

# EMD 融合 PSD、CSP 的脑电特征提取方法

陈启超<sup>1,2</sup>, 张学军<sup>1</sup>, 黄婉露<sup>1</sup>

(1. 南京邮电大学 电子与光学工程学院 微电子学院, 江苏 南京 210023;  
2. 南京邮电大学 射频集成与微组装技术国家地方联合工程实验室, 江苏 南京 210023)

**摘要:**为了提高运动想象分类精确度,提出一种基于经验模式分解(EMD),并结合功率谱密度(PSD)和公共空间模式(CSP)的特征提取算法。首先将采集的脑电信号进行预处理,再对信号使用 EMD 算法得到多个固有模态函数(IMFs)。通过计算每次实验原始脑电信号与各阶 IMF 分量之间的相关系数,并计算所有实验得出的相关系数的绝对值的平均数,选择具有较大相关系数绝对值平均数的固有模态函数,计算其功率谱密度作为特征,经公共空间模式投影映射再提取相应的特征向量,并用支持向量机(SVM)进行分类。对9名受试者的运动想象进行分类结果分析,得到的平均分类正确率在96%以上。最后将该方法与其他方法做比较,证明了该算法的可行性。

**关键词:**脑电信号;经验模式分解;相关系数;功率谱密度;公共空间模式

**中图分类号:**R318;TN911.7

**文献标识码:**A

**文章编号:**1673-629X(2019)05-0126-05

**doi:**10.3969/j.issn.1673-629X.2019.05.027

## An EEG Feature Extraction Method of EMD Fusing PSD and CSP

CHEN Qi-chao<sup>1,2</sup>, ZHANG Xue-jun<sup>1</sup>, HUANG Wan-lu<sup>1</sup>

(1. School of Electronics and Optical Engineering & School of Microelectronics, Nanjing University of Posts and Telecommunications, Nanjing 210023, China;  
2. Nation-Local Joint Project Engineering Lab of RF Integration & Micropackage, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

**Abstract:** In order to improve the classification accuracy of motor imagery, we propose a feature extraction algorithm based on empirical mode decomposition (EMD) combined with power spectral density (PSD) and common space pattern (CSP). First, the collected EEG signal is preprocessed, and then EMD algorithm is used to obtain multiple natural modal functions (IMFs). The correlation coefficient between the original EEG signal and the IMF components are calculated, and the average of the absolute values of the correlation coefficients derived from all experiments is computed. The intrinsic modal function with an average absolute number of large correlation coefficients is selected, and its power spectral density is calculated as a feature. The corresponding feature vector is extracted by the common space pattern projection mapping and classified by a support vector machine (SVM). The classification of the motor imagery of 9 subjects is analyzed, and the average classification accuracy is above 96%. Finally, this method is compared with other methods to prove its the feasibility.

**Key words:** brain electrical signals; empirical mode decomposition; correlation coefficient; power spectral density; common spatial pattern

## 0 引言

脑机接口(brain computer interface, BCI)技术的发展可以帮助有运动障碍的病人提高自理能力和生活质量,可以通过大脑控制外部辅助设备,如计算机、语音合成器、辅助应用和神经假肢<sup>[1]</sup>等来加强他们与外界环境的交流和交互。脑机接口技术还被广泛地应用于

其他一些领域,如游戏应用和导航<sup>[2]</sup>。

经过多年的研究,脑机接口技术主要包含信号采集、预处理、特征提取、特征分类和接口设备控制等五个步骤。特征提取可以识别不同想象运动的脑电信号的判别信息,因此在脑机接口的研究领域备受关注。由于 EEG 信号的非线性和非平稳性,传统的时间-频

收稿日期:2018-05-06

修回日期:2018-09-12

网络出版时间:2018-12-21

基金项目:国家自然科学基金(61271334)

