

基于局部均值分解和多尺度熵的运动想象脑电信号特征提取方法^①

邹晓红^② * * * * * 张轶勃 * * * * * 孙廷贞 * * * * *

(* 燕山大学信息科学与工程学院 秦皇岛 066004)

(** 河北省计算机虚拟技术与系统集成重点实验室 秦皇岛 066004)

(*** 河北省软件工程重点实验室 秦皇岛 066004)

摘 要 研究了脑电信号特征的提取。考虑到传统的脑电信号特征提取方法不能够很好地刻画脑电信号特征,因而会给不同意识任务下运动想象脑电信号的分类带来困难,该研究提出了一种基于局部均值分解(LMD)和多尺度熵(MSE)的脑电信号特征提取方法。该方法首先把脑电信号自适应地分解为一系列具有物理意义的乘积函数(PF)分量;然后选取有效的PF分量并计算多尺度熵,将多尺度熵组成特征向量;最后将其作为支持向量机(SVM)的输入来对脑电信号进行分类识别。实验表明该方法能够有效地提取脑电信号的特征,从而验证了该方法的有效性和可行性。

关键词 脑电信号(EEG),特征提取,局部均值分解(LMD),多尺度熵(MSE),支持向量机(SVM)

0 引 言

脑电信号(electroencephalogram, EEG)反映了大脑功能状态和大脑组织电活动,当进行运动任务或运动想象时,大脑中负责运动感知的 μ 节律(8 ~ 12Hz)和 β 节律(18 ~ 25Hz)会发生改变,发生事件相关去同步(event-related desynchronization, ERD)事件或事件相关同步(event-related synchronization, ERS)事件^[1]。运动想象脑电信号被广泛地用来进行脑机接口(BCI)控制,所以运用特征提取方法对运动想象脑电信号进行特征提取是脑机接口技术的关键。

脑电信号特征提取常见算法有共空间模式滤波法^[2]、自回归模型法^[3]、小波变换法^[4]等。共空间模式(common spatial pattern, CSP)算法是将脑电信号在空间域滤波,从而提取出不同运动模式下的脑电特征。CSP算法在二分类脑电信号上取得了很好

的效果,但是它需要针对特定的频段和大量的电极。自回归(adaptive regressive, AR)模型法是利用AR模型或者AR谱特征来反映脑电信号的时变特性。该方法适合分析平稳信号,而脑电信号是一种典型的非平稳非线性信号。小波变换(wavelet transform, WT)法采用可变的时频窗口逐级对信号进行分解,然后根据先验信息选取特定的小波系数作为特征。但对于有着复杂机理脑电信号,通常无法得到准确的先验信息。局部均值分解(local mean decomposition, LMD)算法是对脑电信号进行自适应的分解以获得若干个调幅-调频乘积函数(product function, PF)分量,这些分量能够反映信号的时频变化^[5]。常用的熵分析有近似熵(approximate entropy, AE)、样本熵(sample entropy, SE)、多尺度熵(multi-scale entropy, MSE)。多尺度熵是在样本熵的基础上提出的,能够衡量不同尺度下的样本熵值,从而可以对脑电信号进行全面的度量。针对不同运动想象模式

① 国家自然科学基金(No. 61472340),国家自然科学基金青年基金(No. 61602401)资助项目。
② 女,1967年生,教授,硕士生导师;研究方向:数据挖掘,数字图像处理;联系人,E-mail: xiaohongzou@ysu.edu.cn
(收稿日期:2017-07-10)

下的非线性非平稳脑电信号特征提取问题,本研究将 LMD 和多尺度熵(MSE)结合用于运动想象脑电信号特征提取,即将原始脑电信号进行分解,计算分解后的有效乘积函数(PF)分量的多尺度熵值作为特征向量,然后利用支持向量机(support vector machine, SVM)分类识别。

