

基于弹性变分模态分解的癫痫脑电信号分类方法

景鹏 张学军 孙知信

eEpileptic electroencephalogram signal classification method based on elastic variational mode decomposition

Jing Peng Zhang Xue-Jun Sun Zhi-Xin

引用信息 Citation: *Acta Physica Sinica*, 70, 018702 (2021) DOI: 10.7498/aps.70.20200904

在线阅读 View online: <https://doi.org/10.7498/aps.70.20200904>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于变分模态分解与多尺度排列熵的生物组织变性识别

Recognition of denatured biological tissue based on variational mode decomposition and multi-scale permutation entropy

物理学报. 2019, 68(2): 028702 <https://doi.org/10.7498/aps.68.20181772>

优化递归变分模态分解及其在非线形信号处理中的应用

Application of the proposed optimized recursive variational mode decomposition in nonlinear decomposition

物理学报. 2019, 68(23): 238401 <https://doi.org/10.7498/aps.68.20191005>

基于变分模态分解-相干分析的肌间耦合特性

Intermuscular coupling characteristics based on variational mode decomposition-coherence

物理学报. 2017, 66(6): 068701 <https://doi.org/10.7498/aps.66.068701>

不同磁场构型下Richtmyer-Meshkov不稳定性的数值研究及动态模态分解

Numerical investigations of Richtmyer-Meshkov instability in different magnetic field configurations and the corresponding dynamic mode decomposition

物理学报. 2019, 68(16): 165201 <https://doi.org/10.7498/aps.68.20190410>

基于多尺度熵的Duffing混沌系统阈值确定方法

Threshold determination method of Duffing chaotic system based on multi-scale entropy

物理学报. 2020, 69(16): 160501 <https://doi.org/10.7498/aps.69.20191642>

静息态功能磁共振成像评估健康老年人认知行为的多尺度熵模型研究

Study of multiscale entropy model to evaluate the cognitive behavior of healthy elderly people based on resting state functional magnetic resonance imaging

物理学报. 2020, 69(10): 108703 <https://doi.org/10.7498/aps.69.20200050>

基于弹性变分模态分解的癫痫脑电信号分类方法*

景鹏¹⁾ 张学军^{1)2)†} 孙知信³⁾⁴⁾

1) (南京邮电大学电子与光学工程学院, 微电子学院, 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, 南京 210023)

3) (南京邮电大学, 江苏省邮政大数据技术与应用工程研究中心, 南京 210003)

4) (南京邮电大学, 国家邮政局邮政行业技术研发中心(物联网技术), 南京 210003))

(2020 年 6 月 14 日收到; 2020 年 8 月 26 日收到修改稿)

癫痫脑电信号分类对于癫痫诊治具有重要意义. 为了实现病灶性与非病灶性癫痫脑电信号的分类, 本文利用弹性网回归重构变分模态分解算法, 提出弹性变分模态分解算法并将其应用到所提癫痫脑电信号分类方法中. 该方法先将原信号分割成多个子信号, 并对各子信号进行弹性变分模态分解, 然后从分解后的不同变分模态函数中提取精细复合多尺度散布熵作为特征, 最后利用支持向量机进行分类. 针对癫痫脑电的公共数据集, 最终的实验结果表明, 准确率、灵敏度和特异度三个性能指标分别达到 92.54%, 93.22% 和 91.86%.

关键词: 弹性变分模态分解, 精细复合多尺度散布熵, 癫痫脑电

PACS: 87.85.-d, 87.85.Ng, 87.15.A-

DOI: 10.7498/aps.70.20200904

1 引言

癫痫是一种由于人脑神经系统紊乱造成脑功能障碍的疾病, 据世界卫生组织 2019 年统计, 全世界有超 5000 万人患有癫痫^[1], 对具有抗药性的癫痫患者需要进行侵入性手术治疗, 切除大脑中的部分致痫区域, 因此手术前准确识别大脑中的致痫区域至关重要. 脑电图 (electroencephalograph, EEG) 记录了人脑皮层发出的轻微电流信号, 癫痫发作时人脑中枢神经系统会出现异常使得同步神经元突然放电, 相应的脑电图就会呈现出异常波动, 从人脑致痫区域获得的脑电信号称为病灶性脑电信号, 从非致痫区域获得的脑电信号称为非病灶性脑电信号^[2], 因此对病灶性脑电信号和非病灶性脑电信号进行分类, 有助致痫区域的识别.

目前国内外对于癫痫脑电信号的分类技术主

要由特征提取与特征分类组成, 其本质可归结为模式识别问题^[3]. 特征提取先要选择合适的特征, 如高阶矩^[4]、概率密度函数^[5]、自回归模型系数^[6]、各类熵^[7-10]等均已成功应用到癫痫脑电信号的分类中. 文献^[11]提出了一种新的复杂性度量特征, 即精细复合多尺度散布熵 (refined composite multiscale dispersion entropy, RCMDE), 其作为特征在处理非线性不平稳生物信号时效果较好. 此外, 特征提取通常在时域、频域或时频域内进行, 其中又以时频域分析法应用较多, 如短时傅里叶变换^[12]、小波变换^[13]、经验模态分解等, 经验模态分解作为一种自适应时频处理方法应用较广, 但存在端点效应、模态混叠等缺陷^[14]. 由文献^[14]提出的变分模态分解 (variational mode decomposition, VMD) 是一种新的基于完全非递归思想的时频域信号分析方法, 其主要思想是将信号分解成若干个围绕在各个中心频率附近的窄带变分模态分量且中心频

* 国家自然科学基金 (批准号: 61972208, 61672299) 资助的课题.

