

文章编号: 1000-8608(2013)06-0898-05

基于序贯似然比检验的运动想象脑电信号分类方法研究

刘 蓉^{*1}, 李春月¹, 王永轩², 王媛媛¹, 李 响³

(1. 大连理工大学 电子信息与电气工程学部, 辽宁 大连 116024;

2. 大连大学附属中山医院, 辽宁 大连 116001;

3. 大连医科大学附属第二医院, 辽宁 大连 116027)

摘要: 快速准确地对脑电信号进行特征分类是脑-机接口研究的关键问题之一。从人脑决策模型出发,结合自适应小波基特征提取方法,提出了一种基于序贯似然比检验的运动想象脑电信号动态分类方法。该方法在分类中无须预先固定样本量,而是逐次取样,累积分类信息,有利于解决脑-机接口的实时控制问题。为了更好地衡量该方法的有效性,进行了10次10折交叉验证,实验结果表明3个运动想象数据集共8位受试者的平均正确率达到87%以上,互信息和分类时间等指标也表明该方法能够有效提高脑-机接口系统的性能,具有较好的实用性。

关键词: 脑-机接口; 运动想象; 自适应特征提取; 动态分类

中图分类号: R318 **文献标志码:** A

0 引言

当人进行单侧肢体运动想象时,大脑对侧运动感觉区的 μ 节律和 β 节律能量降低,而同侧运动感觉区的 μ 节律和 β 节律能量增大,这种现象称为事件相关去同步/同步(event-related desynchronization/synchronization, ERD/ERS)^[1]。基于ERD/ERS的脑-机接口(brain-computer interface, BCI)系统在方向控制等方面具有独特的优势,用于控制的脑电信号(electroencephalogram, EEG)只是通过运动想象产生,如想象左右手运动,而不依赖于任何感觉刺激,因此被广泛应用于BCI系统中。运动想象EEG的分类方法主要有线性判别分析(linear discriminant analysis, LDA)、人工神经网络(artificial neural network, ANN)和支持向量机(support vector machine, SVM)等^[2]。这些方法在各自适用的领域取得了较好的分类结果,但都是对特定时间段的EEG特征进行分类,虽然该时间段可能在统计分析上具有最好的区分度,但EEG是非平稳信号,若只考虑某一时间段的EEG信息,则无法充分利用单次观测信号的全部信息。此外,对特定时间段的EEG特征

进行分类会导致所有试验的分类时间都相同,即便有些试验的特征非常明显也不能提前判决,无法实现准确率与判决时间的折中。而研究表明在进行过程选择时准确率更重要,如用EEG信号选择电视频道、选择菜单等^[3];在基于BCI的游戏 中速度则更为重要^[4]。近年来,研究者逐渐将目光转向了能连续反映用户意图的动态分类算法,使分类器能够对单次实验的EEG进行分段并有效地累积分类信息。文献[5-7]假设各时段数据的分类区分度是相等的,直接对各段的分类结果进行累积,但由于疲劳或注意力减退等原因,这种假设通常并不满足。文献[8-9]根据贝叶斯分类误差 Chernoff 上限估计出区分度权值,然后对各段分类结果进行加权累积。该方法使Lemm等在2003年和2005年的BCI竞赛中获胜,然而,该方法得出的区分度权值只能增强实验中期获得的信息,减弱了实验初期及后期的信息^[10]。

本文首先基于自适应小波基进行运动想象EEG的特征提取,然后依据心理学家Bogacz等提出的决策扩散模型^[11],使用序贯似然比检验(sequential probability ratio testing, SPRT)^[12]

收稿日期: 2012-12-09; 修回日期: 2013-09-24。

基金项目: 国家自然科学基金资助项目(61005088);“十一五”国家科技支撑计划资助项目(2012BAJ18B06);中央高校基本科研业务费专项资金资助项目(DUT13JB02)。