1 基于局部均值分解和多尺度熵的运动想象脑电信号特征提取方法

局部均值分解(LMD)方法于 2005 年提出,它首先是被应用在脑电信号特征提取中,之后在其他领域也被广泛使用,比如机械故障诊断^[6,7]。在信号特征提取中,常提取 LMD 的 PF 分量的时域统计量作为特征,这容易受到运动想象脑电信号中噪声的影响而不能提取出有效的特征向量。

1.1 局部均值分解

局部均值分解是由 Smith^[8]提出的一种适用于非平稳非线性信号的自适应时频分析方法。LMD 的本质是自适应地将脑电信号分解为多个乘积函数(PF)分量和一个残余量 R,每个 PF 分量都是一个纯调频信号和一个包络信号的乘积,由调频信号可以求出 PF 分量的瞬时频率,由包络信号得到 PF 分量的瞬时幅值。对于脑电信号 $x(t)$,其具体分解过程如下:

(1) 找出脑电信号 $x(t)$ 所有的局部极大值点和极小值点 $n_i (i = 1, 2, \dots)$, 求取相邻两个极值点 n_i 和 n_{i+1} 的均值 $m_i = (n_i + n_{i+1})/2$; 将相邻平均值点 m_i 用直线连接起来,然后采用滑动平均法进行平滑处理,得到局部均值函数 $m_{11}(t)$ 。

(2) 利用局部均值点 n_i 来求取包络估计值 $a_i = (|n_i + n_{i+1}|)/2$; 将相邻包络估计值点 a_i 用直线连接起来,然后采用滑动平均法进行平滑处理,得到包络估计曲线 $a_{11}(t)$ 。

(3) 从原始信号 $x(t)$ 中分离出局部均值函数 $m_{11}(t)$, 得到 $h_{11}(t) = x(t) - m_{11}(t)$ 。

(4) 将分离得到的 $h_{11}(t)$ 除以 $a_{11}(t)$ 进行解调,得到调频信号 $s_{11}(t) = h_{11}(t)/a_{11}(t)$; 判断 $s_{11}(t)$ 是否为纯调频函数,即 $s_{11}(t)$ 的包络函数

$a_{12}(t)$ 是否等于 1。若 $a_{12}(t) = 1$, $s_{11}(t)$ 为纯调频函数;若 $a_{12}(t) \neq 1$, 继续重复以上步骤 k 次得到 $s_{1k}(t)$, 直到包络函数满足 $a_{1(k+1)}(t) = 1$ 。即为 $h_{1n}(t) = s_{1(k-1)}(t) - m_{1k}(t)$, 迭代终止条件为 $\lim_{n \rightarrow \infty} a_{1k}(t) = 1$;

(5) 将所得的包络估计函数 $a_{1d}(t)$ 相乘得到第一个分量的包络 $a_1(t) = a_{11}(t)a_{12}(t)\cdots a_{1k}(t) = \prod_{d=1}^k a_d(t)$ 。

(6) 原始信号的第一个 PF 分量为包络信号 $a_1(t)$ 和纯调频信号 $s_{1k}(t)$, 即 $PF_1 = a_1(t)s_{1k}(t)$ 。

(7) 从原始信号中分离出 $PF_1(t)$, 将得到的新信号 $u_1(t)$ 作为源信号重复以上步骤,迭代 q 次,直到 $u_q(t)$ 为单调函数,迭代方程为 $u_q(t) = u_{q-1}(t) - PF_q(t)$; 最后,原始信号 $x(t)$ 表示为 $x(t) = \sum_{j=1}^q PF_j(t) + u_q(t)$, $u_q(t)$ 为残余分量。

脑电信号经过 LMD 的分解,原信号的特征信息会被分布到不同的时间特征尺度上,从而放大了原信号隐含的特征信息。接下来采用多尺度熵可以更容易地对 PF 分量合理性定量评估。

1.2 多尺度熵

近似熵(approximate entropy, ApEn)是由 Pincus^[9]提出的一种用于衡量时间序列复杂性的方法。Richman 等^[10]为了解决近似熵中存在自身匹配引起偏差的不足,在近似熵的基础上提出了样本熵(sample entropy, SampEn)的度量方法。近似熵和样本熵都是基于单尺度的分析,没有衡量尺度大于 1 的特性。为了衡量信号不同尺度的特性,Costa^[11]在样本熵的基础上引入了尺度因子,提出了多尺度熵(MSE)的方法,用来描述时间序列在不同尺度因子下的复杂性。对 LMD 分解后得到的 PF 分量计算其多尺度熵特征,其具体算法如下。

