



# (12)发明专利申请

(10)申请公布号 CN 110516711 A

(43)申请公布日 2019. 11. 29

(21)申请号 201910677300.6

A61B 5/04(2006.01)

(22)申请日 2019.07.25

A61B 5/00(2006.01)

(71)申请人 安徽大学

地址 230601 安徽省合肥市经济技术开发区  
九龙路111号

(72)发明人 吴小培 陈文伟 周蚌艳 陈瑶  
吕钊 张超 郭晓静 张磊  
高湘萍

(74)专利代理机构 合肥市上嘉专利代理事务所  
(普通合伙) 34125

代理人 郭华俊

(51)Int. Cl.

G06K 9/62(2006.01)

G06K 9/00(2006.01)

A61B 5/0476(2006.01)

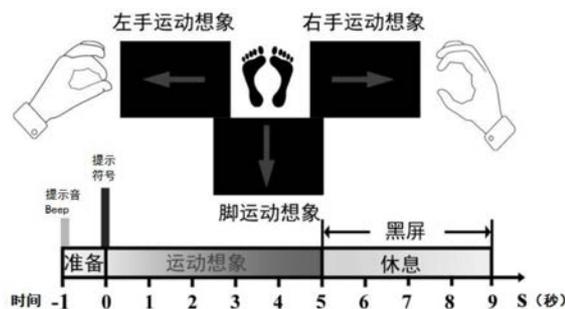
权利要求书2页 说明书10页 附图10页

## (54)发明名称

MI-BCI系统的训练集质量评估方法和单次训练样本的优化方法

## (57)摘要

本发明公开了一种MI-BCI系统的训练集质量评估方法和单次训练样本的优化选择方法,采集单次试验MI-EEG数据;设计相应的ICA空域滤波器;计算单次识别率Acc(k)和总体识别率Acc;确定用于单次试验数据质量评估的量化指标;计算训练样本质量的综合评分Cs;根据Cs的数值对训练集的质量进行评估。再根据所得Acc、量化指标和综合评分Cs,对训练集中的单次试验样本进行优化选择。本发明的训练集质量评估方法和单次训练样本的优化方法,具有较好的稳定性、准确性和运算复杂度低等优点。



1. MI-BCI系统的训练集质量评估方法,其特征是,包括如下步骤:

步骤1: 试验开始,运动想象EEG的单次试验数据采集,获得单次试验的MI-EEG数据;

步骤2: 利用所述单次试验的MI-EEG数据,设计ICA空域滤波器;

步骤3: 根据ICA空域滤波器来计算单次识别率Acc(k)和总体识别率Acc;

步骤4: 确定单次试验数据的质量评估的量化指标;

步骤5: 根据所述步骤4的量化指标计算训练样本质量的综合评分Cs;

步骤6: 根据所述步骤5的综合评分Cs的数值,对训练集的质量进行评估。

2. 根据权利要求1所述的MI-BCI系统的训练集质量评估方法,其特征是,所述步骤1中,单次试验数据采集过程为:在每次试验开始时计算机发出提示音,提醒受试者单次试验即将开始,1秒种后计算机显示器会出现箭头状运动想象的类型提示符号,左/右方向箭头提示被试进行左/右手运动想象,向下箭头则提示双脚运动想象;MI类型提示箭头在计算机屏幕上显示时间 $T_M$ 为5-6秒,然后计算机显示器黑屏3-4秒,表示单次试验结束;完整的单次试验时间 $T$ 为10秒,相邻单次试验的时间间隔2-3秒。

3. 根据权利要求1所述的MI-BCI系统的训练集质量评估方法,其特征是,所述步骤2中,所述ICA空域滤波器的设计过程为:选择单次试验数据,进行独立分量分析,获取构成头皮EEG的源信号成分及ICA混合模型A和分离矩阵W;利用ICA混合模型A所包含的独立分量空域特征,从N个输出独立分量中,确定运动相关的独立分量成分;选择与所述独立分量成分对应分离矩阵W的行向量作为MI空域滤波器,分别记为: $w_l, w_r, w_f$ ,并保存。

4. 根据权利要求3所述的MI-BCI系统的训练集质量评估方法,其特征是,所述步骤3中,利用MI空域滤波器对训练集 $D = \{[x_i, y_i], i = 1, L, I\}$ 中的全部I个单次试验数据进行空域滤波,然后进行特征提取和分类;并结合单次试验的运动想象类型和单次试验对应的MI类型标签计算单次试验的识别率Acc(k),进而得到总体识别率Acc。

5. 根据权利要求4所述的MI-BCI系统的训练集质量评估方法,其特征是,所述步骤4中,所述量化指标包括Med、Max和Min;其中, $Med = \text{median}(\text{Acc})$ , $Max = \text{Max}(\text{Acc})$ , $Min = \text{Min}(\text{Acc})$ 。

6. 根据权利要求5所述的MI-BCI系统的训练集质量评估方法,其特征是,所述步骤5中,综合评分Cs的计算公式为:

$$C_s = \frac{(w_1 \times Med + w_2 \times Max + w_3 \times Min)}{3} \times 100 \quad (5)$$

其中 $w_1, w_2, w_3$ 为权系数。

7. 一种根据权利要求1-6的评估方法的单次训练样本优化方法,其特征是,包括如下步骤:

步骤01: 试验开始,运动想象EEG的单次试验数据采集,获得单次MI-EEG数据;

步骤02: 利用所述单次MI-EEG数据,设计ICA空域滤波器;

步骤03: 根据ICA空域滤波器来计算总体识别率Acc;

步骤04: 确定单次试验数据的质量评估的量化指标;

步骤05: 根据所述步骤4的量化指标计算训练样本质量的综合评分Cs;

步骤06: 根据所述步骤5的综合评分Cs的数值,对训练集的质量进行评估;

步骤07: 根据所得 $\text{Acc} = \{\text{Acc}(k), k = 1, \dots, I\}$ 、量化指标 $\{\text{Med}, \text{Max}, \text{Min}\}$ 和综合评分Cs,

对训练集中的单次试验样本进行优化选择。

8. 根据权利要求7所述的单次训练样本的优化方法,其特征是,所述优化选择的步骤包括:

步骤07-1:从训练集D中剔除Acc(k)明显低于中值Med的单次试验样本;

步骤07-2:从训练集D中选择Acc(k)值明显高于中值Med的单次试验样本构造新的训练集;

步骤07-3:如果训练集D的综合评分Cs值明显偏低,则可标注此训练集整体无效。

## MI-BCI系统的训练集质量评估方法和单次训练样本的优化方法

### 技术领域

[0001] 本发明涉及一种数据的处理方法,尤其是一种用于MI-BCI系统的训练集质量评估方法和单次训练样本的优化方法。

### 背景技术

[0002] 脑-机接口(brain computer interface,BCI)是一种以脑电(electroencephalography,EEG)为信息载体的新型人-机交互技术,近年来受到了广泛的关注。脑-机接口技术的实现过程是通过对多导联EEG信号进行分析和处理,提取出EEG中任务相关的特征模式并转化为命令,进而实现大脑对外部设备的直接控制。BCI技术的终极目标是为肢体残疾(或运动功能障碍)群体提供一个控制外部设备的直接信息通道。同时,BCI技术也可应用于运动功能康复训练、脑-机控制游戏的开发、精神和心理状态评估以及特殊环境下的人-机交互等。运动想象BCI(Motor Imagery BCI,MI-BCI)是一种内源型脑-机通信实现模式,其基本原理是运动想象诱发的事件相关同步和去同步现象(event-related synchronization/desynchronization,ERS/ERD)。MI-BCI系统的优点是使用者不借助外部刺激,而是通过自身思维活动实现对外部设备的直接控制。MI-BCI系统的设计和训练这两个过程中,都需要大量的运动想象EEG数据。但是,在运动想象EEG(Motor Imagery EEG,MI-EEG)数据的采集过程中,受试者精神状态(如疲劳,精力不集中和缺乏经验等)以及随机出现的非神经活动伪迹干扰,均会影响MI-EEG的数据质量。

[0003] 根据现有的研究资料,针对MI-EEG数据质量的自动评估方法研究还未见报道。在当前MI-BCI系统设计和实现过程中,通常采用人工检查方式剔除一些被严重干扰的EEG数据。但人工方式的效率低,误差大,不具有实用价值。

[0004] 申请号为201610066356.4的中国发明专利“一种脑电信号质量实时判定方法”,包括以下步骤:获取与生成的时间窗口相对应的脑电信号片段;计算出当前时间窗口内的脑电信号片段的平均值,利用该平均值计算出当前时间窗口内的脑电信号片段的方差;通过计算出的方差得出当前时间窗口内的脑电信号片段的波动性判定结果;利用当前的脑电信号片段的波动性判定结果和之前的脑电信号片段的波动性判定结果,共同判定出当前脑电信号的质量。本发明能够快速、有效地对脑电信号质量做出实时判定分析,有效保障了后续脑电应用分析的准确性。但是,由于该方法采用EEG信号的时域统计特征(均值和方差)对其质量进行评估,而时域特征只能大致描述EEG被噪声污染的程度,而无法对任务相关EEG(例如MI-EEG)的质量进行全面评价。因为,MI-EEG数据质量不仅会受到噪声伪迹的影响,还与受试者的精神状态和实验经验等因素有关。因此该方法不适合MI-EEG的质量评估。

[0005] 纵观生物医学信号处理的其他研究领域,有关心电和脉搏信号质量评估研究报道相对较多。然而,由于心电和脉搏等生物信号数据与EEG信号的时-频-空三域特征以及生成机理等方面均有显著差异,因此现有的心电和脉搏等信号质量评估方法无法直接应用于EEG信号。

[0006] 因此,需要建立一种有效实用的MI-EEG的数据质量评估方法,以便为MI-BCI系统设计提供高质量的训练样本。

### 发明内容

[0007] 本发明是为避免上述已有技术中存在的不足之处,提供一种具有较好的稳定性、准确性和运算复杂度低的MI-BCI系统的训练集质量评估方法和单次训练样本的优化方法。

[0008] 本发明为解决技术问题采用以下技术方案。

[0009] MI-BCI系统的训练集质量评估方法,其包括如下步骤:

[0010] 步骤1: 试验开始,运动想象EEG的单次试验数据采集,获得单次试验的MI-EEG数据;

[0011] 步骤2: 利用所述单次试验的MI-EEG数据,设计ICA空域滤波器;

[0012] 步骤3: 根据ICA空域滤波器来计算单次识别率Acc(k)和总体识别率Acc;