† 通信作者. E-mail: xjzhang@njupt.edu.cn

率不断变化,与经验模态分解通过不断循环筛分获取本征模态函数不同,VMD通过寻找约束变分模型的最优解自适应获取变分模态函数(variational mode function, VMF). VMD在构建初始是针对盲信号,使用Tikhonov正则化思想去构建一个约束变分方程^[14],但本质是通过最小二乘回归加残差平方和二次惩罚项,从而达到收缩估计系数的目的,这也可以理解为使用了线性规划中的岭回归(ridge regression)来构建约束方程,虽然岭回归与最小二乘回归相比降低了估计值的方差,提高了估计精度,但是它的回归结果中包含所有的预测变量,没有进行变量选择,因此会影响模型的准确性^[15]. 而套索回归(LASSO regression)使用一次惩罚项来对变量进行收缩,相较于岭回归收缩程度要小,能选出更精确的模型,但是如果预测变量具有群组效应或者预测变量相关性较强时,则会导致估计不稳定^[15]. Zou和Hastie^[16]提出的弹性网回归(elastic net regression)方法综合了岭回归和套索回归的思想,兼有套索回归和岭回归的优点.

本文利用弹性网回归对变分模态分解算法进行改进,构建一种新的时频域信号分析方法,称为弹性变分模态分解(elastic variational mode decomposition, EVMD),并通过在EVMD域中提取精细复合多尺度散布熵作为癫痫脑电信号的特征,然后利用支持向量机(support vector machine, SVM)实现病灶性脑电信号和非病灶性脑电信号的分类. 此外,将本文所提方法应用到实际癫痫脑电数据集中进行仿真分析,以证明该方法的有效性.

2 方法

应用本文所提方法处理癫痫脑电信号的步骤如图1所示. 先利用弹性变分模态分解算法对分割后的癫痫脑电信号进行分解,从分解后的各VMF中提取RCMDE作为特征,最后利用支持向量机对特征进行分类,进而完成对病灶性脑电信号与非病灶性脑电信号的区分. 本节将对所提方法中使用到的主要算法进行叙述.

2.1 弹性变分模态分解

本文提出的弹性变分模态分解是一种完全非递归式的时频域信号分析与处理方法,与原始变分模态分解算法不同的是, EVMD利用弹性网回归

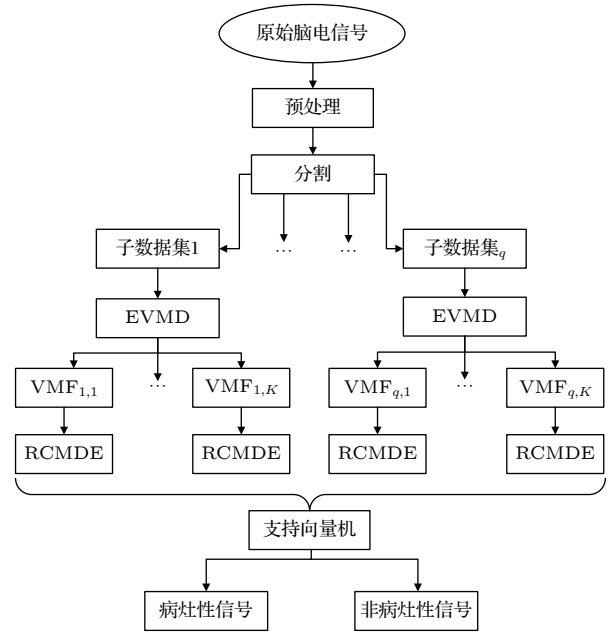


图1 弹性变分模态分解处理癫痫脑电信号的流程图

Fig. 1. Architecture of processing epileptic EEG by EVMD.

代替VMD中构建约束变分方程时使用的岭回归,即EVMD算法构建的约束变分方程中既存在L2正则化项,也存在L1正则化项. 本节将对EVMD算法的构造与求解进行详细叙述.

首先,通过以下步骤建立一个约束变分模型:

1) 通过希尔伯特变换求出每个模态函数的解析信号,进而可以获得信号的单边频谱; 2) 通过一个混合指数项将每个模态函数调制到对应中心频率的基频带上; 3) 通过弹性网回归方法施加L1正则化与L2正则化来估计带宽,即求信号梯度二范数的平方与梯度一范数的和. 建立的约束变分模型为

$$\begin{aligned} \min_{\{u_k\}, \{\omega_k\}} & \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right. \\ & \left. + \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k t} \right\|_1 \right\} \\ \text{s.t. } & \sum_k u_k(t) = x, \end{aligned} \quad (1)$$

其中, x 为原脑电信号, $\{u_k\} = \{u_1, u_2, \dots, u_K\}$ 为所有模态的集合, $\{\omega_k\} = \{\omega_1, \omega_2, \dots, \omega_K\}$ 为对应中心频率的集合, $\sum_k = \sum_{k=1}^K$, $\delta(t)$ 为冲激函数, ∂_t 为对 t 求偏导.

