



(12)发明专利申请

(10)申请公布号 CN 108875604 A

(43)申请公布日 2018.11.23

(21)申请号 201810552305.1

(22)申请日 2018.05.31

(71)申请人 中国地质大学(武汉)

地址 430000 湖北省武汉市洪山区鲁磨路
388号

(72)发明人 吴敏 万雄波 方泽林 万婷
杜玉晓

(74)专利代理机构 武汉知产时代知识产权代理
有限公司 42238

代理人 郝明琴

(51)Int.Cl.

G06K 9/00(2006.01)

A61B 5/0476(2006.01)

A61B 5/00(2006.01)

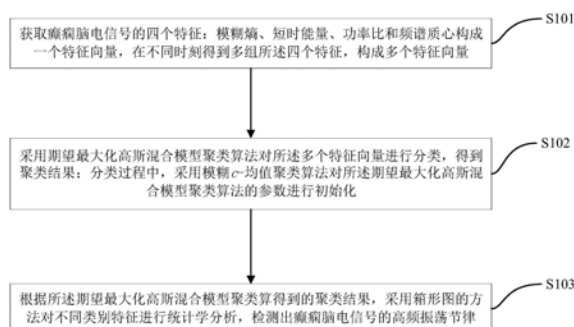
权利要求书3页 说明书10页 附图2页

(54)发明名称

基于模糊聚类优化高斯混合模型的高频振荡节律检测方法

(57)摘要

本发明提供了基于模糊聚类优化高斯混合模型的高频振荡节律检测方法,基于聚类分析的方法对高频振荡节律进行检测,选取模糊熵、短时能量、功率比和频谱质心为癫痫脑电信号的特征,将其构成特征向量作为聚类算法的输入,采用期望最大化高斯混合模型聚类算法对该特征向量进行分类,并采用模糊c-均值聚类算法得到期望最大化高斯混合模型聚类算法的初始化参数;选取中位数和四分位距分析每类的统计学特征,检测出高频振荡节律。本发明的有益效果:提高癫痫脑电信号的高频振荡节律的检测速度,帮助医生进行癫痫诊断和致痫灶切除。



1. 基于模糊聚类优化高斯混合模型的高频振荡节律检测方法, 其特征在于: 包括以下步骤:

S101: 获取癫痫脑电信号的四个特征: 模糊熵、短时能量、功率比和频谱质心构成一个特征向量, 在不同时刻得到多组所述四个特征, 构成多个特征向量;

S102: 采用期望最大化高斯混合模型聚类算法对所述多个特征向量进行分类, 得到聚类结果;

分类过程中, 采用模糊c-均值聚类算法对所述期望最大化高斯混合模型聚类算法的参数进行初始化; 初始化的具体步骤为:

S201: 随机产生初始聚类中心 $\mathbf{v}_c, c=1, 2, \dots, C, C$ 是聚类数目;

S202: 根据聚类中心 $\mathbf{v}_c$, 计算第 i 个特征向量属于第 c 类的隶属度值 μ_{ic} :
$$\mu_{ic} = \left(\sum_{k=1}^C \left(\frac{\|\mathbf{V}_i - \mathbf{v}_c\|}{\|\mathbf{V}_i - \mathbf{v}_k\|} \right)^{\frac{2}{p-1}} \right)^{-1},$$

其中, μ_{ic} 满足条件: $\mu_{ic} \in [0, 1], \forall i, c, \sum_{c=1}^C \mu_{ic} = 1, \forall i, 0 < \sum_{i=1}^{N_f} \mu_{ic} < N_f, \forall c, C$ 是聚类数目, p 是加权模糊参数, $\mathbf{V}_i$ 是所述 D 维特征向量, 且由提取的所述模糊熵、所述短时能量、所述功率比和所述频谱质心构成;

S203: 根据公式 (1) 更新聚类中心 $\mathbf{v}_c$:

$$\mathbf{v}_c = \frac{\sum_{i=1}^{N_f} \mu_{ic}^m \mathbf{V}_i}{\sum_{i=1}^{N_f} \mu_{ic}^m} \quad (1)$$

其中, $\mathbf{V}_i$ 是所述 D 维特征向量, 由所述模糊熵、所述短时能量、所述功率比和所述频谱质心构成, N_{f1} 是所述 D 维特征向量 $\mathbf{V}_i$ 的长度, μ_{ic} 是第 i 个特征向量属于第 c 类的隶属度值;