[0013] 步骤4: 确定单次试验数据的质量评估的量化指标;

[0014] 步骤5: 根据所述步骤4的量化指标计算训练样本质量的综合评分Cs;

[0015] 步骤6: 根据所述步骤5的综合评分Cs的数值,对训练集的质量进行评估。

[0016] 本发明的单次训练样本的优化方法的特点也在于:

[0017] 所述步骤1中,单次试验数据采集过程为:在每次试验开始时计算机发出提示音,提醒受试者单次试验即将开始,1秒种后计算机显示器会出现箭头状运动想象的类型提示符号,左/右方向箭头提示被试进行左/右手运动想象,向下箭头则提示双脚运动想象;MI类型提示箭头在计算机屏幕上显示时间 $T_M$ 为5-6秒,然后计算机显示器黑屏3-4秒,表示单次试验结束;完整的单次试验时间 $T$ 为10秒,相邻单次试验的时间间隔2-3秒。

[0018] 所述步骤2中,所述ICA空域滤波器的设计过程为:选择单次试验数据,进行独立分量分析,获取构成头皮EEG的源信号成分及ICA混合模型A和分离矩阵W;利用ICA混合模型A所包含的独立分量空域特征,从N个输出独立分量中,确定运动相关的独立分量成分;选择与所述独立分量成分对应分离矩阵W的行向量作为MI空域滤波器,分别记为: $w_l, w_r, w_f$ ,并保存。

[0019] 所述步骤3中,利用MI空域滤波器对训练集 $D = \{[x_i, y_i], i = 1, L, I\}$ 中的全部I个单次试验数据进行空域滤波,然后进行特征提取和分类;并结合单次试验的运动想象类型和单次试验对应的MI类型标签计算单次试验的识别率Acc(k),进而得到总体识别率Acc。

[0020] 所述步骤4中,所述量化指标包括Med、Max和Min;其中, $Med = \text{median}(\text{Acc})$ , $Max = \text{Max}(\text{Acc})$ , $Min = \text{Min}(\text{Acc})$ 。

[0021] 所述步骤5中,综合评分Cs的计算公式为:

$$[0022] \quad C_s = \frac{(w_1 \times Med + w_2 \times Max + w_3 \times Min)}{3} \times 100 \quad (5)$$

[0023] 其中 $w_1, w_2, w_3$ 为权系数。

[0024] 本发明还公开了一种根据上述的评估方法的单次训练样本优化方法,包括如下步骤:

[0025] 步骤01: 试验开始,运动想象EEG的单次试验数据采集,获得单次MI-EEG数据;

[0026] 步骤02: 利用所述单次MI-EEG数据,设计ICA空域滤波器;

- [0027] 步骤03:根据ICA空域滤波器来计算总体识别率Acc;
- [0028] 步骤04:确定单次试验数据的质量评估的量化指标;
- [0029] 步骤05:根据所述步骤4的量化指标计算训练样本质量的综合评分Cs;
- [0030] 步骤06:根据所述步骤5的综合评分Cs的数值,对训练集的质量进行评估;
- [0031] 步骤07:根据所得 $Acc = \{Acc(k), k=1, \dots, I\}$ 、量化指标 $\{Med, Max, Min\}$ 和综合评分Cs,对训练集中的单次试验样本进行优化选择。
- [0032] 所述优化选择的步骤07中,根据 $Acc = \{Acc(k), k=1, \dots, I\}$ 、量化指标 $\{Med, Max, Min\}$ 和综合评分Cs优化的方法包括:
- [0033] 步骤07-1:从训练集D中剔除Acc(k)明显低于中值Med的单次试验样本;
- [0034] 步骤07-2:从训练集D中选择Acc(k)值明显高于中值Med的单次试验样本构造新的训练集;
- [0035] 步骤07-3:如果训练集D的综合评分Cs值明显偏低,则可标注此训练集整体无效。
- [0036] 与已有技术相比,本发明有益效果体现在:
- [0037] 本发明的一种MI-BCI系统的训练集质量评估方法,包括如下步骤:
- [0038] 步骤1:试验开始,运动想象EEG的单次试验数据采集,获得单次MI-EEG数据;
- [0039] 步骤2:利用所述单次MI-EEG数据,设计ICA空域滤波器;
- [0040] 步骤3:根据ICA空域滤波器来计算总体识别率Acc;
- [0041] 步骤4:确定单次试验数据的质量评估的量化指标;
- [0042] 步骤5:根据所述步骤4的量化指标计算训练样本质量的综合评分Cs;
- [0043] 步骤6:根据所述步骤5的综合评分Cs的数值,对训练集的质量进行评估。
- [0044] 本发明的单次训练样本的优化方法,是在上述训练集质量评估方法的基础上进行的。除了上述6个步骤外,还包括如下步骤07:根据所得 $Acc = \{Acc(k), k=1, \dots, I\}$ 、量化指标 $\{Med, Max, Min\}$ 和综合评分Cs,对训练集中的单次试验样本进行优化选择。
- [0045] 本发明的MI-BCI系统的训练集质量评估方法和单次训练样本的优化方法,采集单次试验MI-EEG数据;设计相应的ICA空域滤波器;计算单次识别率Acc(k)和总体识别率Acc;确定用于单次试验数据质量评估的量化指标;计算训练样本质量的综合评分Cs;根据Cs的数值对训练集的质量进行评估。根据所得Acc、量化指标和综合评分Cs能够确定描述MI-EEG数据质量的量化指标,并以此为依据,对训练集中的单次试验样本进行筛选,为MI-BCI系统关键模块(如空域滤波器和分类器)的设计和训练提供优质的训练样本,具有较好的稳定性、准确性和运算复杂度低等技术效果。
- [0046] 本发明的一种MI-BCI系统的训练集质量评估方法和单次训练样本的优化方法,具有可使得使用者能够通过非口头语言和肢体语音的方式交流、应用范围广、扩展性强、使用舒适、交互性好、鲁棒性强等优点。

## 附图说明

- [0047] 图1a为本发明的MI单次试验时间范式。
- [0048] 图1b为本发明的26导联的EEG电极分布位置图。
- [0049] 图1c为本发明的26导联单次MI-EEG数据。
- [0050] 图2为本发明的基于单次试验的MI-EEG数据质量评估示意图。

- [0051] 图3a为本发明的数据集D<sub>1</sub>对应的识别率。
- [0052] 图3b为本发明的数据集D<sub>2</sub>对应的识别率。
- [0053] 图3c为本发明的数据集D<sub>3</sub>对应的识别率。
- [0054] 图3d为本发明的数据集D<sub>4</sub>对应的识别率。
- [0055] 图4a为数据集D<sub>1</sub>中的单次MI-EEG的x<sub>43</sub>对应的波形图。
- [0056] 图4b为数据集D<sub>1</sub>中的单次MI-EEG的x<sub>46</sub>对应的波形图。
- [0057] 图5a为本发明的数据集D<sub>1</sub>对应的差异矩阵Y<sub>D</sub>(j)的可视化展示(二值图像)。
- [0058] 图5b为本发明的数据集D<sub>2</sub>对应的差异矩阵Y<sub>D</sub>(j)的可视化展示(二值图像)。
- [0059] 图5c为本发明的数据集D<sub>3</sub>对应的差异矩阵Y<sub>D</sub>(j)的可视化展示(二值图像)。
- [0060] 图5d为本发明的数据集D<sub>4</sub>对应的差异矩阵Y<sub>D</sub>(j)的可视化展示(二值图像)。
- [0061] 以下通过具体实施方式,并结合附图对本发明作进一步说明。

### 具体实施方式

- [0062] 参见图1~5d,MI-BCI系统的训练集质量评估方法,其特征是,包括如下步骤:
- [0063] 步骤1:试验开始,运动想象EEG的单次试验数据采集,获得单次试验的MI-EEG数据;
- [0064] 步骤2:利用所述单次试验的MI-EEG数据,设计ICA空域滤波器;
- [0065] 步骤3:根据ICA空域滤波器来计算单次识别率Acc(k)和总体识别率Acc;
- [0066] 步骤4:确定单次试验数据的质量评估的量化指标;
- [0067] 步骤5:根据所述步骤4的量化指标计算训练样本质量的综合评分Cs;
- [0068] 步骤6:根据所述步骤5的综合评分Cs的数值,对训练集的质量进行评估。
- [0069] 所述步骤1中,单次试验数据采集过程为:在每次试验开始时计算机发出提示音,提醒受试者单次试验即将开始,1秒种后计算机显示器会出现箭头状运动想象的类型提示符号,左/右方向箭头提示被试进行左/右手运动想象,向下箭头则提示双脚运动想象;MI类型提示箭头在计算机屏幕上显示时间T<sub>M</sub>为5-6秒,然后计算机显示器黑屏3-4秒,表示单次试验结束;完整的单次试验时间T为10秒,相邻单次试验的时间间隔2-3秒。
- [0070] MI-EEG训练集通常是由若干单次试验(a trial)组成,单次试验数据采集的时间范式如图1所示。在每次试验开始时,计算机发出提示音“Beep”,提醒受试者单次试验即将开始,1秒种后,计算机显示器会出现箭头状运动想象(MI)类型提示符号,左/右方向箭头提示被试进行左/右手运动想象,向下箭头则提示双脚运动想象。MI类型提示箭头在计算机屏幕上显示时间(T<sub>M</sub>)大约为5-6秒,然后计算机显示器黑屏3-4秒,表示单次试验结束。完整的单次试验时间T大约为10秒,相邻单次试验的时间间隔2-3秒。在单次运动想象试验期间采集的N导联EEG数据称之为单次MI-EEG数据。训练集D由I个单次MI-EEG数据组成,即: $D = \{[x_i, y_i], i = 1, L, I\}$ 。其中 $X_i = [x_1, \dots, x_N]^T$ 是由N导联的EEG信号x<sub>j</sub>, j=1, ..., N组成的MI-EEG信号矩阵,大小为N×L, L=T×F<sub>s</sub>表示单导联EEG样本长度, F<sub>s</sub>为信号的采样频率, T为单次试验总时长。y<sub>i</sub> ∈ {‘l’, ‘r’, ‘f’}为单次试验对应的MI类型标签, ‘l’, ‘r’和‘f’分别对应左手、右手和脚运动想象。图1(b)为一26导联的单次MI-EEG数据。
- [0071] 所述步骤2中,所述ICA空域滤波器的设计过程为:选择单次试验数据,进行独立分量分析,获取构成头皮EEG的源信号成分及ICA混合模型A和分离矩阵W;利用ICA混合模型A