接着对约束变分模型求最优解. 先将拉格朗日乘法算子以及损失项加入到约束变分模型中,从而将约束变分模型转化为一个非约束变分模型,得到的增广拉格朗日函数为

$$\begin{aligned} & \Gamma(\{u_k\}, \{\omega_k\}, \lambda) \\ &= \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \times u_k(t) \right) e^{-j\omega_k t} \right] \right\|_2^2 \\ &+ \beta \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k t} \right\|_1 \\ &\|x(t) - \sum_k u_k(t)\|_2^2 + [\lambda, x(t) - \sum_k u_k(t)], \quad (2) \end{aligned}$$

其中 λ 为拉格朗日乘法算子, α 为二次惩罚因子, β 为一次惩罚因子, $\|x(t) - \sum_k u_k(t)\|_2^2$ 为损失项。

然后利用交替方向乘子法来求解 (2) 式的最优解, 迭代公式为:

$$u_k^{n+1} \leftarrow \arg \min_{u_k} \Gamma(\{u_{i < k}^{n+1}\}, \{u_{i \geq k}^n\}, \{\omega_i^n\}, \lambda^n), \quad (3)$$

$$\omega_k^{n+1} \leftarrow \arg \min_{\omega_k} \Gamma(\{u_i^{n+1}\}, \{\omega_{i < k}^{n+1}\}, \{\omega_{i \geq k}^n\}, \lambda^n), \quad (4)$$

$$\lambda^{n+1} \leftarrow \lambda^n + \eta \times \left(x - \sum_k u_k^{n+1} \right). \quad (5)$$

迭代停止条件为:

$$\sum_k \|u_k^{n+1} - u_k^n\|_2^2 / \|u_k^n\|_2^2 < \varepsilon, \quad (6)$$

其中, η 为拉格朗日乘子更新参数, ε 为收敛容限。

下面介绍模态 u_k 与中心频率 ω_k 的详细更新求解步骤, 即迭代式 (3) 式与 (4) 式的具体推导过程。

A 模态 u_k 更新

步骤 1 将 (3) 式改写为如下等值公式:

$$\begin{aligned} u_k^{n+1} = \arg \min_{u_k} & \left\{ \alpha \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right. \\ & + \beta \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k t} \right\|_1 \\ & \left. + \left\| x(t) - \sum_i u_i(t) + \frac{\lambda(t)}{2} \right\|_2^2 \right\}; \quad (7) \end{aligned}$$

步骤 2 将 (7) 式进行整体傅里叶变换:

$$\begin{aligned} \hat{u}_k^{n+1} = \arg \min_{\hat{u}_k} & \left\{ \alpha \|j\omega[(1 + \text{sgn}(\omega + \omega_k))\hat{u}_k(\omega + \omega_k)]\|_2^2 \right. \\ & + \beta \|j\omega[(1 + \text{sgn}(\omega + \omega_k))\hat{u}_k(\omega + \omega_k)]\|_1 \\ & \left. + \left\| \hat{x}(\omega) - \sum_i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right\|_2^2 \right\}; \quad (8) \end{aligned}$$

步骤 3 将 (8) 式中 $\omega + \omega_k$ 替换为 ω , 即将带有惩罚因子的项中的 ω 用 $\omega - \omega_k$ 代替, 得:

$$\begin{aligned} \hat{u}_k^{n+1} = \arg \min_{\hat{u}_k} & \left\{ \alpha \|j\omega[(1 + \text{sgn}(\omega))\hat{u}_k(\omega)]\|_2^2 \right. \\ & + \beta \|j\omega[(1 + \text{sgn}(\omega))\hat{u}_k(\omega)]\|_1 \\ & \left. + \left\| \hat{x}(\omega) - \sum_i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right\|_2^2 \right\}; \quad (9) \end{aligned}$$

步骤 4 因为实信号具有埃米尔特对称性, 因此将 (9) 式改为非负频域上的半空间积分:

$$\begin{aligned} \hat{u}_k^{n+1} = \arg \min_{\hat{u}_k} & \int_0^{+\infty} 2\alpha(\omega - \omega_k)^2 |\hat{u}_k(\omega)|^2 \\ & + \beta j(\omega - \omega_k) \hat{u}_k(\omega) \\ & + \left| \hat{x}(\omega) - \sum_i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right|^2 d\omega; \quad (10) \end{aligned}$$

步骤 5 将 (10) 式看为一个二次优化问题进行求解, 解得:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{x}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2 + \beta j(\omega - \omega_k)}. \quad (11)$$

B 中心频率 ω_k 更新

步骤 1 因为中心频率 ω_k 不出现在重构保真项中, 因此可将 (4) 式重新写为

$$\begin{aligned} \omega_k^{n+1} = \arg \min_{\omega_k} & \left\{ \alpha \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right. \\ & \left. + \beta \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k t} \right\|_1 \right\}; \quad (12) \end{aligned}$$

步骤 2 与求解模态 u_k 的操作类似, 将 (12) 式进行傅里叶变换到频域, 并最终变换成非负频域上的半空间积分:

$$\begin{aligned} \omega_k^{n+1} = \arg \min_{\omega_k} & \left\{ \int_0^{+\infty} 2\alpha(\omega - \omega_k)^2 |\hat{u}_k(\omega)|^2 \right. \\ & \left. + \beta j(\omega - \omega_k) \hat{u}_k(\omega) d\omega \right\}; \quad (13) \end{aligned}$$

步骤 3 求解 (13) 式二次优化问题, 可得:

$$\omega_k^{n+1} = \frac{\int_0^{+\infty} \omega |\hat{u}_k(\omega)|^2 + \frac{\beta j}{2\alpha} |\hat{u}_k(\omega)| d\omega}{\int_0^{+\infty} |\hat{u}_k(\omega)|^2 d\omega}. \quad (14)$$