首先,对 PF 分量离散时间序列 $\{x_1, x_2, \dots, x_i, \dots, x_n\}$ 按照公式 $y_i^\tau = 1/\tau \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \leq j \leq n/\tau$ 进行多尺度化形成粗粒化时间序列 $\{y^{(\tau)}\}$, 序列长度 $M = \text{int}(n/\tau)$, 其中 τ 为尺度因子。

然后,计算粗粒化的时间序列在每个尺度因子下的样本熵值。

(1) 对离散时间序列构造一组 m 维的向量 $\mathbf{X}(i) = [x(i), x(i+1), \dots, x(i+m+1)], i = 1, 2, L, N-m+1$ 。

(2) 计算向量 $\mathbf{X}(i)$ 和向量 $\mathbf{X}(j)$ 之间的距离：
$$d[\mathbf{X}(i), \mathbf{X}(j)] = \max_{k=0, \dots, m-1} [|x(i+k) - x(j+k)|]$$
。

(3) 设定公差阈值 $r(r > 0)$ ，统计 $d[\mathbf{X}(i), \mathbf{X}(j)]$ 小于 r 的数目 $B^m(i)$ ， $B^m(i)$ 是模板匹配数目。计算 $B^m(i)$ 与距离总数 $N-m$ 的比值，记作 $C_r^m(r) = B^m(i)/(N-m)$ 。

(4) 计算 $C_r^m(r)$ 的平均值 $C_r^m(r) = (M-m+1)^{-1} \times \sum_{i=1}^{N-m+1} C_r^m(r)$ 。

(5) 将维数增加 1 维，并重复以上过程，计算 $C_r^{m+1}(r)$ ，并求平均值 $C_r^{m+1}(r) = (M-m)^{-1} \times \sum_{i=1}^{N-m} C_r^{m+1}(r)$ 。

(6) 当 N 为有上限值时，得出序列长度为 N 时样本熵估计值 $SampEn(m, r, N) = -\ln[C^{m+1}(r)/C^m(r)]$ 。

重复以上过程，即可以得到不同尺度下的样本熵值。

1.3 LMD 和 MSE 的运动想象脑电信号特征提取

不同的运动想象模式会激发大脑皮层不同区域的电活动，脑电采集系统的分布式电极记录下脑电信号的变化。当想象单侧肢体运动时，大脑对侧运动感觉区域的 μ 节律和 β 节律频谱震荡的幅度减低或者阻滞为事件相关同步事件； μ 节律和 β 节律频谱波幅增高的电活动事件去相关去同步事件。其中 μ 节律集中在 $8 \sim 12\text{Hz}$ ， β 节律集中在 $18 \sim 25\text{Hz}$ ，一

般采用 $8 \sim 30\text{Hz}$ 的脑电信号对这两种节律进行特征提取。

针对运动想象脑电信号的非线性、非平稳特点，本文提出了基于 LMD 和 MSE 结合的特征提取方法。在 LMD 的分解过程中，信号被逐次分解，保留了原信号的本质特征，同时使得原信号的特征信息在不同的分辨率下表现出来，通过多个 PF 分量更易提取出特征信息。MSE 方法是一种有效描述信息特征的方法，通过对 PF 分量计算多尺度熵，从而定量地描述脑电信号。将两种方法结合，可以有效地分析和提取运动想象脑电信号的特征，利于对运动想象类别进行识别。