作者简介:陈启超(1992-),男,硕士研究生,研究方向为电路与系统;张学军,博士,教授,研究方向为智能信息处理、认知网络频谱感知、无线射频识别技术等。

网络出版地址: <http://cnki.net/kcms/detail/61.1450.TP.20181221.1445.014.html>

率方法有:短时傅里叶变换(short-time Fourier transform,STFT)、小波变换(wavelet transform,WT)等,但是这些方法不可能同时得到时间-频率的良好分辨率。近年来,希尔伯特黄变换(Hilbert-Huang transform,HHT)作为另一种时间-频率分析法变得越来越流行,同时它也十分适合分析非线性和非平稳信号。原信号经过经验模式分解<sup>[3]</sup>(empirical mode decomposition,EMD)生成固有模态函数(intrinsic mode functions,IMFs),随后对每个固有模态函数进行希尔伯特黄变换,求其相应的能量谱和边际谱作为特征进行分类。HHT不涉及海森堡不确定性原理,可以获得时域和频域的高分辨率,目前广泛应用于多个领域,如雷达探测、地震信号和生物医学信号等。

在此基础上,由于 EEG 信号的低空间分辨率,EEG 信号构成的 BCI 系统需要进行有效的空间滤波,从而确保能从受试者的相关脑域中提取特征信息。常用算法有:共空间模式<sup>[4]</sup>(common spatial pattern,CSP)、独立主成分分析(independent component analysis,ICA)、共域空间谱模式(CSSP)、滤波器 CSP(filter bank common spatial pattern,FBCSP)、判别滤波 CSP(discriminant filtering common spatial pattern,DFBCSP)等<sup>[5]</sup>。

在上述研究的基础上,文中提出一种 EMD 结合 PSD<sup>[6]</sup>和 CSP 算法的特征提取方法。首先将采集到的脑电信号进行预处理<sup>[7]</sup>,利用小波去噪方法对原始脑电信号进行滤波去噪<sup>[8]</sup>。接着将处理后的信号通过 EMD 分解成多个 IMF 分量,分别计算每次试验原始信号和各个 IMF 分量之间相关系数的绝对值,再计算所有实验的平均值,选择具有较大相关系数绝对值平均数的固有模态函数,计算其功率谱密度作为特征进行 CSP 滤波,再选择合适的特征向量,使用支持向量机进行分类。

## 1 EMD 结合 PSD 和 CSP 算法

### 1.1 算法步骤

文中采用 EMD 结合 PSD 和 CSP 算法对脑电信号进行特征提取。具体过程为:信号采集;预处理;EMD 分解并选取合适 IMF 分量为特征值;计算特征值的 PSD;CSP 滤波;SVM 分类。

### 1.2 经验模式分解

经验模式分解是一种数据处理算法<sup>[9]</sup>,具有以下限制条件:数据最小的两个极值,最小值和最大值;根据极值间的间隔确定特征时间尺度;在数据只有拐点没有极值点的情况下通过一次或多次微分得到极值,再由积分得到分解结果。

具体步骤如下:

(1)根据  $x(t)$  的极大值点外加极小值点,拟合成原信号的上包络线和下包络线,分别为  $v_{\max}(t)$  和  $v_{\min}(t)$ ,可得包络线的平均值  $m(t)$ 。

$$m(t) = \frac{1}{2} [v_{\max}(t) + v_{\min}(t)] \tag{1}$$

(2)计算原信号与均值的差值  $c(t)$ 。

$$c(t) = x(t) - m(t) \tag{2}$$

如果  $c(t)$  不符合 IMF 的截止条件,反复上述步骤,否则会将它作为固有模态函数,剩余量  $r(t)$  为:

$$r(t) = x(t) - c(t) \tag{3}$$

(3)假设原信号  $x(t)$  被分解为  $n$  个固有模态函数和一个剩余量  $r(t)$ ,重构信号为:

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \tag{4}$$

### 1.3 相关系数

一般情况下,使用相关系数用来反映变量之间相关关系精密程度的标准<sup>[10]</sup>。相关系数的定义:假设有两个随机变量  $X$  和  $Y$ ,则它们的相关系数为:

$$r = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E\{[X - E(X)][Y - E(Y)]\}}{\{E[X - E(X)]^2\}^{\frac{1}{2}} \{E[Y - E(Y)]^2\}^{\frac{1}{2}}} \tag{5}$$

