



(12)发明专利申请

(10)申请公布号 CN 107174203 A

(43)申请公布日 2017.09.19

(21)申请号 201710323803.4

(22)申请日 2017.05.10

(71)申请人 东华大学

地址 201620 上海市松江区人民北路2999号

(72)发明人 任立红 李嘉伟 丁永生 郝矿荣 陈磊

(74)专利代理机构 上海统摄知识产权代理事务所(普通合伙) 31303

代理人 金利琴

(51)Int.Cl.

A61B 5/00(2006.01)

A61B 5/0476(2006.01)

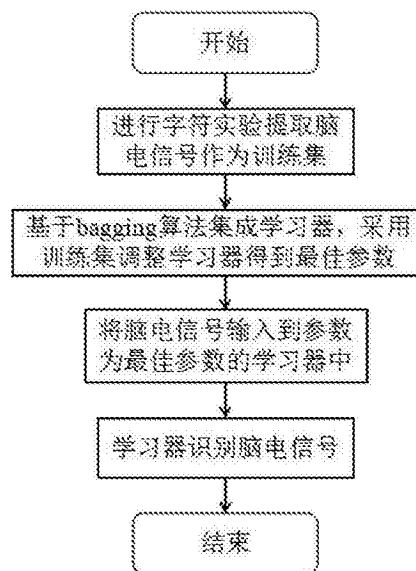
权利要求书2页 说明书5页 附图3页

(54)发明名称

一种脑电信号的识别方法

(57)摘要

本发明涉及一种脑电信号的识别方法,首先进行字符实验提取包含P300信号的脑电信号作为训练集,然后利用bagging算法集成多个混合核的支持向量机作为学习器并采用训练集基于免疫算法自适应调整学习器的参数得到最佳参数,最后利用参数为最佳参数的学习器对脑电信号中的P300信号进行识别,其中最佳参数是使得学习器能够准确识别P300信号的参数,准确识别是指12次以上重复实验的准确率为96~98%。本发明的一种脑电信号的识别方法能够根据优化内容智能选择参数,弥补了传统的学习器需要不断调优、交叉验证过程的不足,提高了集成学习器的智能性,且识别性能好、准确率高、整体泛化能力强,具有很好的推广应用价值。



1. 一种脑电信号的识别方法,其特征是:首先进行字符实验提取包含P300信号的脑电信号作为训练集,然后利用bagging算法集成多个混合核的支持向量机作为学习器并采用训练集基于免疫算法自适应调整学习器的参数得到最佳参数,最后将脑电信号输入到参数为最佳参数的学习器中,学习器对脑电信号中的P300信号进行识别;

所述最佳参数是使得学习器能够准确识别P300信号的参数,所述准确识别是指12次以上重复试验的准确率为96~98%。

2. 根据权利要求1所述的一种脑电信号的识别方法,其特征在于,所述多个混合核的支持向量机为17个混合核的支持向量机。

3. 根据权利要求2所述的一种脑电信号的识别方法,其特征在于,所述混合核为高斯核与多项式核的组合,组合方式为:

$$\alpha * \text{poly} \cdot + \cdot (1-\alpha) * \text{Gaussian};$$

其中,poly为多项式核,Gaussian为高斯核, α 的取值为0.1~0.9。

4. 根据权利要求1所述的一种脑电信号的识别方法,其特征在于,所述字符实验提取出的脑电信号由 6×6 P300字符拼写器的行或列闪烁产生,当选定的字符所在的行或列闪烁时产生P300信号。

5. 根据权利要求4所述的一种脑电信号的识别方法,其特征在于,所述训练集的采集步骤如下:

(1) 在 6×6 P300字符拼写器上选定一个字符,然后字符拼写器逐行或逐列随机地闪烁产生脑电信号,所有行和所有列闪烁一遍为一次实验,每个字符连续进行15次重复实验形成一个实验组,一个实验组提取 12×15 个脑电信号;