运动想象脑电信号的特征提取方法流程如图 1 所示。

基于 LMD 和 MSE 的运动想象脑电信号特征提取方法具体实现步骤如下。

(1) 输入脑电信号，对每个样本进行 LMD 分解，得到一系列的 PF 分量。

(2) 脑电信号依次被分解，特征信息主要分布在前几个 PF 分量之中，所以选取前几个 PF 分量来提取特征信息。

(3) 计算每个训练样本 PF 分量的多尺度熵值，将它们组成特征向量。

(4) 把特征向量作为 SVM 分类器的输入，对分类器进行训练。

(5) 将待测样本的特征向量输入训练好的 SVM 分类器，以 SVM 分类器的输出确定运动想象的类别。

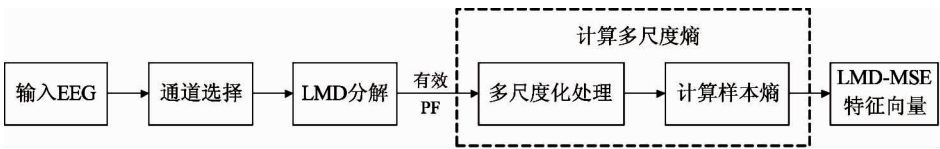


图 1 脑电信号特征提取流程

1.4 支持向量机特征识别

支持向量机(SVM)是一种效果优异的分类算法，它主要是建立在统计学习理论的结构风险最小理论和 VC(Vapnik-Chervonenkis)维基础之上。支

持向量机把输入的样本通过一种映射变换到高维空间，然后在这个高维空间中找出最佳分类面，从而将样本分开。

若训练集样本为 d 维向量， n 个样本，每个样本

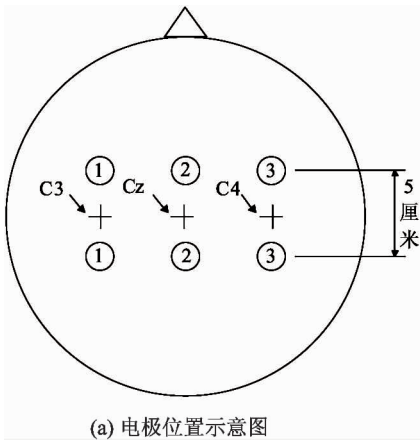
的类别表示为 (x_i, y_i) , $y_i \in \{-1, 1\}$, $i = 1, 2, \dots, n$ 。在 d 维空间中,线性判别函数的形式为 $g(x) = \mathbf{w} \cdot \mathbf{x} + b$, 那么超平面方程为 $\mathbf{w} \cdot \mathbf{x} + b = 0$ 。将判别函数进行归一化,使两类所有样本都满足 $y_i(\mathbf{w}x_i + b) \geq 1$ 在线性不可分情况下,同时将式中加入松弛项 $\eta_i \geq 0$, 则目标函数为

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + C \left(\sum_{i=1}^n \eta_i \right) \quad (1)$$

式中 $\mathbf{w}$ 为超平面法向量,常数 C 是惩罚因子。对于测试集数据,其判别函数为

$$f(x) = \text{sgn} \left(\sum_{i=1}^{n_s} \mathbf{a}_i^* y_i K(x_i, x) + b \right) \quad (2)$$

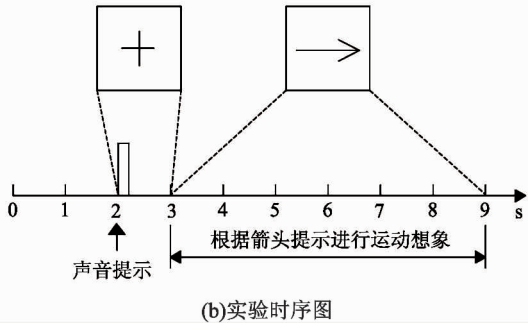
式中 $\text{sgn}(\cdot)$ 是求符号函数, $\mathbf{a}_i^*$ 对应的向量 $\mathbf{x}_i$ 是支持向量, n_s 是支持向量的个数, $K(x_i, x)$ 是核函数, b 是截距。



