

文章编号: 1006-3080(2020)06-0792-08

DOI: 10.14135/j.cnki.1006-3080.20190901002

基于多相关性的导联前向搜索算法用于运动想象分类

殷飞宇, 金晶, 王行愚

(华东理工大学信息科学与工程学院, 化工过程先进控制和优化技术教育部重点实验室, 上海 200237)

摘要: 针对基于运动想象(Motor Imagery, MI)的脑-机接口(Brain-Computer Interface, BCI)系统中导联过多的问题, 提出了一种多相关性导联前向搜索(Multi-correlation Forward Searching, MCFS)算法来优化导联集, 改善系统性能。首先基于训练集对导联集进行前向搜索, 同时以验证集分类精度更新对3种相关性算法的信任值; 然后根据3种相关性方法的信任值, 选择优质导联组合, 采用共空间模式(Common Spatial Pattern, CSP)获得运动想象特征, 通过线性核的支持向量机(Support Vector Machine, SVM)训练分类模型。对该算法在两个数据集(BCI竞赛IV中的data set I数据集I和BCI竞赛III中的data set IVa)上进行验证, 分别得到了81%和87%的平均分类精度。此外, 与其他3种常用导联选择方法相比, MCFS算法获得了最高的平均分类精度, 性能优越, 为基于运动想象的BCI系统的应用提供了技术参考。

关键词: 脑-机接口; 运动想象; 多相关性算法; 导联选择; 共空间模式

中图分类号: TP273

文献标志码: A

脑-机接口(BCI)系统是指直接跳过大脑外周神经和肌肉的参与, 通过意念实现对外部装置的控制^[1]。由于脑电图(EEG)信号拥有较短的时间常数和相对较高的时间分辨率, 因此成为最广泛应用于BCI系统中的脑电信号^[2-3]。对于不同的受试者, 其在进行同一个思维任务时所影响的皮层区域范围各不相同, 所以仅仅通过生理知识和实验经验不能完全确定有效导联位置和导联使用数量^[4], 因此实验中采取多导联的方式记录脑电数据^[5]。事实上, 多导联采集脑电数据虽然可以获得较全的脑电信息, 但同时也会引入冗余信息, 增加数据维度, 影响模型性能并且不利于实际使用^[6], 因此, 选择优质的导联组合对提升分类精度非常重要。

传统的导联优化算法主要分为3类: (1)基于神经学的先验知识。分布在运动想象感知皮层的导联往往会被认为包含更多有效信息而被选为最终导联组合, 例如C3和C4等^[7]。(2)基于Wrapper的导联

优化算法。其核心思想以分类器的分类精度作为导联搜索策略的评价指标, 根据分类结果不断修正所选导联子集, 选择的导联分类准确率高但计算开销大^[8]。(3)基于导联分布的优化策略。根据导联之间相关性进行排序选择, 选择强相关导联作为最优导联子集^[9]。对于同一个范式下的实验数据, 已有的算法往往根据单个相关性进行导联选择, 但单个相关性会由于受试者之间的差异性而无法对所有受试者的导联分布进行准确排序选择。因此本文提出了结合神经学的先验知识和基于导联分布优化策略相结合的导联选择算法, 在选择位于运动感知皮层的C3和C4导联后, 通过多相关性集成算法前向搜索选择剩余导联, 从而获得有效导联组合, 减少数据维度和计算量, 提高分类精度。同时, 以BCI竞赛IV中的dataset I和BCI竞赛III中的dataset IVa两个公共竞赛数据集来检验算法的可行性。实验结果表明, 运用MCFS导联优化算法不仅大幅减少了导联

收稿日期: 2019-09-01

基金项目: 国家重点研发计划(2017YFB13003002); 国家自然科学基金(61573142, 61773164, 91420302); 111引智计划(B17017)

作者简介: 殷飞宇(1994—), 男, 江苏人, 硕士生, 研究方向为脑-机接口。E-mail: yinfy0822@163.com

通信联系人: 金晶, E-mail: jinjingat@gmail.com

引用本文: 殷飞宇, 金晶, 王行愚. 基于多相关性的导联前向搜索算法用于运动想象分类[J]. 华东理工大学学报(自然科学版), 2020, 46(6): 792-799.

Citation: YIN Feiyu, JIN Jing, WANG Xingyu. Channel Selection Based on Multi-Correlation Forward Searching Algorithm for MI Classification[J]. Journal of East China University of Science and Technology, 2020, 46(6): 792-799.