其中,  $\text{cov}(X,Y)$  为  $X$  和  $Y$  的协方差函数;  $\sigma_X$ 、 $\sigma_Y$  表示  $X$ 、 $Y$  的标准差;  $E(X)$ 、 $E(Y)$  为平均值。相关系数  $r$  的取值为  $[-1,1]$ ,  $r$  的绝对值越大,说明相关程度<sup>[11]</sup>越高。

### 1.4 功率谱密度

计算步骤如下:

(1)  $x(n)$  为无限长随机序列,截取长度  $N$  变为有限长序列称为  $x_N(n)$ 。

(2) 计算  $x_N(n)$  在  $2m - 1$  点的自相关函数  $\hat{R}_x(m)$ 。

$$\hat{R}_x(m) = \frac{1}{N} \sum_{n=0}^{N-1} x_N(n) x_N(n+m) \tag{6}$$

其中,  $m = -(M - 1), \dots, -1, 0, 1, \dots, M - 1$ ,  $M \leq N$ 。

(3)求相关函数的傅氏变换得到功率谱,即:

$$\hat{S}_x(e^{j\omega}) = \sum_{m=-(M-1)}^{M-1} \hat{R}_x(m) e^{-j\omega m} \tag{7}$$

### 1.5 公共空间模式

CSP 是对输入的脑电数据进行投影映射<sup>[12-13]</sup>,得到区分度最大的特征向量,是基于对两个不同类别的协方差矩阵来完成的,对两个类别的协方差矩阵同时进行对角化来设计最优的空间滤波器,将原始脑电数据投影映射,得到分离度最大的特征。具体步骤为:

(1) 设  $X$  为运动想象单次任务的脑电信号矩阵,  $X$  为  $N * T$  维,  $N$  为采集脑电信号的通道数,  $T$  为采样点数。  $X$  归一化后的协方差矩阵为:

$$R_d = \frac{X_d X_d^T}{\text{tr}(X_d X_d^T)}$$

(8)

其中,  $X_d$  表示脑电信号的分类;  $X_d^T$  表示转置;  $\text{tr}(\cdot)$  表示迹。

(2) 分解步骤 1 得到的矩阵为:

$$R = R_1 + R_2 = U_0 \Sigma U_0^T$$

(9)

其中,  $\Sigma$  为特征值对角矩阵;  $U_0$  为其对应的特征向量矩阵。

(3) 求白化矩阵  $P$ 。

$$P = \sqrt{\Sigma^{-1}} U_0^T$$

(10)

(4) 白化协方差矩阵  $R_1$  和  $R_2$ 。

$$\begin{cases} S_1 = P R_1 P^T \\ S_2 = P R_2 P^T \end{cases}$$

(11)

(5) 主成分分解。

$$\begin{cases} S_1 = U_1 \Sigma_1 U_1^T \\ S_2 = U_2 \Sigma_2 U_2^T \end{cases}$$

(12)

其中,  $\Sigma_1 + \Sigma_2 = I$ 。取  $\Sigma_1$  和  $\Sigma_2$  中最大的  $m$  个特征值对应的特征向量组成投影矩阵  $U, U = (U_1 U_2)$ , 其中  $U_1、U_2$  为最大的  $m$  个特征值对应的特征向量。

(6) 求 CSP 空间滤波器  $W$ 。

$$W = U^T P$$

(13)

(7) 将信号  $Z_p(p = 1, \cdots, 2m)$  取对数, 增加特征差异的明显程度。

$$f_p = \log(\text{var}(Z_p) / \sum_{i=1}^{2m} \text{var}(Z_i))$$

(14)