(a) 电极位置示意图

2 实验分析

本文采用第二届 BCI 竞赛 dataset III^[12] 数据集对上述提出的算法进行验证。实验数据记录了一个 25 岁正常女性的想象左右手运动脑电信号。实验在同一天进行,包含 7 组实验,每组 40 次,一共 280 组。实验记录 C3、C4、Cz (如图 2(a)) 三通道数据的 EEG,采样频率为 128Hz,带通滤波为 0.5 ~ 30Hz。实验如图 2(b) 所示,每次实验持续 9s, 0 ~ 2s 为静息状态,在从第 2s 开始有声音提示实验开始,并同时在屏幕上显示一个十字架“+”,显示 1s 结束;从第 3 秒开始屏幕中箭头提示实验者想象运动,在第 9s 结束。



(b) 实验时序图

图 2 电极位置分布和实验时序图

实验选取 C3、C4 通道的数据进行特征提取。运动想象的时间是 3 ~ 9s,所以提取每个通道 3 ~ 9s 间的数据,每组实验的每个通道共有 $128 \times 6 = 768$ 个数据点,然后将每个通道的数据进行 LMD 自适应的分解。图 3 是在想象右手运动时, C4 通道进行 LMD 分解的时频图。

从图 3 的左列图中可以看出,脑电信号被自适应地分解为 4 个 PF 分量和 1 个残余分量 $u(t)$ 。图 3 的右列图所示为 PF 分量和残余分量的频谱。从图中可以明显看出在 10Hz 和 20Hz 附近有大的

波动,说明了事件相关同步/事件相关去同步 (ERS/ERD) 现象。通过实验分析表明,包含左右手运动想象的 μ 节律和 β 节律信息主要分布在前 3 阶 PF 分量。所以选择前 3 个 PF 分量作为特征提取的信号。

运动想象的时间共有 6s,考虑到处理时间和分类的准确率,对提取最优特征所用时间段进行选择,采用的滑动时间窗为 2s,滑动步长为 1s。结果发现在 4 ~ 6s 和 5 ~ 7s 时分类效果较好,所以选取 4 ~ 7s 作为特征提取数据段。

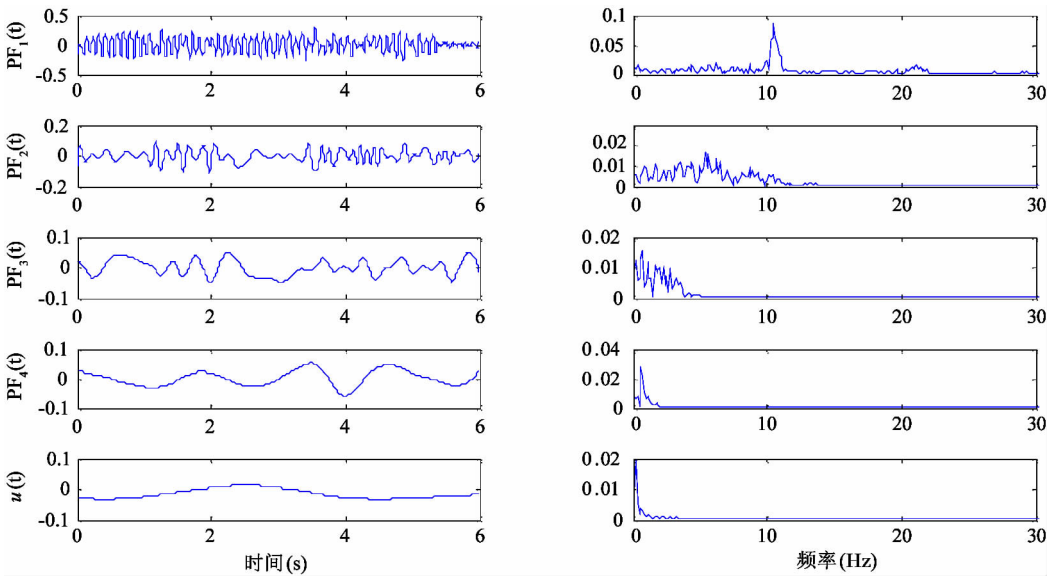


图 3 脑电信号 (C4 通道) 经 LMD 分解后的各 PF 分量的时频图

根据多次实验选定相似容限 $r=0.35SD$ (SD 是原始信号的标准偏差), 计算多组脑电信号在不同尺度下的熵值, 从而选定多尺度熵的参数。如图 4 所示, PF_1 和 PF_2 的熵值随着尺度因子 τ 的增大呈现先上升后下降的趋势, PF_3 的熵值随着尺度因子 τ