集的数目,而且分类精度提升明显,因此选择策略具有可行性。

1 运动想象脑电信号数据集

数据集 1(Dataset 1)采用的是 2008 年第四届国际脑-机接口大赛的数据(BCI competition IV dataset I),由柏林 BCI 研究组提供。实验选择 4 名健康受试者(分别命名为 a、b、f、g),实验中计算机屏幕上提示进行左手、右手或脚 3 种运动想象思维任务,但单次实验只进行两类运动想象任务。实验共包含 200 个运动想象任务,降采样频率为 100 Hz。BCI 研究组使用 59 个有效导联进行脑电数据的采集。实验步骤如下:

(1)0~2 s,此时计算机黑屏,受试者处于空闲放松的状态;

(2)2 s,短暂的蜂鸣声提示受试者实验即将开始;
(3)2~6 s,屏幕出现向左、向右或向下的箭头,受试者根据屏幕箭头提示,完成相应的运动想象任务。

数据集 2(Dataset 2)采用 2005 年第三届国际脑-机接口大赛的脑电数据(BCI competition III dataset IVa),5 名健康受试者(分别命名为 aa、al、av、aw、ay)执行右手和脚两类运动想象任务。实验采用 118 导联电极帽,按国际标准 10-20 导联放置,BrainAmp 放大器为脑电采集设备,降采样频率为 100 Hz,共执行 280 个运动想象任务。实验步骤如下:

(1)0~3.50 s,受试者根据屏幕提示执行运动想象任务;

(2)3.50~5.75 s,受试者在 1.75~2.25 s 内的不固定时间段放松。

Dataset 1、Dataset 2 的实验范式如图 1 所示。

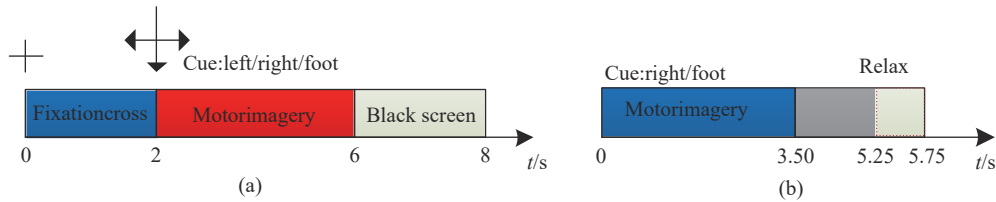


图 1 Dataset 1(a)和 Dataset 2(b)的实验范式

Fig. 1 Experimental paradigms of dataset 1 (a) and dataset 2 (b)

2 分析方法

2.1 数据预处理

对 Dataset 1 和 Dataset 2 都采用提示想象信号出现后到提示信号消失的时间窗,分别将 2~6 s 和 0~3.50 s 时间序列作为分析数据。由于运动想象具有相关同步/去同步现象,其脑电特征在 Mu 节律段(主要分布在 8~13 Hz)^[10]和 Beta 节律段(主要分布在 13~30 Hz)表现最为明显,同时脑电信号弱,易混入噪声,信噪比低,因此本文采用 8~30 Hz 带通滤波器对多导联 EEG 信号进行滤波,去除眼电、肌电等干扰信号对脑电信号的影响,同时保留运动想象事件相关去同步/事件相关同步发生的 Mu 和 Beta 节律段,得到运动想象思维的有效信号。

2.2 MCFS 算法

MCFS 算法通过更新 3 种相关性的信任值进行导联优化,因此计算导联之间的相似性是导联选择的一个重要环节。相关性计算有很多方法,本文选用了互信息、余弦相似性和皮尔森相关分析法作为

相似性度量方法。设滤波后的单次实验的脑电数据为 $E = (s, c, n)$, 其中 $s = \{0, 1\}$ 为单次实验的类别标签; $c = 1, 2, \dots, C$, C 为采集脑电信号的总导联数; $n = 1, 2, \dots, N$, N 为脑电信号的时间序列。

2.2.1 互信息 互信息^[10](Mutual Information, MI)是一种常用的信息度量方法,可以有效地衡量两个随机变量之间的相关程度。导联 c_i 和 c_j 之间的互信息 $I(c_i, c_j)$ 计算公式如下:

$$I(c_i, c_j) = \iint p(c_i, c_j) \lg \frac{p(c_i, c_j)}{p(c_i)p(c_j)} dc_i dc_j \quad (1)$$

其中: $p(c_i)$ 和 $p(c_j)$ 分别为导联 c_i 和 c_j 的边缘概率分布; $p(c_i, c_j)$ 为导联 c_i 和 c_j 的联合概率密度。

2.2.2 余弦相似性 余弦相似性(Cosine Similarity, CS)被广泛应用于文本分类、模式识别和神经生理学数据分析^[11]。通过式(2)计算导联 c_i 和 c_j 之间的余弦相似程度:

$$\cos(c_i, c_j) = \frac{\sum_{n=1}^N (c_{in} \cdot c_{jn})}{\sqrt{\sum_{n=1}^N c_{in}^2} \sqrt{\sum_{n=1}^N c_{jn}^2}}, \quad (2)$$

$$i, j = 1, 2, \dots, C$$

其中: $\cos(\cdot)$ 表示导联 c_i 和 c_j 的角度, 角度越小, 则两个导联的相似程度越高, 反之亦然。 c_{in} 和 c_{jn} 分别表示导联 c_i 和 c_j 的时间序列长度。