至此, 完整的 EVMD 算法已经介绍完毕, 其具体流程可以参考算法 1 所示。

Algorithm 1: EVMD

Initialize $\hat{u}_k^1, \omega_k^1, \hat{\lambda}^1, n \leftarrow 0$

repeat

$n \leftarrow n + 1$

for $k = 1: K$ do

Update $\hat{u}_k : \hat{u}_k^{n+1}(\omega) \leftarrow$

$$\frac{\hat{x}(\omega) - \sum_{i < k} \hat{u}_i^{n+1}(\omega) - \sum_{i > k} \hat{u}_i^{n+1}(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2 + \beta j(\omega - \omega_k^n)} \quad (15)$$

Update $\omega_k : \omega_k^{n+1} \leftarrow$

$$\frac{\int_0^{+\infty} \omega |\hat{u}_k^{n+1}(\omega)|^2 + \frac{\beta j}{2\alpha} |\hat{u}_k^{n+1}(\omega)| d\omega}{\int_0^{+\infty} |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (16)$$

end for

for all $\omega \geq 0$ do

$$\hat{\lambda}^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + \eta \times \left(\hat{x}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right) \quad (17)$$

end for

until convergence: $\sum_k \frac{\|u_k^{n+1} - u_k^n\|_2^2}{\|u_k^n\|_2^2} < \varepsilon$

2.2 精细复合多尺度散布熵

精细复合多尺度散布熵是文献 [11] 在散布熵的基础上对信号进行多尺度量化得到的, 避免了散布熵因单一尺度上处理信号会出现复杂性特征提取不完全等问题. 文献 [11] 详细介绍了 RCMDE 的原理, 这里仅简述其对脑电信号处理时的计算步骤:

步骤 1 假设脑电信号 x 经弹性变分模态分解后得到的信号 u_k 是长度为 Ψ 的时间序列, 则 u_k 的第 h 个粗粒度近似信号为

$$a_{h,\gamma}^\tau = \frac{1}{\tau} \sum_{\chi=h+(\gamma-1)\tau}^{h+\gamma\tau-1} u_{k,\chi}, \quad (18)$$

其中, $\gamma = \{1, 2, \dots, N\}$, $N = \Psi/\tau$, τ 为尺度因子.

步骤 2 使用正态累积分布函数将粗粒度近似信号 a 映射到 $b = \{b_1, b_2, \dots, b_N\}$ 上:

$$b_\gamma^\tau = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{a_\gamma^\tau} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt, \quad (19)$$

其中, μ 表示均值, σ 表示粗粒度近似 a 的标准差. b 的范围从 0 到 1.

步骤 3 为了实现信号的多尺度化, 将 b 以线性变换的形式映射到 $\{1, 2, \dots, c\}$ 中, 记为 z , 即

$$z_\gamma^c = R(c \cdot b_\gamma + 0.5), \quad (20)$$

其中 $R(\cdot)$ 是取整运算, c 代表类别个数.

步骤 4 若嵌入维数标记为 m 和时间延迟标记为 d , 则时间序列 $z_i^{m,c}$ 定义为

$$z_i^{m,c} = \{z_i^c, z_{i+d}^c, \dots, z_{i+(m-1)d}^c\}, \quad (21)$$

其中 $i = \{1, 2, \dots, N - (m-1)d\}$.

步骤 5 设每一个时间序列 $z_i^{m,c}$ 对应一个散布模式 $\Theta_{v_0 v_1 \dots v_{m-1}}$, 其中 $v = 1, 2, \dots, c$. 若 $z_i^c = v_0$, $z_{i+d}^c = v_1$, $z_{i+(m-1)d}^c = v_{m-1}$, 则 $z_i^{m,c}$ 对应的散布模式为 $\Theta_{v_0 v_1 \dots v_{m-1}}$.

步骤 6 计算每个散布模式 $\Theta_{v_0 v_1 \dots v_{m-1}}$ 的概率

$$p(\Theta_{v_0 v_1 \dots v_{m-1}}) = \frac{\#(\Theta_{v_0 v_1 \dots v_{m-1}})}{N - (m-1)d}, \quad (22)$$

其中 $\#(\Theta_{v_0 v_1 \dots v_{m-1}})$ 为 $z_i^{m,c}$ 对应散布模式的个数.

步骤 7 根据香农熵的定义, 信号 u_k 的精细复合多尺度散布熵为

$$\begin{aligned} \text{RCMDE}(u_k, m, c, d, \tau) \\ = - \sum_{\Theta=1}^{c^m} \bar{p}(\Theta_{v_0 v_1 \dots v_{m-1}}) \ln \bar{p}(\Theta_{v_0 v_1 \dots v_{m-1}}), \end{aligned} \quad (23)$$

其中, $\bar{p}(\Theta_{v_0 v_1 \dots v_{m-1}}) = \frac{1}{\tau} \sum_{h=1}^{\tau} p_h^\tau$ 表示粗粒度近似信号 a 的散布模式 Θ 的平均值.