S204: 当满足 $\max_{1 \leq c \leq C} \|\mathbf{v}_{c, \text{new}} - \mathbf{v}_{c, \text{old}}\| < \varepsilon$ 时, 所述模糊c-均值聚类算法结束; 其中, $\mathbf{v}_{c, \text{new}}$ 是更新后的聚类中心, $\mathbf{v}_{c, \text{old}}$ 是更新前的聚类中心, ε 是设定的精度误差;

根据所述模糊c-均值聚类算法, 通过公式 (2)、公式 (3) 和公式 (4) 分别得到所述基于期望最大化高斯混合模型的聚类算法的初始化参数: 权重 W_c 、聚类中心 $\mathbf{u}_c$ 和协方差矩阵 Σ_c :

$$W_c = \frac{N_c}{N_f} \quad (2)$$

$$\mathbf{u}_c = \mathbf{v}_c \quad (3)$$

$$\Sigma_c = \text{cov}(\mathbf{V}_i^c) \quad (4)$$

其中, N_c 是第 c 类特征向量的数目之和, $\mathbf{V}_i^c$ 是属于第 c 类的特征向量构成的矩阵;

S103: 根据所述期望最大化高斯混合模型聚类算得到的聚类结果, 采用箱形图的方法对不同类别特征进行统计学分析, 检测出癫痫脑电信号的高频振荡节律。

2. 如权利要求1所述的基于模糊聚类优化高斯混合模型的高频振荡节律检测方法, 其特征在于: 步骤S101中, 所述模糊熵的具体计算步骤如下:

S301: 获取癫痫脑电信号时间序列 $X = [x(1), x(2), \dots, x(N)]$;

S302:将所述癫痫脑电信号时间序列按顺序构造 m 维向量 $X_i^m = [x(i), x(i+1), \dots, x(i+m-1)] - u_0(i) \cdot \mathbf{I}$, 其中, $i=1, 2, \dots, N-m+1$, m 为正整数, 且 $m \geq 1$, $\mathbf{I}$ 是元素全为1的 m 维行向量, $u_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} x(i+j)$;

S303:计算出向量 X_i^m 与 X_j^m 之间的距离 $d_{ij}^m, d_{ij}^m = \max |x(i+k) - u_0(i) - (x(j+k) - u_0(j))|$, 其中, $i, j \in 1, 2, \dots, N-m+1, j \neq i, k=0, 1, \dots, m-1$;

S304:根据所述距离 d_{ij}^m , 采用模糊隶属度函数 $\mu(d_{ij}^m, n, r)$, 计算所述 X_i^m 与所述 X_j^m 之间的相似度 $D_{ij}^m, D_{ij}^m = \mu(d_{ij}^m, n, r) = \exp \left[-\left(\frac{d_{ij}^m}{r} \right)^n \right]$, 其中, D_{ij}^m 为 X_i^m 与 X_j^m 之间的相似度, r 为相似容忍度;

S305:根据相似度 D_{ij}^m , 得到函数 $\Phi^m(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m-1} \sum_{j=1, j \neq i}^{N-m} D_{ij}^m \right)$, 其中, D_{ij}^m 为 X_i^m 与 X_j^m 之间的相似度, N 为所述癫痫脑电信号时间序列的长度, m 为所述向量 $\mathbf{I}$ 的维数;

S306:将所述向量 $\mathbf{I}$ 的维数增加到 $m+1$, 得到函数 $\Phi^{m+1}(n, r)$;

S307:根据所述函数 $\Phi^m(n, r)$ 和所述函数 $\Phi^{m+1}(n, r)$, 得到所述模糊熵:FuzzyEn(n, r, N) $= \ln \Phi^m(n, r) - \ln \Phi^{m+1}(n, r)$;

采用均方法计算得到所述短时能量的估计值: $E^*(t) = \frac{1}{N} \sum_{k=t-N+1}^t x^2(k)$, 其中, $E^*(t)$ 为所述短时能量的估计值, $x(k)$ 为脑电信号的幅值;

所述功率比 R 的计算公式为: $R = \frac{P_{[80-200]}}{P_{[250-500]}}$, 其中, $P_{[80-200]}$ 是疑似高频振荡节律在ripples带宽内的功率, $P_{[250-500]}$ 是疑似高频振荡节律在fast ripples带宽内的功率;