2.2.3 皮尔森相关分析 皮尔森相关分析(Pearson's Correlation Analysis, PCA)是一种判别变量之间线性相关性的分析方法^[12], 具体定义如下:

$$\rho(c_i, c_j) = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{c_{in} - \bar{c}_{in}}{\sigma_{c_{in}}} \right) \left(\frac{c_{jn} - \bar{c}_{jn}}{\sigma_{c_{jn}}} \right) \quad (3)$$

其中: \bar{c}_{in} 和 \bar{c}_{jn} 分别表示 c_{in} 和 c_{jn} 的时间序列均值; $\sigma_{c_{in}}$ 和 $\sigma_{c_{jn}}$ 分别表示 c_{in} 和 c_{jn} 的时间序列方差; $\rho(c_i, c_j)$ 表示导联的相关程度。

2.2.4 信任值更新 信任值的更新基于 3 种相关性对精度下降的惩罚和对精度提升的奖励。

$$\sigma = \begin{cases} \sigma_p = \tau \lg \frac{1 - |J_{(f^+)} - J_{(f)}|}{|J_{(f^+)} - J_{(f)}|}, & \text{if } J_{(f^+)} - J_{(f)} < 0 \\ \sigma_r = \frac{1}{\sigma_p}, & \text{if } J_{(f^+)} - J_{(f)} > 0 \\ \sigma_n = \mathbb{I}(|J_{(f^+)} - J_{(f)}|), & \text{if } J_{(f^+)} - J_{(f)} = 0 \end{cases} \quad (4)$$

其中: σ 为每种相关性的信任值; $J_{(i)}$ 表示导联子集 10 次交叉验证后的精度; f 表示当前的导联组合, 则 f^+ 为在前向搜索时, 添加导联后的新导联组合; σ_p 、 σ_r 和 σ_n 分别表示对分类精度下降时的惩罚、对分类精度提升时的奖励以及当精度没有变化的参数; τ ($\tau = 100$) 是一个常数, 防止随着惩罚的次数增多, 出现信任值弥散的情况; $\mathbb{I}(\cdot)$ 为示性函数, 表示 0 和 1 取值的情况。各个受试者脑电数据对 3 种相关性的信任值如表 1 所示。

2.2.5 基于多相关性的导联搜索算法 MCFS 算法

表 1 各受试者对于 3 种相关性算法的信任值

Table 1 Trust values of three correlation algorithms for each subject

| Subject | Trust value | | |
|---------|-----------------------|----------------------|-----------------------|
| | MI | CS | PCA |
| a | 1.9×10^9 | 1.8×10^{-9} | 2×10^{-13} |
| b | 3.7×10^{-9} | 1.1 | 2.5×10^{-13} |
| f | 8.3×10^{24} | 1.1×10^{-2} | 9.9×10^{-1} |
| g | 1.1×10^{11} | 2.7×10^{10} | 1.1×10^5 |
| aa | 4.5×10^{66} | 7.7×10^{37} | 2.3×10^{32} |
| al | 1.4×10^{157} | 1.5×10^{15} | 1.5×10^{15} |
| av | 4×10^{119} | 2.2×10^3 | 2.2×10^3 |
| aw | 10^{113} | 2.4×10^{11} | 4.1×10^{15} |
| ay | 7.7×10^{198} | 2.2×10^{20} | 4.6×10^{21} |

主要分为训练更新信任值和通过信任值进行导联优化两个阶段:

(1) 分别计算导联之间在互信息、余弦相似性和皮尔森相关分析下的导联之间的相关性, 然后选择位于运动感知皮层的 C3 和 C4 导联, 根据随机初始化的导联排序前向搜索, 通过训练集训练模型, 在验证集上交叉验证, 以分类精度作为评价标准。一旦新增加的导联提高分类准确率, 则该导联被选用, 否则被放弃, 与此同时根据式(4)更新对 3 种相关性的信任值。

(2) 随机初始化导联排序后, 根据 3 种相关性的信任值选择导联组合。此时导联组合 f 由 P_1 ($P_1 = 0.9$) 概率的信任值和 P_2 ($P_2 = 0.1$) 概率的投票共同决定, 然后对导联组合 f 进一步优化。前向搜索导联组合 f , 提取共空间模式(CSP)特征后输入以线性核 SVM 作为分类器训练模型, 以 10 次交叉验证的分类精度作为评价标准, 输出最优导联组合 f_{best} 。

MCFS 算法流程图如图 2 所示。

2.3 共空间模式

CSP 算法是一种广泛使用的脑电信号预处理算法, 能够有效地提取差异性特征^[13]。该算法通过设计空间滤波器对脑电信号进行空间投影, 使得两类信号的方差差异性最大, 进而实现对两类信号的分类识别。

假设滤波后的脑电数据为 $X_i \in R^{N \times T}$ (N 为有效导联数目; T 为每个通道的采样点), 则规定单次脑电信号 EEG 的空间协方差矩阵求解如下:

$$\Sigma_i = \frac{X_i X_i^T}{\text{trace}(X_i X_i^T)} \quad (5)$$

其中: X_i 为左、右手运动想象的第 i 次实验的脑电数据; $\text{trace}(X_i X_i^T)$ 为 $X_i X_i^T$ 的对角元素之和; Σ_i 为第 i 类空间协方差矩阵。