3 实验结果与分析

3.1 实验数据集

本文使用的癫痫脑电信号数据集是来自瑞士伯尔尼大学神经科提供的一个公共数据集 [17]. 该数据集数据采自 5 位患有颞叶病灶性癫痫的病人, 共包括 3750 对病灶性数据段和 3750 对非病灶性数据段, 每对数据段均包含两列数据, 分别采自同一个区域相邻的两个通道. 数据段的采样频率为 512 Hz, 时间为 20 s [18]. 本文选用该数据集中全部脑电数据段作为实验数据, 即选用 3750 对病灶性脑电数据段与 3750 对非病灶性脑电数据段, 每对数据段均包含两列数据.

3.2 实验

3.2.1 EEG 数据段分割

先将原始脑电数据经过 0.5—100 Hz 的带通滤波器进行预处理, 从而去除部分伪迹与噪声. 然

后将每对时长为 20 s 的数据段分割为时长均为 1 s 的子数据段, 且相邻两个子数据段之间的时间重叠率为 50%, 因此每个数据段共分割为 39 个子数据段. 之所以需要对原始脑电数据进行分割, 第一是因为 EVMD 算法与 VMD 一样在处理长时间非平稳非线性的脑电信号时模态的谱带会随时间剧烈变化影响分解效果, 而将信号分割为短时间的子信号, 可以一定程度规避这种局限性. 第二是因为可以减小脑电数据中随机波动点的影响, 比如假设某数据段中由于噪声等原因出现数个非预期的随机波动点, 若对整个数据段进行分析显然这些点会影响最终分析结果, 而对若干子数据段进行分析, 这些点只会影响部分子数据段的分析结果, 而对最终所求结果的影响将降低.

3.2.2 EVMD 分解

因为测试时间和被试的不同, 各频带段的脑电信号的中心频率会随着神经活动而发生轻微的变化, 简单地固定中心频率的带通滤波无法消除这种变化的影响, 而 EVMD 旨在将信号分解为围绕在中心频率附近的变分模态分量, 即变分模态函数 VMF, 且其中心频率是不断迭代变化的, 故可以捕捉到这种变化.

根据 2.1 节对 EVMD 的介绍, 可以看出 EVMD 算法在对信号进行分解之前需要分别设定: 分解模态的个数 K , 二次惩罚因子 α , 一次惩罚因子 β 三个参数. 对于惩罚因子 α , β , 其值过大会引起模态重叠, 较小会引入噪声, 本文根据经验建议将 α 设定为 2000, β 设定为 20. 对于分解模态的个数 K , 若 K 值过大会造成过分解, 产生无用分量, 同时增加计算复杂度, 若 K 值过小会造成欠分解, 使部分带限信号分解不出来造成原信号信息的丢失. 为寻求合适的 K 值, 本文先随机选择一对病灶性脑电数据段, 并分别进行 3 重、4 重、5 重变分模态分解, 即将 K 值分别设为 3, 4, 5. 图 2 所示为不同 K 值下, 各个变分模态函数分量其中心频率 ω 随迭代次数的变化曲线. 可以看出, 当 $K = 3$ 时, 变分模态函数 VMF1—VMF3 的中心频率均在 40 Hz 以下. 当 $K = 4$ 时, 变分模态函数 VMF1—VMF4 的中心频率仍在 40 Hz 以下. 但当 $K = 5$ 时, 可以看到 VMF5 的中心频率超过了 50 Hz. 一方面由于脑电活动的信息大都包含在低频带 (频率 < 40 Hz), 另一方面为了防止信号的过分解或欠分解, 本文选取 $K = 4$.

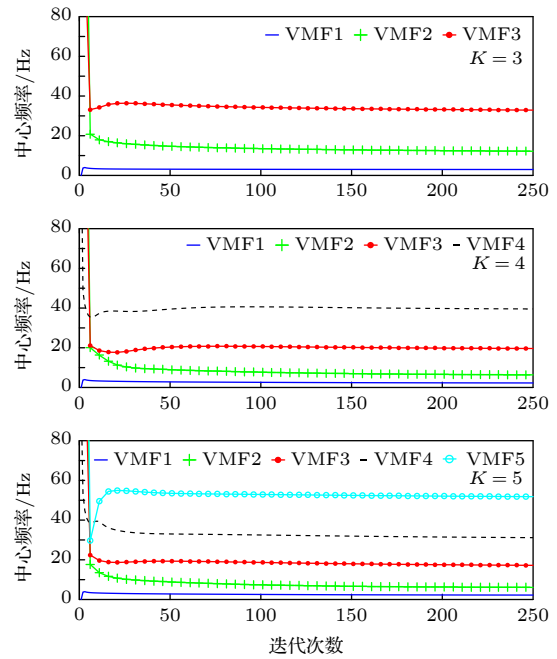


图 2 不同 K 值下中心频率随迭代次数的变化曲线

Fig. 2. The curves of the center frequency with the number of iterations under different K values.

3.2.3 RCMDE 提取特征

根据 (23) 式, 计算 RCMDE 需要分别设置: 嵌入维度 m , 类别个数 c , 时间延迟 d 以及尺度因子 τ 四个参数. 类别个数 c 若过大会导致具有较大差异的两个量被归为同一类, 过小则导致具有较小差异的两个量被归为不同类, 根据文献 [11] 的建议, 将 c 值设为 6. 对于嵌入维度 m , 为了保证统计可靠性, 文献 [12] 中建议 $c^m < \psi$, 其中 ψ 为数据长度, 本文将采样率 512 Hz, 采样时间 20 s 的数据段分割为时长为 1 s 的子数据段, 因此 $\psi = 512$, 而 $6^4 = 1296 > \psi$, 故将 m 值设为 3. 时间延迟 d 为正整数, 但 $d > 1$ 会导致模态混叠, 故取 1. 尺度因子 τ 决定信号粗粒化程度, 为了选取合适的 τ 值, 先将 τ 最大值设为 15.