所包含的独立分量空域特征,从N个输出独立分量中,确定运动相关的独立分量成分;选择与上述独立分量成分对应分离矩阵W的行向量作为MI空域滤波器,分别记为: $w_l, w_r, w_f$ ,并保存。

[0072] 如图2所示,利用训练集 $D = \{[x_i, y_i], i = 1, L, I\}$ 中的单次MI-EEG数据 $x$ 设计ICA空域滤波器,然后测试其应用于训练集D的分类效果(识别率)。鉴于ICA滤波器性能与其设计样本 $x$ 的质量密切相关,因此,所得识别率可用于对单次样本 $x$ 的质量评估。方法的具体实现过程如图2所示。

[0073] 设置变量 $k=1$ ,从训练集D中的选择一单次试验数据 $x \in \{x_i, i = 1, \dots, I\}$ 进行独立分量分析(independent component analysis, ICA),获取构成头皮EEG的源信号成分 $s = [s_1, \dots, s_N]^T$ 以及ICA混合模型A和分离矩阵 $W = A^{-1}$ 。ICA混合模型A、分离矩阵W、选择一单次试验数据 $x$ 、EEG的源信号成分 $s$ 之间的关系如(1)和(2)式所示,分别表示多导EEG的混合模型和分离模型, $a_j = [a_{1,j}, \dots, a_{N,j}]^T$ 为混合矩阵A的第j列, $w_j = [w_{j,1}, \dots, w_{j,N}]^T$ 为分离矩阵W的第j行。

$$[0074] \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \mathbf{M} \\ x_N \end{bmatrix} = \begin{bmatrix} a_{11} & L & a_{1N} \\ \mathbf{M} & \mathbf{O} & \mathbf{M} \\ a_{N1} & L & a_{NN} \end{bmatrix} \begin{bmatrix} s_1 \\ \mathbf{M} \\ s_N \end{bmatrix} = \sum_{j=1}^N \mathbf{a}_j s_j = \mathbf{A} \mathbf{s} \quad (1)$$

$$[0075] \quad \mathbf{s} = \begin{bmatrix} s_1 \\ \mathbf{M} \\ s_N \end{bmatrix} = \begin{bmatrix} w_{11} & L & w_{1N} \\ \mathbf{M} & \mathbf{O} & \mathbf{M} \\ w_{N1} & L & w_{NN} \end{bmatrix} \begin{bmatrix} x_1 \\ \mathbf{M} \\ x_N \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1^T \mathbf{x} \\ \mathbf{M} \\ \mathbf{w}_N^T \mathbf{x} \end{bmatrix} = \mathbf{W} \mathbf{x} \quad (2)$$

[0076] 利用ICA混合模型A所包含的独立分量(IC)空域特征,从N个输出独立分量 $s = [s_1, \dots, s_N]^T$ 中(即EEG的源信号成分 $s$ ),确定运动相关的独立分量成分(以左右手和脚三类MI为例),记为 $s_l, s_r, s_f$ ,其中的 $s_l, s_r, s_f$ 分别表示左手、右手和脚的独立分量成分。选择与 $s_l, s_r, s_f$ 对应的分离矩阵行向量作为MI空域滤波器,分别记为: $w_l, w_r, w_f$ ,并保存, $w_l, w_r, w_f$ 分别表示与 $s_l, s_r, s_f$ 对应的分离矩阵行向量。

[0077] 目前,存在多种形式的ICA算法,例如Infomax、FastICA、Sobi和Jade等。本发明设计了一种简化的信息极大ICA算法,如下表1所示,算法的初始值设定为 $W = \text{eye}/100$ ,学习率 $\text{lrate} = 0.02$ ,迭代次数 $\text{Num} = 300$ ;带通滤波范围7-40Hz。基于对大量实测数据的分析和测试,本发明所设计的ICA算法相比经典ICA算法具有更好的稳定性和适用性。针对不同受试者以及不同时间采集的MI-EEG数据集,前述初始化参数的设置具有较好的通用性。

表 1 ICA 算法伪代码

```

initializing:  $\mathbf{W}$ ,  $lrate$ ,  $Num$ ,  $i = 0$ ;
1    $\mathbf{x} \leftarrow \mathbf{x}$  is bandpass filtered between  $[f_L, f_H]$ Hz
2   While  $i \leq Num$  do
3      $i = i + 1$ 
[0078] 4    $\mathbf{y} = \mathbf{W}\mathbf{x}$ 
5      $\Delta\mathbf{W} = \{\mathbf{I} - E[\mathbf{y}\mathbf{y}^T + \text{sign}(\mathbf{y})\mathbf{y}^T]\} \mathbf{W}$ 
6      $\mathbf{W} \leftarrow \mathbf{W} + lrate * \Delta\mathbf{W}$ 
7   End while
8    $\mathbf{A} = \text{inv}(\mathbf{W})$ 
9   Return  $\mathbf{A}$ ,  $\mathbf{W}$ 

```

[0079] 所述步骤3中,利用MI空域滤波器对训练集 $D = \{[x_i, y_i], i = 1, L, I\}$ 中的全部I个单次试验数据进行空域滤波,然后进行特征提取和分类;并结合单次试验的运动想象类型和单次试验对应的MI类型标签计算单次试验的识别率 $Acc(k)$ ,进而得到总体识别率 $Acc$ 。

[0080] 利用所得 $w_l, w_r, w_f$ ,对训练集 $D = \{[x_i, y_i], i = 1, L, I\}$ 中的全部I个单次试验数据进行空域滤波,然后进行特征提取和分类,I为单次试验的总次数。其中,特征为 $s_l, s_r, s_f$ 在MI执行时段(参见图1(a))的方差。基于运动相关mu节律成分的ERD现象,通过比较方差的大小,可估计出单次试验的运动想象类型: $\hat{y}_i^k, i = 1, L, I$ ,  $\hat{y}_i$ 为MI类型标签的估值;结合训练集提供的真实标签值 $y_i, i = 1, L, I$ ,计算识别率 $Acc(k)$ 并保存,同时将 $\hat{y}_i^k, i = 1, L, I$ 保存于估计标签矩阵 $\hat{\mathbf{Y}}$ 的第k行,即:

$$[0081] \quad Acc(k) = \frac{\text{正确识别的单次试验数}}{\text{训练集D中单次试验总数}I} \quad (3)$$

$$[0082] \quad \hat{\mathbf{Y}}(k,:) = [\hat{y}_1^k, L, \hat{y}_I^k] \quad (4)$$

[0083]  $k \leftarrow k + 1$ 重复步骤1,2。即,从训练集D中选择下一个单次试验数据进行ICA空域滤波器设计及其相应的识别率测试,直至 $k = N$ ,保存全部所得识别率 $Acc = \{Acc(k), k = 1, \dots, I\}$ 和估计标签矩阵 $\hat{\mathbf{Y}}$ 。

[0084] 所述步骤4中,所述量化指标包括Med、Max和Min;其中, $Med = \text{median}(Acc)$ ,  $Max = \text{Max}(Acc)$ ,  $Min = \text{Min}(Acc)$ 。

[0085] 所述步骤5中,综合评分 $C_s$ 的计算公式为:

$$[0086] \quad C_s = \frac{(w_1 \times Med + w_2 \times Max + w_3 \times Min)}{3} \times 100 \quad (5)$$

[0087] 其中 $w_1, w_2, w_3$ 为权系数。

[0088] 根据Acc,对MI-EEG训练集中的I个单次试验数据进行质量评估。表2为本发明提出的各项量化指标,它们分别从不同的角度反映训练样本的质量。公式(5)是对训练样本质量的综合评分,其取值范围:0-100。

[0089] 表2基于ACC的质量评估量化指标

[0090]

指标名称	识别率中值	最大识别率	最小识别率
计算公式	Med=median (Acc)	Max=max (Acc)	Min=min (Acc)
理想值	1	1	1

[0091] 公式(5)中,权系数取值要求满足: $w_1+w_2+w_3=3$ ,且 $w_1 \geq 1, 0 < w_2 \leq 1, 0 < w_3 \leq 1$ ;典型取值: $w_1=2, w_2=0.5, w_3=0.5$ 。经过多次的试验证明,取值为 $w_1=2, w_2=0.5, w_3=0.5$ 的时候,评估的准确性最高。

[0092] 一种根据上述的评估方法的单次训练样本优化方法,包括如下步骤:

[0093] 步骤01:试验开始,运动想象EEG的单次试验数据采集,获得单次MI-EEG数据;

[0094] 步骤02:利用所述单次MI-EEG数据,设计ICA空域滤波器;

[0095] 步骤03:根据ICA空域滤波器来计算总体识别率Acc;

[0096] 步骤04:确定单次试验数据的质量评估的量化指标;

[0097] 步骤05:根据所述步骤4的量化指标计算训练样本质量的综合评分Cs;

[0098] 步骤06:根据所述步骤5的综合评分Cs的数值,对训练集的质量进行评估;

[0099] 步骤07:根据所得 $Acc = \{Acc(k), k=1, \dots, I\}$ 、量化指标 $\{Med, Max, Min\}$ 和综合评分Cs,对训练集中的单次试验样本进行优化选择。

[0100] 所述优化选择的步骤包括:

[0101] 步骤07-1:从训练集D中剔除Acc(k)明显低于中值Med的单次试验样本;

[0102] 步骤07-2:从训练集D中选择Acc(k)值明显高于中值Med的单次试验样本构造新的训练集;

[0103] 步骤07-3:如果训练集D的综合评分Cs值明显偏低,则可标注此训练集整体无效。

[0104] 影响MI-EEG数据质量的因素包括以下2种。

[0105] 1、噪声伪迹干扰。

[0106] 运动想象诱发的mu/beta节律成分通常被淹没在自发EEG中。此外,非神经活动的电生理伪迹(眼电,肌电和心电等)、环境电磁干扰以及电极与头皮之间的连接故障等,均会对MI-EEG数据质量产生较大的负面影响。

[0107] 2、无效单次MI-EEG训练样本。

[0108] MI-EEG的采集过程一般时间较长,因此受试者易出现精神疲劳或精力不集中等问题,这会导致受试者在进行单次试验时,不能准确地完成指定类型的运动想象任务,此种情况下采集到MI-EEG数据属于无效训练样本。