CSP 通过解决优化式(6)来学习空间滤波器(w), 使得两类信号差异最大化:

$$\max_w J(w) = \frac{w^T \Sigma_1 w}{w^T \Sigma_2 w}, \quad \text{s.t. } \|w\|^2 = 1 \quad (6)$$

其中, $\|\cdot\|^2$ 表示 L_2 范数。同时, 式(6)可以通过等效求解以下广义特征值问题来实现^[13]:

$$\Sigma_1 w = \lambda \Sigma_2 w \quad (7)$$

然后收集对应于式(7)中的 M 个最大和 M 个最小特征值的特征向量以形成所学习的空间滤波器。将滤波后的脑电数据经过式(8)转变为投影矩阵 Z_p :

$$Z_p = WX \quad (8)$$

对脑电信号进行特征提取, 将投影后的信号

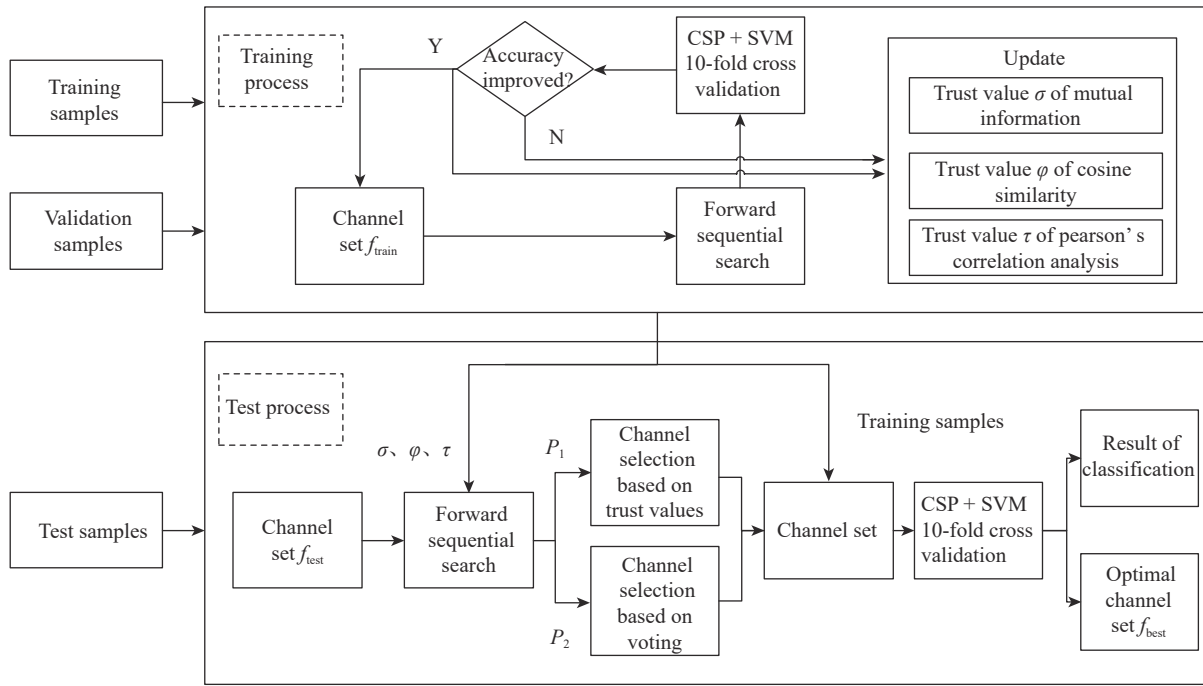


图 2 MCFS 算法导联选择流程图

Fig. 2 Flow chart of channel selection with MCFS algorithm

Z_p 做如下变换得到 f_p, f_p 即为两类运动想象任务的特征值。

$$f_p = \lg \frac{\text{var}(Z_p)}{\sum_{i=1}^{2m} \text{var}(Z_p)} \quad (9)$$

2.4 交叉验证

为了减少模型过拟合的可能性, 交叉验证是一个有效的方法^[14]。为防止信任值的参数信息被提前学到, 本文将数据集分为训练集 (80%)、验证集 (10%) 和测试集 (10%)。在训练集上创建模型, 通过验证集 10 次交叉验证修正对 3 种相关性的信任值, 最终在测试集上测试模型的性能。

3 结果与分析

3.1 信任值优化

MCFS 导联优化算法通过更新 3 种相关性的信任值优化最终的导联组合。从表 1 中不难看出, 相比其他两种相关性, 互信息的信任值越高, 在选择导联上的效果越好。对于 Dataset 1 的 4 名受试者而言, 受试者 g 的互信息和余弦相似性的信任值非常接近; 受试者 b 的余弦相关性的信任值高于互信息。对于 Dataset 2 中的 5 名受试者而言, 互信息总是值得信赖的, 而余弦相似性和皮尔森相关分析的信任值更为接近, 仍有一定的参考价值。因此, 本文提出以 P_1 概率信任某一种相关性, 同时仍然以 P_2 概率保证其余