所述频谱质心 f_c 的计算公式为: $f_c = \frac{\sum_{k=0}^{L/2} \frac{k}{LT} |M(k)|^2}{\sum_{k=0}^{L/2} |M(k)|^2}$, 其中, T 是采样周期, L 是窗长, $M(k)$

是多窗口功率谱密度估计值, 且 $M(k) = \sum_{l=0}^{L-1} w(l)x(l)e^{-j(2\pi/L)lk}$, $w(l)$ 是长度为 L 的汉明窗。

3. 如权利要求1所述的基于模糊聚类优化高斯混合模型的高频振荡节律检测方法, 其特征在于:在步骤102中, 基于期望最大化高斯混合模型的聚类算法进行分类的具体步骤为:

S401:采用模糊 c -均值聚类算法得到所述基于期望最大化高斯混合模型的聚类算法的初始化参数;

S402:根据所述初始化参数, 通过公式(5)计算 D 维特征向量 V_i 属于第 c 类的概率 P_{ic}^t :

$$P_{ic}^t = \frac{W_c^t P_c(V_i | \mathbf{u}_c^t, \Sigma_c^t)}{P(V_i)^t}, t=1, 2, \dots, T \quad (5)$$

其中, t 是迭代次数, $T \geq 1$, W_c^t 、 u_c^t 和 Σ_c^t 分别是第 t 次迭代的权重、聚类中心和协方差矩阵, $P(V_i)^t$ 是第 t 次迭代所述 D 维特征向量 V_i 对应的概率;

计算所述特征向量 V_i 对应的概率的公式是:

$$P(V_i) = \sum_{c=1}^C W_c P_c(V_i | u_c, \Sigma_c), i=1, 2, \dots, N_f \quad (6)$$

其中, C 是聚类数目, W_c 是第 c 个高斯成分的权重, 并且满足 $\sum_{c=1}^C W_c = 1$, $c=1, 2, \dots, C$, u_c 为聚类中心, $u_c \in R^D$, Σ_c 为协方差矩阵, $\Sigma_c \in R^{D \times D}$;

S403: 根据计算出的所述 D 维特征向量 V_i 属于第 c 类的概率, 通过公式 (7)、公式 (8) 和公式 (9), 更新高斯混合模型的参数, 分别得到第 $t+1$ 次迭代的权重 W_c^{t+1} 、聚类中心 u_c^{t+1} 和协方差矩阵 Σ_c^{t+1} :

$$W_c^{t+1} = \frac{N_c^t}{N_f} \quad (7)$$

$$u_c^{t+1} = \frac{1}{N_c^t} \sum_{i=1}^{N_f} P_{ic}^t V_i \quad (8)$$

$$\Sigma_c^{t+1} = \frac{1}{N_c^t} \sum_{i=1}^{N_f} P_{ic}^t (V_i - u_c^{t+1}) (V_i - u_c^{t+1})' \quad (9)$$

其中, P_{ic}^t 是所述 D 维特征向量 V_i 属于第 c 类的概率, N_c^t 是所述 D 维特征向量 V_i 属于第 c 类的总数, 且 $N_c^t = \sum_{i=1}^{N_f} P_{ic}^t$;

S404: 根据第 $t+1$ 次迭代的权重 W_c^{t+1} 、聚类中心 u_c^{t+1} 和协方差矩阵 Σ_c^{t+1} , 通过公式 (5) 得到第 $t+1$ 次迭代所述 D 维特征向量 V_i 对应的概率 $P(V_i)^{t+1}$, 再通过公式 (10) 计算得到第 $t+1$ 次迭代的对数似然函数:

$$L^{t+1} = \sum_{i=1}^{N_f} \log P(V_i)^{t+1} \quad (10)$$

其中, $P(V_i)^{t+1}$ 为第 $t+1$ 次迭代时所述 D 维特征向量 V_i 对应的概率;

当 $|L^{t+1} - L^t| < E$ 且 $t < T-1$, 或者当 $t = T$ 时, 结束所述基于期望最大化高斯混合模型的聚类算法, 得到聚类结果; 其中, E 是设定的误差阈值。

4. 如权利要求1所述的基于模糊聚类优化高斯混合模型的高频振荡节律检测方法, 其特征在于: 在步骤S103中, 使用箱形图得到中位数和四分位距两个指标, 对不同类别的统计学特征进行分析, 确定出每个类别随时间变化的波形状态, 从而检测出高频振荡节律。