选用数据集中全部 3750 对病灶性脑电信号和 3750 对非病灶性脑电信号, 对分割后的子数据段进行 4 重 EVMD 分解, 从分解后得到的 4 个 VMF (VMF1—VMF4) 中分别提取 15 个尺度因子下的 RCMDE, 其特征熵值的均值加减标准差随尺度因子变化曲线如图 3 所示. 图 3 中不同尺度因子下的 RCMDE 均值用曲线相连, 每个均值点上下两个点即表示加减标准差. 从图 3 可以看出, 从非病灶性脑电信号提取的 RCMDE 特征熵值基本比病

灶性脑电信号提取的 RCMDE 熵值大, 这说明人脑非致痫区域的神经活动相对致痫区域更活跃, 也说明非致痫区域提取的脑电信号相较致痫区域提取的脑电信号更随机、更不平稳, 这一结果也与文献 [18] 的研究结果一致. 此外, 之所以会出现从不同 VMF 中提取的两类信号 RCMDE 特征熵值的均值差异性不同的现象, 是因为不同 VMF 是体现不同中心频率上的脑电信号数据, 而人脑非致痫区域与致痫区域在不同频带段上的神经活动会不同, 比如经 4 重 EVMD 分解下 VMF3 代表的是中心频率为 20 Hz 附近频带段的信号, 此时两类脑电信号 RCMDE 熵值的均值的差值与其他 VMF 相比较小, 表明在 20 Hz 附近人脑非致痫区域与致痫区域的神经活动更相近, 产生的脑电信号的差异性较其他频带段较小. 但是随着尺度因子的增大, 总体非病灶性脑电信号与病灶性脑电信号提取的

RCMDE 特征熵值的均值的差值会有一个先增大再减小的过程, 尤其是 VMF2 尺度因子等于 15 时, VMF3 尺度因子大于 12 时, 会出现病灶性脑电信号提取的 RCMDE 特征熵值的均值更大的反常现象, 这是因为信号被过度粗粒化从而出现失真, 说明尺度因子不能过大. 本文从计算量的角度综合考虑将尺度因子 τ 设为 7, 即每个数据段提取 7 个尺度上的 RCMDE 值.

为了更好评估从非病灶性脑电信号提取的 RCMDE 特征与从病灶性脑电信号提取的 RCMDE 特征的区分度, 这里使用学生检验的 p 值来评估其统计学上的差异性, 其结果如表 1 所列. 从表 2 可以看出数据经 4 重 EVMD 分解后得到的 4 个 VMF 中提取的 RCMDE 特征 p 值均小于 0.05, 因此在统计学上具有显著差异, 说明 RCMDE 可以作为区别病灶性和非病灶性癫痫脑电数据的分类特征.

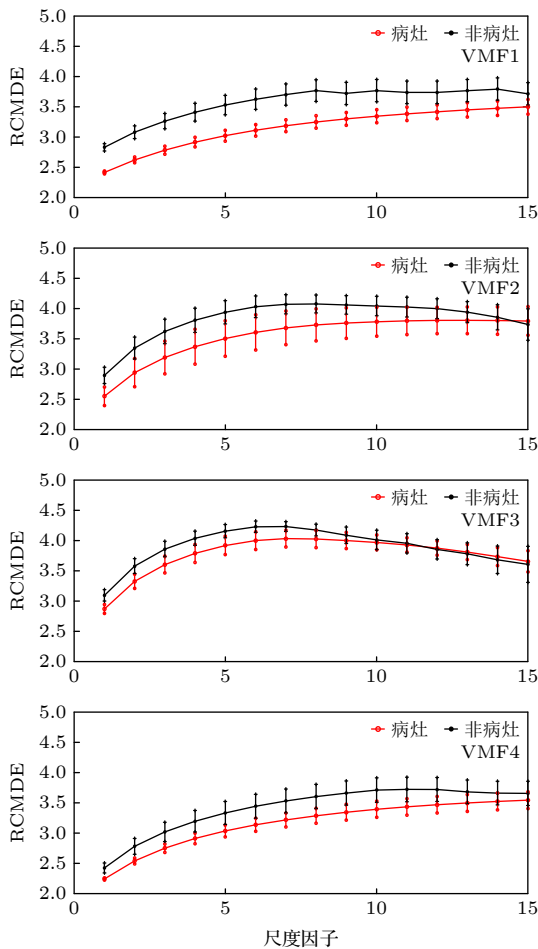


图 3 从各 VMF 中提取的 RCMDE 熵值的均值 (\pm 标准差) 随尺度因子变化曲线

Fig. 3. The curve of mean value (\pm SD) of RCMDE computed from VMF.

表 1 从各 VMF 中提取的 RCMDE 特征 p 值

Table 1. The p values of RCMDE computed from VMF.

	VMF1	VMF2	VMF3	VMF4
p 值	1.17×10^{-4}	2.38×10^{-2}	4.42×10^{-2}	7.50×10^{-3}

表 2 EVMD 与 VMD 实验结果对比

Table 2. Comparison of experimental result between EVMD and VMD.