(2) 对每个实验组提取出来的脑电信号进行采样得到7794个点作为样本特征,以85个实验组提取出的 85×7794 个样本特征作为训练集。

6. 根据权利要求5所述的一种脑电信号的识别方法,其特征在于,采用训练集自适应调整学习器的参数时,首先采用自主采样法对训练集中的85组样本特征进行采样分组,每5组样本特征为一个采样集,分成17个采样集,然后使用17个采样集训练学习器得到学习器的最佳参数,所述最佳参数包括误差惩罚参数、高斯核函数参数和多项式核参数。

7. 根据权利要求6所述的一种脑电信号的识别方法,其特征在于,所述使用17个采样集训练学习器得到学习器的最佳参数的具体步骤如下:

(1) 确定免疫算法的抗原与抗体,将采样集视为抗原,将学习器的参数误差惩罚参数 V_1 、高斯核函数参数 V_2 和多项式核参数 V_3 的组合 (V_1, V_2, V_3) 视为抗体;

(2) 生成初始抗体群,抗体初始化生成 s 个抗体构成初始抗体群;

(3) 计算初始抗体群中每个抗体的适应度值,即将采样集输入到参数 $V_t = (1, 2, 3)$ 确定的学习器中得到学习器识别P300信号的准确率,调控参数 V_t ,使得每个抗体的适应度值增大,得到新的抗体群,设定抗体完成一次克隆、变异和交叉操作为一次迭代,当前的迭代次数为 g ;

(4) 判断当前的迭代次数 g 是否达到 g_{\max} 次,如果是,则终止算法,输出适应度值最大的抗体即为学习器的最佳参数;如果不是,则执行步骤(5);

(5) 计算抗体的选择概率,对当前抗体群中的抗体进行促进和抑制,得到更新后的抗体群 $A(k)$,高适应度值低浓度的抗体受到促进,低适应度值高浓度的抗体受到抑制,抗体的选

择概率公式如下：

$$P_i = \mu \cdot P_{f_i} + (1 - \mu)P_{st} = \mu \cdot \frac{T(i)}{\sum_{j=1}^N T(j)} + (1 - \mu) \frac{1}{N} e^{-\frac{C_i}{v}}$$

式中， $\mu=0.8$ ， $v=12$ ， N 为抗体群A(k)中的抗体总数， i 为抗体的编号， $i=1, 2, 3 \dots N$ ， $T(i)$ 为抗体的适应度值， C_i 为抗体浓度，表示当前种群中适应度值相近的抗体的浓度， C_i = 适应度值在 $[T(i), T(i) + \Delta T]$ 之间的抗体个数/抗体总数， $\Delta T = (T_{\max} - T_{\min}) / 3$ ， T_{\max} 为抗体群A(k)中抗体适应度值的最大值， T_{\min} 为抗体群A(k)中抗体适应度值的最小值；

(6) 克隆，对当前的第k代父本种群A(k)进行克隆，得到种群A1(k)；

(7) 变异，对A1(k)以变异概率0.7进行变异操作得到种群A2(k)；

(8) 交叉，对A2(k)以交叉概率0.4进行交叉操作得到种群A3(k)；

(9) 计算种群A3(k)中抗体的适应度值，选择适应度值最大的m个抗体组成下一代抗体种群A(k+1)，返回步骤(3)。

8. 根据权利要求7所述的一种脑电信号的识别方法，其特征在于，所述学习器识别P300信号的准确率以误差率ER作为最终评判结果，计算如下：

$$ER = \frac{ED}{CD + ED}$$

其中ED为错误决策数目，CD为正确决策数目。

9. 根据权利要求7所述的一种脑电信号的识别方法，其特征在于，所述初始抗体群中抗体的个数s为50，所述 g_{\max} 的取值为50，所述m的取值为5~10。

一种脑电信号的识别方法

技术领域

[0001] 本发明属于脑电信号识别领域,涉及一种脑电信号的识别方法,尤其涉及一种基于免疫算法的集成支持向量机的脑电信号的识别方法。

背景技术

[0002] 脑-机接口(Brain-Computer Interface,BCI)技术形成于20世纪70年代,是一种涉及神经科学、信号检测、信号处理、模式识别等多领域的交叉技术。20多年来,随着人们对神经系统功能认识的提高和计算机技术的发展,BCI技术的研究呈明显的上升趋势,特别是1999年和2002年两次BCI国际会议的召开为BCI技术的发展指明了方向。目前,BCI技术已引起国际上众多学科科技工作者的普遍关注,成为生物医学工程、计算机技术、通信等多领域的一个新的研究热点。

