



(12)发明专利申请

(10)申请公布号 CN 105740887 A

(43)申请公布日 2016.07.06

(21)申请号 201610051948.9

(22)申请日 2016.01.26

(71)申请人 杭州电子科技大学

地址 310018 浙江省杭州市下沙高教园区2
号大街

(72)发明人 马玉良 丁晓慧 孟明 高云园
罗志增

(74)专利代理机构 杭州君度专利代理事务所
(特殊普通合伙) 33240

代理人 杜军

(51)Int.Cl.

G06K 9/62(2006.01)

G06K 9/00(2006.01)

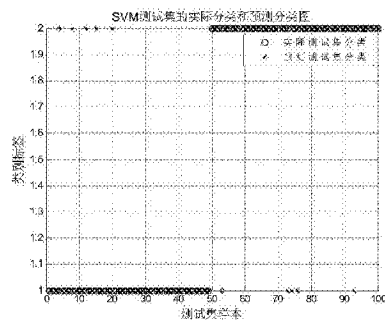
权利要求书2页 说明书5页 附图1页

(54)发明名称

一种基于PSO-SVM的脑电信号特征分类方法

(57)摘要

本发明涉及一种基于PSO-SVM的脑电信号特征分类方法。本发明首先利用正则化的CSP(R-CSP)算法进行脑电信号特征提取;其次利用粒子群算法对支持向量机的惩罚因子C和核参数g进行优化。最后,用得到的最优参数对SVM分类器进行训练,利用训练好的分类器对样本进行分类预测。本发明与传统的SVM分类识别进行对比,结果表明基于PSO-SVM的分类识别算法能有效的提高脑电信号的分类识别率,较传统的分类识别方法有明显的优势。



1. 一种基于PSO-SVM的脑电信号特征分类方法,其特征在于,该方法包括以下步骤:

步骤1. 利用正则化的CSP算法对脑电信号进行特征提取,得到样本特征向量Y;

步骤2. 利用粒子群算法对支持向量机的核函数参数进行迭代寻优;

步骤3. 利用PSO优化后的最优参数对SVM分类器进行训练,利用训练好的分类器对样本进行分类预测。

2. 根据权利要求1所述的一种基于PSO-SVM的脑电信号特征分类方法,其特征在于:其中步骤1中脑电信号特征提取得到特征向量具体步骤如下:

采用正则化的CSP算法对脑电信号进行特征提取,从两类特征中提取出首尾各q个具有代表性特征向量,得到空间滤波器为W,设运动想象脑电信号为 $E_c^i, c \in \{1, 2\}$,表示类别为c的第i个样本的脑电信号,得到投影后的矩阵 $X = W^T E$,对X进行对数化处理得到的特征系数表示如下:

$$y_j = \log \left(\frac{\text{var}(x_j)}{\sum_{k=1}^{k=2m} \text{var}(x_k)} \right), j=1, 2, \dots, 2q$$

其中 x_j 为x的第j行, $\text{var}(x_q)$ 为方差。

3. 根据权利要求1所述的一种基于PSO-SVM的脑电信号特征分类方法,其特征在于:步骤2中的核函数 $K(x, x_i)$ 选用径向基核函数(RBF),公式如下:

$$K(x, x_i) = \exp(-|x - x_i|^2 / g^2)$$

其中 $(\cdot)$ 为内积, $x, x_i \in R^n$, x_i 为特征向量,g为核参数,则支持向量机的最优决策函数公式转换为:

$$f(x) = \text{sgn}(\sum_{i=1}^l a_i^* y_i K(x, x_i) + b^*), 0 \leq a_i \leq C$$

式中C为惩罚因子, a_i 为对应的拉格朗日系数, b^* 为分类阈值。

4. 根据权利要求3所述的一种基于PSO-SVM的脑电信号特征分类方法,其特征在于:步骤2中所述的粒子群算法对支持向量机的核函数参数,即核参数g和惩罚因子C,进行迭代寻优的具体步骤如下:

(1) 初始化:在D维空间中,初始化M个粒子的初始位置和速度,包括设定粒子群初始参数 c_1, c_2 ,确定每个粒子的初始速度及每个初始粒子的最优位置和全局最优位置;

其中第i个粒子的位置和速度分别为: $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$, $v_i = (v_{i1}, v_{i2}, \dots, v_{id})^T$, 式中 $i=1, 2, \dots, M$;第i个粒子最优位置是 $P_i = (P_{i1}, P_{i2}, \dots, P_{id})$,种群全局最优位置是 $P_g = (P_{g1}, P_{g2}, \dots, P_{gd})$;

(2) 计算适应度:利用初始化的参数对支持向量机建模,根据训练样本对该模型进行训练,并且利用适应度函数计算得到各个粒子适应度函数值;

(3) 调整:根据粒子的适应度值调整粒子的个体最优位置和全局最优位置;

(4) 更新:根据粒子的速度和位置迭代更新公式,对粒子的位置和速度进行更新,获得新的参数pbest和gbest;

其中粒子的速度和位置迭代公式为:

$$v_{id}^{t+1} = \omega v_{id}^t + c_1 r_1 (p_{id}^t - x_{id}^t) + c_2 r_2 (p_{gd}^t - x_{id}^t)$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1}$$

其中位置 $x_{id}^t \in [L_d, U_d]$, L_d, U_d 分别为D维空间的下限和上限, 速度 $v_{id}^t \in [v_{\min,d}, v_{\max,d}]$, $v_{\min,d}, v_{\max,d}$ 分别是粒子速度的下限和上限, r_1, r_2 是均匀分布, 在(0,1)的随机数, c_1 和 c_2 为常数, 称为学习因子; ω 为惯性权重, 其值越大表示全局搜索能力越强, 反之则局部搜索能力越强; t 为迭代次数, 当达到终止条件时即为最优解;

(5) 判断: 当达到满足要求的误差条件或达到最大的迭代次数时, 终止迭代, 否则返回步骤(3)。

一种基于PSO-SVM的脑电信号特征分类方法

技术领域

[0001] 本发明涉及脑电信号的特征提取与分类方法,特别涉及基于PSO-SVM的脑电信号特征分类方法。

背景技术

[0002] 脑-接口(BCI)是一种基于脑电信号(EEG)实现人脑与计算机或其他设备之间通信和控制接口。BCI技术指的是能够使人在无外周神经系统和肌肉组织的参与条件下,通过计算机等电子设备输出控制信号,进而与外界环境进行交流的一种新型人机交互手段。近年来,基于脑电的脑-机接口技术成为域的研究热点,并逐步发展成为一门新兴的多学科交叉技术。

[0003] BCI的关键技术是如何快速有效的提取脑电信号特征和提高识别准确率。EEG分类识别的主要方法有线性判别分析(LDA)、K-最邻近分析以及人工神经网络等。SVM是支持向量机(SVM)是1995年首次提出的一种分类技术,常用于模式分类和非线性回归。其思想是将向量映射变换到一个更高维的空间里,通过计算得到一个最优分类面,从而使样本线性分开。但由于SVM使用过程中由于脑电信号的随机、非平稳性,以及研究人员对信号的分布特性缺少一点先验知识,导致SVM的最优核函数的选取存在不确定性。而支持向量机分类效果的好坏主要取决于核函数和参数的选择。

发明内容

[0004] 本发明的目的是在正则化CSP(R-CSP)进行特征提取的基础上,结合粒子群算法(PSO)对支持向量机的核函数参数进行迭代寻优,提出了基于PSO-SVM的脑电信号特征分类方法。

[0005] 本发明的目的可以通过以下技术方案实现:

[0006] 基于PSO-SVM的脑电信号特征分类方法,该方法包括以下步骤:

[0007] 步骤1.利用正则化的CSP算法(R-CSP)对脑电信号进行特征提取,得到样本特征向量Y。

[0008] 步骤2.利用粒子群算法对支持向量机的核函数参数进行迭代寻优。

[0009] 步骤3.利用PSO优化后的最优参数对SVM分类器进行训练,利用训练好的分类器对样本进行分类预测。

[0010] 其中步骤1中脑电信号特征提取得到特征向量具体步骤如下:

[0011] 采用正则化的CSP算法对脑电信号进行特征提取,从两类特征中提取出首尾各q个具有代表性特征向量,得到空间滤波器为W,设运动想象脑电信号为 $E_{c,i}$, $c \in \{1, 2\}$,表示类别为c的第i个样本的脑电信号,得到投影后的矩阵 $X = W^T E$,对X进行对数化处理得到的特征系数表示如下:

$$[0012] \quad y_j = \log \left(\frac{\text{var}(x_j)}{\sum_{k=1}^{k=2m} \text{var}(x_k)} \right), j=1,2,\dots,2q$$

[0013] 其中 x_j 为 x 的第 j 行, $\text{var}(x_q)$ 为方差。

[0014] 作为优选,步骤2中的核函数 $K(x, x_i)$ 选用径向基核函数(RBF),公式如下:

$$[0015] \quad K(x, x_i) = \exp(-|x - x_i|^2) / g^2$$

[0016] 其中 $(\cdot)$ 为内积, $x, x_i \in \mathbb{R}^n$, x_i 为特征向量, g 为核参数,则支持向量机的最优决策函数公式转换为:

$$[0017] \quad f(x) = \text{sgn}(\sum_{i=1}^l a_i^* y_i K(x, x_i) + b^*), 0 \leq a_i \leq C$$

式中 C 为惩罚因子, a_i 为对应的拉格朗日系数, b^* 为分类阈值。

[0018] 作为优选,步骤2中所述的粒子群算法对支持向量机的核函数参数,即核参数 g 和惩罚因子 C ,进行迭代寻优的具体步骤如下:

[0019] (1)初始化:在 D 维空间中,初始化 M 个粒子的初始位置和速度,包括设定粒子群初始参数 c_1, c_2 ,确定每个粒子的初始速度及每个初始粒子的最优位置和全局最优位置。

[0020] 其中第 i 个粒子的位置和速度分别为: $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$, $v_i = (v_{i1}, v_{i2}, \dots, v_{id})^T$, 式中 $i=1, 2, \dots, M$ 。第 i 个粒子最优位置是 $P_i = (P_{i1}, P_{i2}, \dots, P_{id})$, 种群全局最优位置是 $P_g = (P_{g1}, P_{g2}, \dots, P_{gd})$ 。

[0021] (2)计算适应度:利用初始化的参数对支持向量机建模,根据训练样本对该模型进行训练,并且利用适应度函数计算得到各个粒子适应度函数值。

[0022] (3)调整:根据粒子的适应度值调整粒子的个体最优位置和全局最优位置。

[0023] (4)更新:根据粒子的速度和位置迭代更新公式,对粒子的位置和速度进行更新,获得新的参数 p_{best} 和 g_{best} 。

[0024] 其中粒子的速度和位置迭代公式为:

$$[0025] \quad v_{id}^{t+1} = \omega v_{id}^t + c_1 r_1 (p_{id}^t - x_{id}^t) + c_2 r_2 (p_{gd}^t - x_{id}^t)$$

$$[0026] \quad x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1}$$

[0027] 其中位置 $x_{id}^t \in [L_d, U_d]$, L_d, U_d 分别为 D 维空间的下限和上限,速度 $v_{id}^t \in [v_{\min,d}, v_{\max,d}]$, $v_{\min,d}, v_{\max,d}$ 分别是粒子速度的下限和上限, r_1, r_2 是均匀分布,在 $(0, 1)$ 的随机数, c_1 和 c_2 为常数,称为学习因子。 ω 为惯性权重,其值越大表示全局搜索能力越强,反之则局部搜索能力越强。 t 为迭代次数,当达到终止条件时即为最优解。