作者简介: 刘 蓉*(1980-),女,副教授,硕士生导师,E-mail: rliu@dlut.edu.cn。

进行分类。SPRT 以似然比作为单段分类信息，并以连乘方式得到累积分类信息。该方法无须计算区分度权值，并且能够实现准确率与分类时间的折中。

1 特征提取方法

设 $E_D = \{(x_i, D)\}_{i=1}^{N_D}$, 其中 $D \in \{L, R\}$, 表示左、右手两个类别, 为包含 N_D 次运动想象实验的数据集, x_i 为第 i 次实验的数据。将单次实验数据分段表示为 $x_i = (x_{i1} \dots x_{ij} \dots x_{iJ})$, 其中 x_{ij} 代表第 i 次实验中的第 j 段数据, J 是单次实验数据的总段数。由于小波变换具有时频谱分析的作用, 本文采用 Morlet 小波来提取 EEG 的时变频带能量。若采样频率为 f_s , 则离散化 Morlet 小波为

$$\psi(n) = \pi^{1/4} a^{1/2} \exp\left(-\frac{n^2}{2a^2 f_s^2}\right) \exp\left(jK \frac{n}{af_s}\right) \quad (1)$$

其中 a 为尺度因子, $K(K \geq 5)$ 为调制系数。中心频率为

$$f = K/2\pi a f_s \quad (2)$$

因此, 尺度因子 a 可由中心频率 f 和调制系数 K 表示为

$$a = K/2\pi f_s f \quad (3)$$

有效时宽

$$\Delta n = \sqrt{2} a f_s \quad (4)$$

对于给定的一组参数 (f, K) , 由式(1)和(3)可以确定某个频段的小波基 $\psi_f(n)$, $f \in \{\mu, \beta\}$, 其中 ERD/ERS 现象发生在 $\mu(8 \sim 12 \text{ Hz})$ 和 $\beta(18 \sim 25 \text{ Hz})$ 节律。用 $x_c(n)$, $c \in \{C_3, C_4\}$ 表示 x_{ij} 的单导联数据, 向小波基投影并取模平方即可得频带能量为

$$z_{cf} = |\langle x_c(n), \psi_f(n) \rangle|^2; c \in \{C_3, C_4\}, f \in \{\mu, \beta\} \quad (5)$$

为了得到不同频段的最优参数 (f, K) , 计算两类信号的 t -检验统计量 (t -test statistic):

$$t = \frac{m_L - m_R}{\sqrt{\frac{(N_L - 1)S_L^2 + (N_R - 1)S_R^2}{N_L + N_R - 2}} \cdot \left(\frac{1}{N_L} + \frac{1}{N_R}\right)} \quad (6)$$

其中 $m_D(D \in \{L, R\})$ 为两类信号频带能量 z_{cf} 的均值, S_D 为标准差。使用网格搜索法 (grid search method, GSM) 得到不同频段内最大的 $|t|$, 对应的 (f, K) 即为所求的最优参数, 从而确定最优小波基 $\psi_f(n)$, 这样得到的最优小波基称为自适应小波基 (adaptive wavelet base, AWB)。中心频率

f 的搜索范围为 $6 \sim 30 \text{ Hz}$, 调制系数 K 的搜索范围为 $5 \sim 30$ 。

根据式(6)确定最优参数, 将两个导联的数据向两个频段的自适应小波基上投影得到的频带能量即可组成一个四维特征向量:

$$\mathbf{z}_{ij} = (z_{C_3\mu} \ z_{C_3\beta} \ z_{C_4\mu} \ z_{C_4\beta})^\top \quad (7)$$

为了使用 SPRT 方法对运动想象信号进行分类, 需要确定两类数据各段特征向量的概率密度函数 (probability density function, PDF) $f_{j|D}(\mathbf{z}_j)$ 。假设特征向量 \mathbf{z}_j 符合四维高斯分布, 则均值向量 $\mathbf{m}_{j|D}$ 和协方差矩阵 $\mathbf{S}_{j|D}$ 可由训练集计算得到:

$$\begin{aligned} \mathbf{m}_{j|D} &= E(\mathbf{z}_{j|D}) = \sum_{i=1}^{N_D} \mathbf{z}_{ij|D} / N_D; \\ \mathbf{S}_{j|D} &= E((\mathbf{z}_{j|D} - \mathbf{m}_{j|D}) \cdot (\mathbf{z}_{j|D} - \mathbf{m}_{j|D})^\top) = \\ &\quad \sum_{i=1}^{N_D} (\mathbf{z}_{ij|D} - \mathbf{m}_{j|D}) \cdot (\mathbf{z}_{ij|D} - \mathbf{m}_{j|D})^\top / N_D; \\ f_{j|D}(\mathbf{z}_j) &= f(\mathbf{z}_j | \mathbf{m}_{j|D}, \mathbf{S}_{j|D}) = (2\pi)^{-2} \times \\ &\quad |\mathbf{S}_{j|D}^{-1}|^{1/2} \cdot \exp\left(-\frac{1}{2}(\mathbf{z}_j - \mathbf{m}_{j|D})^\top \cdot \right. \\ &\quad \left. \mathbf{S}_{j|D}^{-1}(\mathbf{z}_j - \mathbf{m}_{j|D})\right) \end{aligned} \quad (8)$$

2 SPRT 动态分类方法

SPRT 以似然比作为单段数据的分类信息, 并以连乘的方式进行累积, 当达到预先设定的阈值时即可进行判决, 若未达到就继续累积。设第 j 段数据的似然比为

$$\lambda_j(\mathbf{z}_j) = \log\left(\frac{f_{j|L}(\mathbf{z}_j)}{f_{j|R}(\mathbf{z}_j)}\right) \quad (9)$$

则 k 段数据的累积信息为

$$\lambda_{1,2,\dots,k} = \log\left(\frac{f_L(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k)}{f_R(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k)}\right) \quad (10)$$

若 $\mathbf{z}_j (j = 1, 2, \dots, k)$ 相互独立, 则有

$$\begin{aligned} \lambda_{1,2,\dots,k} &= \log\left(\frac{f_L(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k)}{f_R(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k)}\right) = \\ &\quad \log\left(\frac{f_{1|L}(\mathbf{z}_1)}{f_{1|R}(\mathbf{z}_1)} \cdot \frac{f_{2|L}(\mathbf{z}_2)}{f_{2|R}(\mathbf{z}_2)} \cdots \frac{f_{k|L}(\mathbf{z}_k)}{f_{k|R}(\mathbf{z}_k)}\right) = \\ &\quad \sum_{j=1}^k \log\left(\frac{f_{j|L}(\mathbf{z}_j)}{f_{j|R}(\mathbf{z}_j)}\right) = \sum_{j=1}^k \lambda_j(\mathbf{z}_j) \end{aligned} \quad (11)$$

其中 $f_D(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k) (D \in \{L, R\})$ 为 $(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k)$ 的联合 PDF。

给定两个阈值 $\Gamma_1, \Gamma_2 (\Gamma_1 < \Gamma_2)$, 则 SPRT 的判决准则为

$$\begin{cases} \lambda_{1,2,\dots,k} \in [\Gamma_2, \infty); & \hat{D} = L \\ \lambda_{1,2,\dots,k} \in (-\infty, \Gamma_1]; & \hat{D} = R \\ \lambda_{1,2,\dots,k} \in (\Gamma_1, \Gamma_2); & \text{继续累积} \end{cases}$$

减小 Γ_1 或增大 Γ_2 可以减小两类错误概率,但数据的有限会使无判决的概率增大,因此阈值的设定非常重要.本文通过 GSM 方法得到使训练数据集分类正确率最高的阈值并应用于测试数据集的判决中.

3 实验数据及结果分析

为客观评价算法的有效性,本文对 3 个运动想象数据集:BCI 竞赛 DatasetIII(2003)、BCI 竞赛 DatasetIIIb(2005) 及大连理工大学生物医学工程实验室采集的数据进行了分析. DatasetIII 包含受试者 S2003 的 280 个实验数据, DatasetIIIb 包含 3 位受试者(O3、S4、X11), 其中 O3 有 640 个数据,S4 有 1 080 个数据,X11 有 1 080 个数据. BCI 竞赛数据的实验范式详见文献[13-14]. 本实验室自行采集的数据来自 4 位受试者(L1、L2、L3、L4), 采样频率为 125 Hz, 实验次数均为 100 次. 实验任务是根据屏幕上出现的箭头想象左手或右手运动,每次实验持续 6 s. $t=0 \sim 2$ s 时,受试者保持安静; $t=2$ s 时,显示器上出现一个向左或向右的箭头,同时要求受试者被箭头引导着想象手的运动,直至实验结束. 由于想象手运动的 ERD/ERS 现象发生在初级感觉运动皮层的运动功能区,选用 C₃ 和 C₄ 导联的数据进行分析.