其中,  $\text{var}$  表示求向量的方差。

1.6 支持向量机

支持向量机(SVM)<sup>[14]</sup>通过将低维度的信号映射到高维度的特征空间,从而实现将特征向量更好的分离。采用超平面进行数据分类,例如二维平面数据使用一条直线即可划分,三维立体数据需使用二维平面进行划分,即  $N$  阶高维数据空间采用  $N - 1$  阶超平面进行划分。选取最优分类面时,既要求能够将数据正确分离,又要求分类间隔最大,即离超平面最近的数据点到超平面的距离应该尽可能大。文中采用网格搜索法确定惩罚因子和核参数,通过 3 折交叉验证的方法

对训练数据进行网格搜索。

2 数据集描述

实验采用的数据是来自 BCI Competition 2008 data sets 2b 的数据,采集了 9 名受试者的 EEG 数据,并且均是右利手。实验开始时,要求受试者全身放松坐在舒适的椅子上,眼睛到屏幕的直线距离为 1 m。屏幕上会出现相应的方向箭头作为提示,受试者根据提示完成相应的手运动想象,两天完成两个 sessions,其中每个 session 包含 6 个 run,每个 run 包含 20 次实验,10 次左手,10 次右手,则一个 session 共包含 60 次左手,60 次右手,共 120 次。具体过程为:每次实验持续 8 ~ 9 s,前 2 s 是放松状态;在第 2 s 时,会有提示音提示实验即将开始,同时屏幕上出现十字光标,进入准备状态,持续 1 s;在接下来的 1.25 s 屏幕上会随机持续出现方向箭头;在第 4 s,受试者根据提示进行相应的想象运动,持续 3 s;在第 7 s 受试者停止想象,进入休息状态。采用 Ag/AgCl 电极,使用三个电极记录采样的脑电 EEG 数据,分别为  $C_3、C_z$  和  $C_4$ ,其中  $C_z$  为参考电极,而  $C_3、C_4$  电极位置则包含了想象手运动最丰富的信息。文中使用  $C_3、C_4$  通道的数据作为分析,采样频率为 250 Hz,采样数据通过 0.5 ~ 100 Hz 带通滤波器进行预处理,50 Hz 陷波滤波器。

3 实验

3.1 EMD 分解

对原信号进行 EMD 分解后,得到  $n$  阶 IMF 分量和残波,实验中  $n$  取 8。相关系数是用以反映变量之间相关关系密切程度的统计指标。通过计算每次实验各个 IMF 分量与原始信号的相关系数,并计算所有实验各阶相关系数绝对值的平均数,选取相关系数绝对值平均数最大的 IMF 分量作为特征信号。与传统频谱分析的方法进行 IMF 特征选择相比,求解相关系数选取特征更加直观。表 1 显示的是所有实验 EMD 分解后  $C_3、C_4$  通道的左右手各阶 IMF 与原始信号的相关系数的绝对值的平均数。从中可以看出,前两个 IMF 分量  $IMF_1$  和  $IMF_2$  的相关系数绝对值平均数较大,剩余的 IMF 则为低频信号,多为伪迹信号与噪声信号。文中选取第一阶 IMF 作为脑电信号,将其余的 IMF 分

表 1 各阶 IMF 的相关系数

		IMF <sub>1</sub>	IMF <sub>2</sub>	IMF <sub>3</sub>	IMF <sub>4</sub>	IMF <sub>5</sub>	IMF <sub>6</sub>	IMF <sub>7</sub>	IMF <sub>8</sub>
C <sub>3</sub>	左手	0.806 9	0.615 9	0.063 1	0.019	0.005 7	0.001 9	0.000 3	0.000 8
	右手	0.803 8	0.562 2	0.069	0.016 1	0.008 1	0.001 9	0.003 5	0.007 3
C <sub>4</sub>	左手	0.837 8	0.521 7	0.044 6	0.023 4	0.011 6	0.002 6	0.001	0.000 5
	右手	0.803 8	0.562 2	0.069	0.016 1	0.008 1	0.009 7	0.003 5	0.007 3