[0109] 为了改善多通道EEG的信噪比,空域滤波技术在MI-BCI系统实现中起到了极其关键的作用。目前常用的空域滤波技术有ICA和共同空间模式(common spatial pattern, CSP)。ICA和CSP均需采用训练样本进行设计,其性能与训练样本的质量密切相关。ICA是不依赖MI类型标签的无监督设计方法,理论上任意EEG数据段均可用于ICA空域滤波器的设计,因此数据中的伪迹干扰是影响ICA滤波器性能的主要原因。CSP则是一种依赖MI标签信

息的有监督空域滤波设计方法,因此MI-EEG数据中的伪迹干扰以及标签信息的准确性对CSP滤波器的性能均存在影响。

[0110] 由于ICA与CSP空域滤波技术对训练样本质量有不同需求,同时考虑到ICA/CSP空域滤波技术自身具有一定的抗伪迹干扰的能力,因此,仅仅依赖MI-EEG信号的时域统计特征无法准确判断数据质量,例如,眼动伪迹虽然幅度很大,且在MI-EEG信号中频繁出现,但经过常规的时域和空域滤波预处理后,均可有效缓解其对BCI系统性能的负面影响,因此不能简单地将含有眼动伪迹干扰的MI-EEG数据段定性为低质量样本,并加以剔除。相反,因受试者疲劳或精力不集中造成的无效单次试验数据,其时域波形及特征并无明显异常,但如果将这类单次训练样本用于CSP滤波器的设计,则可能会导致CSP空域滤波性能的明显下降。

[0111] 基于上述分析,仅仅依靠时域特征分析方法,难以全面准确地评估MI-EEG数据质量。为此,本发明提出了一种新的MI-EEG数据质量评估方法:利用训练集 $D = \{[x_i, y_i], i = 1, L, I\}$ 中的单次MI-EEG数据 $x$ 设计ICA空域滤波器,然后测试其应用于训练集 $D$ 的分类效果(识别率)。鉴于ICA滤波器性能与其设计样本 $x$ 的质量密切相关,因此,所得识别率可用于对单次样本 $x$ 的质量评估。方法的具体实现过程如图2所示。

[0112] 表2为本发明提出的各项量化指标。根据所得 $Acc = \{Acc(k), k = 1, \dots, I\}$ 及其相应的量化指标 $\{Med, Max, Min\}$ 和综合评分 $C_s$ ,可对训练集中的单次试验样本进行优化选择。

[0113] 选择策略有以下几种:

[0114] 1) 从训练集 $D$ 中剔除 $Acc(k)$ 明显低于中值 $Med$ 的单次试验样本。

[0115] 2) 从训练集 $D$ 中选择 $Acc(k)$ 值明显高于中值 $Med$ 的单次试验样本构造新的训练集。

[0116] 3) 如果训练集 $D$ 的综合评分 $C_s$ 值明显偏低,则可标注此训练集整体无效。

[0117] 如果MI-BCI系统设计采用CSP滤波器,则需要进一步判断所选单次试验样本的有效性,即判断单次试验数据 $x_i$ 的原始类型标签值 $y_i$ 与估值 $\hat{y}_i$ 是否一致。为此,本发明通过比较标签矩阵 $\hat{Y}$ 的每一行 $[\hat{y}_i^k, L, \hat{y}_i^k]$ 与原始标签 $[y_1, L, y_k]$ 的差异性,构建识别矩阵 $Y_D$ 。

[0118]  $Y_D = \{d_{k,i}, k = 1, L, I; i = 1, L, I\}$  (6)

[0119] 公式(6)中:  $d_{k,i} = \begin{cases} 0, & \text{if } \hat{y}_i^k = y_i \\ 1, & \text{if } \hat{y}_i^k \neq y_i \end{cases}$  (7)

[0120] 根据图2所述数据质量的评估过程,为了检测单次试验数据受伪迹干扰的程度,训练集中的每一单次试验数据都分别用于设计不同的ICA滤波器,并应用于训练集中全部 $I$ 个单次试验的运动类型识别。在完成步骤走1至步骤3的全部测试过程后,训练集中的每个单次试验均被检测了 $I$ 次,所得类型识别结果被放置在 $Y_D$ 的每列中。 $Y_D$ 的一行则是某一ICA滤波器应用于整个训练集的分类结果。根据前述伪迹干扰和精神状态等因素与数据质量的关系, $Y_D$ 矩阵列向量可用于分析和判断MI标签的可靠性。行向量则反映了单次试验数据受伪迹干扰的程度。因此,基于识别矩阵 $Y_D$ ,除了可以获得若干量化质量指标外,还可通过对 $Y_D$ 的可视化处理,直观展示每一单次试验数据的质量信息,为训练样本的合理选择提供参考。

[0121] 针对图2所示MI-EEG数据质量评估方法,图3a、图3b、图3c、图3d给出了对四个MI-EEG训练集 $D_1, D_2, D_3, D_4$ 的质量测试结果: $Acc = \{Acc(k), k = 1, \dots, I\}$ 及各量化指标。

[0122] 训练集 $D_1$ 和 $D_2$ 含有 $I = 75$ 个三类MI单次试验(左右手和脚运动想象),每类单次试验

样本数量为25。训练集 $D_3$ 和 $D_4$ 含有 $I=150$ 个三类MI单次试验(左右手和脚运动想象),每类单次试验样本数量为50。

[0123] 根据图3所示结果可以看到,训练集 $D_1$ 的总体质量指标 $C_s$ 值只有65.8,并且Acc中含有明显低于中值的识别率,表明对应的单次MI-EEG数据受到了较严重的干扰。此外,对应于单次试验样本的最大识别率也只有78.7%。因此,如果将 $D_1$ 用于空域滤波器设计和分类器训练,那么MI-BCI系统性能难以保证。相比而言, $D_3$ 和 $D_4$ 的总体质量指标均超过90,最高识别率达到99%,表明这两组数据集的质量非常高。在剔除少量识别率较低的单次试验样本后, $D_3, D_4$ 可应用于基于ICA或CSP的MI-BCI系统设计。 $D_2$ 的各项指标适中,各单次试验对应的识别率差距不大,比较突出的是Acc中极低识别率较少,表明各单次试验数据质量比较均匀。类似 $D_2$ 性能指标的训练集适合ICA空域滤波设计。

[0124] 在图3所示的识别率结果中,异常识别率对应的单次MI-EEG数据通常含有严重的伪迹干扰,例如,图3a中, $Acc(43)=34.7\%$ 和 $Acc(46)=33.3\%$ 明显低于识别率中值70.7%。它们对应的两个单次MI-EEG( $x_{43}$ 和 $x_{46}$ )波形如图4a和图4b所示,可以看到 $x_{43}$ 的全部导联均出现了突发干扰,而 $x_{46}$ 的 $C_3$ 导联则出现了电极连接故障。因此,采用 $x_{43}$ 和 $x_{46}$ 设计的两组ICA空域滤波器的性能受到了严重的影响,进而导致异常分类结果的出现。

[0125] 图5是四个训练集( $D_1, D_2, D_3, D_4$ )的差异矩阵 $\{Y_D(j), j=1, \dots, 4\}$ 的可视化展示(二值图像),灰度“0”/“1”(黑/白)分别表示原始标签与识别标签一致/不一致。根据前面对识别矩阵 $Y_D$ 的定义及其特性的介绍,可以对四组训练集的质量进行更全面地分析。总体上看,训练集 $D_1$ 的质量最差,因为其识别矩阵图像中含有最多的“1”,其中,43和46两行含有“1”元素最多,表明第43和46次单次试验数据( $x_{43}$ 和 $x_{46}$ )受到了较严重的干扰。另外观察 $Y_D(1)$ 图像的列,也可以发现,较多单次试验的原始类型标签与估计结果不一致,具有类似特征的训练集,通常不能直接用于ICA/CSP空域滤波器设计。相比而言,训练集 $D_2$ 的质量要优于 $D_1$ 。通过观察图像的行和列,可发现 $D_2$ 中不含被严重干扰的单次试验数据,但有若干(10个左右)原始标签与估计标签可能不一致的单次试验。同理不难得出, $D_3$ 和 $D_4$ 的数据质量要明显优于 $D_1$ 和 $D_2$ 。

[0126] 在MI-BCI系统实现研究中,对训练样本的质量和数量均有较高要求,但由于MI训练样本的采集实验受限于诸多因素(例如,测试环境,受试者经验和精神状态等),因此训练样本的质量往往难以保证,每次实验能够采集的样本数量也相对较少。在已报到的MI-BCI研究中,通常采用人工方式检查MI-EEG数据波形,以剔除含有明显伪迹干扰的低质量数据段,但人工方式不仅效率低、误差大,而且也难以检测出因受试者精神状态和经验等因素导致的“无效”训练样本。

[0127] 本发明所提出的解决方案能自动实现对训练集中各单次试验样本的质量评估,并给出了反映训练集质量的各项量化指标和总体评分。所提方法所涉及的运算量很低,因此具有很好的实用性。本发明的特色是将“训练样本质量的直接评估问题”转化成对“基于单次试验设计的ICA滤波器的性能评估问题”,这一思路能综合考虑了影响数据质量的不同因素,所得质量评估结果具有较好的准确性和全面性。

[0128] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其他的具体形式实现本发明。因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权

利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化囊括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。

[0129] 此外,应当理解,虽然本说明书按照实施方式加以描述,但并非每个实施方式仅包含一个独立的技术方案,说明书的这种叙述方式仅仅是为清楚起见,本领域技术人员应当将说明书作为一个整体,各实施例中的技术方案也可以经适当组合,形成本领域技术人员可以理解的其他实施方式。

