

基于感兴趣脑区 LASSO-Granger 因果关系的脑电特征提取算法

余青山* 陈希豪 高发荣 罗志增

(杭州电子科技大学机器人研究所 杭州 310018)

摘 要: 该文将脑功能网络引入到脑电特征提取的研究中, 提出一种基于感兴趣脑区 LASSO-Granger 因果关系的新方法, 克服了当前基于孤立脑区的研究方法的不足。先利用主成分分析提取各感兴趣区的最大主成分, 然后计算它们之间的 LASSO-Granger 因果度量, 并将其作为特征向量, 最后输入支持向量机分类器, 对 BCI Competition IV dataset 1 中的 4 组数据进行分类识别。结果表明, 基于感兴趣脑区间 LASSO-Granger 因果关系分析和支持向量机分类器的方法对不同的运动想象任务识别率较高, 提供了新的研究思路。

关键词: 脑功能网络; LASSO-Granger; 感兴趣脑区; 特征提取

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2016)05-1266-05

DOI: 10.11999/JEIT150851

Feature Extraction of Electroencephalography Based on LASSO-Granger Causality Between Brain Region of Interest

SHE Qingshan CHEN Xihao GAO Farong LUO Zhizeng

(Institute of Intelligent Control and Robotics, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: Brain functional network is introduced to feature extraction of ElectroEncephaloGraphy (EEG), and a novel method is proposed based on Least Absolute Shrinkage and Selection Operator (LASSO)-Granger causality between Region Of Interest (ROI) in the brain, in order to overcome the inherent deficiencies of research methods based on isolated brain region. Firstly, the maximum principal component of ROIs is extracted by Principal Component Analysis (PCA), and then causality values between ROIs are calculated by LASSO-Granger. Finally, the values are used as the input vector for Support Vector Machine (SVM), and then four datasets of BCI Competition IV Dataset 1 are used for classification. Experimental results show that different motor imagery tasks are successfully identified by the method of SVM classifier combined with feature extraction which is based on LASSO-Granger causality between the brain region of interest (ROIs). This method provides a new idea for the study of extracting EEG features.

Key words: Brain functional network; LASSO-Granger; Region Of Interest (ROI); Feature extraction

1 引言

随着对人脑结构和功能的不断探索, 越来越多的研究者意识到脑功能网络对运动想象 (Motor Imagery, MI) 研究的重要性。MI 是指仅仅在大脑中反复进行运动模拟训练的一种心理作业, 在神经科学和临床医学上扮演越来越重要的角色。即便是一项非常简单的认知任务也需要多个不同的功能区域相互作用才能发挥功能^[1,2], 当人们在进行 MI 时,

大脑的主运动区(M1)、辅助运动区(SMA)、感觉皮层(S1)等感兴趣区(Region Of Interest, ROI)被激活, 并且不同 ROI 之间会产生相应的脑功能网络效应^[3]。

研究者用相关性、互信息、欧氏距离等特征来表征脑功能网络效应, 但这些特征仅仅只描述了脑区之间的连接强度, 不能体现出脑区之间的连接方向。近年来人们将推测变量间信息因果流向(cause flow)的 Granger 因果分析引入脑功能研究^[4], 以期同时计算出连接的方向和强度。研究者一般通过对多通道的脑电^[5,6](ElectroEncephaloGraphy, EEG)信号进行分析来研究脑区间的连接效应, 并且由于 EEG 信号具有高时间分辨率, 能在很短的时间内记录到大量数据, 有利于后续的 Granger 因果分析^[7]。Granger 因果分析在脑功能的应用中增长迅速, 如

收稿日期: 2015-07-16; 改回日期: 2016-01-29; 网络出版: 2016-03-29

*通信作者: 余青山 qsshe@hdu.edu.cn

基金项目: 国家自然科学基金(61201302, 61172134), 国家留学基金(201308330297), 浙江省自然科学基金(LY15F010009)

Foundation Items: The National Natural Science Foundation of China (61201302, 61172134), State Scholarship Fund of China (201308330297), Natural Science Foundation of Zhejiang Province (LY15F010009)