两种相关性对导联选择的作用。

3.2 导联优化

对于 Dataset 1 和 Dataset 2, 实验中分别采用 59 和 118 个导联记录脑电数据。事实上, 多导联采集脑电数据虽然可以提供充足的脑电信息, 但也会带入冗余信息和引入噪声, 对脑电分类精度造成影响。本文采用 MCFS 算法来选取有效导联, 不但可以减少数据复杂度和计算量, 还可以提升分类精度。图 3 示出了选择全部导联 (All Channels, AC) 和所有受试者通过 MCFS 导联搜索算法各自选择的导联数。从图中可以清晰地看出, 经过 MCFS 算法选择后, 导联数大幅减少 (Dataset 1: 59vs13, Dataset 2: 118vs35), 各受试者经过 AC 和 MCFS 算法选择的导联数如表 2 所示。

如表 2 所示, 对于 Dataset 1 中的 4 名受试者而言, 通过 MCFS 算法优化导联组合后, 导联数目大幅减少, 其中受试者 a 保留 11 导联, 减幅 81%; 受试者 b 保留 4 导联, 减幅 93%; 受试者 f 保留 13 导联, 减幅 78%; 受试者 g 保留 24 导联, 减幅 59%; 4 名受试者平均选择 13 导联, 减幅接近 78%。

对于 Dataset 2 中的 5 名受试者, 采用 MCFS 算法选取导联后, 受试者 aa 仅从 118 导联中选取 25 导联, 减幅 79%; 受试者 al 选取 28 导联, 减幅 76%; 受试者 av 选取 52 导联, 减幅 56%; 受试者 aw 选取 51 导联, 减幅 57%; 受试者 ay 选取 19 导联, 减幅 84%; 5 名受试者平均从 118 个电极中选取 35 导联,

减幅约为 70%。

由此可见, MCFS 算法在降低导联数目上具有明显的优势, 因此可以通过 MCFS 算法选取导联数目, 节约建立模型时间, 有利于在线系统的使用。图 4

示出了所选导联的分布情况。对于 Dataset 1, 圆圈的颜色表示导联被选中的次数, 其中 α 表示该圆圈内导联被选中的次数; 对于 Dataset 2, β 表示该圆圈内导联被选中的次数。如图 4(a) 所示, 选择的导联 ($\alpha \geq 2$)

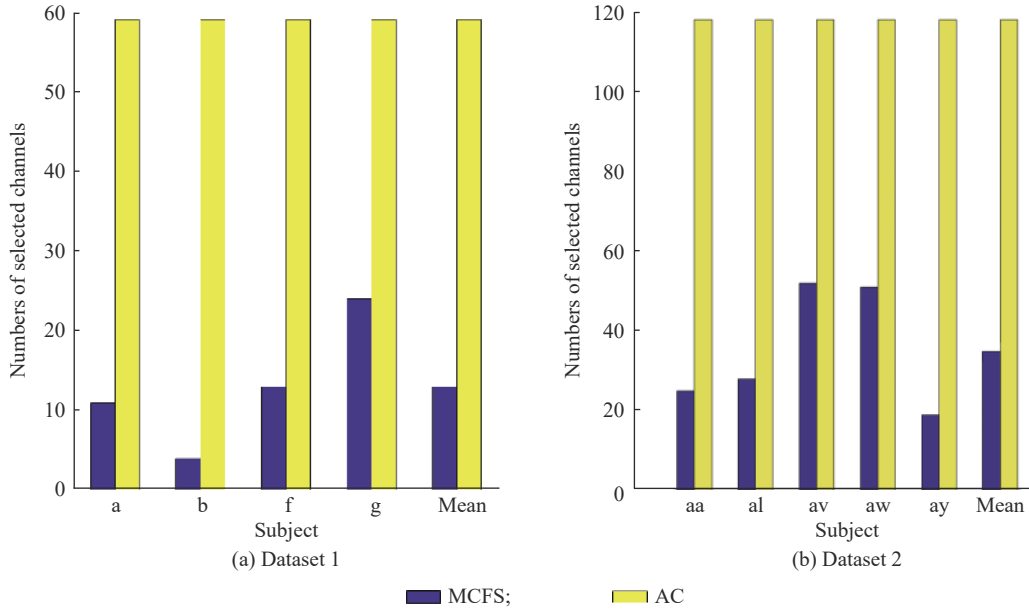


图 3 各个受试者在 AC 和 MCFS 算法下的导联选择数

Fig. 3 Numbers of channels selected by each subject under AC and MCFS algorithm

表 2 各个受试者在 AC 和 MCFS 算法下的导联选择数

Table 2 Numbers of channels selected by each subject under AC and MCFS algorithm

| Method | Dataset 1 | | | | | Datasets 2 | | | | | |
|--------|-----------|----|----|----|------|------------|-----|-----|-----|-----|------|
| | a | b | f | g | Mean | aa | al | av | aw | ay | Mean |
| AC | 59 | 59 | 59 | 59 | 59 | 118 | 118 | 118 | 118 | 118 | 118 |
| MCFS | 11 | 4 | 13 | 24 | 13 | 25 | 28 | 52 | 51 | 19 | 35 |