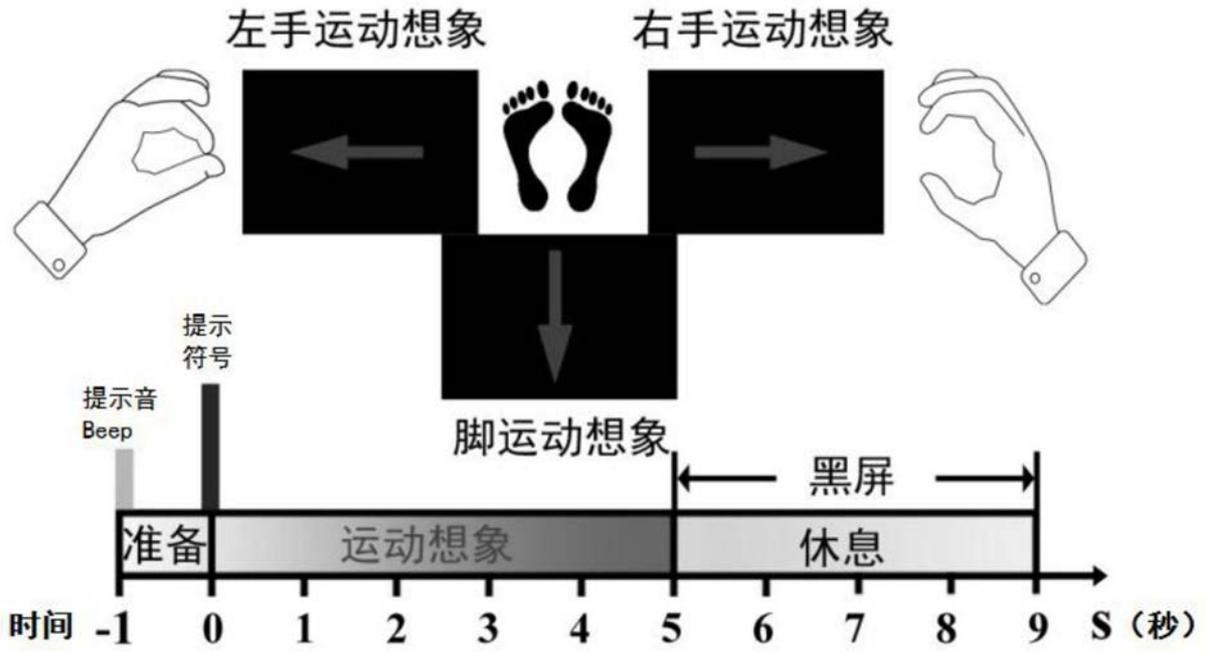


图1a

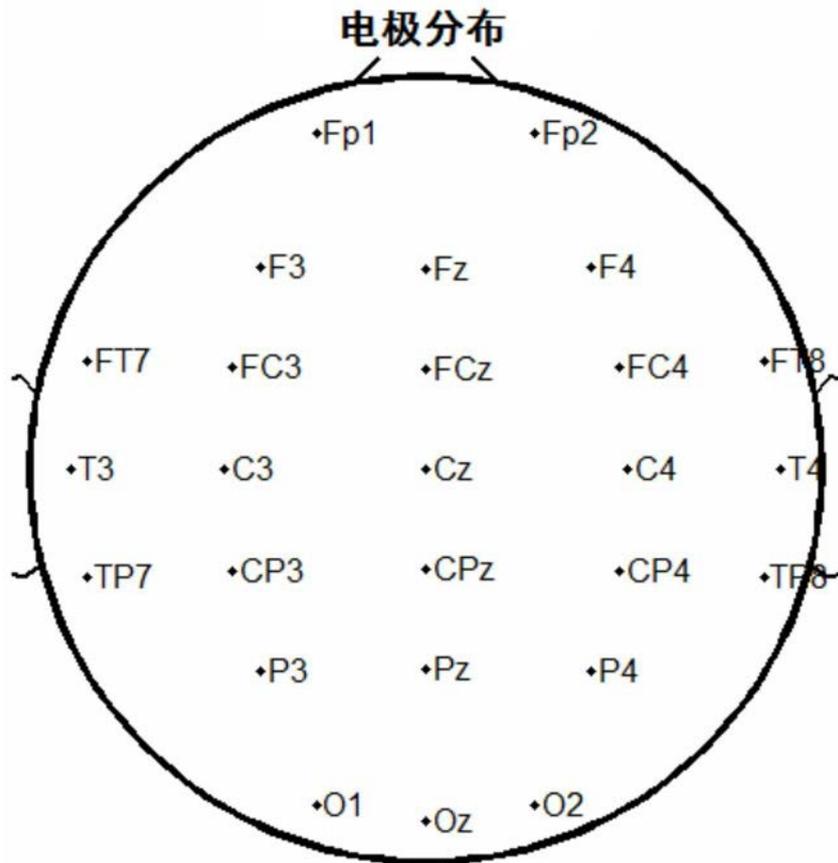


图1b

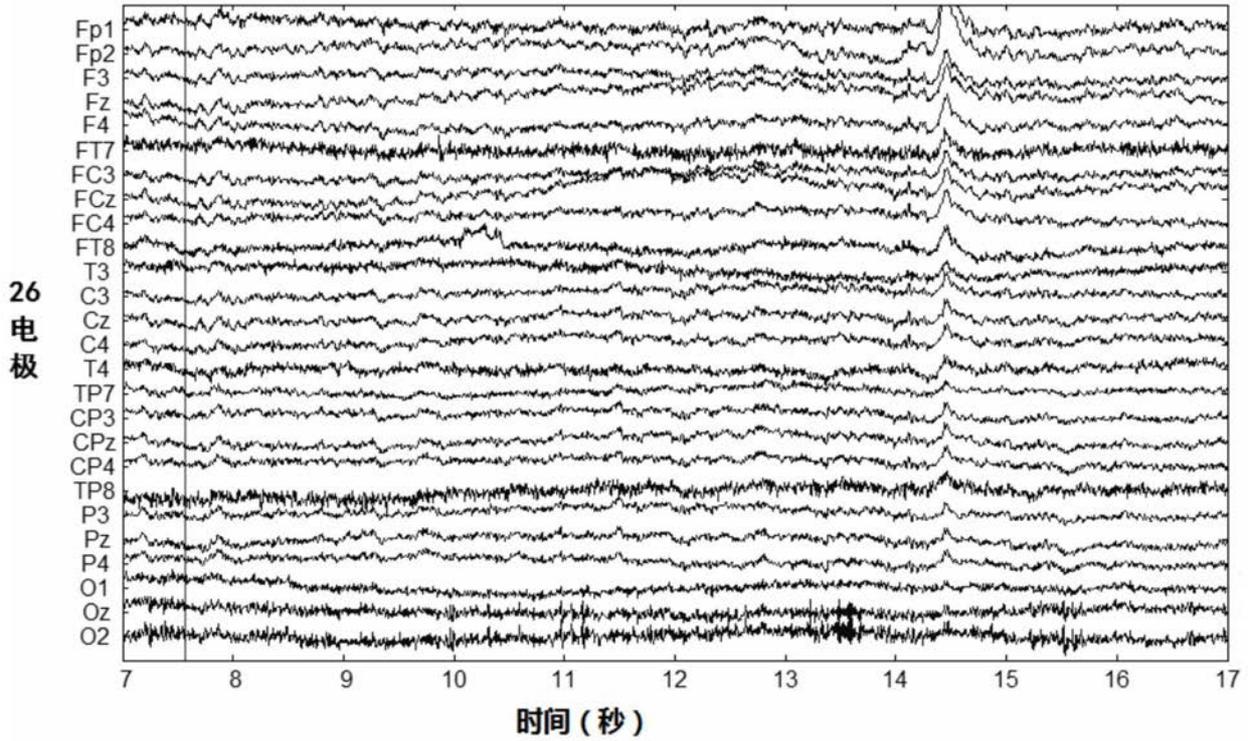


图1c

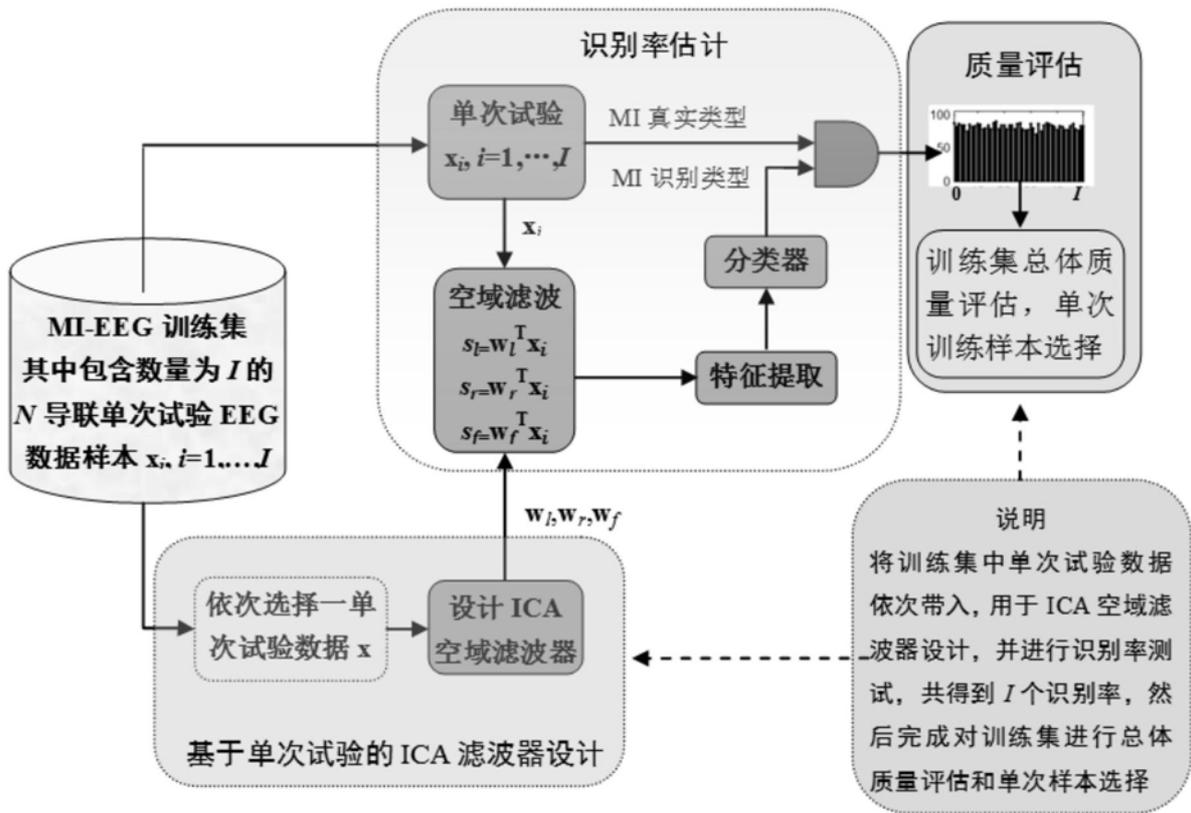


图2

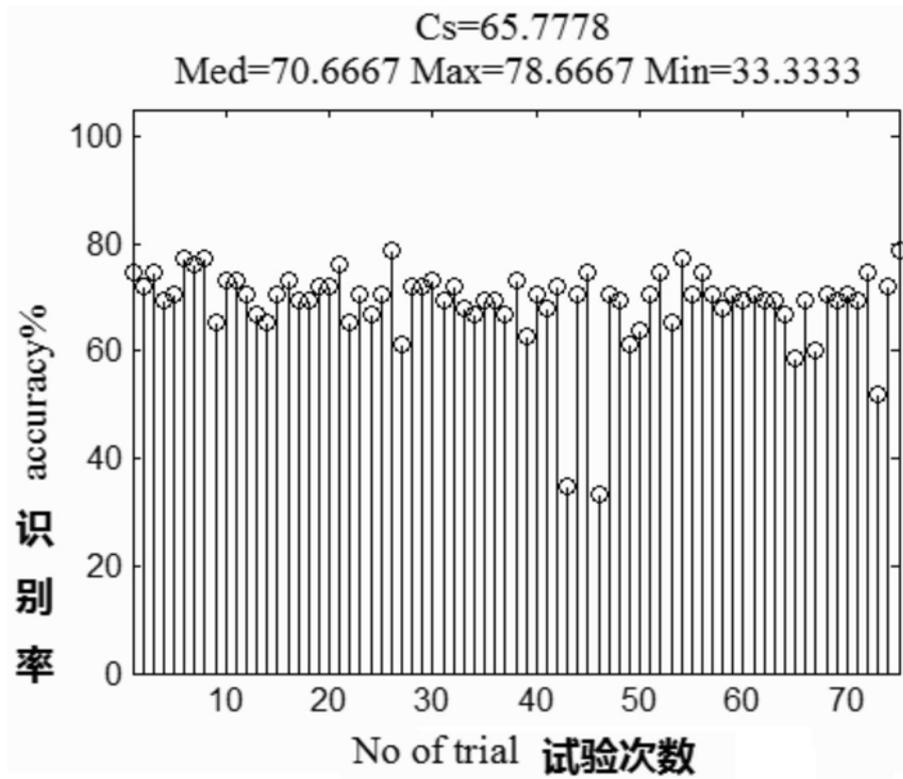


图3a

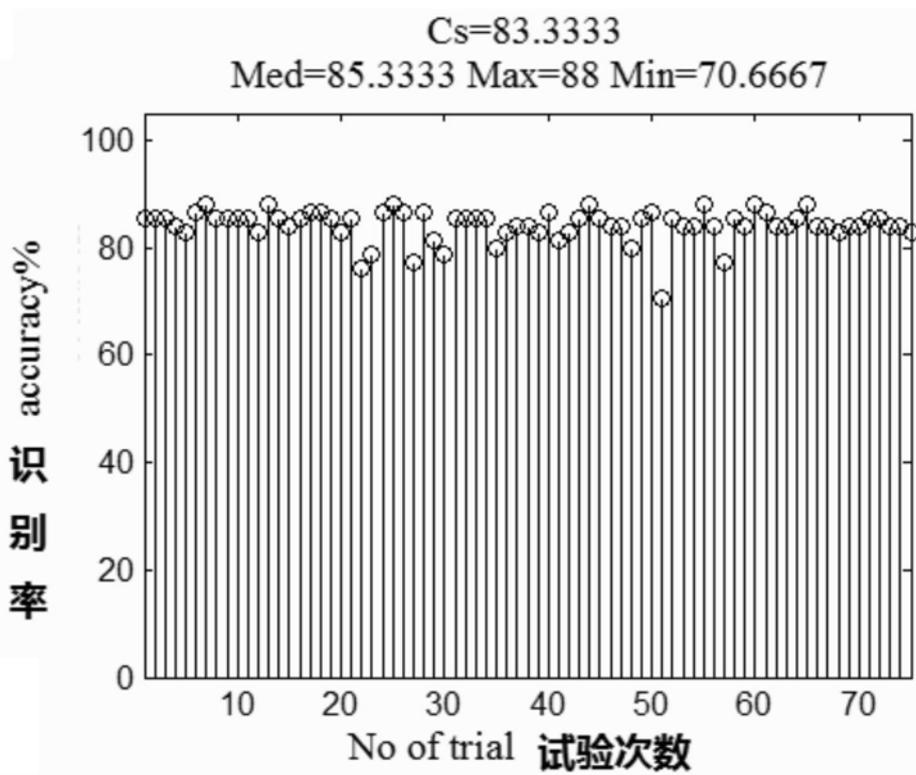


图3b

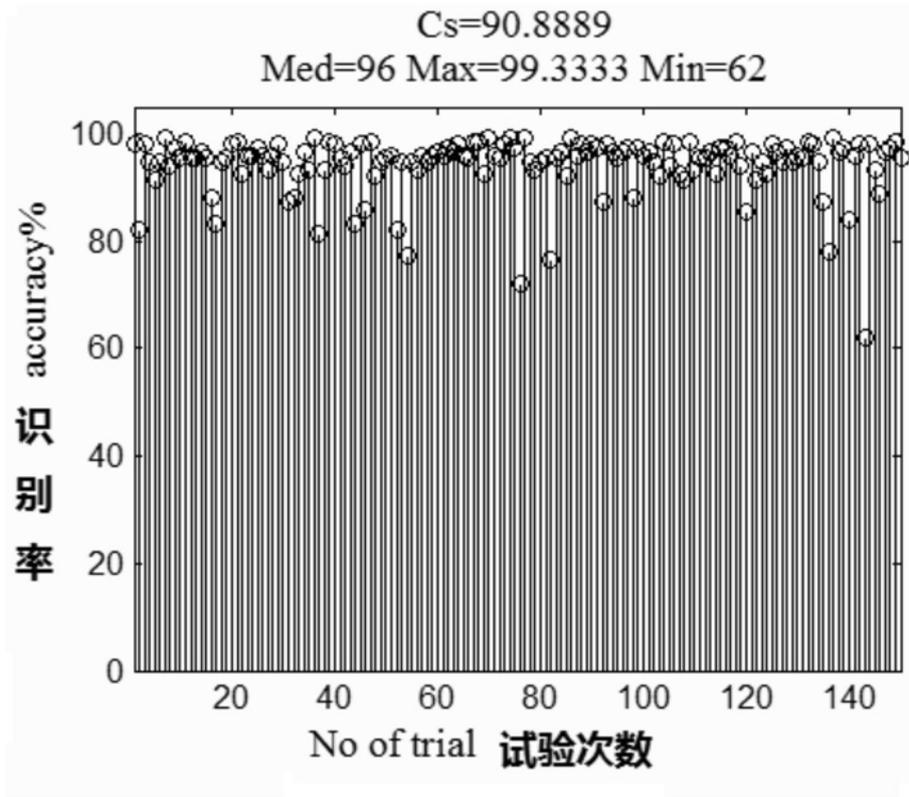


图3c