对癫痫发作期间头皮脑电信号进行致痫灶定侧^[8],在全身麻醉的情况下区分清醒状态和昏迷状态^[9]等。

已有的 EEG 特征提取方法大多是基于孤立脑区的信号研究,未考虑到脑区之间的功能联通性关系,因此存在一定的局限性和不足。本文以 Granger 因果理论为基础,利用主成分分析(Principal Component Analysis, PCA)提取大脑的 ROI 内的最大主成分代表 ROI 的信息,引用 LASSO(Least Absolute Shrinkage and Selection Operator)算法^[10]求解 ROI 之间的因果关系,分析 ROI 之间的因果流向,提取因果度量作为特征向量,并以支持向量机(Support Vector Machine, SVM)作为分类器。采用 BCI Competition IV dataset 1 数据进行实验验证,同一受试者进行不同运动想象任务时,本文方法提取的脑电特征都具有较为明显的区分度,取得了较好的识别结果,验证了所提方法的有效性。

2 实验数据

本文采用 BCI Competition IV dataset 1^[11]进行实验分析。实验范式为 7 名健康受试者(记为 A~F)面对电脑屏幕的相应提示执行左手、右手、脚运动想象中的两类,其中 2 名受试者选择左手和脚运动想象,另外 5 名受试者选择左手和右手运动想象。实验得到了 7 组标定数据,每组数据均包含了 59 个通道的 EEG 信号,采样频率为 100 Hz,带通滤波范围是 0.05~200 Hz。值得注意的是,该竞赛数据中 Data Set 1 包含实际数据和人工生成数据,本文选用受试者 A, B, F, G 提供的实际 EEG 数据。

对于标定数据,每个受试者执行 200 次实验,单次实验持续 8 s。首先给受试者 2 s 安静准备时间,同时屏幕中心显示十字符号;第 2 s 结束时,显示一个向左或向右或向下方向的箭头,受试者按照箭头的指向执行相应的运动想象任务,持续时间为 4 s;紧接着是 2 s 黑屏,表示一次实验结束。提示“Stop”停止运动想象,安静地准备 1.5~8 s 进行下一次的实验。

3 方法及实验结果分析

当人在进行运动想象的时候,主运动区、辅助运动区、感觉皮层等 ROI 被激活,每个 ROI 包含多个脑电通道,需要对其进行整合提取。本文采用 PCA 提取大脑各 ROI 的最大主成分,之后采用 Granger 因果模型,对最大主成分进行因果分析。本文选用 LASSO 算法构建因果模型,原因是该模型会自动地选择一部分变量构造模型,剔除那些对因变量影响较小的自变量,实现变量的可选择性,从而提高模型的预测精度。

3.1 LASSO-Granger

设 $X_1(t)$ 和 $X_2(t)$ 是两个时间变量序列,若联合 $X_1(t)$ 和 $X_2(t)$ 过去的信息对 $X_1(t)$ 预测比仅使用 $X_1(t)$ 过去的信息更加精确,则称 $X_2(t)$ 对 $X_1(t)$ 具有 Granger 因果关系。 $X_1(t)$ 和 $X_2(t)$ 的自回归模型如式(1)、式(2)所示。

$$X_{1,t} = \sum_{j=1}^m a_{11,j} X_{1,t-j} + \varepsilon_{1,t} \quad (1)$$

$$X_{2,t} = \sum_{j=1}^m a_{22,j} X_{2,t-j} + \varepsilon_{2,t} \quad (2)$$

联合回归模型如式(3)、式(4)所示。

$$X_{1,t} = \sum_{j=1}^m a_{11,j} X_{1,t-j} + \sum_{j=1}^m a_{12,j} X_{2,t-j} + \eta_{1,t} \quad (3)$$

$$X_{2,t} = \sum_{j=1}^m a_{21,j} X_{1,t-j} + \sum_{j=1}^m a_{22,j} X_{2,t-j} + \eta_{2,t} \quad (4)$$

式中, $t = 0, 1, \dots, N$, m 为自回归阶数,矩阵 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

为回归系数, ε_i 和 η_i 为回归估计残差,互不相关,均值为 0,方差分别为 $\sigma_{\varepsilon_i}^2$ 和 $\sigma_{\eta_i}^2$ ($i = 1, 2$), η_1 和 η_2 的协方差定义为

$$\sigma_{\eta_1\eta_2} = \text{cov}(\eta_1, \eta_2) \quad (5)$$

如果在相同统计条件下, $\sigma_{\eta_1}^2$ 的值小于 $\sigma_{\varepsilon_1}^2$, 则称 $X_2(t)$ 对 $X_1(t)$ 有因果效应。此时,式(3)对 $X_1(t)$ 的估计比式(1)更精确。如果 $\sigma_{\eta_1}^2 = \sigma_{\varepsilon_1}^2$, 则称 $X_2(t)$ 对 $X_1(t)$ 没有因果效应。此时,式(1)等于式(3)。