图 1 为通过 AWB 方法得到的受试者 S2003 随时间变化两类信号的四维特征向量均值.由图

1 可以看出,在实验准备阶段(前 3 s)两类信号的特征向量的均值很接近,之后想象差异开始逐渐增大,在 4 s 左右时达到高峰,此后逐渐降低直至结束阶段.

本文将 SPRT 算法与 Lemm 等采用的 Bayesian^[8]方法得到的结果进行了比较,包括分类信息累积过程、分类准确率、互信息和分类时间. 图 2 为 S2003 分别使用 SPRT 和 Bayesian 方法的分类信息平均累积过程. Bayesian 和 SPRT 算法分别以后验概率和似然比为累积信息. 累积似然比是单调上升且没有上界的,所以在线 BCI 系统中无法对其进行归一化. 但为了比较不同分类算法的信息累积过程,本文在离线分析中对 SPRT 算法得到的累积信息 I 进行归一化,将范围限制在 [-1, 1]. 由图 2(a)可知,在使用 Bayesian 算法的情况下,累积 Bayesian 后验概率会在某个时刻达到峰值,之后反而开始下降,这意味着最优判决时刻出现在峰值点附近,而之后的分类信息无法再供判决使用. 图 2(b)表明在使用 SPRT 算法的情况下,随着新信息的不断加入,两类想象运动似然比的差值不断增大,并无减小现象. 一旦累积对数似然比超出阈值,则立即进行判决,对于分类特征较强的信号可以显著减少判决时间. 另外由判决准则可知, SPRT 算法可以实现分类准确率与判决时间的折中,在一定范围内增加阈值可以提高分类准确率,降低阈值可以减少判决时间.

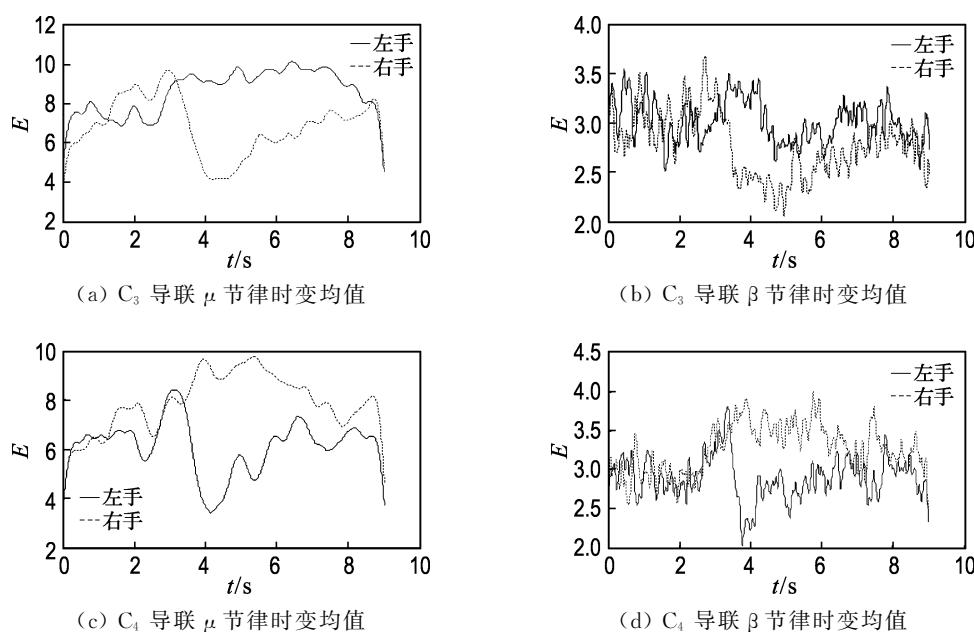


图 1 受试者 S2003 两类信号四维特征向量时变均值

Fig. 1 The mean of the four-dimensional feature vectors of the two classes for subject S2003