的增大呈现增大的趋势。经过实验尺度因子 $\tau = 9$ 时, 各个分量的熵值都有区分度, 分类准确率较高, 所以选取 $\tau = 9$ 。表 1 所示为不同类别的 C3、C4 通道在 $r=0.35SD$ 、 $\tau = 9$ 下想象右手运动时的平均多尺度熵值。

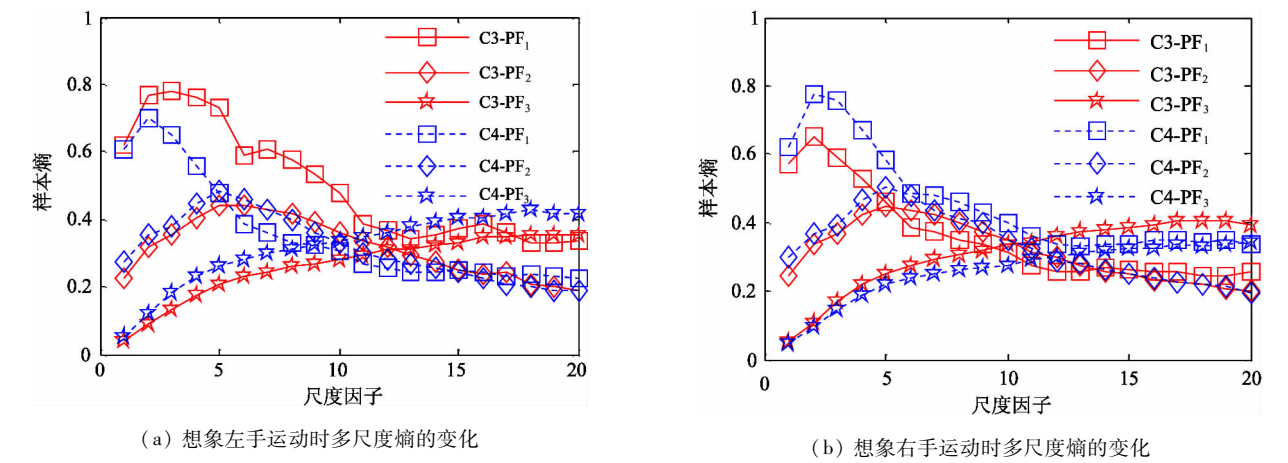


图 4 PF 分量在两个类别中不同的尺度下的样本熵

表 1 想象右手运动时的平均多尺度熵值

尺度因子	C3			C4		
	PF ₁	PF ₂	PF ₃	PF ₁	PF ₂	PF ₃
3	0.5885	0.3660	0.1670	0.7533	0.3933	0.1444
6	0.3843	0.4342	0.2747	0.4849	0.4776	0.2387
9	0.3356	0.3710	0.3175	0.4318	0.3967	0.2705
12	0.2546	0.2974	0.3596	0.3336	0.2886	0.3015
15	0.2600	0.2491	0.3888	0.3375	0.2483	0.3268
18	0.2436	0.2165	0.4054	0.3407	0.2182	0.3332

选取 C3、C4 通道的第一个 PF 分量计算上述条件的多尺度熵,如图 5 所示。从图 5 可知,想象左右手运动的脑电信号的多尺度熵虽有部分重叠,但仍有区分度。所以,多尺度熵能够反映出想象左右手