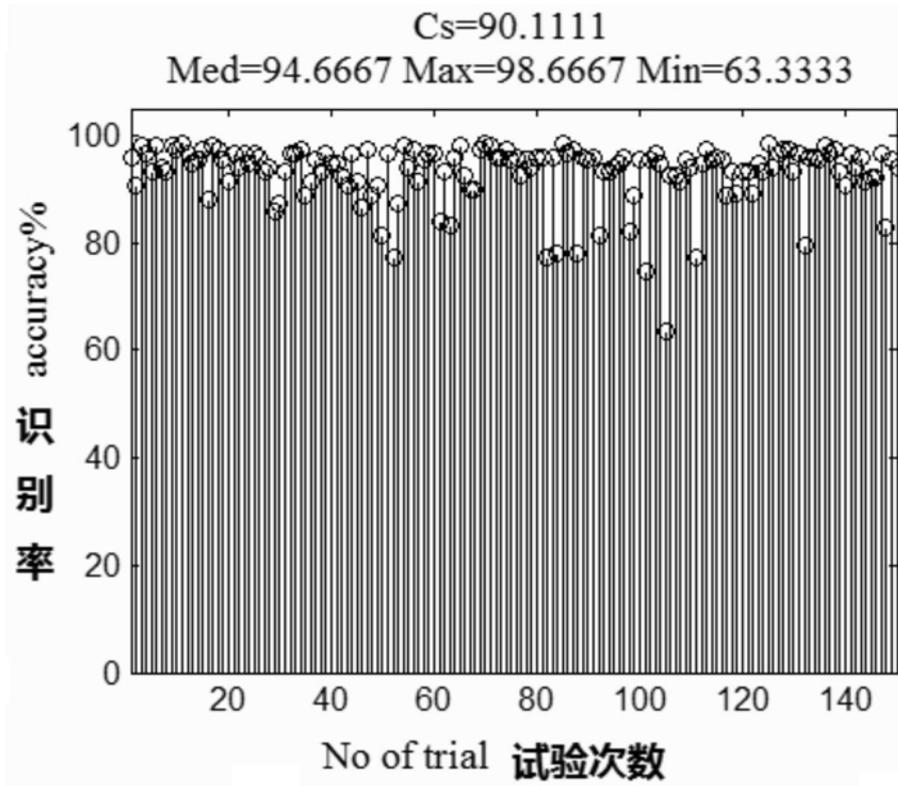


图3d

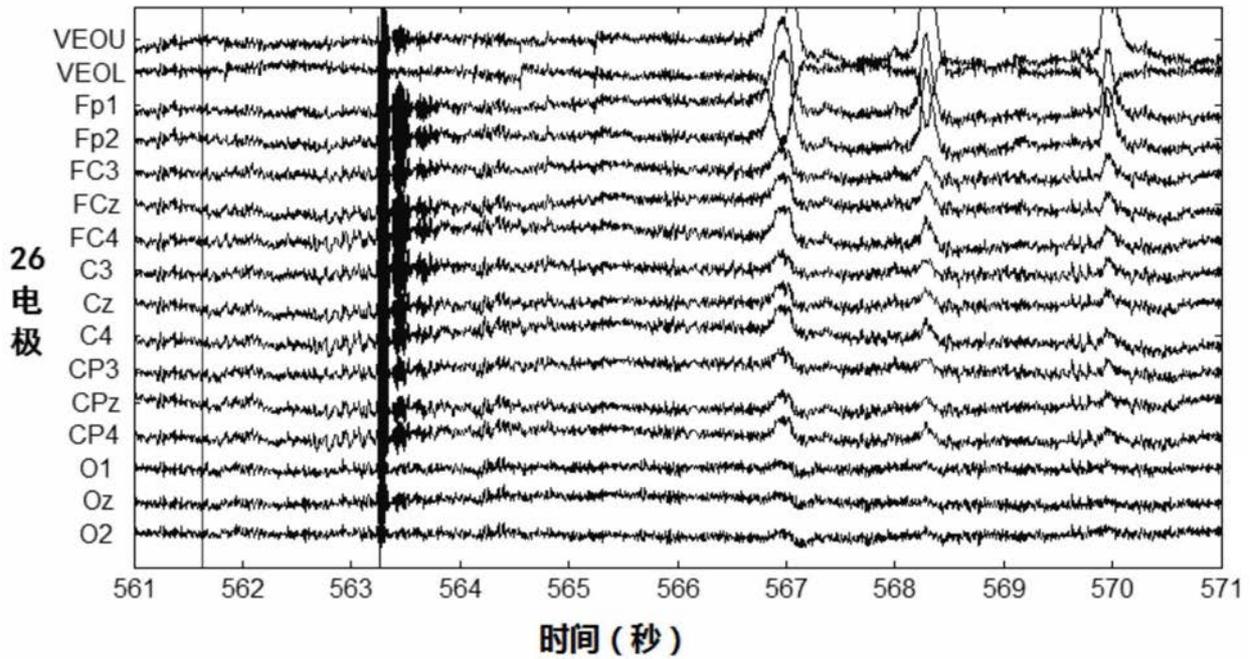


图4a

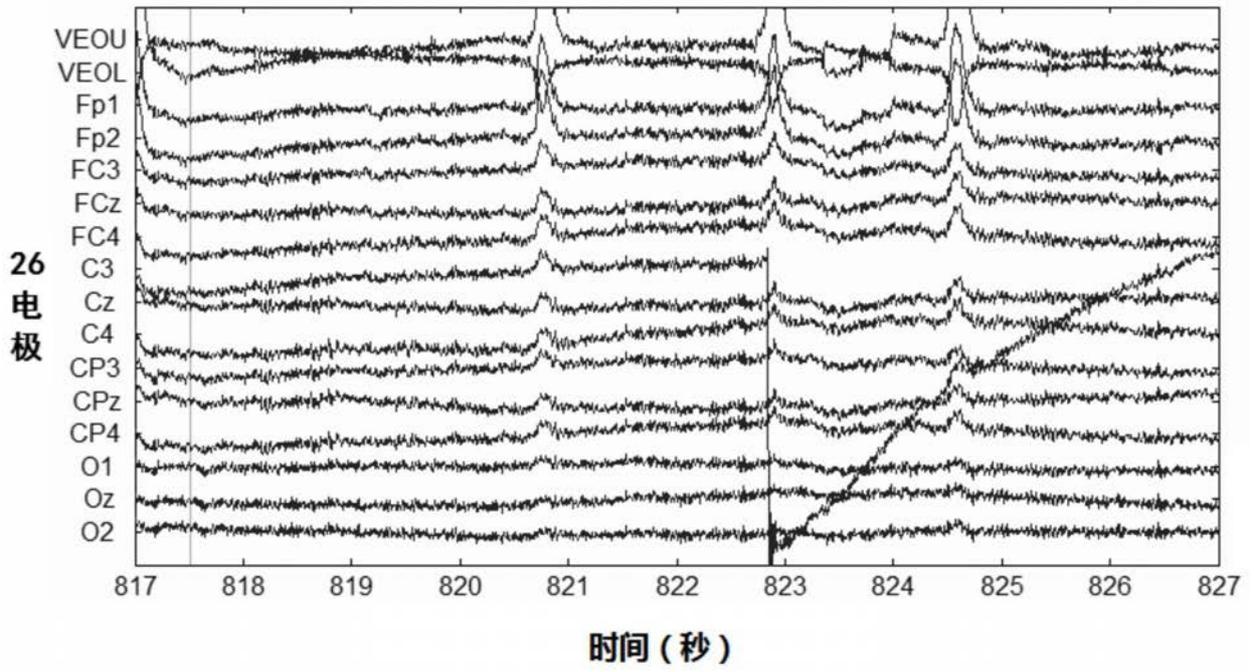


图4b

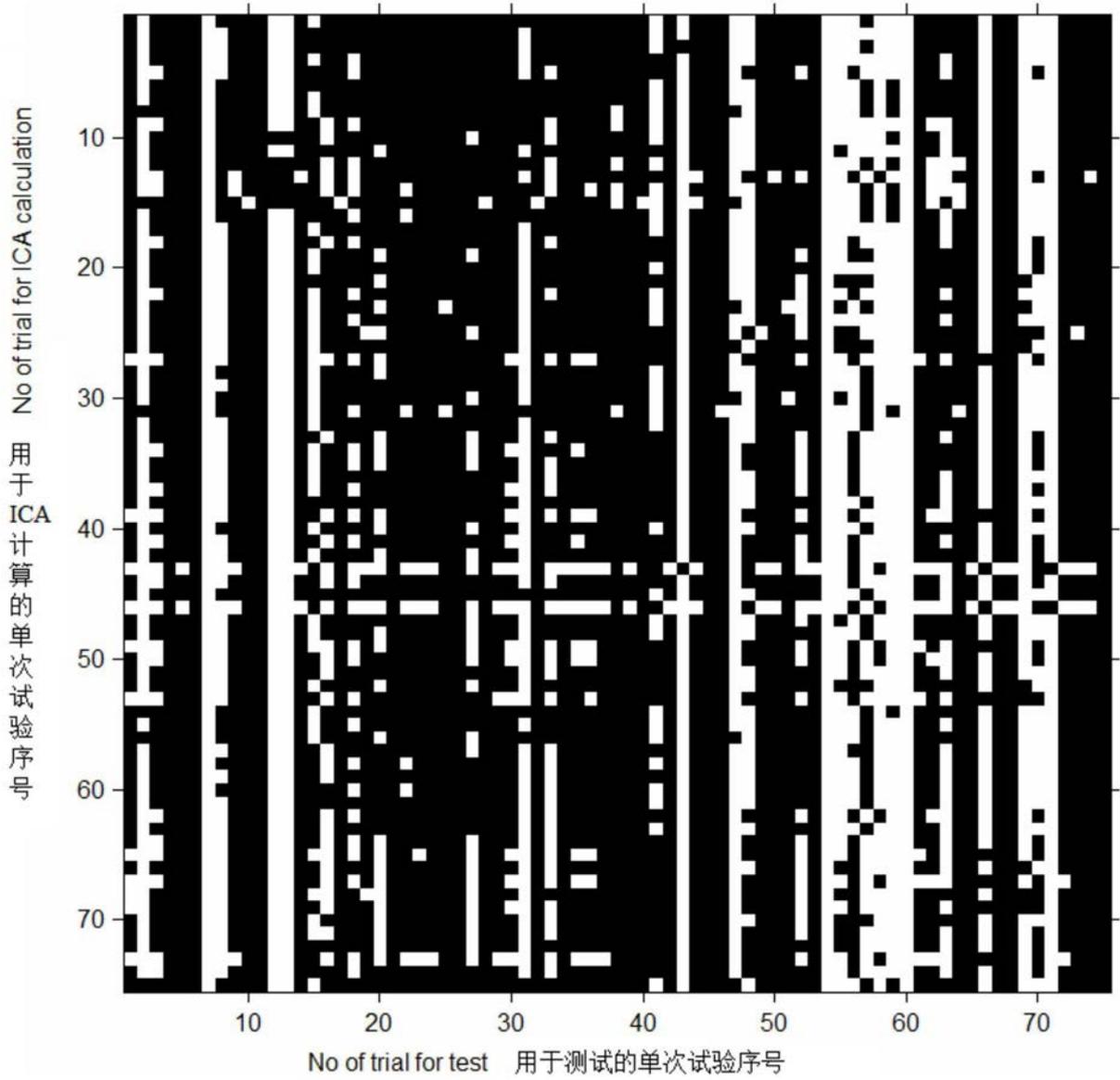


图5a

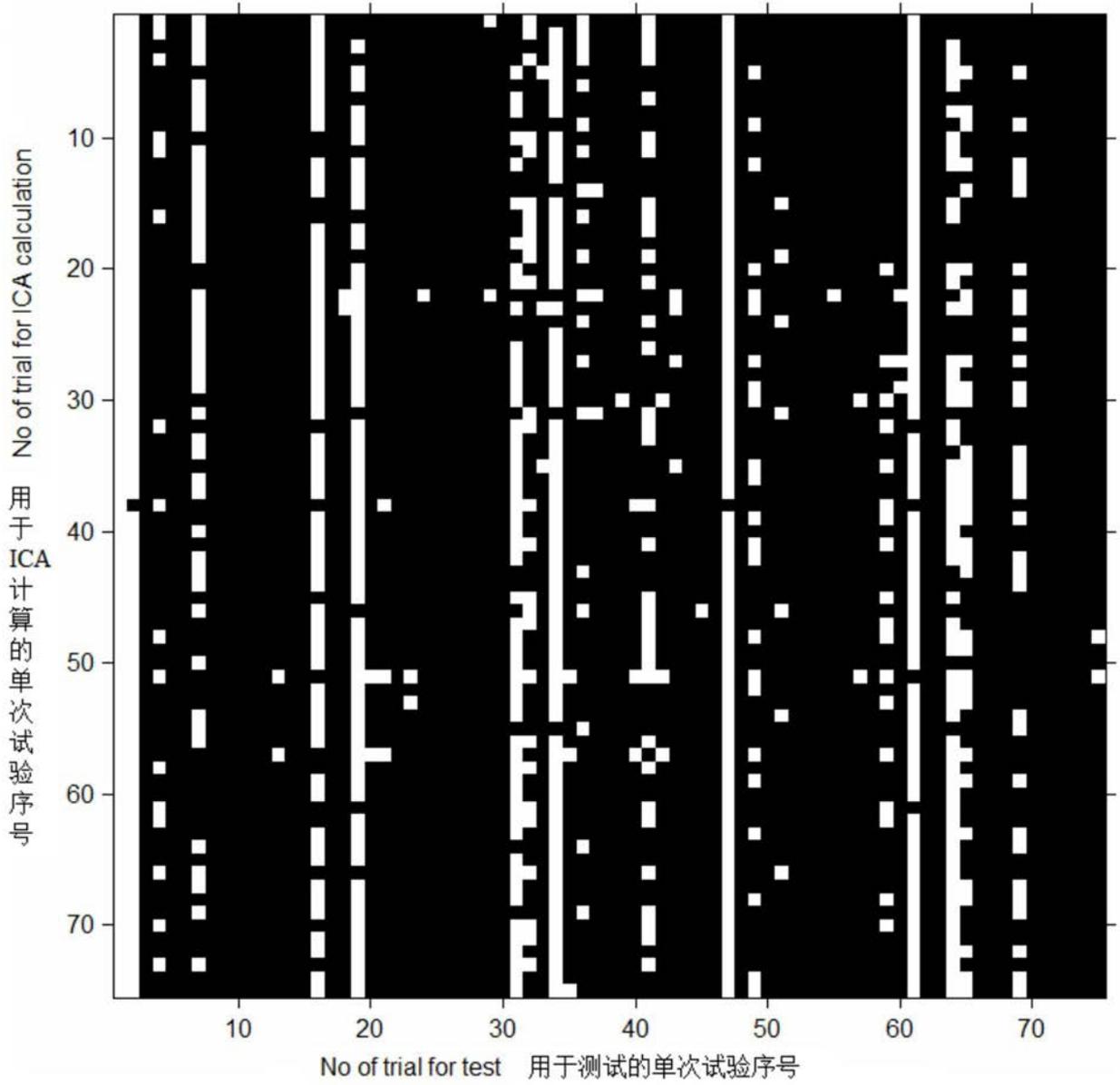


图5b

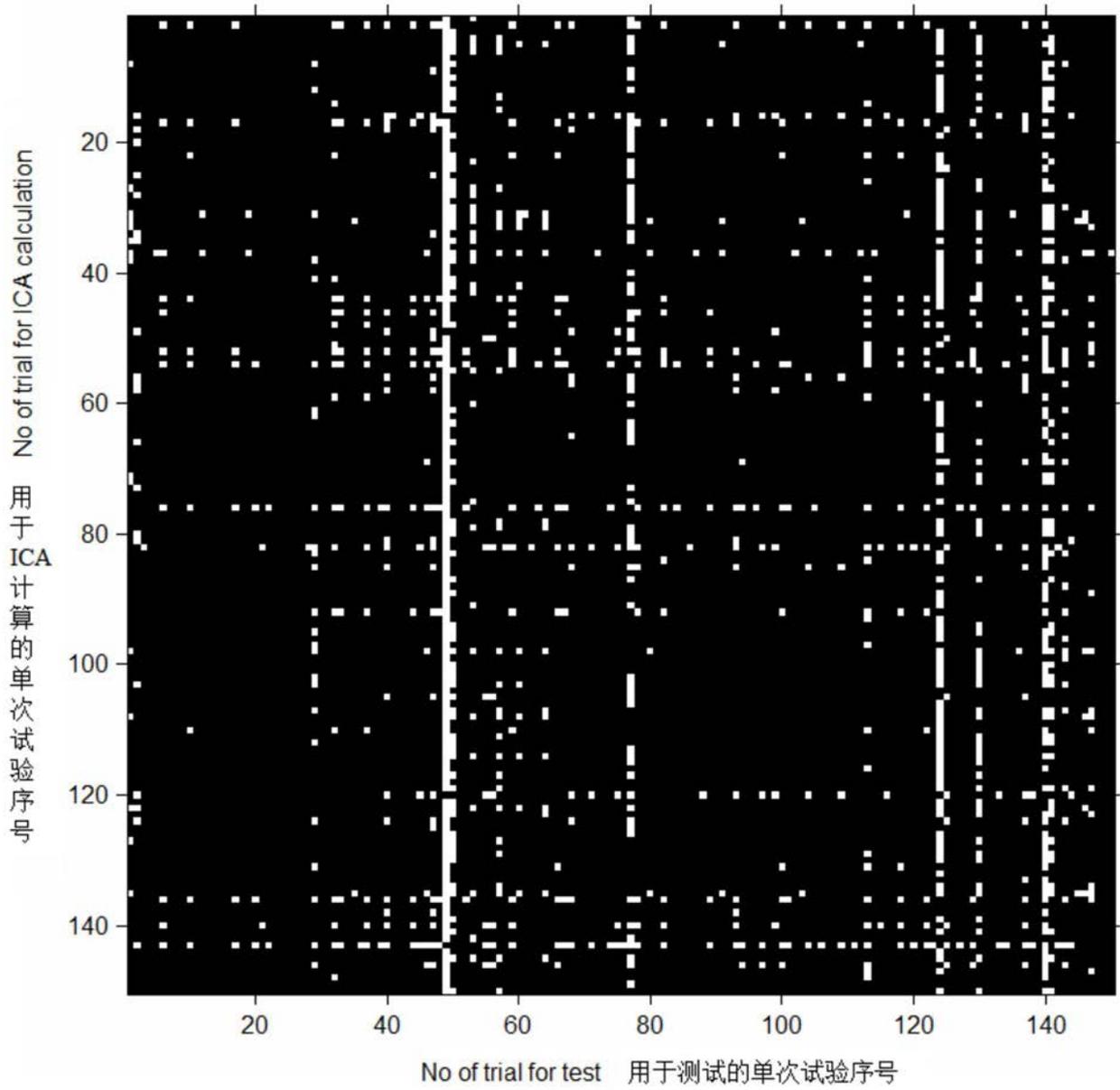


图5c

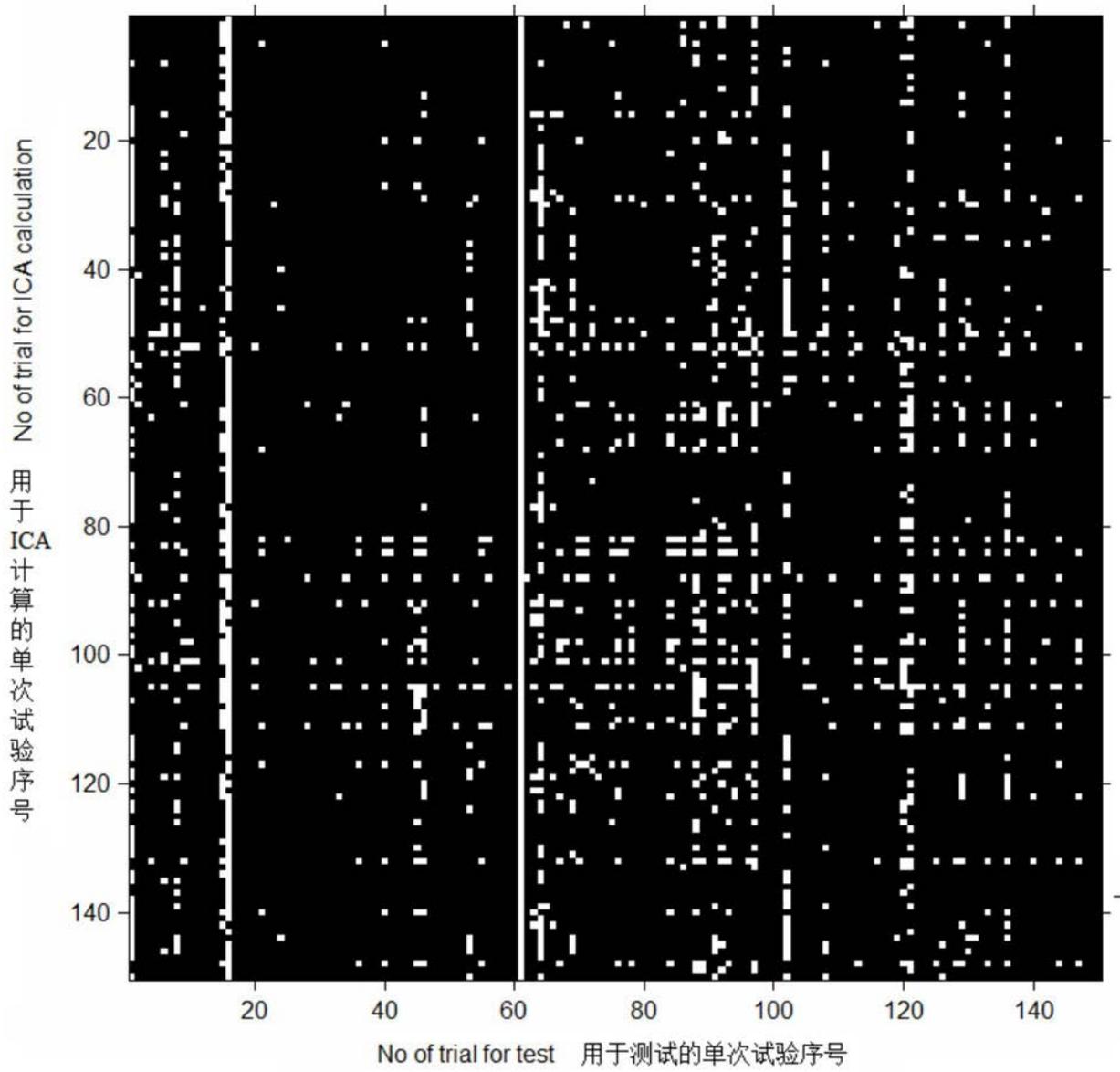


图5d

专利名称(译)	MI-BCI系统的训练集质量评估方法和单次训练样本的优化方法		