[0028] (5)判断:当达到满足要求的误差条件或达到最大的迭代次数时,终止迭代,否则返回步骤(3)。

[0029] 本发明的有益效果:利用R-CSP进行特征提取后,将基于PSO-SVM分类识别结果与使用单一SVM进行分类识别、传统的LDA分类识别结果进行对比,结果表明,使用PSO优化后的SVM分类器对脑电信号进行分类的正确率更高,有效的提高了分类识别率。

附图说明

- [0030] 图1为PSO寻找最优参数的曲线；
 [0031] 图2为优化前的SVM分类准确率示意图；
 [0032] 图3为优化后的SVM分类准确率示意图；

具体实施方式

- [0033] 以下结合附图对本发明作进一步说明。
 [0034] 本发明包括以下步骤：
 [0035] 步骤1.利用正则化的CSP算法(R-CSP)对脑电信号进行特征提取,得到样本特征向量Y。
 [0036] 步骤2.利用粒子群算法对支持向量机的核函数参数进行迭代寻优。
 [0037] 步骤3.利用PSO优化后的最优参数对SVM分类器进行训练,利用训练好的分类器对样本进行分类预测。
 [0038] 其中步骤1中脑电信号特征提取得到特征向量具体步骤如下：
 [0039] 采用正则化的CSP算法对脑电信号进行特征提取,从两类特征中提取出首尾各q个具有代表性特征向量,得到空间滤波器为W,设运动想象脑电信号为 $E_c^i, c \in \{1, 2\}$,表示类别为c的第i个样本的脑电信号,得到投影后的矩阵 $X = W^T E$,对X进行对数化处理得到的特征系数表示如下：

$$[0040] \quad y_j = \log \left(\frac{\text{var}(x_j)}{\sum_{k=1}^{k=2m} \text{var}(x_k)} \right), j = 1, 2, \dots, 2q$$

- [0041] 其中 x_j 为x的第j行, $\text{var}(x_q)$ 为方差。
 [0042] 作为优选,步骤2中的核函数核函数 $K(x, x_i)$ 选用径向基核函数(RBF),公式如下：
 [0043] $K(x, x_i) = \exp(-|x - x_i|^2) / g^2$
 [0044] 其中 $(\cdot)$ 为内积, $x, x_i \in R^n, x_i$ 为特征向量,g为核参数,则支持向量机的最优决策函数公式转换为：

$$[0045] \quad f(x) = \text{sgn} \left(\sum_{i=1}^l a_i^* y_i K(x, x_i) + b^* \right), 0 \leq a_i \leq C$$

式中C为惩罚因子, a_i 为对应的拉格朗日系数, b^* 为分类阈值。

- [0046] 作为优选,步骤2中所述的粒子群算法对支持向量机的核函数参数,即核参数g和惩罚因子C,进行迭代寻优的具体步骤如下：
 [0047] (1)初始化:在D维空间中,初始化M个粒子的初始位置和速度,包括设定粒子群初始参数 c_1, c_2 ,确定每个粒子的初始速度及每个初始粒子的最优位置和全局最优位置等。
 [0048] 其中第i个粒子的位置和速度分别为: $x_i = (x_{i1}, x_{i2}, \dots, x_{id}), v_i = (v_{i1}, v_{i2}, \dots, v_{id})^T$,式中 $i = 1, 2, \dots, M$,第i个粒子最优位置是 $P_i = (P_{i1}, P_{i2}, \dots, P_{id})$,种群全局最优位置是 $P_g = (P_{g1}, P_{g2}, \dots, P_{gd})$ 。