X_2 对 X_1 的因果度量为

$$F_{X_2 \rightarrow X_1} = \ln(\sigma_{\varepsilon_1}^2 / \sigma_{\eta_1}^2) \quad (6)$$

相似地, X_1 对 X_2 的因果为

$$F_{X_1 \rightarrow X_2} = \ln(\sigma_{\varepsilon_2}^2 / \sigma_{\eta_2}^2) \quad (7)$$

本文在求解 Granger 因果模型时引入 LASSO 算法^[10],称为 LASSO-Granger。这种算法通过构造一个惩罚函数获得一个精炼的模型,最终确定一些指标的系数为零,其巨大的优势在于它所构造的模型是稀疏的,它会自动地选择一部分变量构造模型,剔除那些对因变量影响较小的自变量,实现变量的可选择性,从而提高模型的解释性和预测精度。假设时间序列的长度为 T ,对每一组时间序列 $\mathbf{X}^{(i)}$, LASSO 模型为

$$\min_{\{\mathbf{a}_i\}} \sum_{t=L+1}^T \left\| \mathbf{X}_t^{(i)} - \sum_{j=1}^P \mathbf{a}_{i,j}^T \mathbf{X}_{t-L}^{(j)} \right\|_2^2 + \lambda \|\mathbf{a}_i\|_1 \quad (8)$$

其中, $\mathbf{X}_{t-L}^{(j)}$ 是一组延迟时间序列, L 是滞后阶数, $\mathbf{X}_{t-L}^{(j)} = [\mathbf{X}_{t-L}^{(j)}, \dots, \mathbf{X}_{t-1}^{(j)}]$, $\mathbf{a}_{i,j}$ 表示时间序列 j 对时间序列 i 的影响因子, λ 是惩罚系数,决定 $\mathbf{a}_i$ 的稀疏性,通过梯度下降法进行求解。

3.2 特征提取

特征提取包括以下两步:

步骤 1 依据布罗德曼分区系统(Brodmann Area)^[12]对脑电极划分大脑的感兴趣区域。左半球区域: LSMA(F1,F3), LM1(Fc3), LS1(Cp1, Cp3, C1, C3), LPMD(Fc1), LIPL(P1,P3); 右半球区域: RSMA(F2,F4), RM1(Fc4), RS1(Cp2, Cp4, C2, C4), RPMD(Fc2), RIPL(P2,P4)。

步骤 2 利用 PCA 提取各个感兴趣区域的最大主成分, 计算它们之间的 Granger 因果度量作为特征值。为了使特征向量维数不至于过大, 这里选取 $F_{LSMA \rightarrow LIPL}$, $F_{RSMA \rightarrow RIPL}$, $F_{LPMD \rightarrow RPMD}$, $F_{RPMD \rightarrow LPMD}$, $F_{LS1 \rightarrow LSMA}$, $F_{RS1 \rightarrow RSMA}$, $F_{LSMA \rightarrow RM1}$, $F_{RSMA \rightarrow LM1}$ (这里 L 表示左(Left), R 表示右(Right))8 个因果度量值组成 8 维特征向量。

3.3 实验结果与分析

本次仿真实验的环境: 硬件平台为宏基笔记本电脑, CPU 为酷睿 i3, RAM 为 4 GB, Windows 7 系统, 软件运行平台为 Matlab 2014a。

3.3.1 因果流向图分析 计算 ROI 之间的因果度量, 受篇幅所限, 仅列出左手运动想象的因果度量, 如

表 1 所示。表中带“*”的值经过 F 检测之后得到的 P 值(P-Value)小于 0.05, 即表示该值具有显著性意义, 从而构建两个区域之间的因果流, 用箭头表示, 如图 1 所示。

由图 1 所知, 在进行左手、右手以及脚的运动想象任务时, 感兴趣区具有比较丰富的因果连接效应。在进行左手运动想象时, 流向大脑右侧区域的信息流多于左侧, 尤其是 RM1, RS1 等与其他区域之间的效应比较明显; 在进行右手运动想象时, 流向大脑左侧区域的信息流多于右侧, 尤其是 LSMA, LM1 等与其他区域的效应比较明显; 而在进行脚运动想象时, 各区域间的双侧因果效应比较突出。根据这些特性, 从因果流向图的角度, 可比较直观地区分出不同类别的运动想象任务。文献[13]对运动想象任务时大脑区域之间的因果关系进行了研究, 其实验结果显示在进行左手运动想象, 大脑右侧区域的信息交流多于左侧, 在进行右手运动想象, 大脑左侧区域的信息交流多于右侧, 与本文相印证。

