

心算和想象运动二分类的近红外光谱脑机接口范式重测信度研究^①

马建爱^② 郑燕春 王 玲 汪待发^③

(北京航空航天大学生物与医学工程学院 北京 100191)

摘 要 本文对心算和想象动右手这两个经典范式的二分类问题的重测信度在分类正确率、特征参数以及脑激活强度三个方面进行了研究。针对 10 名被试在长达一个月的时间跨度内进行了三次重复实验。提取氧合血红蛋白均值特征后,利用支持向量机进行分类。结果显示在分类正确率层面,心算和想象动右手范式的组内分类和组间分类正确率均具有较高的重测信度;接着通过组间相关系数,激活通道比和激活通道相同比指标对特征参数和脑激活强度的重测信度进行量化研究,发现其重测信度与分类正确率的重测信度高度相关。本文研究表明对于大部分被试,心算和想象动右手范式具有较高的重测信度,为后续 fNIRS-BCI 的实际应用提供了很好的理论支撑。

关键词 近红外光谱,脑机接口(BCI),重测信度,心算,想象动右手,氧合血红蛋白均值

0 引 言

脑机接口(brain computer interface, BCI)可通过对脑信号的解析实现对外部设备的控制,在神经康复等方面具有重要的应用价值。近年来,近红外光谱技术(functional near-infrared spectroscopy, fNIRS)因其抗电磁干扰、抗运动干扰以及高空间分辨率等优点,在脑网络连接^[1,2]、神经康复^[3,4]、BCI 领域^[5-7]得到了越来越广泛的关注。

在基于 fNIRS 的 BCI 研究中,研究者们多关注运动区和前额区这两个脑功能区。运动区的典型范式是运动想象(包括想象动手指、手臂、脚趾、旋转等)^[7]。前额区的典型范式包括心算和想象听音乐等^[7]。Hwang^[8]等人对 8 种想象任务进行了研究,发现利用氧合血红蛋白(oxygenated hemoglobin, HBO)均值特征可以实现对心算和想象旋转任务 71% 的分类结果。Sitaram^[9]等人通过提取 HBO 和脱氧血红蛋白(deoxygenated hemoglobin, HBR)均值

特征组合,使用 SVM 分类器实现了想象动左右手 73% 的分类正确率。Stangl^[10]等人提取 HBO 均值特征并使用 LDA 分类器实现了对心算和想象运动 65% 的分类正确率。

在这些研究中,fNIRS-BCI 实验均需花数十分钟进行数据采集然后再建立分类模型,耗时较长。如可通过前期采集数据训练得到分类模型,并直接用于后期实验,将有望大大节约时间并提高 fNIRS-BCI 的实用性。而这需要 fNIRS-BCI 范式具有高的重测信度。

重测信度问题在 BCI 领域得到了广泛关注^[5,11,12]。Friedrich^[12]等人在两周内对 9 名被试采集了 4 批数据,然后对 7 种脑电 BCI 范式在不同频率带、不同脑区以及不同时间窗下,对时间相关(去)同步化特征进行了重测信度的研究,采用克朗巴哈系数法对特征进行重测信度量化研究。结果表明上述所提特征在 α 波段具有高的重测信度,而在 β 波段重测信度较低。fNIRS-BCI 相对脑电 BCI 起步较晚,目前对其重测信度的相关研究较少。Power

① 国家自然科学基金(61675013, 31771071)和国家重点研发计划(2016YFF0201002)资助项目
② 女,1992 年生,研究生;研究方向:近红外光谱脑机接口;E-mail: jianaima@buaa.edu.cn
③ 通信作者,E-mail: daifa.wang@buaa.edu.cn
(收稿日期:2017-12-27)

等人^[5]对 10 名被试在一个月内采集了 5 批数据,然后利用斜率特征对心算和休息态在分类正确率层面的重测信度进行了探索。通过分析组内和组间的分类正确率,发现心算和休息态在组内分类结果无显著差异,组间分类结果略低于组内分类。然而,心算任务与休息态的响应本身便存在显著的脑激活强度差异,其重测信度因而相对较好。同时,该文也没有对斜率特征进行定量研究,给出重测信度评价指标。