公开(公告)号	<a href="#">CN110516711A</a>	公开(公告)日	2019-11-29
申请号	CN201910677300.6	申请日	2019-07-25
[标]申请(专利权)人(译)	安徽大学		
申请(专利权)人(译)	安徽大学		
当前申请(专利权)人(译)	安徽大学		
[标]发明人	吴小培 陈文伟 周蚌艳 陈瑶 吕钊 张超 郭晓静 张磊 高湘萍		
发明人	吴小培 陈文伟 周蚌艳 陈瑶 吕钊 张超 郭晓静 张磊 高湘萍		
IPC分类号	G06K9/62 G06K9/00 A61B5/0476 A61B5/04 A61B5/00		
CPC分类号	A61B5/04012 A61B5/0476 A61B5/7221 G06K9/00523 G06K9/00536 G06K9/6256		
代理人(译)	郭华俊		
外部链接	<a href="#">Espacenet</a> <a href="#">SIPO</a>		

摘要(译)

本发明公开了一种MI-BCI系统的训练集质量评估方法和单次训练样本的优化选择方法，采集单次试验MI-EEG数据；设计相应的ICA空域滤波器；计算单次识别率Acc(k)和总体识别率Acc；确定用于单次试验数据质量评估的量化指标；计算训练样本质量的综合评分Cs；根据Cs的数值对训练集的质量进行评估。再根据所得Acc、量化指标和综合评分Cs，对训练集中的单次试验样本进行优化选择。本发明的训练集质量评估方法和单次训练样本的优化方法，具有较好的稳定性、准确性和运算复杂度低等优点。