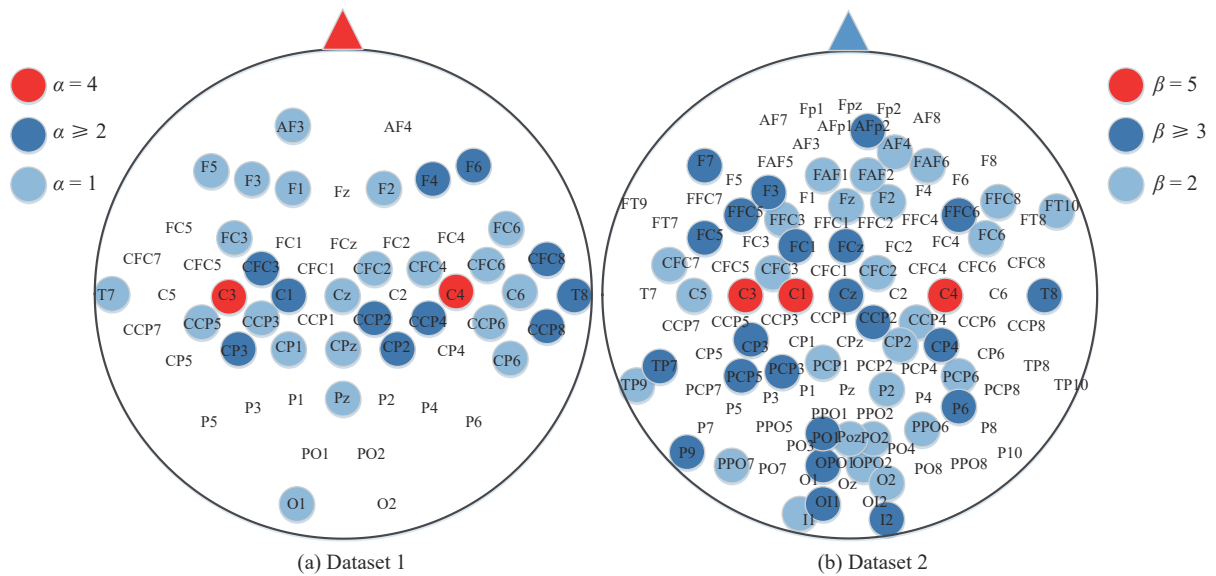


图 4 Dataset 1 和 Dataset 2 的导联选择分布图

Fig. 4 Distribution of selected channels for dataset 1 and dataset 2

主要分布在运动感觉区,符合生理学的认知,对于图 4(b)而言,选择的导联($\beta \geq 3$)也仍然符合这一生理学认知。不仅如此,从图 4 中也可以看出,由于受试者之间的差异,每个受试者所选择的导联仍有一定的区别,因此也会出现导联相对分散的情况。

3.3 导联优化算法

将 MCFS 算法与其他 3 种算法进行比较,结果如表 3 和表 4 所示。其中 AC 算法将所有导联采集的脑电数据用于运动想象分类任务中。3C 算法仅将 3 个通道(即 C3、Cz 和 C4)所采集的 EEG 数据应用于运动想象分类任务中,可以减少因准备时间不足而导致精度降低的问题。CSP-rank 算法是基于 CSP 滤波器排序的导联选择算法,该算法首先根据滤波器系数绝对值对滤波器分类排序,然后从两个滤波器中依次选择次最大滤波器系数值,如遇到重复导联,则在该滤波器下重新选择次最大滤波器系数值,直至选到新的导联。

从表 3 和表 4 可以看出,无论是单个受试者的分类精度还是不同数据集的平均分类精度,MCFS 算法都有最好的表现。

表 3 Dataset 1 在不同导联选择算法下的分类精度

Table 3 Classification accuracy of dataset 1 with different algorithms

| Subject | Accuracy/% | | | |
|----------|------------|---------|-----------|---------|
| | AC | 3C | CSP-rank | MCFS |
| a | 68 | 75 | 79.75 | 83 |
| b | 49.5 | 56 | 53.55 | 63 |
| f | 62.5 | 67.5 | 74.1 | 83.5 |
| g | 83.5 | 87 | 81 | 94.5 |
| Mean±std | 65.9±14.1 | 71.4±13 | 72.1±12.7 | 81±13.1 |

表 4 Dataset 2 在不同导联选择算法下的分类精度

Table 4 Classification accuracy of dataset 2 with different algorithms

| Subject | Accuracy/% | | | |
|----------|------------|---------|----------|-----------|
| | AC | 3C | CSP-rank | MCFS |
| aa | 62.14 | 66.43 | 72.36 | 80.71 |
| al | 90 | 89.29 | 92.54 | 97.14 |
| av | 61.43 | 57.85 | 61.61 | 72.86 |
| aw | 84.29 | 78.93 | 83.18 | 93.21 |
| ay | 76.79 | 89.29 | 84.11 | 92.86 |
| Mean±std | 74.9±12.9 | 76.4±14 | 78.8±12 | 87.4±10.2 |