3.3.2 滞后阶数 L 及惩罚参数 λ 的选取 本文采用局部最小互信息方法^[14]对 L 进行选取, L 从 1 开始取, 求时间序列 $\mathbf{X}(t)$ 和 $\mathbf{X}(t+L)$ 之间的互信息值,

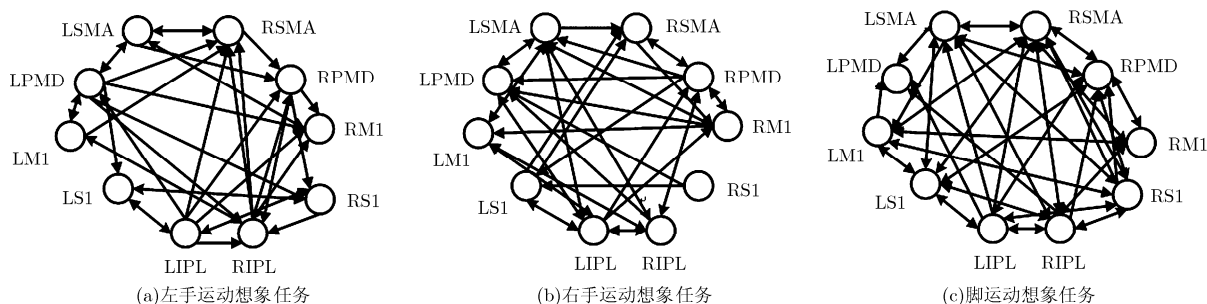


图1 Granger 因果关系流向图

表 1 左手运动想象因果度量

	原因感兴趣区									
	LSMA	LIPL	RSMA	RIPL	LM1	LPMD	RPMD	RM1	LS1	RS1
结果感兴趣区	LSMA	\	0.0127*	0.0065*	0.0045*	0.0031	0.0107*	0.0033	0.0025	0.0031
	LIPL	0.0026	\	0	0.0150*	0.0125*	0.0106*	0.0033	0.0022	0.0130*
	RSMA	0.0179*	0.0144*	\	0.0051*	0.0055*	0.0141*	0.0032	0.0020	0.0011
	RIPL	0.0020	0.0242*	0.0035	\	0.0185*	0.0145*	0.0151*	0.0148*	0.0040
	LM1	0.0024	0.0012	0.0016	0.0089*	\	0.0096*	0.0050*	0.0012	0.0108
	LPMD	0.0105*	0.0071*	0.0022	0.0036	0.0054*	\	0.0020	0.0016	0.0027
	RPMD	0.0083*	0.0156*	0.0042*	0.0073*	0.0080*	0.0111*	\	0.0026	0
	RM1	0.0041*	0.0212*	0.0033	0.0241*	0.0147*	0.0139*	0.0130*	\	0.0012
	LS1	0.0023	0.0293*	0.0019	0.0275*	0.0186*	0.0148*	0.0077*	0.0086*	\
	RS1	0.0021	0.0270*	0.0029	0.0380*	0.0159*	0.0131*	0.0116*	0.0128	0.0074*

取第 1 个最小互信息值出现时对应的 L 作为滞后阶数。由图 2 可知, 当 $L = 3$ 时, 互信息值第 1 次出现最小值, 故取 $L = 3$ 。

确定好 L 后, 再利用交叉验证方法选择最优的 λ 值。 λ 取值范围为交叉验证均方误差最小值对应的 $\lambda_{\min}$ 到标准误差范围内的最大 $\lambda_{\max}$ 。如图 3 所示, 左边虚线对应的横坐标值表示 $\lambda_{\min}$, 右边虚线对应的横坐标值表示 $\lambda_{\max}$ 。图 3 的横坐标值是对 λ 取以 e 为底的对数值。

3.3.3 不同运动想象任务下的特征值 任意选取受试者在不同运动想象任务时的几组 8 维脑电特征向量, 分析比较差异性, 如图 4 所示。图中, 纵坐标表示特征值的大小, 横坐标表示特征值的编号, 不同的线型表示不同运动想象任务对应的特征向量:

“+”表示左手运动想象, “*”表示脚运动想象, “o”表示右手运动想象。

实验数据对应的特征值较容易区分, 而且特征向量的变化趋势具有一定的相似性, 这表明同一受试者进行不同运动想象任务时, 特征值区分度较大。

3.3.4 分类结果 本文分类器采用支持向量机, 以 65%/35% 的比例将数据集随机分成训练集与测试集, 采用 5 倍交叉验证对参数进行择优和网格搜索法寻找最佳参数, 利用最佳参数对训练集进行训练

