hz)。采用梯形数值积分法计算每个EEG活动带的总功率作为特征值,可以获得4个单位的特征值,所以30个EEG通道一共获得120个特征单位。

[0102] 下面通过分类器及分类指标验证本发明的一种基于加权主成分分析的EEG信号特征降维方法:

[0103] 支持向量机是一种高效的有监督二分类器,在实现小样本、非线性和高维模式的数据分类中具有重要作用。本发明选用SVM作为分类器,径向基(RBF)作为核函数。

[0104] 为评估分类结果,采用以下三个评估指标:

$$[0105] \quad \text{准确率: Accuracy} = \frac{TP+TN}{TP+FN+FP+TN} * 100\% \quad (11)$$

$$[0106] \quad \text{灵敏度: Sensitivity} = \frac{TP}{TP+FN} * 100\% \quad (12)$$

$$[0107] \quad \text{特异性: Specificity} = \frac{TN}{TN+FP} * 100\% \quad (13)$$

[0108] 其中,TP为正确识别出正样本的数量,即将疲劳驾驶正确识别出的样本数量;TN为正确识别出负样本数量,即将正常驾驶识别为正常驾驶的样本数量;FN为未识别出正样本的数量,即将疲劳驾驶识别为正常驾驶的样本数量;FP为未识别出负样本的数量,即将正常驾驶识别为疲劳驾驶的样本数量。准确率反映了分类正确的样本占总体样本的比例,灵敏度反映了正样本的分类准确率,特异性反映了负样本的分类准确率。

[0109] Matlab实验设置:

[0110] 在5名受试者的30路脑电信号数据集中,每个受试者都获得1200组样本数据。本发明采用AR模型对数据集提取特征,阶数分别为3阶、4阶和5阶,同时也采用了功率谱密度作为特征提取方式。

[0111] 本发明实施例采用WPCA对特征数据降维,累计贡献率为0.95,选取SVM作为分类器。为了验证实验结果的可靠性,使用了10折交叉验证法:将数据集分为10份,1份作为测试集,其余9份为训练集。交叉验证重复10次,使得每份都可以作为测试集测试一次,最后取10次测试数据的平均值作为最终结果。