3.4 导联优化算法比较

为了进一步体现 MCFS 导联优化算法的优越性,将 MCFS 算法与同一数据集的其他方法相比较,结果如表 5 所示。文献 [15] 提出的智能多目标粒子群导联选择算法获得了 60.47% 的平均分类精度,相比于 MCFS 算法准确率降低了 20.53%。Qiu 等^[16] 通过改进序列前向搜索导联优化算法 (Improved Sequential Floating Forward Selection, Improved SFFS),将分类精度提高至 67.25%,但是该算法平均精度仍比 MCFS 算法低 13.75%。文献 [12] 提出的基于互信息的导联选择算法(仅选用文章中的 CCS-CSP 算法),得到平均分类精度 78%,比本文算法低 3%。

表 5 Dataset 1 在不同导联选择算法下的分类精度

Table 5 Comparison of classification accuracy with different algorithms for dataset 1

| Methods | Accuracy/% | | | | |
|--------------------------------|------------|------|-------|-------|-------|
| | a | b | f | g | Mean |
| Binary-SDMOPSO ^[15] | 62.31 | 55.5 | 60.07 | 64 | 60.47 |
| Improved SFFS ^[16] | 69 | 63 | 65 | 72 | 67.25 |
| CCS-CSP ^[12] | 80.50 | 60 | 78 | 93.50 | 78 |
| MCFS | 83 | 63 | 83.50 | 94.50 | 81 |

表 6 示出了 MCFS 算法与其他算法在 Dataset 2 上的比较结果。Barachant 等^[17] 提出的基于瑞利熵的方法,得到 78% 的平均分类精度,比 MCFS 算法低 9.4%。Arvaneh 等^[18] 提出空间稀疏共空间模式 (Spatially Sparsed Common Spatial Pattern),提升运动皮层区导联选择的权重,减弱非运动皮层区导联权重,其平均精度比 MCFS 算法低 5.12%。Belwafi 等^[19] 提出对脑电信号实现动态滤波的 WOLA-CSP 算法得到 62.79% 的平均分类精度,与 MCFS 算法相比低 20.11%。Feng 等^[20] 提出基于多频带的 CSP 特征提

表 6 Dataset 2 在不同导联选择算法下的分类精度

Table 6 Comparison of classification accuracy with different algorithms for dataset 2

| Methods | Accuracy/% | | | | | |
|--------------------------|------------|-------|-------|-------|-------|-------|
| | aa | al | av | aw | ay | Mean |
| RD-CSP ^[17] | 74.10 | 98.20 | 59.20 | 77.70 | 80.60 | 78 |
| SCSP ^[18] | 80.71 | 97.14 | 57.14 | 85 | 91.42 | 82.28 |
| WOLA-CSP ^[19] | 66.79 | 96.07 | 52.14 | 71.43 | 50 | 67.29 |
| CSP-R-MF ^[20] | 81.43 | 92.41 | 70 | 83.57 | 85 | 82.48 |
| MCFS | 80.71 | 97.14 | 72.86 | 93.21 | 92.86 | 87.40 |

取算法,通过 Lasso 算法进一步提取有效特征组合,获得 82.48% 的平均分类精度,与 MCFS 算法相比低 4.92%。

综上所述,本文提出的 MCFS 算法对提升模型的分类性能具有良好的效果。

4 结 论

针对运动想象脑-机接口导联优化问题,本文提出一种基于多相关性导联前向搜索(MCFS)算法。该算法可以根据受试者的不同,选择与运动想象任务相关的脑区所对应的导联,从而减少无关脑区导联所采集的脑电数据对分类的干扰。实验结果表明,对两个竞赛数据集运用 MCFS 算法进行导联选择时,不仅仅大幅减少导联数,减少计算量,同时相比传统算法分类精度也得到较大提升。此外,将 MCFS 算法与其他成功的算法相比,仍然可以获得最高的平均分类精度,同时也为基于运动想象的 BCI 系统提供了技术参考。

参考文献:

- [1] JIN J, SELLERS E W, ZHOU S, *et al.* A P300 brain-computer interface based on a modification of the mismatch negativity paradigm[J]. *International Journal of Neural Systems*, 2015, 25(3): 1550011.
- [2] SUN H, ZHANG Y, GLUCKMAN B J, *et al.* Optimal-channel selection algorithms in mental tasks based brain-computer interface[C]// Proceedings of the 2018 8th International Conference on Bioscience, Biochemistry and Bioinformatics. USA: ACM, 2018: 118-123.
- [3] SI Y, WU X, LI F, *et al.* Different decision-making responses occupy different brain networks for information processing: A study based on EEG and TMS[J]. *Cerebral Cortex*, 2019, 29(10): 4119-4129.
- [4] AL-ANI A, AL-SUKKER A. Effect of feature and channel selection on EEG classification[C]// 2006 International Conference of the IEEE Engineering in Medicine and Biology Society. USA: IEEE, 2006: 2171-2174.
- [5] POPESCU F, FAZLI S, BADOWER Y, *et al.* Single trial classification of motor imagination using 6 dry EEG electrodes[J]. *Plos One*, 2007, 2(7): e637.
- [6] 冯建奎, 金晶, 王蓓, 等. 基于 SVM_RFE 的多任务导联选择算法建模[J]. *系统仿真学报*, 2018, 30(12): 4506-4512.
- [7] GE S, WANG R, YU D. Classification of four-class motor imagery employing single-channel electroencephalography [J]. *Plos One*, 2014, 9(6): e98019.
- [8] 张德明, 殷国栋, 金贤建, 等. 基于 CSP 和 SFBS-SFBS 的两级双向脑电导联特征选取方法[J]. *东南大学学报(自然科学版)*, 2019, 49(1): 125-132.
- [9] 陈书立, 李新建, 胡玉霞, 等. 互信息引导下的前向搜索脑-机接口导联选择算法[J]. *计算机应用研究*, 2018, 35(4): 1080-1083, 1087.
- [10] GAUR P, PACHORI R B, WANG H, *et al.* A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry[J]. *Expert Systems with Applications*, 2018, 95: 201-211.
- [11] XU Y, WEI Q, ZHANG H, *et al.* Transfer learning based on regularized common spatial patterns using cosine similarities of spatial filters for motor-imagery BCI[J]. *Journal of Circuits, Systems and Computers*, 2019, 28(7): 1-19.
- [12] JIN J, MIAO Y, DALY I, *et al.* Correlation-based channel selection and regularized feature optimization for MI-based BCI[J]. *Neural Networks*, 2019, 118: 262-270.
- [13] JIAO Y, ZHANG Y, CHEN X, *et al.* Sparse group representation model for motor imagery EEG classification[J]. *IEEE Journal of Biomedical and Health Informatics*, 2018, 23(2): 631-641.
- [14] KOHAVI R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]// 14 th International Joint Conference on Artificial Intelligenc. Montreal, Canada: ACM, 1995: 1137-1145.
- [15] AL MOUBAYED N, HASAN B A S, GAN J Q, *et al.* Binary-SDMOPSO and its application in channel selection for brain-computer interfaces[C]//2010 UK Workshop on Computational Intelligence (UKCI). Colchester, UK: IEEE, 2010: 1-6.
- [16] QIU Z, JIN J, LAM H K. *et al* Improved SFBS method for channel selection in motor imagery based BCI[J]. *Neuro-computing*, 2016, 207: 519-527.
- [17] BARACHANT A, BONNET S. Channel selection procedure using Riemannian distance for BCI applications[C]// 2011 5th International IEEE/EMBS Conference on Neural Engineering. Cancun, Mexico: IEEE, 2011: 348-351.
- [18] ARVANEH M, GUAN C, ANG K K, *et al.* Optimizing the channel selection and classification accuracy in EEG-based BCI[J]. *IEEE Transactions on Biomedical Engineering*, 2011, 58(6): 1865-1873.
- [19] BELWAFI K, ROMAIN O, GANNOUNI S, *et al.* An embedded implementation based on adaptive filter bank for brain-computer interface systems[J]. *Journal Neuroscience Method*, 2018, 305: 1-16.
- [20] FENG J K, JIN J, DALY I. *et al* An optimized channel selection method based on multifrequency CSP-rank for motor imagery-based BCI system[J]. *Computation Intelligence and Neuroscience*, 2019(6): 1-10.

Channel Selection Based on Multi-Correlation Forward Searching Algorithm for MI Classification

YIN Feiyu, JIN Jing, WANG Xingyu

(Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China)

Abstract: Aiming at the shortcoming that there exist too many channels in motor imagery (MI)-based brain-computer interface (BCI) systems, this paper proposes a channel selection algorithm based on multi-correlation forward searching (MCFS) algorithm such that the performance of BCI systems can be improved via the optimized the channel set. First, a forward searching algorithm is performed on the channel set via the training set. Meanwhile, the trust values of three correlation algorithms are updated with the classification accuracy of the validating set. Then, according to the above trust values, the high-quality channel set is selected, the common spatial pattern (CSP) algorithm is adopted to obtain the motor imagery related features, and the classification model is trained by means of support vector machine (SVM) with a linear kernel. Finally, the proposed algorithm is implemented on two datasets (BCI competition IV dataset I and BCI competition III datasets IVa), by which the average classification accuracy of 81% and 87% are achieved, respectively. Moreover, compared with three other common channel selection algorithms, the proposed MCFS algorithm obtains the highest average classification accuracy. These results show that the proposed MCFS algorithm has superior performance and provides a technical reference for the application of MI-based BCI system. .

Key words: brain-computer interface; motor imagery; multi-correlation algorithm; channel selection; common spatial pattern