以建立分类模型, 最后对测试集进行预测。重复上述分类过程 30 次。

采用盒形图(box plot)对分类结果进行统计分析, 如图 5。其中, 纵坐标表示分类准确率, 横坐标表示受试者的编号。从图中的中位数、四分位数以及四分位数间距可以看出, 采用 SVM 时, 中位数均超过 0.84, 这说明有一半以上的分类结果大于 84%。

为了进一步验证所述方法的有效性, 将本文方法与文献[15]中采用的公共空间模式(CSP)、滤波器组共空间模式(FBCSP)、时-频域滤波网络(OSSFN)方法以及文献[16]采用的多元经验模态分解(MEMD)方法进行对比, 两篇文献中都用了该数据, 结果如表 2 所示。在比较受试者 A 和 B 的分类准确率时, 本文方法的分类效果要好于其他方法, 尤其比 CSP 有较明显的提高; 而对于 F 和 G, 本文方法也表现出了较高的识别率, 在平均分类正确率方面高于 CSP 和 FBCSP 方法, 但 G 略低于 OSSFN 和 MEMD 方法, F 略低于 OSSFN 方法。这说明不同受试者的差异对分类结果有一定的影响。本文方法能够比较正确地区分出两类运动想象任务, 体现了本文方法在 EEG 特征提取上的有效性。

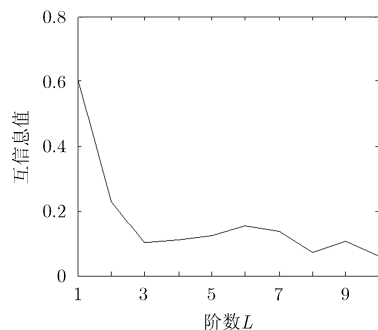


图 2 L 与互信息值的关系

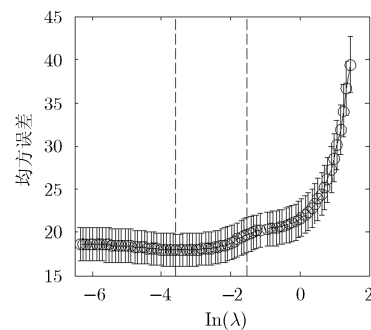


图 3 λ 与均方误差间的关系

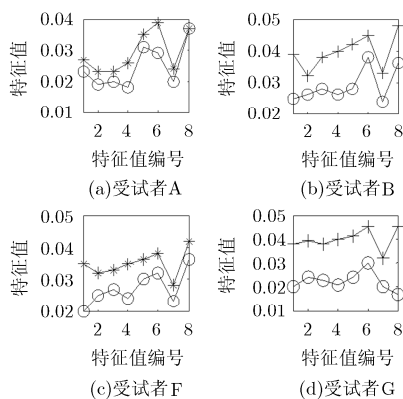


图 4 不同运动想象任务下的特征向量比较

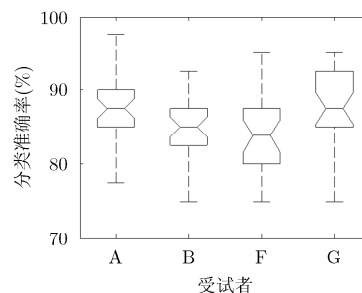


图 5 不同受试者的分类准确率

表2 Data2 中两类运动想象任务的分类正确率(%)比较

受试者	CSP	FBCSP	OSSFN	MEMD	本文
A	67.1	66.5	89.8	85.7	91.7
B	76.6	87.3	86.8	73.9	91.7
F	65.3	87.2	93.0	77.8	87.8
G	74.6	82.5	92.4	91.9	91.0

4 结论

本文基于 LASSO-Granger 因果关系提出了一种分析脑功能连接效应的新方法,应用于脑电的特征提取与分类。本文主要对不同运动想象的脑功能网络效应进行了研究,先用 PCA 提取大脑各个激活区域的最大主成分信息,再运用 LASSO-Granger 因果关系求出激活区域之间的因果度量,提取出 8 维特征向量。最后,通过 SVM 进行分类识别,并与其它特征提取方法进行了对比,结果表明本文方法能较好地对脑电信号进行识别,因此,本文方法可以作为 EEG 数据信号分析中一种有效的特征处理方法。