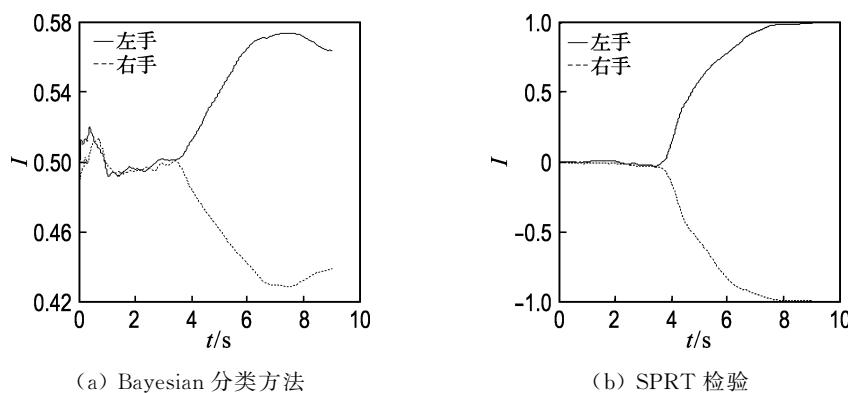


图 2 分类信息平均累积过程

Fig. 2 Average accumulative process of classification information

在分类中采用 10 次 10 折交叉验证, SPRT 和 Bayesian 方法得到的分类结果列于表 1 中。对比表中的结果可以看出, SPRT 算法平均分类时间为 2.02 s, 也就是说, 在开始想象后 2 s 左右即可做出判决, 比 Bayesian 算法的平均分类时间缩短了

1 s. SPRT 算法的平均分类准确率比 Bayesian 算法提高了 3.6%, 平均互信息提高了 17.5%。对两种算法的 3 种评价指标进行统计分析, 结果表明, 在显著性水平为 0.05 的情况下, 3 种指标均具有显著性差异。

表 1 SPRT 及 Bayesian 方法的平均分类准确率、互信息和分类时间

Tab. 1 The mean of accuracy, mutual information and time for SPRT and Bayesian algorithm

| 受试者 | SPRT | | | Bayesian | | |
|-------|---------|---------|--------|----------|---------|--------|
| | 分类准确率/% | 互信息/bit | 分类时间/s | 分类准确率/% | 互信息/bit | 分类时间/s |
| S2003 | 90.7 | 0.428 | 2.16 | 89.0 | 0.375 | 3.18 |
| O3 | 88.2 | 0.379 | 1.53 | 84.2 | 0.302 | 2.48 |
| S4 | 79.4 | 0.218 | 3.00 | 79.5 | 0.222 | 3.74 |
| X11 | 85.0 | 0.306 | 2.35 | 83.8 | 0.300 | 3.66 |
| L1 | 89.1 | 0.433 | 1.50 | 84.0 | 0.344 | 2.81 |
| L2 | 87.4 | 0.406 | 1.80 | 83.2 | 0.325 | 2.94 |
| L3 | 89.2 | 0.438 | 1.91 | 86.3 | 0.388 | 2.74 |
| L4 | 87.2 | 0.400 | 1.90 | 81.9 | 0.301 | 2.63 |
| 平均 | 87.0 | 0.376 | 2.02 | 84.0 | 0.320 | 3.02 |

4 结 论

针对 BCI 系统的分类问题, 本文从人脑决策模型出发, 结合自适应小波基特征提取方法, 提出了一种基于 SPRT 的运动想象 EEG 分类算法。该方法首先针对每个受试者得到最优小波参数, 从而得到自适应小波基, 以频带能量为分类特征; 然后使用 SPRT 算法对特征进行信息累积分类, 将单段分类信息按时间顺序进行累积, 若累积信息达到预先设定的阈值则进行判决分类, 否则继续累积信息, 各次试验的判决时间可以信号自身特点进行调整。适当地降低/升高阈值可以减少/增加判决时间, 准确率也相应地降低/升高。因此, 本文使用的 SPRT 分类算法可以根据不同的应用目的来调整参数, 以满足不同的要求, 具有较好的实用性。

参 考 文 献 :

- [1] Formaggio E, Storti S F, Galazzo I B, et al. Modulation of event-related desynchronization in robot-assisted hand performance: brain oscillatory changes in active, passive and imagined movements [J]. *Journal of Neuroengineering and Rehabilitation*, 2013, **10**(1):24.
- [2] Lotte F, Congedo M, Lécuyer A, et al. A review of classification algorithms for EEG-based brain-computer interfaces [J]. *Journal of Neural Engineering*, 2007, **4**(2):R1-R13.
- [3] Iturrate I, Antelis J M, Kubler A, et al. A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation [J]. *IEEE Transactions on Robotics*, 2009, **25**(3):614-627.