指标	准确度/%	灵敏度/%	特异度/%
EMVD	92.54	93.22	91.86
VMD	89.49	87.56	88.12

3.3 分类结果

选用数据集中全部 3750 对病灶性癫痫脑电数据段和 3750 对非病灶性癫痫脑电数据段, 按上述方法进行处理, 并将得到的每段脑电数据的 RCMDE 特征送入 SVM 进行特征分类, SVM 选用线性核函数, 通过网格搜索法将惩罚参数设定为 0.52, 并重复进行 10 次 5 折交叉验证实验, 采用准确度、灵敏度和特异度这三个指标对最终分类结果进行度量, 结果如图 4 所示. 由图 4 可知 10 次 5 折交叉验证实验的平均准确度、灵敏度和特异度分别可达 92.54%, 93.22%, 91.86%.

为了比较本文提出的 EVMD 算法与原始 VMD 在处理病灶性癫痫脑电信号与非病灶性信号分类中的性能, 将本文实验中的 EVMD 换成原始 VMD, 其余步骤不变, 重复进行实验, 对比实验的结果如

表 2 所列. 可以看出, EVMD 算法相较原始 VMD 算法, 三个性能指标均有 3 个百分点以上的提高, 因此按本文所提方法进行病灶性癫痫脑电信号与非病灶性癫痫脑电信号的分类实验中, 本文所提 EVMD 算法相较原始 VMD 算法性能具有一定优势.

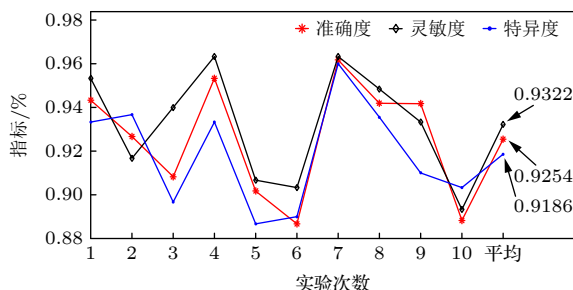


图 4 10 次 5 折交叉验证实验结果折线图

Fig. 4. The line chart of the results by 5-fold cross validation for 10 times.

对相同的实际癫痫脑电信号数据集, 将本文方法得到的结果与其他文献的结果相比较, 结果如表 3 所列, 可以看出本文所提方法性能优于其他方法.

表 3 本文方法与其他方法对比

Table 3. Comparison between proposed method and previously published methods.

	提取方法	特征	分类器	灵敏度、特异度、准确度/%
文献[7]	EMD	平均香农熵、样本熵等	LSSVM	90, 84, 87
文献[8]	TQWT	模糊熵	LSSVM	83.86, 85.46, 84.67
文献[19]	MFDFA	多重分形	SVM	92.5, 92.7, 92.2
文献[20]	EWT	RPS, CTM	LSSVM	88, 92, 90
本文	EVMD	RCMDE	SVM	93.22, 91.86, 92.54

4 结 论

本文利用弹性网回归重构了变分模态分解的约束方程, 提出一种新的时频域信号分析与处理算法, 称为弹性变分模态分解, 并给出了具体推导过程. 提出一种基于弹性变分模态分解, 从分解后的 VMF 中提取 RCMDE 特征, 并利用支持向量机对其进行分类的病灶性癫痫脑电信号与非病灶性癫痫脑电信号的分类方法. 应用本文提出的方法对公共癫痫脑电数据集进行处理, 得到最终准确率为 92.54%, 灵敏度 93.22%, 特异度 91.86%, 并与针对同一公共数据集的其他处理方法进行比较, 证明了该方法的有效性. 然而, 弹性变分模态分解的参数

(分解模态的个数 K 、惩罚因子 α 与 β) 只能通过经验设定或者根据处理信号不断试错设定, 无法自动获取最优参数, 因为弹性变分模态分解与原始变分模态分解结构类似, 所以现在针对变分模态分解参数的优化方法, 如相关性分析优化^[21]、粒子群算法优化^[22,23]等方法, 理论上应该也可以应用到本文提出的弹性变分模态算法上, 因此, 对弹性变分模态分解的参数优化将是后续研究的一个可能方向.