运动的脑电信号特征。前三阶 PF 分量的多尺度熵对分类贡献最大,将前三阶的多尺度熵结合起来作为特征向量,可能提高分类正确率。

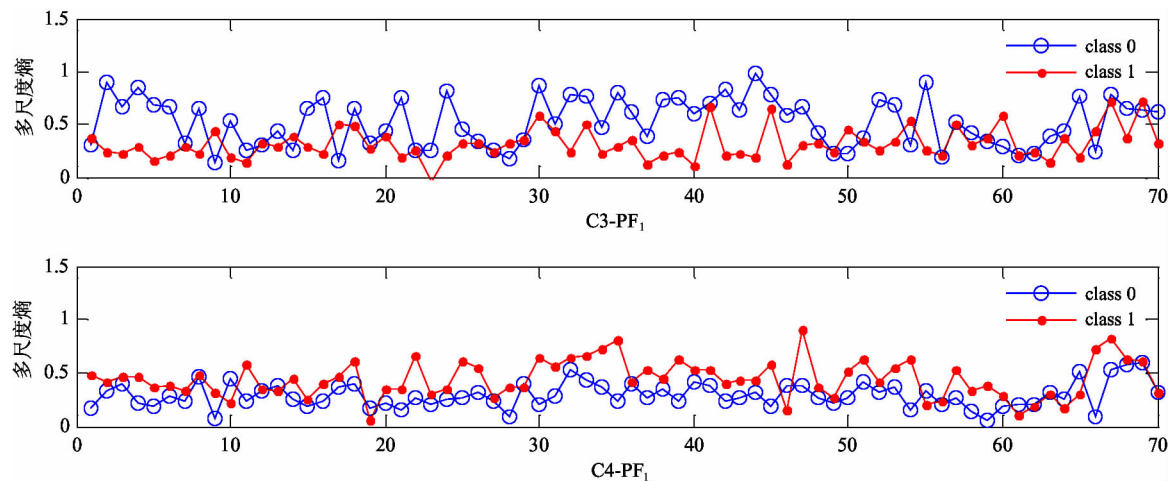


图 5 多尺度熵值分布

接下来,采用支持向量机分类器对提取到的特征向量进行分类。支持向量机是一种基于统计学习理论的机器学习方法,在解决小样本分类、非线性问题和高维模式识别中表现出了很多优势。本文利用支持向量机^[13]对左右手运动想象脑电信号进行分类。为了加快训练网络的收敛性,首先对训练集和测试集进行[0~1]归一化处理,目的是归纳统一数据的统计分布性。然后,将训练集送入 SVM 分类器,得到支持向量机模型。最后,输入测试集特征向量进行分类后输出运动想象类别。

实验数据集一共有 280 组实验,随机将其分为两份,每份有 140 组试验。将用多尺度熵提取的 6 维特征向量输入 SVM 分类器中进行训练分类,得到了最优分类识别率为 85.21%。从表 2 可以看出,本文的特征提取算法的分类准确率比文献[12]、[14]和[15]高了 0.92%~4.13%。本文使用了 6 维特征向量,特征维数的降低,可以使分类器的模型更简单,减少分类时间。相对已有文献特征数量大幅减少,进行了 100 次测试,该算法运行一次平均消耗的时间为 73ms。由此可见,本文算法减少了特征数量,且改善了分类精度。

表 2 不同特征提取方法下的识别率

特征	特征维数	识别率
HHT + AR ^[14]	14	81.08
CSP ^[15]	-	82.86
AR ^[12]	-	84.29
本文	6	85.21

3 结 论

本文提出了一种 LMD 和多尺度熵的特征提取方法。使用 LMD 对左右运动想象的脑电信号进行自适应分解,将分解后的 PF 分量提取多尺度熵的特征组成特征向量,输入 SVM 中进行分类识别,实现了运动想象脑电信号的分类。实验表明,采用 LMD 和多尺度熵的特征提取方法对脑电信号的识别优于采用传统的 CSP 特征提取算法和 AR 特征提取算法,最优识别率能达到 85.21%。研究结果说明,提出的方法可以有效地对运动想象脑电信号进行特征提取,有助于在 BCI 应用的特征提取环节提取出有效特征,为在线 BCI 的应用打下基础。