[0003] BCI是一种连接大脑和外部设备的实时通信系统,BCI系统可以把大脑发出的信息直接转换成能够驱动外部设备的命令,并代替人的肢体或语言器官,实现人与外界的交流以及对外部环境的控制,换言之,BCI系统可以代替正常外围神经和肌肉组织,实现人与计算机之间或人与外部环境之间的通信。BCI技术的核心是把用户输入的脑电信号转换成输出控制信号或命令的转换算法,BCI研究工作中相当重要的部分就是调整人脑和BCI系统之间的相互适应关系,也就是寻找合适的信号处理与转换算法,使得神经电信号能够实时、快速、准确地通过BCI系统转换成可以被计算机识别的命令或操作信号。

[0004] 近年来,基于脑电(electroencephalogram,EEG)信号的脑机接口系统受到了广泛的关注。脑机接口是一种新的人机交互界面,可以通过识别人脑电信号来获取被试者意图,这项技术可以为神经肌肉损伤的患者提供不依赖外为神经的新的通讯工具,将来可以为人类带来极大的便利。在脑电信号中,P300诱发电位由于它的稳定性以及不需要训练等特性越来越被研究人员所接受,P300诱发电位是一种事件相关电位(event related potential,ERP),是对发生概率相对较小的外界事件或刺激的反应,由于P300诱发电位一般出现在刺激发生后300ms左右,故称为P300信号。为了准确分辨出EEG信号中是否含有P300信号,以往的学者提出了很多算法,例如基于小波变换的识别算法、基于独立分量分析的识别算法、基于支持向量机的分类算法等。由于P300分类为二分类问题,所以适用于二分类的支持向量机分类算法被广泛应用,很多学者采用支持向量机对脑电信号进行分类并且取得了不错的成果,然而单支持向量机普遍存在分类能力有限、分类准确率不高、整体泛化能力差等问题,阻碍了能够识别P300信号的支持向量机的进一步推广应用。

[0005] 因此,研究一种识别性能好、准确率高、整体泛化能力强的利用支持向量机识别P300信号的方法具有重要意义。

发明内容

[0006] 本发明针对现有技术的不足,提供了一种识别性能好、准确率高、整体泛化能力强的P300信号的识别方法。

[0007] 为了达到上述目的,本发明采用的技术方案为:

[0008] 一种脑电信号的识别方法,首先进行字符实验提取包含P300信号的脑电信号作为训练集,然后利用bagging算法集成多个混合核的支持向量机作为学习器并采用训练集基于免疫算法自适应调整学习器的参数得到最佳参数,最后将脑电信号输入到参数为最佳参数的学习器中,学习器对脑电信号中的P300信号进行识别,学习器识别脑电信号的过程与支持向量机识别脑电信号的过程相同;

[0009] 所述最佳参数是使得学习器能够准确识别P300信号的参数,所述准确识别是指12次以上重复实验的准确率为96~98%。

[0010] 作为优选的技术方案:

[0011] 如上所述的一种脑电信号的识别方法,所述多个混合核的支持向量机为17个混合核的支持向量机。

[0012] 如上所述的一种脑电信号的识别方法,所述混合核为高斯核与多项式核的组合,组合方式为:

[0013] $\alpha * \text{poly} \cdot + \cdot (1-\alpha) * \text{Gaussian}$;

[0014] 其中,poly为多项式核,Gaussian为高斯核, α 的取值为0.1~0.9, α 优选的取值为0.5,即混合核中高斯核或多项式核的占比为10~90%,优选为50%,该占比可以根据脑电信号不同而调整。

[0015] 如上所述的一种脑电信号的识别方法,所述字符实验提取出的脑电信号由BC1 Competition lll Dataset ll提供的6×6P300字符拼写器的行或列闪烁产生,当选定的字符所在的行或列闪烁时产生P300信号。

[0016] 如上所述的一种脑电信号的识别方法,所述训练集的采集步骤如下:

[0017] (1) 在6×6P300字符拼写器上选定一个字符,然后字符拼写器逐行或逐列随机地闪烁产生脑电信号,可以是所有行闪烁后列再闪烁,所有列闪烁后行再闪烁,也可以是行与列交替闪烁,所有行和所有列闪烁一遍为一次实验,每个字符连续进行15次重复实验形成一个实验组,一个实验组提取12×15个脑电信号;

[0018] (2) 对每个实验组提取出来的脑电信号进行采样得到7794个点作为样本特征,以85个实验组提取出的85×7794个样本特征作为训练集。

[0019] 如上所述的一种脑电信号的识别方法,采用训练集自适应调整学习器的参数时,首先采用自主采样法对训练集中的85组样本特征进行采样分组,每5组样本特征为一个采样集,分成17个采样集,然后使用17个采样集训练学习器得到学习器的最佳参数,所述最佳参数包括误差惩罚参数、高斯核函数参数和多项式核参数。

[0020] 如上所述的一种脑电信号的识别方法,本发明以免疫算法的免疫识别模块模拟人体特异性免疫过程,所述使用17个采样集训练学习器得到学习器的最佳参数的具体步骤如下:

[0021] (1) 确定免疫算法的抗原与抗体,将采样集视为抗原,将学习器的参数误差惩罚参数 V_1 、高斯核函数参数 V_2 和多项式核参数 V_3 的组合 (V_1, V_2, V_3) 视为抗体;

[0022] (2) 生成初始抗体群,抗体初始化生成s个抗体构成初始抗体群,一个抗体以一维数组的形式存在,初始抗体群以s维数组的形式存在;

[0023] (3) 计算初始抗体群中每个抗体的适应度值,即将采样集输入到参数 $V_t = (1, 2, 3)$

确定的学习器中得到学习器识别P300信号的准确率,调控参数 V_t ,使得每个抗体的适应度值增大,得到新的抗体群,设定抗体完成一次克隆、变异和交叉操作为一次迭代,当前的迭代次数为 g ;

[0024] (4) 判断当前的迭代次数 g 是否达到 g_{\max} 次,如果是,则终止算法,输出适应度值最大的抗体即为学习器的最佳参数;如果不是,则执行步骤(5);

[0025] (5) 计算抗体的选择概率,对当前抗体群中的抗体进行促进和抑制,得到更新后的抗体群 $A(k)$,高适应度值低浓度的抗体受到促进,低适应度值高浓度的抗体受到抑制,抗体的选择概率公式如下:

$$[0026] \quad P_i = \mu \cdot P_{f_i} + (1 - \mu)P_{a_i} = \mu \cdot \frac{T(i)}{\sum_{j=1}^N T(j)} + (1 - \mu) \frac{1}{N} e^{-\frac{C_i}{v}}$$

[0027] 式中, $\mu=0.8$, $v=12$, N 为抗体群 $A(k)$ 中的抗体总数, i 为抗体的编号, $i=1,2,3 \dots N$, $T(i)$ 为抗体的适应度值, C_i 为抗体浓度,表示当前种群中适应度值相近的抗体的浓度, $C_i = \text{适应度值在}[T(i), T(i) + \Delta T]$ 之间的抗体个数/抗体总数, $\Delta T = (T_{\max} - T_{\min}) / 3$, T_{\max} 为抗体群 $A(k)$ 中抗体适应度值的最大值, T_{\min} 为抗体群 $A(k)$ 中抗体适应度值的最小值;

[0028] (6) 克隆,对当前的第 k 代父本种群 $A(k)$ 进行克隆,得到种群 $A_1(k)$;

[0029] (7) 变异,对 $A_1(k)$ 以变异概率0.7进行变异操作得到种群 $A_2(k)$;

[0030] (8) 交叉,对 $A_2(k)$ 以交叉概率0.4进行交叉操作得到种群 $A_3(k)$;