参考文献

- ARBONELL F, NAGANO-SAITO A, LEYTON M, *et al.* Dopamine precursor depletion impairs structure and efficiency of resting state brain functional networks[J]. *Neuropharmacology*, 2014, 84: 90-100.
- 闫铮, 高小榕, 应俊. 基于认知功能连接的信息流增益计算方法及应用[J]. 电子与信息学报, 2014, 36(11): 2756-2761. doi: 10.3724/SP.J.1146.2013.02019.
- YAN Zheng, GAO Xiaorong, and YING Jun. The flow gain methods and applications based on cognition functional connectivity[J]. *Journal of Electronics & Information Technology*, 2014, 36(11): 2756-2761. doi: 10.3724/SP.J.1146.2013.02019.
- GRABBERR L, JOLA C, BERRA G, *et al.* Motor imagery training improves precision of an upper limb movement in patients with hemiparesis[J]. *NeuroRehabilitation*, 2015, 37(2): 263-271.
- DING M, CHEN Y, and BRESSLER S L. Granger causality: basic theory and application to neuroscience[J]. *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, 2006, 17: 437-460.
- 吴明权, 李海峰, 马琳. 单通道脑电信号中眼电干扰的自动分离方法[J]. 电子与信息学报, 2015, 37(2): 367-372. doi: 10.11999/JEIT140602.
- WU Mingquan, LI Haifeng, and MA Lin. Automatic electrooculogram separation method for single channel electroencephalogram signals[J]. *Journal of Electronics & Information Technology*, 2015, 37(2): 367-372. doi: 10.11999/JEIT140602.
- 吕俊, 谢胜利, 章晋龙. 脑-机接口中基于 ERS/ERD 的自适应空间滤波算法[J]. 电子与信息学报, 2009, 31(2): 314-318.
- LÜ Jun, XIE Shengli, and ZHANG Jinlong. Adaptive spatial filter based on ERD/ERS for brain-computer interfaces[J]. *Journal of Electronics & Information Technology*, 2009, 31(2): 314-318.
- 李卫娜, 郑小林, 吴南, 等. 应用因果分析方法对癫痫发作间期头皮脑电信号进行致痫灶定侧[J]. 国际生物医学工程杂志, 2013, 36(5): 261-265.
- LI Weina, ZHENG Xiaolin, WU Nan, *et al.* Epileptic foci lateralization from interictal scalp EEG by applying causal analysis[J]. *International Journal of Biomedical Engineering*, 2013, 36(5): 261-265.
- EPSTEIN C M, ADHIKARI B M, GROSS R, *et al.* Application of high-frequency Granger causality to analysis of epileptic seizures and surgical decision making[J]. *Epilepsia*, 2014, 55(12): 2038-2047.
- NICOLAOU N, HOURR S, ALEXANDROU P, *et al.* EEG-based automatic classification of 'awake' versus 'anesthetized' state in general anesthesia using Granger causality[J]. *PLoS One*, 2012, 7(3): e33869.
- ARNOLD A, LIU Y, and ABE N. Temporal causal modeling with graphical granger methods[C]. *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, 2007: 66-75.
- BENJAMIN B, DORNHERE G, KRAUEDAT M, *et al.* The non-invasive Berlin brain-computer interface: fast acquisition of effective performance in untrained subjects[J]. *NeuroImage*, 2007, 37(2): 539-550.
- JOHN S. Cortical functions[Z]. Routledge, 1999: 30-45.
- GAO Qing, DUAN Xujun, and CHEN Huaifu. Evaluation of effective connectivity of motor areas during motor imagery and execution using conditional Granger causality[J]. *NeuroImage*, 2011, 54(2): 1280-1288.
- FRASER A M and SWINNEY H L. Independent coordinates for strange attractors from mutual information[J]. *Physical Review A*, 1986, 33(2): 1134-1140.
- ZHANG H, CHIN Z Y, ANG K K, *et al.* Optimum spatio-spectral filtering network for brain-computer interface[J]. *IEEE Transactions on Neural Networks*, 2011, 22(1): 52-63.
- PARK C, LOONEY D, AHRABIAN A, *et al.* Classification of motor imagery BCI using multivariate empirical mode decomposition[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2013, 21(1): 10-22.

余青山: 男, 1980年生, 博士, 副教授, 主要研究方向为生物信息处理与分析、机器学习、模式识别等。

陈希豪: 男, 1990年生, 硕士生, 研究方向为模式识别、脑-机接口及相关应用等。