万方数据

量舍弃。使用 EMD 可以达到有效去除无关信号、增大信噪比的目的。

图 1 为  $C_3$ 、 $C_4$  通道左右手运动想象脑电波形及频谱。左边列为脑电波形图,右边列为频谱图。第一行

为  $C_3$  左手的脑电波形图和频谱图;第二行为  $C_3$  右手的脑电波形图和频谱图;第三行为  $C_4$  左手的脑电波形图和频谱图;第四行为  $C_4$  右手的脑电波形图和频谱图。

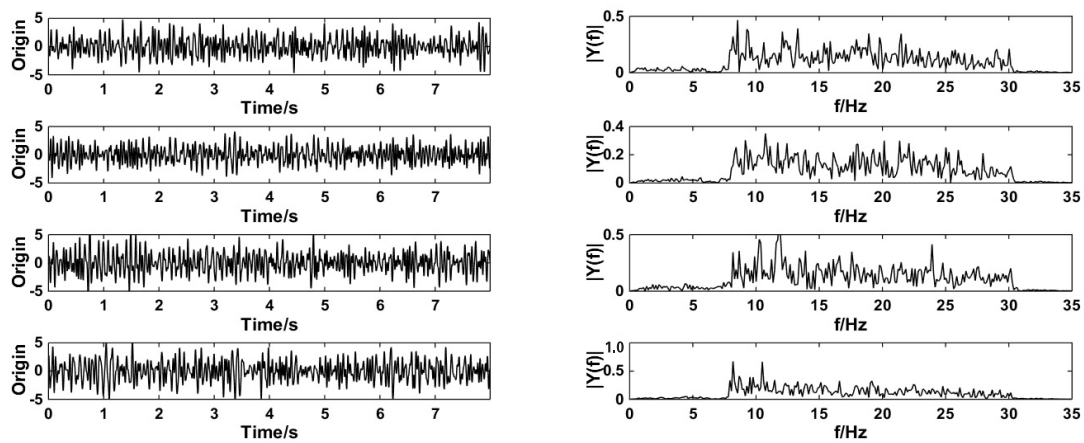


图 1  $C_3$ 、 $C_4$  通道左右手运动想象脑电波形及频谱

3.2 功率谱密度的计算

本次实验功率谱密度计算选择的采样点  $nfft = 1\,000$ 。图 2 为受试者  $S_1$  在某次实验中的功率谱密度与频率的关系。第一行左图为  $C_3$  通道左手的功率谱

密度与频率的关系;第一行右图为  $C_3$  通道右手的功率谱密度与频率的关系;第二行左图为  $C_4$  通道左手的功率谱密度与频率的关系;第二行右图为  $C_4$  通道右手的功率谱密度与频率的关系。