[0031] (9) 计算种群 $A_3(k)$ 中抗体的适应度值,选择适应度值最大的 m 个抗体组成下一代抗体种群 $A(k+1)$,返回步骤(3)。

[0032] 如上所述的一种脑电信号的识别方法,所述学习器识别P300信号的准确率以误差率 ER 作为最终评判结果,计算如下:

$$[0033] \quad ER = \frac{ED}{ED + CD}$$

[0034] 其中 ED 为错误决策数目, CD 为正确决策数目。

[0035] 如上所述的一种脑电信号的识别方法,其特征在于,所述初始抗体群中抗体的个数 s 为50,所述 g_{\max} 的取值为50,所述 m 的取值为5~10。

[0036] 有益效果:

[0037] (1) 本发明的一种脑电信号的识别方法,利用bagging算法集成多个混合核的支持向量机作为学习器,集成学习器可以根据不同样本特点对异常的学习器进行修复,从而弥补异常学习器对结果造成的干扰,具有很强的泛化能力。

[0038] (2) 本发明的一种脑电信号的识别方法,支持向量机采用混合核函数,混合核函数为高斯函数与多项式函数两者按照一定的权重组合成的新的核函数,弥补了单一核函数的不足,提高了支持向量机的识别性能。

[0039] (3) 本发明的一种脑电信号的识别方法,无需对学习器选择参数,根据优化内容智能选择参数,弥补了传统的学习器需要不断调优、交叉验证过程的不足,进一步优化了学习器的识别性能,提高了学习器的智能性。

附图说明

[0040] 图1为本发明脑电信号的识别方法的流程图;

[0041] 图2为P300信号与非P300信号的对比图;

[0042] 图3为基于bagging算法集成学习器的示意图；

[0043] 图4为基于免疫算法自适应调整学习器的参数的流程图；

[0044] 图5为本发明的算法、原始算法和粒子群算法识别脑电信号的对比图，本发明的算法即本发明脑电信号的识别方法，本发明基于bagging算法和免疫优化算法，原始算法与本发明的算法不同之处在于没有采用免疫优化算法，粒子群算法与本发明的算法的不同之处在于采用粒子群优化算法替代免疫优化算法，且学习器的核函数为多项式核函数。

具体实施方式

[0045] 下面结合具体实施方式，进一步阐述本发明。应理解，这些实施例仅用于说明本发明而不用来限制本发明的范围。此外应理解，在阅读了本发明讲授的内容之后，本领域技术人员可以对本发明作各种改动或修改，这些等价形式同样落于本申请所附权利要求书所限定的范围。

[0046] 一种脑电信号的识别方法，如图1所示，步骤如下：

[0047] (1) 进行字符实验提取包含P300信号的脑电信号作为训练集，具体操作为：在 6×6 P300字符拼写器上选定一个字符，然后字符拼写器逐行或逐列随机地闪烁产生脑电信号，当选定的字符所在的行或列闪烁时产生P300信号，其他行或列闪烁时产生非P300信号，P300信号与非P300信号的对比图如图2所示，所有行和所有列闪烁一遍为一次实验，每个字符连续进行15次重复实验形成一个实验组，一个实验组提取 12×15 个脑电信号，对每个实验组提取出来的脑电信号进行采样得到7794个点作为样本特征，以85个实验组提取出的 85×7794 个样本特征作为训练集；

[0048] (2) 利用bagging算法集成17个混合核的支持向量机作为学习器(如图3所示)，混合核为高斯核与多项式核的组合，组合方式为：

[0049] $\alpha * \text{poly} + (1-\alpha) * \text{Gaussian}$ ；

[0050] 其中，poly为多项式核，Gaussian为高斯核， α 的取值为0.1~0.9；

[0051] 同时采用训练集基于免疫算法自适应调整学习器的参数得到最佳参数，最佳参数是指使得学习器进行12次以上重复识别脑电信号试验的准确率为96~98%的参数，包括误差惩罚参数、高斯核函数参数和多项式核参数，具体操作为：首先采用自主采样法对训练集中的85组样本特征进行采样分组，每5组样本特征为一个采样集，分成17个采样集，然后使用17个采样集训练学习器得到学习器的最佳参数；

[0052] (3) 将脑电信号输入到参数为最佳参数的学习器中，学习器对脑电信号中的P300信号进行识别。

[0053] 如图4所示，本发明使用17个采样集训练学习器得到学习器的最佳参数的具体步骤如下：

[0054] (1) 确定免疫算法的抗原与抗体，将采样集视为抗原，将学习器的参数误差惩罚参数 V_1 、高斯核函数参数 V_2 和多项式核参数 V_3 的组合 (V_1, V_2, V_3) 视为抗体；