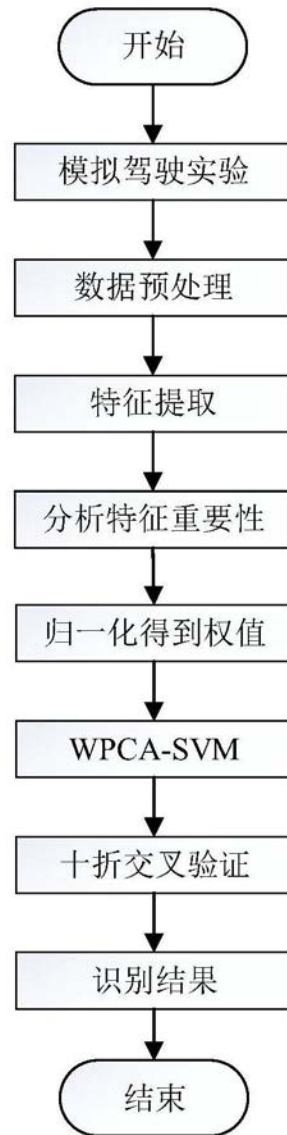


图1

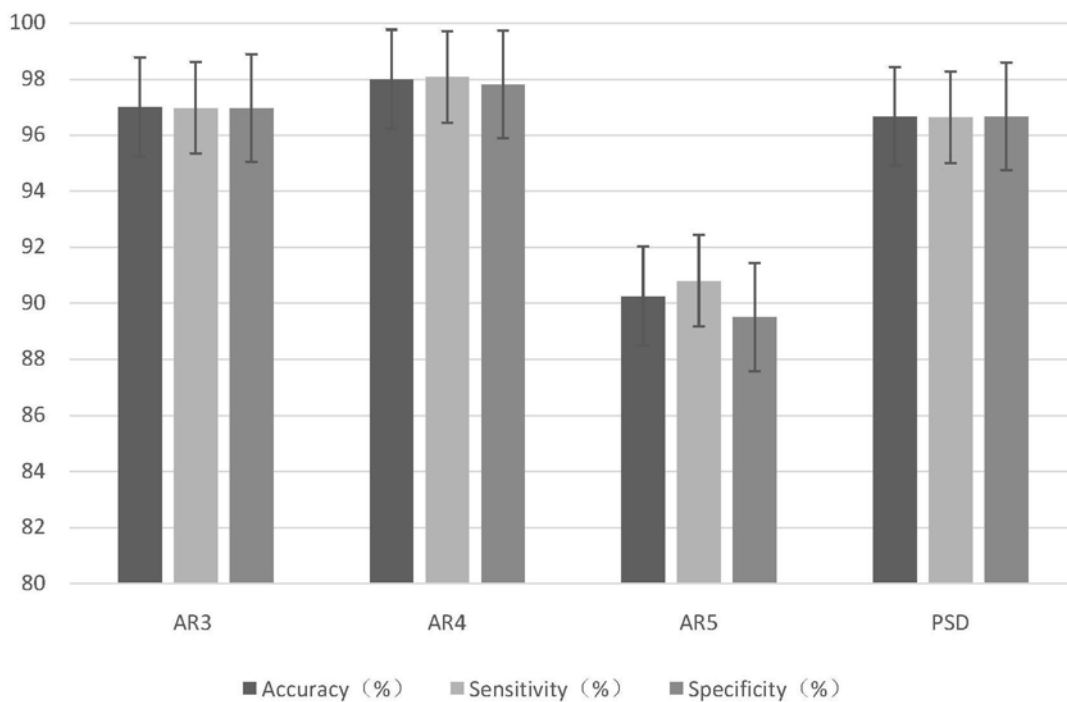


图2a

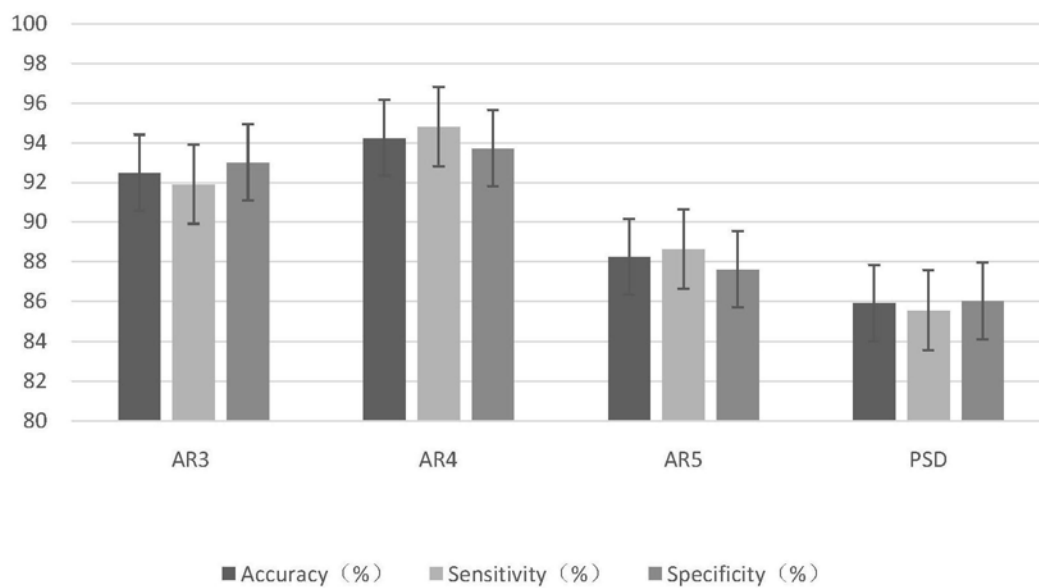


图2b

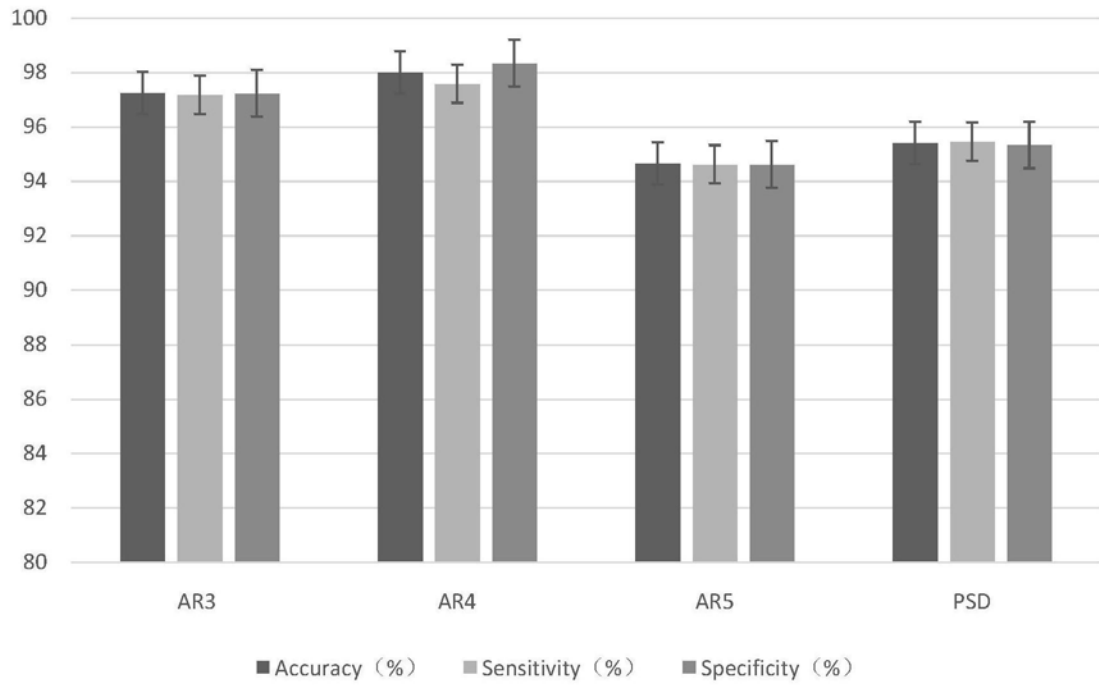


图2c

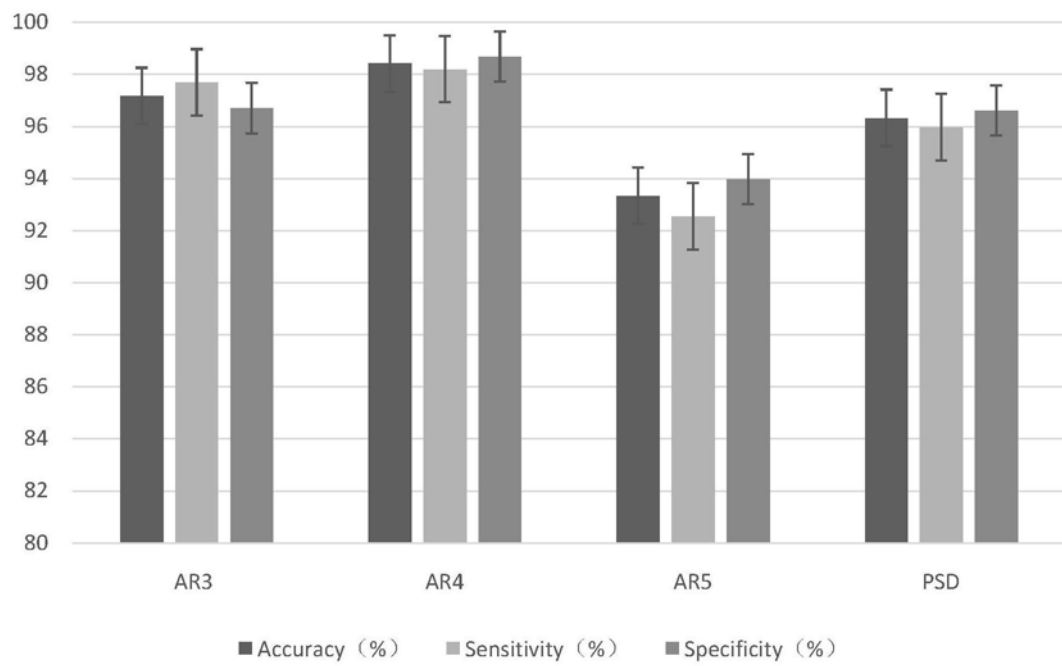


图2d

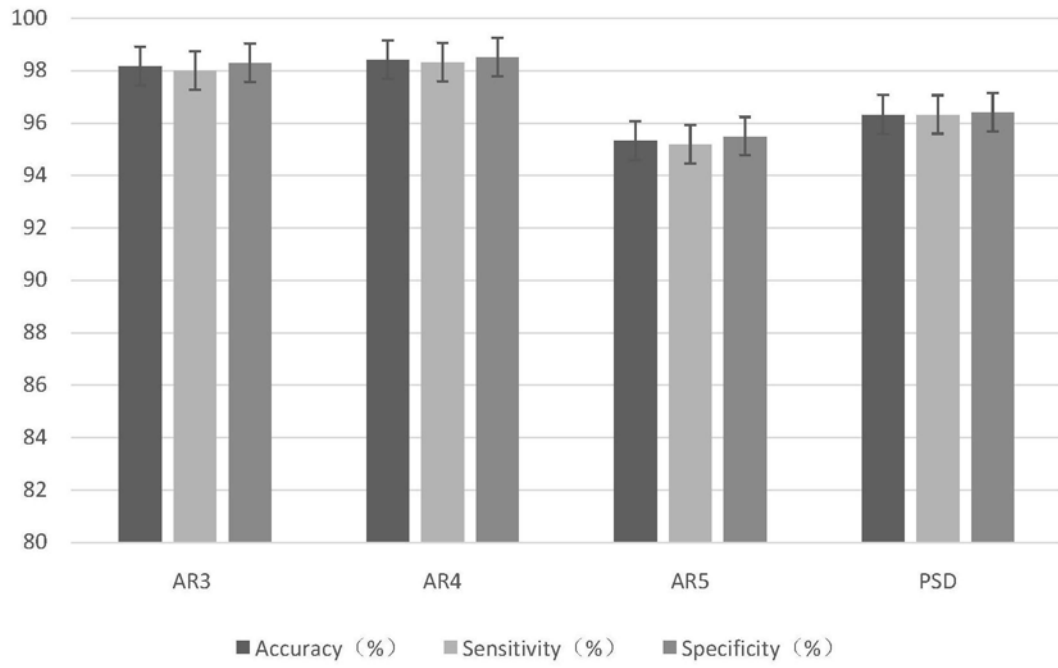


图2e