相对之前对任务态与静息态的研究^[5],本文对两任务态的二分类问题的重测信度进行了研究。由于心算和想象动右手任务是目前 fNIRS-BCI 领域广泛使用的两种任务^[7],且本身已被证明具有较好的二分性^[6,10],故本文对心算和想象动右手这两种经典任务的二分类问题的重测信度开展了研究。首先,对心算和想象动右手范式在三个不同时间阶段开展重复实验得到的二分类正确率的一致性进行分析,包括组内分类和组间分类正确率。其次,利用组间相关系数(intraclass correlations coefficients, ICC)指标对 HBO 均值特征在三阶段的重测信度进行量化研究。再次,利用激活通道比(the quantity of significant activated channels, Rsize)和激活通道相同比(the same location of significant activated channels, Roverlap)指标对大脑激活响应强度在三阶段的一致性进行了分析。

1 实验

1.1 研究对象

本实验共招募了 10 名被试,均为健康男生(25.1 ± 1.7 岁)。他们无心血管疾病、精神疾病、神经性疾病、呼吸疾病、酗酒等,且在实验前三小时不喝咖啡、不吸烟。所有被试视力或者矫正视力正常。本实验由北京航空航天大学伦理委员会批准进行。所有被试均自愿参加并在实验前签署知情同意书。

1.2 实验设备

实验所用设备是多通道的连续波成像系统(NirScan, 慧创, 中国)。该系统使用三波长(分别为 740nm、808nm、850nm)对血红蛋白浓度变化进行检测,采样频率为 13Hz。本研究探头排布如图 1 所

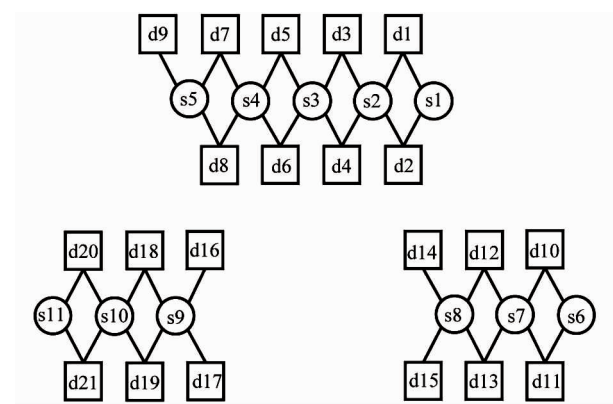
示,11 个发射器和 21 个探测器共形成 37 个通道,覆盖前额区和运动区。

1.3 心算任务

心算任务是指被试在不借助笔、计算器或电脑等外界帮助的情况下,只用大脑对指定数字进行数学运算。本文中,被试需要对给定的三位数连续减去一个给定的个位数(在 5 ~ 9 之间),减的次数尽可能多,例如,583 - 8 = 575, 575 - 8 = 567, 567 - 8 = 559……

1.4 想象动右手任务

在想象动右手任务中,被试需要在大脑中想象大拇指快速地从食指敲到小拇指并重复这个动作。同时需要注意实验过程中手必须保持静止。



圆圈代表发射器,方框代表探测器,实线代表通道;

s1 ~ s5 位于前额区, s6 ~ s8 位于右运动区, s9 ~ s11 位于左运动区

图 1 光极及通道排布

1.5 实验过程

每个被试均会完成 3 个完全相同的阶段实验,这 3 次实验在不同时间进行。其中阶段 1 与阶段 2 间隔 1 周,阶段 2 与阶段 3 间隔 3 周。在每阶段实验中,被试均坐在昏暗的屋子里,看着电脑屏幕,眼睛与电脑屏幕间隔大约 1 米左右且需要完成 6 个实验块(block)。在每个 block 开始前,首先会有 30s 的实验前休息,被试需要保持平静。然后屏幕将从心算和想象动右手两个任务中随机选择一个任务呈现,被试进行 15s 的任务,随后再进行 20s 的休息,一次任务加随后的休息组成一个实验段(trial)。每个 block 包括 10 个 trial 并保证每个任务数量一致,即每阶段结束,每个任务将会有 30 个 trial。单个

block 的实验流程如图 2 所示。

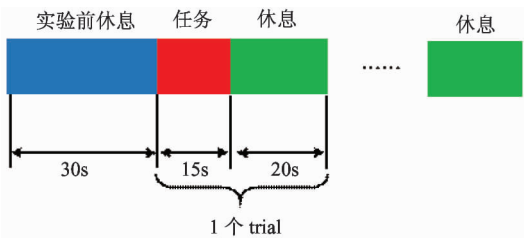


图 2 单个 block 实验流程图

2 数据处理与分析方法

2.1 数据预处理

近红外光谱技术利用近红外光的强穿透能力及其在不同组织成分的吸收光谱差异,采用多个波长的近红外光(光谱),探测脑激活带来的血红蛋白参数(HBO、HBR)变化。在采集信号过程中,近红外光受到各种生理噪声的影响,故先将数据进行 0.01~0.2Hz 的带通滤波,以去除心脏搏动及呼吸的干扰。然后利用修正的 Beer-lambert 定律将数据转化为 HBO 和 HBR 相对变化量如下式:

$$\begin{aligned}\Delta OD^{\lambda_1} &= (\varepsilon_{HbO_2}^{\lambda_1} \Delta C_{HbO_2} + \varepsilon_{Hb}^{\lambda_1} \Delta C_{Hb}) \cdot r \cdot DPF^{\lambda_1} \\ \Delta OD^{\lambda_2} &= (\varepsilon_{HbO_2}^{\lambda_2} \Delta C_{HbO_2} + \varepsilon_{Hb}^{\lambda_2} \Delta C_{Hb}) \cdot r \cdot DPF^{\lambda_2}\end{aligned}\tag{1}$$

其中, ΔOD 为射出光与入射光的光强差, ε_{HbO_2} 和 ε_{Hb} 分别为 HBO 和 HBR 对近红外光的吸收系数, λ_1 和 λ_2 分别代表近红外光的两种波长, r 为探测器与发射器之间的距离, DPF 为差分路径因子,而 ΔC_{HbO_2} 和 ΔC_{Hb} 分别为待求的 HBO 和 HBR 浓度变化量。

2.2 特征提取与分类

在 fNIRS-BCI 领域,研究者们对 fNIRS 信号的均值、斜率、变异性、偏斜度和峭度等特征的可分性已开展了细致的研究^[7,8]。均值特征分类能力较强且使用最为广泛^[8-10],故本文提取 fNIRS 信号(HBO)的均值特征进行后续分类研究。由于 fNIRS 信号存在数秒的血液动力学延迟^[5,8],故本文在特征提取方面加入了时间窗的概念^[5,6,8],即把数据切成不同的时间窗,从不同的时间窗中再提取均值特

征,以尽量减少响应延迟带来的影响。本实验对任务开始时刻后 10~20s 的数据分为 10~15s、10~20s、15~20s 三个时间窗,并对全通道数据提取对应窗的 HBO 均值特征。这样,对于每个任务,共有 37(通道)×3(时间窗)=108 个特征。分类器采用了支持向量机^[7,9,13-15],并使用留一交叉验证法得到分类结果,其中支持向量机使用为台湾大学林智仁等人发布的 libsvm 工具包。

2.3 重测信度评价

本文分 3 个方面对心算和想象动右手两个任务的二分类问题的重测信度进行分析,包括:分类正确率的重测信度,均值特征的重测信度和大脑激活强度的重测信度。

在分类正确率的重测信度分析中,分为组内分类和组间分类。对于组内分类,训练集和测试集均来自同一阶段数据,通过比较三个组内分类结果,可看出分类正确率在三个阶段的重测信度;更严格的组间分类是指,训练集和测试集来自不同阶段数据,若组间分类拥有较高的正确率,则再次说明分类正确率拥有较高的重测信度。

通过对均值特征的重测信度分析可以从更定量的角度探讨分类重测信度产生的原因。本文利用 ICC 指标对均值特征的重测信度进行分析^[16-18]。ICC 指标的定义如式(2)、(3)所示。

$$ICC_{single} = \frac{MS_R - MS_W}{MS_R + (k - 1)MS_W}\tag{2}$$

$$ICC_{average} = \frac{MS_R - MS_W}{MS_R}\tag{3}$$

其中, MS_R 为行变量均方, MS_W 为剩余均方。 ICC_{single} 估计了任意两阶段数据之间的相关度, $ICC_{average}$ 估计了对 k 个阶段数据平均后的可信度^[19]。