[0055] (2) 生成初始抗体群，抗体初始化生成 s 个抗体构成初始抗体群， $s=50$ ；

[0056] (3) 计算初始抗体群中每个抗体的适应度值，即将采样集输入到参数 $V_t = (1, 2, 3)$ 确定的学习器中得到学习器识别P300信号的准确率，调控参数 V_t ，使得每个抗体的适应度值增大，得到新的抗体群，设定抗体完成一次克隆、变异和交叉操作为一次迭代，当前的迭

代次数为 g ,学习器识别P300信号的准确率以误差率ER作为最终评判结果,计算如下:

$$[0057] \quad ER = \frac{ED}{CD + ED};$$

[0058] 其中ED为错误决策数目,CD为正确决策数目;

[0059] (4) 判断当前的迭代次数 g 是否达到 g_{\max} 次, $g_{\max} = 50$,如果是,则终止算法,输出适应度值最大的抗体即为学习器的最佳参数;如果不是,则执行步骤(5);

[0060] (5) 计算抗体的选择概率,对当前抗体群中的抗体进行促进和抑制,得到更新后的抗体群 $A(k)$,高适应度值低浓度的抗体受到促进,低适应度值高浓度的抗体受到抑制,抗体的选择概率公式如下:

$$[0061] \quad P_i = \mu \cdot P_{i_0} + (1 - \mu)P_{ai} = \mu \cdot \frac{T(i)}{\sum_{j=1}^N T(j)} + (1 - \mu) \frac{1}{N} e^{-\frac{C_i}{v}};$$

[0062] 式中, $\mu = 0.8$, $v = 12$, N 为抗体群 $A(k)$ 中的抗体总数, i 为抗体的编号, $i = 1, 2, 3 \dots N$, $T(i)$ 为抗体的适应度值, C_i 为抗体浓度,表示当前种群中适应度值相近的抗体的浓度, $C_i =$ 适应度值在 $[T(i), T(i) + \Delta T]$ 之间的抗体个数/抗体总数, $\Delta T = (T_{\max} - T_{\min}) / 3$, T_{\max} 为抗体群 $A(k)$ 中抗体适应度值的最大值, T_{\min} 为抗体群 $A(k)$ 中抗体适应度值的最小值;

[0063] (6) 克隆,对当前的第 k 代父本种群 $A(k)$ 进行克隆,得到种群 $A1(k)$;

[0064] (7) 变异,对 $A1(k)$ 以变异概率0.7进行变异操作得到种群 $A2(k)$;

[0065] (8) 交叉,对 $A2(k)$ 以交叉概率0.4进行交叉操作得到种群 $A3(k)$;

[0066] (9) 计算种群 $A3(k)$ 中抗体的适应度值,选择适应度值最大的 m ($m = 5 \sim 10$) 个抗体组成下一代抗体种群 $A(k+1)$,返回步骤(3)。

[0067] 分别采用本发明的算法、原始算法和粒子群算法识别脑电信号,结果如图5所示,从图中可以看出本发明相比其他算法能得到更高的识别准确率,进一步提高脑电信号识别的准确性。