[0049] (2)计算适应度:利用初始化的参数对支持向量机建模,根据训练样本对该模型进行训练,并且利用适应度函数计算得到各个粒子适应度函数值。

[0050] (3)调整:根据粒子的适应度值调整粒子的个体最优位置和全局最优位置。

[0051] (4)更新:根据粒子的速度和位置迭代更新公式,对粒子的位置和速度进行更新,获得新的参数pbest和gbest。

[0052] 其中粒子的速度和位置迭代公式为:

$$[0053] \quad v_{id}^{t+1} = wv_{id}^t + c_1r_1(p_{id}^t - x_{id}^t) + c_2r_2(p_{gd}^t - x_{id}^t)$$

$$[0054] \quad x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1}$$

[0055] 其中位置 $x_{id}^t \in [L_d, U_d]$, L_d, U_d 分别为D维空间的下限和上限,速度 $v_{id}^t \in [v_{\min,d}, v_{\max,d}]$, $v_{\min,d}, v_{\max,d}$ 分别是粒子速度的下限和上限, r_1, r_2 是均匀分布,在(0,1)的随机数, c_1 和 c_2 为常数,称为学习因子。 ω 为惯性权重,其值越大表示全局搜索能力越强,反之则局部搜索能力越强。 t 为迭代次数,当达到终止条件时即为最优解。

[0056] (5)判断:当达到满足要求的误差条件或达到最大的迭代次数时,终止迭代,否则返回步骤(3)。

[0057] 从图1中可以看出随着迭代次数的增加,准确率逐渐趋向最佳适应度,当满足迭代终止条件时,准确率为90%左右,输出最优参数值 $c=4.5294, g=0.01$,对于支持向量机分类快速寻找到最佳的惩罚参数 c 和核参数 g ,而且PSO是基于种群的并行全局搜索,调节参数少,收敛速度快,体现出利用PSO-SVM的明显优势。

[0058] 图2和图3为以aa为主受试者,其他四名为辅助受试者将样本数量设为100时,传统SVM和PSO优化后的分类正确率示意图。从图2和图3对比可以明显看出优化后的分类效果比优化前的效果好。分别以aa、al、av、aw为主受试者,其他为辅助受试者,分别利用SVM和PSO-SVM进行分类,其分类结果下表1,表2所示。

[0059] 表1 PSO优化前SVM的分类

[0060]

Subject	aa(100)			al(200)			av(80)			aw(56)		
训练样本	52	54	56	101	108	110	42	44	46	30	32	34
测试样本	48	46	44	99	92	90	38	36	34	26	24	22
识别率	95.8	95.6	97.7	92.5	88.0	87.7	81.5	77.8	73.5	88.4	87.5	90.9
平均识别率	96.4			89.4			77.6			88.9		

[0061] 表2 PSO优化SVM后的分类

[0062]

Subject	aa(100)			al(200)			av(80)			aw(56)		
训练样本集	52	54	56	101	108	110	42	44	46	30	32	34
测试样本集	48	46	44	99	92	90	38	36	34	26	24	22
识别准确率	97.9	97.8	95.4	93.9	90.2	91.1	81.6	80.5	79.4	92.3	91.7	90.9

[0063]

平均识别率	97.0	91.7	80.5	91.6
-------	------	------	------	------

[0064] 从表1和表2中对比优化前后的分类识别率,数据分析表明选择不同数量的样本时PSO优化后的SVM分类相对传统的SVM均有明显提升,PSO-SVM的分类识别率比传统的SVM分类识别率提高了2.2%左右,表明PSO-SVM获取最优的参数的时候有效的提高支持向量机的性能。

[0065] 表3 PSO-SVM与传统的LDA分类准确率对比

[0066]

Subject	aa	al	av	aw	ay
PSO-SVM accuracy(%)	100	98.33	94.26	97.61	100
SVM accuracy(%)	98.48	87.32	86.33	89.78	94.81
LDA accuracy(%)	94.64	96.43	87.86	95.36	98.45

[0067] 将PSO-SVM分类结果与传统的LDA、SVM进行对比如表3所示。实验选取aa、al、av、aw、ay分别作为主试者,其他作为辅助收试者,采用每名实验中前40次实验数据,正则化的参数设为 $\beta=0$, $\gamma=0.01$,利用PSO-SVM、SVM、LDA分别进行分类对比,最高分类准确率达100%,与传统的LDA、SVM分类方法进行对比,PSO-SVM分类方法分类识别率最高,表明PSO-SVM在脑电信号分类表现出明显的优势,能够有效的提高脑电号分类准确率。