本文利用 $Rsize$ ^[18,20] 和 $Roverlap$ ^[17,18,20] 对大脑激活强度在 3 个阶段的重测信度进行量化分析。指标的定义如式(4)、(5)所示。

$$Rsize = 1 - \frac{A1 - A2}{A1 + A2}\tag{4}$$

$$Roverlap = 2 \times Aoverlap / (A1 + A2)\tag{5}$$

其中, $A1, A2$ 分别为两个阶段数据激活的通道数, $Aoverlap$ 为在两个阶段中被共同激活的通道数。

试。

3 结果与讨论

表 1 给出了组内分类结果。3 阶段总体均值为 72.2% ,与目前文献中报告的 fNIRS-BCI 分类正确率(多集中在 65% ~ 75% 之间^[7])相当,且 3 阶段分类正确率均值无显著性差异($p = 0.43$)。可以看出,在组内分类层次,心算和想象动右手两范式具有较高的重测信度。考虑文献报道有些被试本身并不适合做 BCI 实验^[21],而本实验对被试并没有做严格筛选,故本文把实验中执行任务效果较好的被试分为一组,记为 top6(被试 1,3,5,7,8,9),该组组内分类平均结果高达79.2% ;把任务执行较差的被试分为一组,记为 last4(被试 2,4,6,10),该组组内分类平均结果为61.7% 。

表 1 组内分类结果(%)				
被试	阶段 1	阶段 2	阶段 3	平均
1	58.3	78.3	70.0	68.9
2	73.3	71.7	40.0	61.7
3	93.3	93.3	85.0	90.6
4	78.3	61.7	48.3	62.8
5	78.3	75.0	68.3	73.9
6	58.3	53.3	70.0	60.6
7	93.3	66.7	86.7	82.2
8	76.7	83.3	90.0	83.3
9	75.0	83.3	70.0	76.1
10	73.3	65.0	46.7	61.7
平均	75.8	73.2	67.5	72.2
top6	79.2	80.0	78.3	79.2
last4	70.8	62.9	51.3	61.7

表 2 给出了各被试组间分类的平均结果,组间分类平均正确率为65.4% ,虽略低于组内分类结果,但与 Power^[5]分类心算与休息态结果(组间分类结果在 65% 左右)相当。其中,top6 被试的组间分类平均正确率为68.1% ,拥有较好的重测信度,而 last4 被试的平均分类率为61.2% ,可以看出在组内分类表现较好的被试在组间分类上依然高于其他被

表 2 组间分类结果(%)

被试	平均
1	66.6
2	61.7
3	81.7
4	56.4
5	61.8
6	58.3
7	66.2
8	68.4
9	64.1
10	68.3
平均	65.4
top6	68.1
last4	61.2

图 3 展示了组间分类结果,其中横坐标中显示的‘1-3’代表以阶段 1 数据为训练集,以阶段 3 数据为测试集的组间分类结果,其它以此类推。从图中可以看出,top6 被试在任意两阶段配对分类时的二分类结果相当,不存在随时间间隔增大而下降的趋势。而 last4 被试分类结果波动较大。以上结果再次表明在组间分类层次,心算和想象动右手两范式对于 top6 被试具有较高的重测信度。

表 3 给出了 top6 被试和 last4 被试心算和想象动右手任务的 HBO 均值特征以及大脑激活强度的重测信度指标。可以明显看出,top6 被试的各项重测信度量化指标处于适中以上的水平^[22],且均高于 last4 被试,这与组内分类和组间分类结果在这两组被试中的差异高度一致,同时也再次证实 ICC,Rsize 和 Roverlap 可以作为度量任务执行时一致性和有效性的指标。

图 4 则详细给出了被试 3 和被试 6 在心算任务结束时刻的脑激活图(t-map 图)。从图中可以明显看出,分类效果和重测信度均较高的被试 3 的脑激活强度远远高于被试 6。