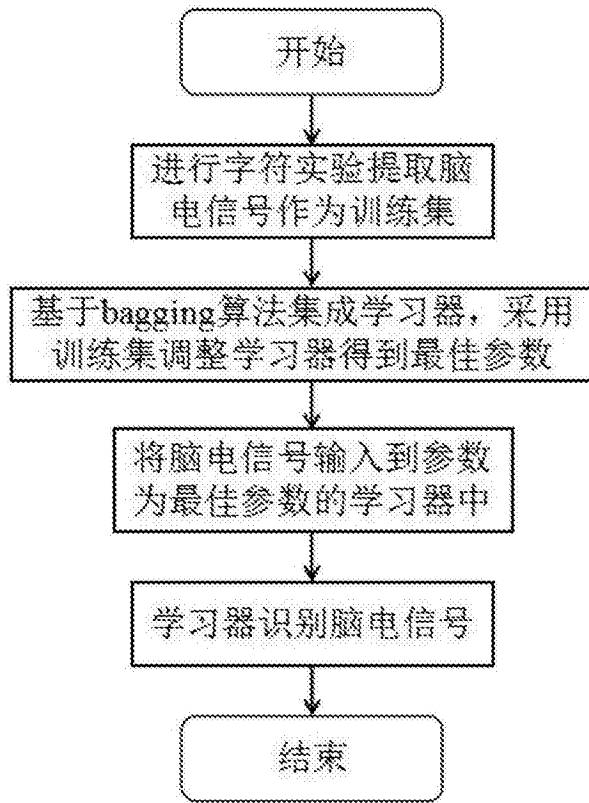


图1

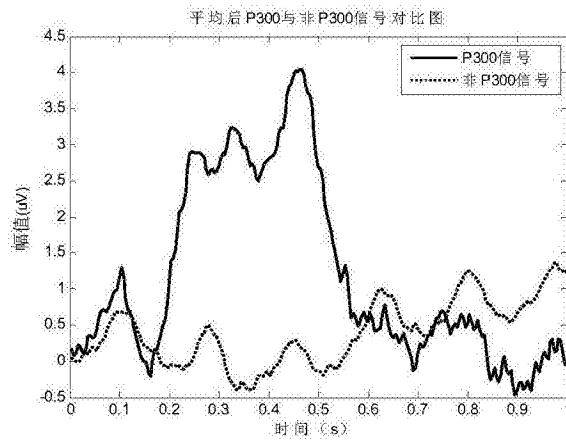


图2

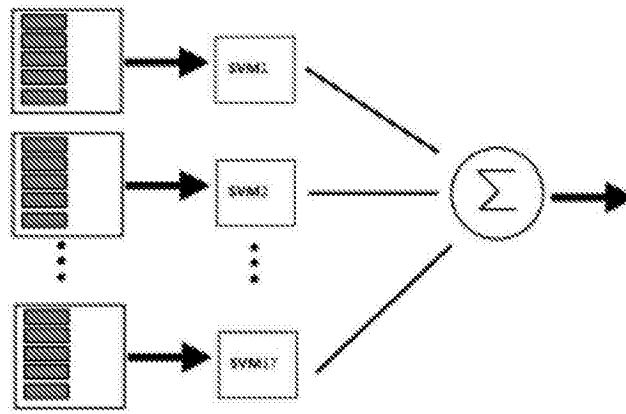


图3

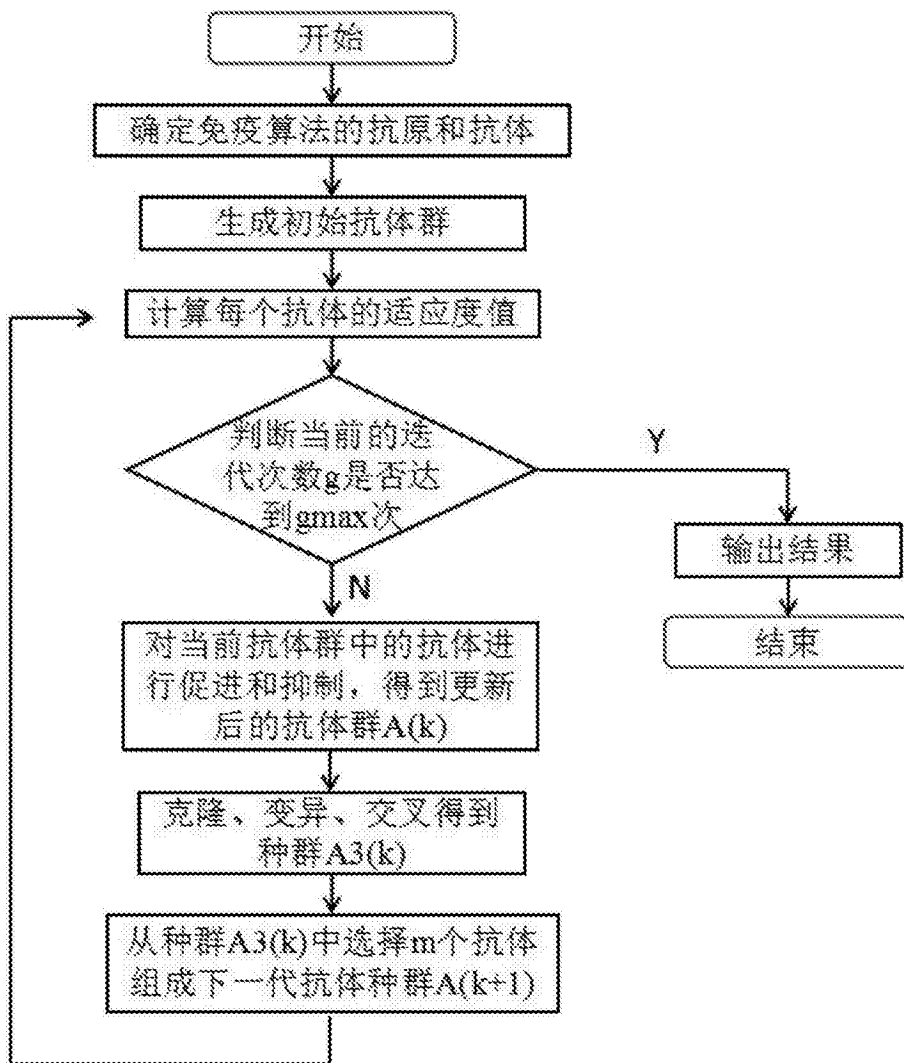


图4

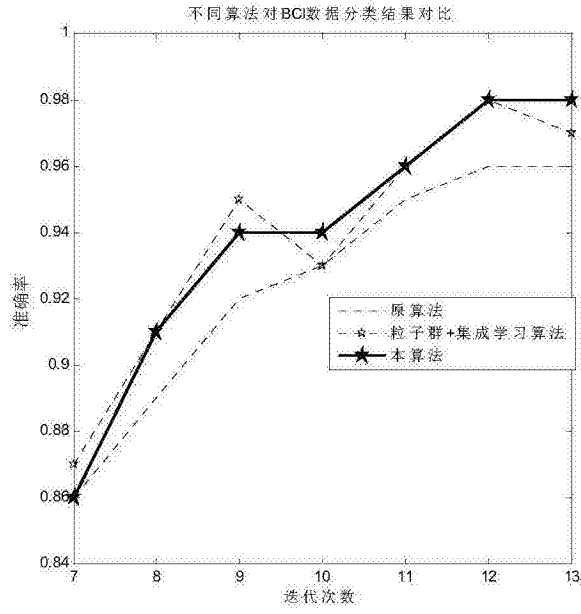


图5

专利名称(译)	一种脑电信号的识别方法		
公开(公告)号	CN107174203A	公开(公告)日	2017-09-19
申请号	CN201710323803.4	申请日	2017-05-10
[标]申请(专利权)人(译)	东华大学		
申请(专利权)人(译)	东华大学		
当前申请(专利权)人(译)	东华大学		
[标]发明人	任立红 李嘉伟 丁永生 郝矿荣 陈磊		
发明人	任立红 李嘉伟 丁永生 郝矿荣 陈磊		
IPC分类号	A61B5/00 A61B5/0476		
CPC分类号	A61B5/0476 A61B5/7264		
其他公开文献	CN107174203B		
外部链接	Espacenet SIPO		

摘要(译)

本发明涉及一种脑电信号的识别方法，首先进行字符实验提取包含P300信号的脑电信号作为训练集，然后利用bagging算法集成多个混合核的支持向量机作为学习器并采用训练集基于免疫算法自适应调整学习器的参数得到最佳参数，最后利用参数为最佳参数的学习器对脑电信号中的P300信号进行识别，其中最佳参数是使得学习器能够准确识别P300信号的参数，准确识别是指12次以上重复实验的准确率为96~98%。本发明的一种脑电信号的识别方法能够根据优化内容智能选择参数，弥补了传统的学习器需要不断调优、交叉验证过程的不足，提高了集成学习器的智能性，且识别性能好、准确率高、整体泛化能力强，具有很好的推广应用价值。