基于模糊聚类优化高斯混合模型的高频振荡节律检测方法

技术领域

[0001] 本发明涉及癫痫脑电信号处理领域,尤其涉及基于模糊聚类优化高斯混合模型的高频振荡节律检测方法。

背景技术

[0002] 癫痫是一种常见的神经系统疾病,患病率大约占世界人口的1%。目前我国有900万以上的癫痫患者,该数目仍在以每年65万-70万的速度增长。大部分癫痫患者通过抗癫痫药物进行治疗。自上个世纪初发现苯巴比妥能治疗癫痫以后,不断有抗癫痫药物如苯妥英、乙琥胺、卡马西平、丙戊酸等问世,它们能使70%癫痫病人的发作获得控制,其中大部分病人经过常规药物治疗后可终生不再发病。但仍有30%的病患被确认为难治性癫痫,需要采取外科治疗等方法的干预。上个世纪80年代后,陆续研制了不少抗癫痫新药如托吡酯、拉莫三嗪、胺己烯酸、加巴喷丁、奥卡西平等,使30%左右的难治性癫痫患者也看到了曙光。在影像学、电生理学等发展基础上,外科治疗的进步使难治性癫痫获得更多控制发作的机会。

[0003] 难治性癫痫是指经过常规的、系统的抗癫痫药物治疗,其血液内抗癫痫药物浓度保持在有效的范围内,但仍不能控制病人的癫痫发作,并严重影响病人的工作、学习或正常生活者;其发作频率在每月2-4次以上,病程在4年以上。

[0004] 癫痫特征波形提取和癫痫病灶的精确定位是术前评估的关键技术性问题。目前对于致痫灶定位方法有临床症状学评估、脑电图、脑磁图、影像学评估等。脑电图是最常用的癫痫病灶定位手段,通过24小时观测脑电记录,可以捕捉患者的发病过程,实现即时观测分析和定位。致痫灶是患者癫痫发作期或者发作间期的病变脑皮质区。通常致痫灶分为发作始发区,激惹灶,致病病变和功能缺失区,其中发作始发区是最为有效和显著的标志,可以看做是致痫灶的替代指标。

[0005] 传统的脑电图方法重点关注的是40Hz以下的脑电信号,通过提取该频段脑电频率成分来定位癫痫发作始发区。但是该方法检测的低频频率成分很容易受到其他信号的干扰,使得定位结果出错,导致切除手术失败。同时,该方法非常耗时,定位时间大约为24-72小时,加大了手术风险。因此,寻找一种新的标志物或者快速定位方法迫在眉睫。

[0006] 近二十年来,越来越多的研究者们开始关注80Hz以上癫痫脑电信号的高频振荡节律。在癫痫发作潜伏期,患者大脑开始发生病理性变化。在病变前后的大脑中,高频振荡节律有着显著的不同。大量研究表明高频振荡节律可以作为癫痫发作始发区的生物标志,它们在癫痫发作始发区的发生率更高。相较于传统的基于痫样放电来定位始发区,高频振荡节律不依赖于记录患者的惯常发作,减少了记录的时间,对某些自发发作稀少或置入电极后很长时间不发作的患者,应用发作间期高频振荡节律定位致痫灶减少了长程记录给患者带来的痛苦和花费,也降低了颅内感染的可能性,较大幅度地提高了手术疗效。

发明内容

[0007] 本发明提供了基于模糊聚类优化高斯混合模型的高频振荡节律检测方法。基于模

糊聚类优化高斯混合模型的高频振荡节律检测方法,主要包括以下步骤:

[0008] S101:获取癫痫脑电信号的四个特征:模糊熵、短时能量、功率比和频谱质心构成一个特征向量,在不同时刻得到多组所述四个特征,构成多个特征向量;

[0009] S102:采用期望最大化高斯混合模型聚类算法对所述多个特征向量进行分类,得到聚类结果;

[0010] 分类过程中,采用模糊c-均值聚类算法对所述期望最大化高斯混合模型聚类算法的参数进行初始化;初始化的具体步骤为:

[0011] S201:随机产生初始聚类中心 $v_c, c=1, 2, \dots, C, C$ 是聚类数目;