- [4] Wolpaw J R, Ramoser H, McFarland D J, et al. EEG-based communication: improved accuracy by response verification [J]. *IEEE Transactions on Rehabilitation Engineering*, 1998, **6**(3):326-333.
- [5] Roberts S J, Penny W D. Real-time brain computer interfacing: A preliminary study using Bayesian learning [J]. *Medical and Biological Engineering and Computing*, 2000, **38**(1):56-61.
- [6] Blanchard G, Blankertz B. BCI competition 2003-Data Set IIa: Spatial patterns of self-controlled brain rhythm modulations [J]. *IEEE Transactions on Biomedical Engineering*, 2004, **51**(6):1062-1066.
- [7] ZHU Xiao-yuan, GUAN Cun-tai, WU Jian-kang, et al. Bayesian method for continuous cursor control in EEG-based brain-computer interface [C] // *Proceedings of the 2005 IEEE Engineering in Medicine and Biology*. Shanghai: IEEE, 2005:7052-7055.
- [8] Lemm S, Schäfer C, Curio G. BCI competition 2003-Data Set III: Probabilistic modeling of sensorimotor μ rhythms for classification of imaginary hand movements [J]. *IEEE Transactions on Biomedical Engineering*, 2004, **51**(6):1077-1080.
- [9] Lemm S, Schäfer C, Curio G. Aggregating classification accuracy across time: Application to single trial EEG [J]. *Advances in Neural Information Processing Systems*, 2007, **19**:825-832.
- [10] 王永轩, 邱天爽, 刘蓉, 等. 基于信号投影能量特征的脑电意识动态分类[J]. 信号处理, 2012, **28**(8):1059-1062.
- WANG Yong-xuan, QIU Tian-shuang, LIU Rong, et al. Dynamic motor imagery classification with signal power projection based feature [J]. *Signal Processing*, 2012, **28**(8):1059-1062. (in Chinese)
- [11] Bogacz R, Wagenmakers E J, Forstmann B U, et al. The neural basis of the speed-accuracy tradeoff [J]. *Trends in Neurosciences*, 2009, **33**(1):10-16.
- [12] Liu R, Newman G, Ying S H, et al. Improved BCI performance with sequential hypothesis testing [C] // *Proceedings of the 2011 IEEE Engineering in Medicine and Biology*. Boston: IEEE, 2011:4215-4218.
- [13] Glassman E L. A wavelet-like filter based on neuron action potentials for analysis of human scalp electroencephalographs [J]. *IEEE Transactions on Biomedical Engineering*, 2005, **52**(11):1851-1862.
- [14] Brodu N, Lotte F, Lécuyer A. Comparative study of band-power extraction techniques for motor imagery classification [C] // *IEEE International Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB)*. Rennes: IEEE, 2011:1-6.

SPRT-based classification method for motor imagery electroencephalogram

LIU Rong^{*1}, LI Chun-yue¹, WANG Yong-xuan², WANG Yuan-yuan¹, LI Xiang³

(1. Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China;
2. Zhongshan Affiliated Hospital of Dalian University, Dalian 116001, China;
3. The Second Affiliated Hospital of Dalian Medical University, Dalian 116027, China)

Abstract: To extract and classify the electroencephalogram (EEG) signal features fast and accurately is a key issue for the brain-computer interface (BCI) systems. Based on the human decision-making model, a motor imagery EEG dynamic classification method is proposed based on the sequential probability ratio testing (SPRT) combined with an adaptive wavelet feature extraction method. Without pre-fixed sample size, this classification method samples and accumulates classification information successively. It is helpful to solve the real-time control problems in the BCI. In order to evaluate the effectiveness of the algorithm, a 10 times 10-fold cross-validation is used. The experimental results show that the average classification accuracy of three motor imagery datasets of eight subjects is above 87%. The results of mutual information and classification time also show that the method can effectively improve the performance of BCI system and has good practicability.

Key words: brain-computer interface (BCI); motor imagery; adaptive feature extraction; dynamic classification