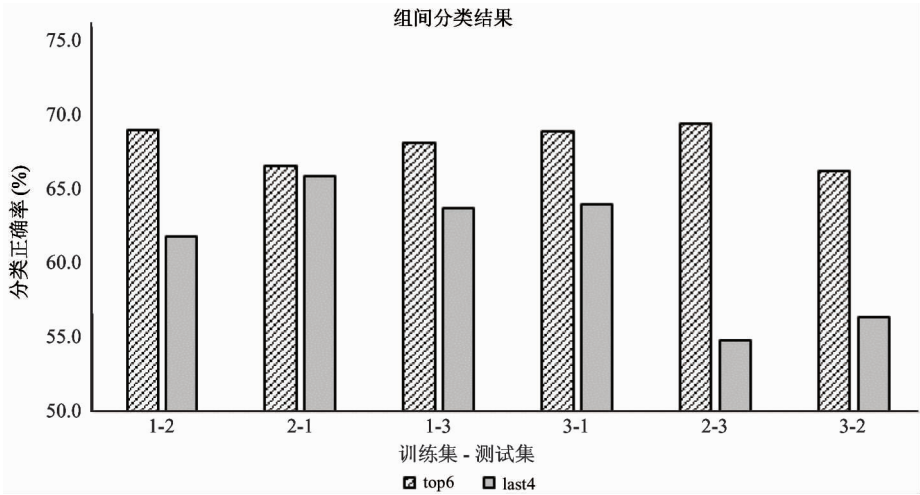


图 3 组间分类结果

表 3 HBO 均值特征及脑激活强度的重测信度指标

	被试	ICC_{single}	$ICC_{average}$	$Rsize$	$Roverlap$
心算	top6	0.46	0.65	0.81	0.53
	last4	0.06	0.12	0.50	0.21
	平均	0.29	0.48	0.69	0.41
想象动右手	top6	0.40	0.57	0.70	0.44
	last4	0.30	0.44	0.39	0.05
	平均	0.33	0.51	0.57	0.28

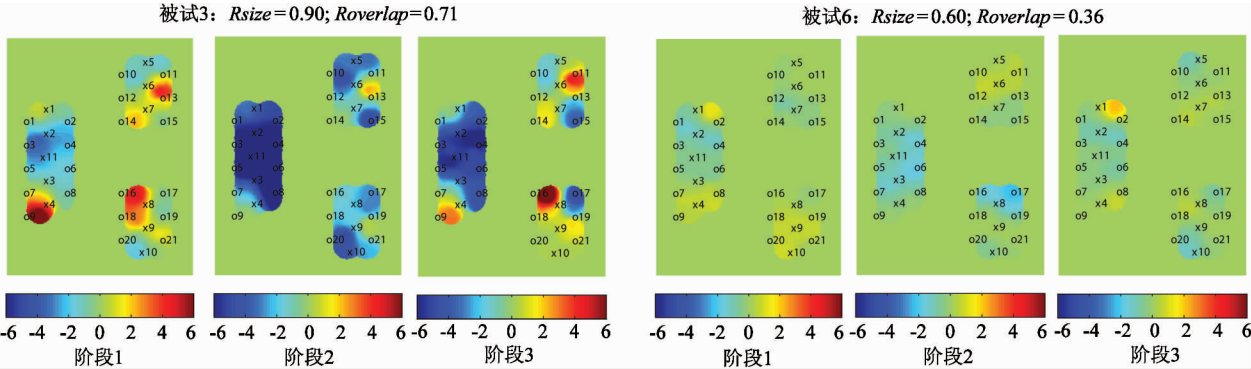


图 4 被试 3 和被试 6 在心算任务 15s 时脑激活图

4 结 论

范式具有较高的重测信度是 BCI 运用到实际环境中非常重要的条件。本文证实了心算和想象动右手范式在分类正确率方面,呈现出较高的可分性和较好的重测信度;在均值特征和脑激活强度方面,自身可分性比较好的被试处于适中以上的水平,且

明显高于可分性较差的被试。本文的结果表明,心算和想象动右手范式对于本身可分性较好的被试具有较高的重测信度,该结果为利用历史数据为训练集、直接用于在线 BCI 的使用提供了很好的理论基础。在后续的研究中,将搭建在线 BCI 系统,运用累积的历史数据直接分类后续数据,以验证其在线实际应用能力。