专利名称(译)	一种基于加权主成分分析的EEG信号特征降维方法		
公开(公告)号	CN110390272A	公开(公告)日	2019-10-29
申请号	CN201910582226.X	申请日	2019-06-30
[标]申请(专利权)人(译)	天津大学		
申请(专利权)人(译)	天津大学		
当前申请(专利权)人(译)	天津大学		
[标]发明人	董娜 李英杰 常建芳 高忠科		
发明人	董娜 李英杰 常建芳 高忠科		
IPC分类号	G06K9/00 G06K9/62 A61B5/00 A61B5/04 A61B5/0476 A61B5/18		
外部链接	Espacenet SIPO		

摘要(译)

一种基于加权主成分分析的EEG信号特征降维方法：提取疲劳驾驶的m个EEG信号的样本，分为训练集和测试集进行训练，得到分类总准确率A和n个不同的分类准确率；将总准确率A分别与不同的分类准确率做差，得到n个差值；对n个差值做归一化处理得到n个权值；对n个权值构造权值对角矩阵；将m个EEG信号的样本写成一个m*n维矩阵；将m*n维矩阵与权值的对角矩阵相乘，得到加权的EEG信号特征数据；计算其协方差矩阵并分解，得到协方差矩阵的特征值和特征值所对应的单位化特征向量；选取前k个协方差矩阵的特征值对应的单位化特征向量进行组合形成映射矩阵；从而得到降维的EEG信号特征数据。本发明有效提高了分类识别精度和降低识别模型的训练时间。