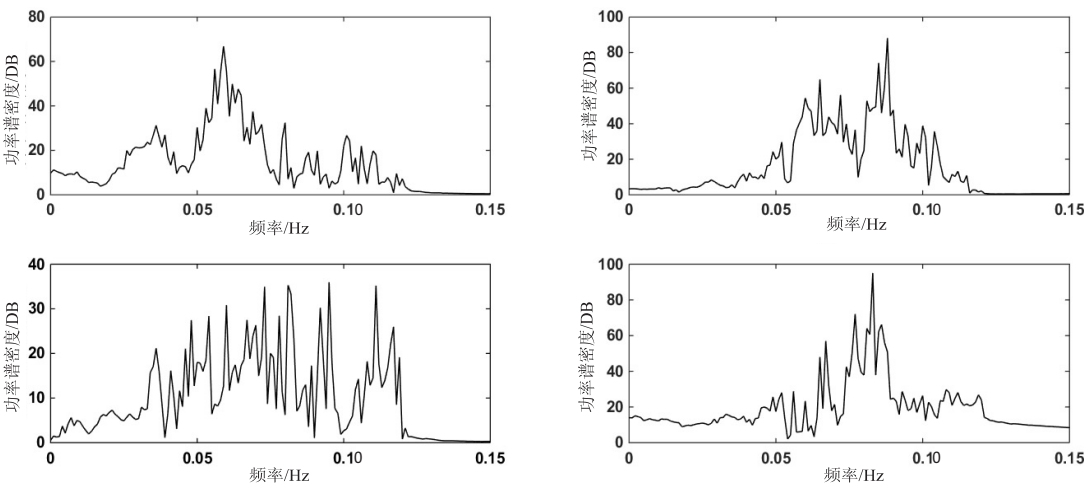


图 2 受试者  $S_1$   $C_3$ 、 $C_4$  左右手通道的功率谱密度

3.3 公共空间模式

在上述研究中,使用经验模式分解将  $C_3$ 、 $C_4$  通道的原始脑电信号进行分解,得到多个本征模态函数 IMFs,通过计算各个 IMF 分量与原始信号的相关系数,并计算所有实验各阶相关系数绝对值的平均数,通过比较可以得到  $IMF_1$  与原始信号的相关系数最大,取左右手的  $IMF_1$  计算其 PSD 作为新的输入信号,构成一个  $2 \times 1\,000$  的矩阵,其中第一行为  $C_3$  通道的数据,第二行为  $C_4$  通道的数据,1 000 为采样个数,最后进行 CSP 滤波。

3.4 结果分析

该方法中将原始脑电信号进行经验模式分解之

后,得到多个固有模态函数 IMF,通过计算相关系数绝对值的平均数并进行比较,选择合适 IMF 分段计算其 PSD 值作为信号,再进行 CSP 空间滤波,采用 SVM 分类器进行分类。图 3、图 4 表示 9 名受试者单次实验,一共 120 组左右手想象运动的分类正确率。由图可见:除了极个别低于 90%,但也都在 80% 以上,绝大多数都在 90% 以上,并且主要集中在 96% 左右。计算 9 名受试者想象运动的平均分类正确率为 96.5%,并且主要集中在 96% ~ 97% 之间。其中,受试者 4 最高达到 98% 以上;受试者 9 最低但也达到 95% 以上,可能造成该现象的原因是数据采集过程中存在误差;选取的 SVM 最优参数对该组数据并不是最优。9 位受试



者所有实验平均分类准确度约为 96.5%,由此可见文中方法可以得到有效且稳定的分类正确率。

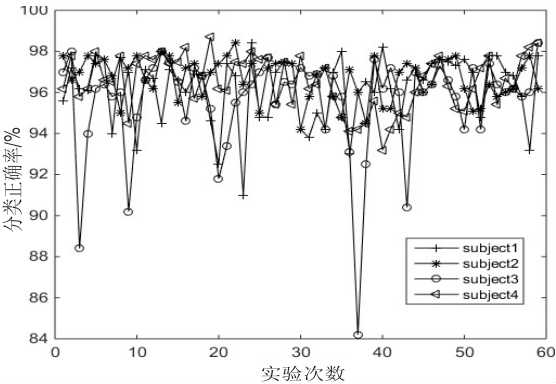


图 3 受试者  $S_1-S_4$  单次实验分类正确率

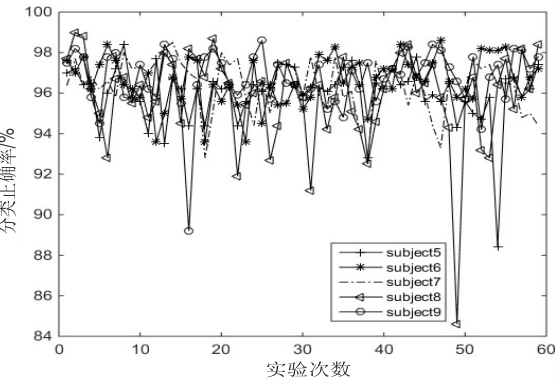


图 4 受试者  $S_5-S_9$  单次实验分类正确率

表 2 第三届 BCI 竞赛成绩和文中结果的比较

方法	分类精度/%
文中	96.5
EMP+CSP	92
CSSD+waveform mean+FDA	91
AR+spectral power+wavelet coefficients	87
offset+spectral power	86