[0012] S202:根据聚类中心 v_c ,计算第 i 个特征向量属于第 c 类的隶属度值 μ_{ic} :

$$\mu_{ic} = \left(\sum_{k=1}^C \left(\frac{\|V_i - v_c\|}{\|V_i - v_k\|} \right)^{\frac{2}{p-1}} \right)^{-1}, \text{其中, } \mu_{ic} \text{ 满足条件: } \mu_{ic} \in [0, 1], \forall i, c, \sum_{c=1}^C \mu_{ic} = 1, \forall i, 0 < \sum_{i=1}^{N_{fl}} \mu_{ic} < N_{fl}, \forall c,$$

C 是聚类数目, p 是加权模糊参数, V_i 是所述 D 维特征向量,且由提取的所述模糊熵、所述短时能量、所述功率比和所述频谱质心构成;

[0013] S203:根据公式(1)更新聚类中心 v_c :

$$v_c = \frac{\sum_{i=1}^{N_{fl}} \mu_{ic}^m V_i}{\sum_{i=1}^{N_{fl}} \mu_{ic}^m} \quad (1)$$

[0015] 其中, V_i 是所述 D 维特征向量,由所述模糊熵、所述短时能量、所述功率比和所述频谱质心构成, N_{fl} 是所述 D 维特征向量 V_i 的长度, μ_{ic} 是第 i 个特征向量属于第 c 类的隶属度值;

[0016] S204:当满足 $\max_{1 \leq c \leq C} \|v_{c, \text{new}} - v_{c, \text{old}}\| < \varepsilon$ 时,所述模糊c-均值聚类算法结束;其中, $v_{c, \text{new}}$ 是更新后的聚类中心, $v_{c, \text{old}}$ 是更新前的聚类中心, ε 是设定的精度误差;

[0017] 根据所述模糊c-均值聚类算法,通过公式(2)、公式(3)和公式(4)分别得到所述基于期望最大化高斯混合模型的聚类算法的初始化参数:权重 W_c 、聚类中心 u_c 和协方差矩阵 Σ_c :

$$W_c = \frac{N_c}{N_{fl}} \quad (2)$$

$$u_c = v_c \quad (3)$$

$$\Sigma_c = \text{cov}(V_i^c) \quad (4)$$

[0021] 其中, N_c 是第 c 类特征向量的数目之和, V_i^c 是属于第 c 类的特征向量构成的矩阵;

[0022] S103:根据所述期望最大化高斯混合模型聚类算得到的聚类结果,采用箱形图的方法对不同类别特征进行统计学分析,检测出癫痫脑电信号的高频振荡节律。

[0023] 进一步地,步骤S101中,所述模糊熵用来衡量癫痫脑电信号的复杂度,所述模糊熵的具体计算步骤如下:

[0024] S301:获取癫痫脑电信号时间序列 $X = [x(1), x(2), \dots, x(N)]$;

[0025] S302:将所述癫痫脑电信号时间序列按顺序构造 m 维向量 $X_i^m = [x(i), x(i+1), \dots,$

$x(i+m-1)]-u_0(i) \cdot \mathbf{I}$, 其中, $i=1, 2, \dots, N-m+1$, m 为正整数, 且 $m \geq 1$, $\mathbf{I}$ 是元素全为1的 m 维向量, $u_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} x(i+j)$;

[0026] S303: 计算出向量 X_i^m 与 X_j^m 之间的距离 d_{ij}^m , $d_{ij}^m = \max |x(i+k) - u_0(i) - (x(j+k) - u_0(j))|$, 其中, $i, j \in 1, 2, \dots, N-m+1$, $j \neq i$, $k=0, 1, \dots, m-1$;

[0027] S304: 根据所述距离 d_{ij}^m , 采用模糊隶属度函数 $\mu(d_{ij}^m, n, r)$, 计算所述 X_i^m 与所述 X_j^m 之间的相似度 D_{ij}^m , $D_{ij}^m = \mu(d_{ij}^m, n, r) = \exp \left[-\left(\frac{d_{ij}^m}{r} \right)^n \right]$, 其中, D_{ij}^m 为 X_i^m 与 X_j^m 之间的相似度, r 为相似容忍度;

[0028] S305: 根据相似度 D_{ij}^m , 得到函数 $\Phi^m(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m-1} \sum_{j=1, j \neq i}^{N-m} D_{ij}^m \right)$, 其中, D_{ij}^m 为 X_i^m 与 X_j^m 之间的相似度, N 为所述癫痫脑电信号时间序列的长度, m 为所述向量 $\mathbf{I}$ 的维数;