参考文献

- [1] Park C, Looney D, Rehman N U, 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
- [2] 杨帮华, 陆文字, 何美燕等. 脑机接口中基于 WPD 和 CSP 的特征提取[J]. *仪器仪表学报*, 2012, 33(11): 2560-2565
- [3] Subasi A, Alkan A, Koklukaya E, et al. Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing[J]. *Neural Networks*, 2005, 18(7):985-997
- [4] 罗志增, 袁飞龙, 高云园. 小波和希尔伯特变换在脑电信号消噪中的对比研究[J]. *计量学报*, 2013, 34(6):567-572
- [5] 孟宗, 李姗姗, 王亚超. 基于 LMD 和局域时频熵的旋转机械故障诊断方法[J]. *计量学报*, 2015, 36(1): 77-81
- [6] Jiang W, Zheng Z, Zhu Y, et al. Demodulation for hydraulic pump fault signals based on local mean decomposition and improved adaptive multiscale morphology analysis [J]. *Mechanical Systems and Signal Processing*, 2015, 58:179-205
- [7] Li Y, Xu M, Zhao H, et al. A new rotating machinery fault diagnosis method based on improved local mean decomposition[J]. *Digital Signal Processing*, 2015, 46(C):201-214
- [8] Smith J S. The local mean decomposition and its application to EEG perception data[J]. *Journal of the Royal Society Interface*, 2012, 11(3):703-704
- [9] Pincus S M. Approximate entropy as a measure of system complexity[C]. In: *Proceedings of the National Academy of Sciences of the United States of America*, 1991, 88(6):2297-301
- [10] Richman J S, Moorman J R. Physiological time-series analysis using approximate entropy and sample entropy [J]. *American Journal of Physiology Heart and Circulatory Physiology*, 2000, 278(6):H2039-H2049
- [11] Costa M. Multiscale entropy analysis of complex physiologic time series[J]. *Physical Review Letters*, 2002, 89(6):705-708
- [12] Blankertz B, Müller K R, Curio G, et al. The BCI Competition 2003: progress and perspectives in detection and discrimination of EEG single trials[J]. *IEEE Transactions on Biomedical Engineering*, 2004, 51(6): 1044-1051
- [13] 陈果, 周伽. 小样本数据的支持向量机回归模型参数及预测区间研究[J]. *计量学报*, 2008, 29(1):92-96
- [14] 孙会文, 伏云发, 熊馨, 等. 基于 HHT 运动想象脑电模式识别研究[J]. *自动化学报*, 2015, 41(9):1686-1692
- [15] 刘冲, 赵海滨, 李春胜, 等. 基于 CSP 与 SVM 算法的运动想象脑电信号分类[J]. *东北大学学报:自然科学版*, 2010, 31(8):1098-110

A method for extraction of motor imagery EEG features based on local mean decomposition and multiscale entropy

Zou Xiaohong * * * * *, Zhang Yibo * * * * *, Sun Yanzhen * * * * *

(* College of Information Science and Engineering, Yanshan University, Qinhuangdao 066004)

(** Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, Qinhuangdao 066004)

(*** Key Laboratory of Software Engineering of Hebei Province, Qinhuangdao 066004)

Abstract

Electroencephalogram (EEG) feature extraction is studied. Considering that traditional EEG feature extraction methods can not depict EEG features accurately, thus causing difficulties to motor imagery EEG classification under different mental tasks, this study proposes a feature extraction method based on local mean decomposition (LMD) and multiscale entropy (MSE). Firstly, the method adaptively decomposes an electroencephalogram (EEG) signal into a series of product function (PF) components with physical significance. Then, it selects effective PF components, calculates multiscale entropy, and combines multiscale entropy as eigenvectors. Finally, eigenvectors are put into the support vector machine (SVM) to identify the type of the electroencephalogram. The experimental results show that the proposed method can effectively extract the features of EEG signal, which verifies the method's effectiveness and feasibility.

Key words: electroencephalogram (EEG), feature extraction, local mean decomposition (LMD), multiscale entropy (MSE), support vector machine (SVM)