参考文献

- [1] World Health Organization <http://www.who.int/news-room/fact-sheets/detail/epilepsy/> [2019-6-20]
- [2] Andrzejak R G, Schindler K, Rummel C 2012 *Phys. Rev. E* **86** 046206
- [3] Zhang R, Song J L, Hu W F 2016 *J. Northwest Univ. (Nat. Sci.)* **46** 781 (in Chinese) [张瑞, 宋江玲, 胡文凤 2016 西北大学学报(自然科学版) **46** 781]
- [4] Alam S, Bhuiyan M 2013 *IEEE J. Biomed. Health Inf.* **17** 312
- [5] Das A, Bhuiyan M, Alam S 2014 *Signal Image Video Process.* **10** 259
- [6] Rahman M, Bhuiyan M, Das A 2019 *Biomed. Signal Process. Control* **50** 72
- [7] Sharma R, Pachori R, Acharya U 2015 *Entropy* **17** 669
- [8] Abhijit B, Ram B P 2017 *Entropy* **19** 99
- [9] Xie P, Yang F M, Li X X, Yang Y, Chen X L, Zhang L T 2016 *Acta Phys. Sin.* **65** 118701 (in Chinese) [谢平, 杨芳梅, 李欣欣, 杨勇, 陈晓玲, 张利泰 2016 物理学报 **65** 118701]
- [10] Wang Y, Hou F Z, Dai J F, Liu X F, Li J, Wang J 2014 *Acta Phys. Sin.* **63** 218701 (in Chinese) [王莹, 侯凤贞, 戴加飞, 刘新峰, 李锦, 王俊 2014 物理学报 **63** 218701]
- [11] Azami H, Rostaghi M, Abasolo D, Escudero J 2017 *IEEE Trans. Biomed. Eng.* **64** 2872
- [12] Kiyim K M, Guler I, Dizibuyuk A, Akin M 2005 *Comput. Biol. Med.* **35** 603
- [13] Zhang T, Chen W Z, Li M Y 2016 *Acta Phys. Sin.* **65** 038703 (in Chinese) [张涛, 陈万忠, 李明阳 2016 物理学报 **65** 038703]
- [14] Dragomiretskiy K, Zosso P 2014 *IEEE Trans. Signal Process.* **62** 531
- [15] Zhang Z, Liang F Z 2013 *J. Harbin Univ. Com. (Nat. Sci.)* **29** 592 (in Chinese) [张哲, 梁冯珍 2013 哈尔滨商业大学学报(自然科学版) **29** 592]
- [16] Zou H, Hastie T 2005 *J. R. Stat. Soc.* **67** 301
- [17] Andrzejak R G <http://www.dtic.upf.edu/~ralph/> [2019-3-20]
- [18] Andrzejak R G, Schindler K, Rummel C 2012 *Physical Review E* **86** 046206
- [19] Chatterjee S, Pratiher S, Bose R 2017 *IET Sci. Meas. Technol.* **11** 1014
- [20] Abhijit B, Manish S, Ram B P, Pradip S Rajendra A 2018 *Neural Comput. Appl.* **29** 47
- [21] Li Z P, Chen J L, Zi Y Y, Pan J 2017 *Mech. Syst. Signal Proc.* **85** 512
- [22] Wang X B, Yang Z X, Yan X A 2018 *IEEE-ASME Trans. Mech.* **23** 68
- [23] Wang Z J, He G F, Du W H, Zhou J, Han X F, Wang J T, He H H, Guo X M, Wang J Y, Kou Y F 2019 *IEEE Access* **7** 44871

Epileptic electroencephalogram signal classification method based on elastic variational mode decomposition^{*}

Jing Peng¹⁾ Zhang Xue-Jun^{1)2)†} Sun Zhi-Xin³⁾⁴⁾

1) (*College of Electronic and Optical Engineering & College of Microelectronics, Nanjing University of Posts and Telecommunications, Nanjing 210023, China*)

2) (*National and Local Joint Engineering Laboratory of RF Integration and Micro-Assembly Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China*)

3) (*Post Big Data Technology and Application Engineering Research Center of Jiangsu Province, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*)

4) (*Post Industry Technology Research and Development Center of the State Posts Bureau (Internet of Things Technology), Nanjing University of Posts and Telecommunications, Nanjing 210003, China*)

(Received 14 June 2020; revised manuscript received 26 August 2020)

Abstract

Epilepsy is an extensive nervous system disease nowadays. Electroencephalogram (EEG) can capture the abnormal discharge of nerves in the brain duration of seizure and provide a non-invasive way to identify epileptogenic sites in the brain. In order to distinguish between focal epilepsy EEG signal and non-focal epilepsy EEG signal, in this paper we propose an automated epileptic EEG detection method based on the elastic variational mode decomposition (EVMD). The proposed EVMD algorithm is a method of analyzing the signals and also a processing method in time-frequency domain, in which the elastic net regression is used to reconstruct a constrained variational model in variational mode decomposition (VMD). Used in the VMD algorithm is the Tikhonov regularization that is also statistically called ridge regression as a solution of recovering the unknown signal and assessing the bandwidth of a mode, namely the variational equation constructed by VMD only has L2 norm. However, the ridge regression cannot select variables when the equation has multiple variables. Another regression method, called lasso regression, only has L1 norm and can select a more accurate model from multiple variables, but it has worse performance when variables have group effect or co-linearity. The elastic net regression has advantages of ridge regression and lasso regression, in other word, the variational equation constructed by EVMD has both L1 regularization item and L2 regularization item, so in this paper we propose the EVMD by elastic net regression. In addition, in this paper the EVMD is used to distinguish between focal epilepsy EEG signal and non-focal epilepsy EEG signal. Firstly, the original EEG signals are divided into several sub-signals where the test signals are divided into sub-signals with shorter durations by time series and a reasonable time overlap is kept between successive sub-signals. After that each sub-signal is decomposed into intrinsic mode functions by using the EVMD. Furthermore, the refined composite multiscale dispersion entropy (RCMDE) as feature is extracted from each intrinsic mode function where a Student's t-test is used to assess the statistical differences between RCMDEs extracted from focal and non-focal EEG signals respectively. Finally, the support vector machine (SVM) is used to classify their features. For an epilepsy EEG signals public data set, the final experimental results show that the performance indices of accuracy, sensitivity, and specificity can reach 92.54%, 93.22% and 91.86% respectively.

Keywords: elastic variational mode decomposition, refined composite multiscale dispersion entropy, epileptic electroencephalogram

PACS: 87.85.-d, 87.85.Ng, 87.15.A-

DOI: 10.7498/aps.70.20200904

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 61972208, 61672299).

[†] Corresponding author. E-mail: xjzhang@njupt.edu.cn