参考文献

- [1] Lu C M, Zhang Y J, Biswal B B, et al. Use of fNIRS to assess resting state functional connectivity[J]. *Journal of Neuroscience Methods*, 2010, 186(2): 242-249
- [2] Niu H, Wang J, Zhao T, et al. Revealing topological organization of human brain functional networks with resting-state functional near infrared spectroscopy[J]. *Plos One*, 2012, 7(9): e45771
- [3] Miyai I, Yagura H, Oda I, et al. Premotor cortex is involved in restoration of gait in stroke[J]. *Annals of Neurology*, 2002, 52(2): 188-194
- [4] Abdalmalak A, Milej D, Norton L, et al. Single-session communication with a locked-in patient by functional near-infrared spectroscopy[J]. *Neurophotonics*, 2017, 4(4): 040501
- [5] Power S D, Kushki A, Chau T. Intersession consistency of single-trial classification of the prefrontal response to mental arithmetic and the no-control state by NIRS[J]. *Plos One*, 2012, 7(7): e37791
- [6] Hong K S, Naseer N, Kim Y H. Classification of prefrontal and motor cortex signals for three-class fNIRS-BCI[J]. *Neuroscience Letters*, 2015, 587: 87-92
- [7] Naseer N, Hong K S. fNIRS-based brain-computer interfaces: a review[J]. *Frontiers in Human Neuroscience*, 2015, 9
- [8] Hwang H J, Lim J H, Kim D W, et al. Evaluation of various mental task combinations for near-infrared spectroscopy-based brain-computer interfaces[J]. *Journal of Biomedical Optics*, 2014, 19(7): 077005-077005
- [9] Sitaram R, Zhang H, Guan C, et al. Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface[J]. *Neuroimage*, 2007, 34(4): 1416-1427
- [10] Stangl M, Bauernfeind G, Kurzmann J, et al. A haemodynamic brain-computer interface based on real-time classification of near infrared spectroscopy signals during motor imagery and mental arithmetic[J]. *Journal of Near Infrared Spectroscopy*, 2013, 21(3): 157-171
- [11] Tonin L, Leeb R, Sobolewski A, et al. An online EEG BCI based on covert visuospatial attention in absence of exogenous stimulation[J]. *Journal of Neural Engineering*, 2013, 10(5): 056007
- [12] Friedrich E V C, Scherer R, Neuper C. Stability of event-related (de-) synchronization during brain: computer interface-relevant mental tasks[J]. *Clinical Neurophysiology*, 2013, 124(1): 61-69
- [13] Naseer N, Hong M J, Hong K S. Online binary decision decoding using functional near-infrared spectroscopy for the development of brain: computer interface[J]. *Experimental Brain Research*, 2014, 232(2): 555-564
- [14] Hu X S, Hong K S, Ge S S. fNIRS-based online deception decoding[J]. *Journal of Neural Engineering*, 2012, 9(2): 026012
- [15] Hai N T, Cuong N Q, Khoa T Q D, et al. Temporal hemodynamic classification of two hands tapping using functional near-infrared spectroscopy[J]. *Frontiers in Human Neuroscience*, 2013, 7
- [16] Zuo X N, Di Martino A, Kelly C, et al. The oscillating brain: complex and reliable[J]. *Neuroimage*, 2010, 49(2): 1432-1445
- [17] Blasi A, Lloyd-Fox S, Johnson M H, et al. Test - retest reliability of functional near infrared spectroscopy in infants[J]. *Neurophotonics*, 2014, 1(2): 025005-025005
- [18] Zhang H, Zhang Y J, Duan L, et al. Is resting-state functional connectivity revealed by functional near-infrared spectroscopy test-retest reliable? [J]. *Journal of Biomedical Optics*, 2011, 16(6): 067008-067008-8
- [19] Johnstone T, Somerville L H, Alexander A L, et al. Stability of amygdala BOLD response to fearful faces over multiple scan sessions[J]. *Neuroimage*, 2005, 25(4): 1112-1123
- [20] Plichta M M, Herrmann M J, Baehne C G, et al. Event-related functional near-infrared spectroscopy (fNIRS): are the measurements reliable? [J]. *Neuroimage*, 2006, 31(1): 116-124
- [21] Allison B Z, Neuper C. Could Anyone Use A BCI? [M]. London: Springer, 2010. 35-54
- [22] Cicchetti D V, Sparrow S A. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior[J]. *American Journal of Mental Deficiency*, 1981. 86(2), 127-137

The test-retest reliability of binary classification between mental arithmetic and motor imagery in functional near-infrared spectroscopy brain computer interface

Ma Jian' ai, Zheng Yanchun, Wang Ling, Wang Daifa

(School of Biological Science and Medical Engineering, Beihang University, Beijing 100191)

Abstract

In this paper, two classical paradigms, mental arithmetic and right-hand motor imagery are utilized in the fNIRS-BCI binary classification and the test-retest reliability is investigated in terms of classification accuracy, classification feature and brain activation response strength. Three repeated experiments sessions are conducted on 10 male subjects over a month span and use the mean value of oxygenated hemoglobin (HBO) as classification feature and SVM as classification method. Results show that in terms of classification accuracy, intra-session and inter-session classification accuracies have good test-retest reliability. Then the intraclass correlations coefficients, the quantity of significant activated channels and the same location of significant activated channels are used to evaluate the test-retest reliability of mean features and brain activation response strength. The results show that the test-retest reliability of the classification accuracy is highly consistent with the reliability of mean features and brain activation response strength for all subjects. This study show that for the majority of subjects, mental arithmetic and right-hand motor imagery paradigm have a good test-retest reliability, which provides a useful theoretical support for the real application of fNIRS-BCI.

Key words: functional near-infrared spectroscopy, brain computer interface (BCI), test-retest reliability, mental arithmetic, right-hand motor imagery, mean oxygenated hemoglobin