将提出方法与其他方法进行比较,如表 2 所示,列出了第三届 BCI 竞赛中成绩较好的三组以及 EMD 组合 CSP 的方法和提出的方法进行比较,再次验证了该算法的可行性。

4 结束语

经过大量研究发现,传统的共空间模式适用于多通道数据分析,并且频域信息缺乏。因此,提出了一种 EMD 结合 PSD 和 CSP 的特征提取方法。原始脑电信号经过 EMD 模拟多通道,在此基础上增加了 EMD 的频域信息。信号经过 EMD 分解得到多个 IMF 分量,通过计算每次实验原始脑电信号与各阶 IMF 分量之间的相关系数,并计算所有实验得出的相关系数的绝对值平均数,选择具有较大相关系数绝对值平均数的固有模态函数并计算其 PSD 对数据进行优化,再组合成

多通道信息簇,经过共空间模式滤波提取特征向量,由支持 SVM 分类后得到所有 9 位受试者平均分类正确率为 96.5%,最高可达到 97.2%。实验结果证明了该方法的可行性和有效性,并且为脑机接口设计开辟了新的思路。

参考文献:

[1] 孙进,张征,周宏甫. 基于脑机接口技术的康复机器人综述[J]. 机电工程技术,2010,39(4):13-16.

[2] YANG Banghua,LI Huarong,WANG Qian,et al. Subject-based feature extraction by using fisher WPD-CSP in brain-computer interfaces[J]. Computer Methods & Programs in Biomedicine,2016,129:21-28.

[3] 徐晓刚,徐冠雷,王孝通,等. 经验模式分解(EMD)及其应用[J]. 电子学报,2009,37(3):581-585.

[4] 刘锦,吴小培,周蚌艳,等. 单次样本对的 CSP 滤波器设计及其在脑电训练样本优化中的应用[J]. 信号处理,2017,33(7):993-1001.

[5] 李明爱,刘净瑜,郝冬梅. 基于改进 CSP 算法的运动想象脑电信号识别方法[J]. 中国生物医学工程学报,2009,28(2):161-165.

[6] 叶竞,石锐,何庆华. 基于 HHT 和改进 CSP 算法的运动想象 BCI 系统[J]. 重庆理工大学学报:自然科学版,2012,26(5):70-73.

[7] 陈悦,张少白. LM 算法在神经网络脑电信号分类中的研究[J]. 计算机技术与发展,2013,23(2):119-122.

[8] 罗志增,李亚飞,孟明,等. 脑电信号的混沌分析和小波包变换特征提取算法[J]. 仪器仪表学报,2011,32(1):33-39.

[9] 王登,苗夺谦,王睿智. 一种新的基于小波包分解的 EEG 特征抽取与识别方法研究[J]. 电子学报,2013,41(1):193-198.

[10] 于路,薄华. 基于改进 EMD 的运动想象脑电信号识别算法研究[J]. 微型机与应用,2016,35(9):58-61.

[11] GUPTA S S,SOMAN S,RAJ P G,et al. Improved classification of motor imagery datasets for BCI by using approximate entropy and WOSF features[C]//International conference on signal processing and integrated networks. Noida, India: IEEE,2014:90-94.

[12] LEEB R,LEE F,KEINRATH C,et al. Brain-computer communication: motivation, aim, and impact of exploring a virtual apartment[J]. IEEE Transactions on Neural Systems & Rehabilitation Engineering,2007,15(4):473-482.

[13] ROBINSON N,VINOD A P,KAI K A,et al. EEG-based classification of fast and slow hand movements using wavelet-CSP algorithm[J]. IEEE Transactions on Biomedical Engineering,2013,60(8):2123-2132.

[14] XU P,DAVOINE F,ZHA H,et al. Evidential calibration of binary SVM classifiers[J]. International Journal of Approximate Reasoning,2016,72:55-70.