[0029] S306: 将所述向量 $\mathbf{I}$ 的维数增加到 $m+1$, 得到函数 $\Phi^{m+1}(n, r)$;

[0030] S307: 根据所述函数 $\Phi^m(n, r)$ 和所述函数 $\Phi^{m+1}(n, r)$, 得到所述模糊熵: $\text{FuzzyEn}(n, r, N) = \ln \Phi^m(n, r) - \ln \Phi^{m+1}(n, r)$;

[0031] 所述短时能量反映了脑电信号在有高频振荡节律时的变化, 采用均方法计算得到所述短时能量的估计值: $E^*(t) = \frac{1}{N} \sum_{k=t-N+1}^t x^2(k)$, 其中, $E^*(t)$ 为所述短时能量的估计值, $x(k)$ 为脑电信号的幅值, N 为用来估计短时能量的样本数;

[0032] 所述功率比用于从不相关脑电信号活动中提取出高频振荡节律, 所述功率比 R 的计算公式为: $R = \frac{P_{[80-200]}}{P_{[250-500]}}$, 其中, $P_{[80-200]}$ 是疑似高频振荡节律在 ripples 带宽内的功率,

$P_{[250-500]}$ 是疑似高频振荡节律在 fast ripples 带宽内的功率;

[0033] 所述频谱质心用来衡量频谱中心对应的频率, 所述频谱质心 f_c 的计算公式为:

$$f_c = \frac{\sum_{k=0}^{L/2} \frac{k}{LT} |M(k)|^2}{\sum_{k=0}^{L/2} |M(k)|^2}, \text{ 其中, } T \text{ 是采样周期, } L \text{ 是窗长, } M(k) \text{ 是多窗口功率谱密度估计值, 且}$$

$$M(k) = \sum_{l=0}^{L-1} w(l)x(l)e^{-j(2\pi/L)lk}, w(l) \text{ 是长度为 } L \text{ 的汉明窗。}$$

[0034] 进一步地, 在步骤102中, 基于期望最大化高斯混合模型的聚类算法对所述特征向量进行分类的具体步骤为:

[0035] S401: 采用模糊 c -均值聚类算法得到所述基于期望最大化高斯混合模型的聚类算法的初始化参数;

[0036] S402: 根据所述初始化参数, 通过公式 (5) 计算 D 维特征向量 V_i 属于第 c 类的概率

P_{ic}^t :

$$[0037] \quad P_{ic}^t = \frac{W_c^t P_c(\mathbf{V}_i | \mathbf{u}_c^t, \Sigma_c^t)}{P(\mathbf{V}_i)^t}, t=1,2,\dots,T \quad (5)$$

[0038] 其中,t是迭代次数(number of the iteration), $T \geq 1$, W_c^t 、 $\mathbf{u}_c^t$ 和 Σ_c^t 分别是第t次迭代的权重、聚类中心和协方差矩阵, $P(\mathbf{V}_i)^t$ 是第t次迭代时所述D维特征向量 $\mathbf{V}_i$ 对应的概率;

[0039] 所述D维特征向量 $\mathbf{V}_i$ 对应的概率分布由C个高斯成分构成,计算所述D维特征向量 $\mathbf{V}_i$ 对应的概率的公式是:

$$[0040] \quad P(\mathbf{V}_i) = \sum_{c=1}^C W_c P_c(\mathbf{V}_i | \mathbf{u}_c, \Sigma_c), i=1,2,\dots,N_{\beta} \quad (6)$$

[0041] 其中,C是聚类数目, W_c 是第c个高斯成分的权重,并且满足 $\sum_{c=1}^C W_c = 1, c=1,2,\dots,C$, $\mathbf{u}_c$ 为聚类中心, $\mathbf{u}_c \in \mathbb{R}^D$, Σ_c 为协方差矩阵, $\Sigma_c \in \mathbb{R}^{D \times D}$,D为所述D维特征向量 $\mathbf{V}_i$ 的维数;

[0042] S403:根据计算出的所述D维特征向量 $\mathbf{V}_i$ 属于第c类的概率,通过公式(7)、公式(8)和公式(9),更新高斯混合模型的参数,分别得到第t+1次迭代的权重 W_c^{t+1} 、聚类中心 $\mathbf{u}_c^{t+1}$ 和协方差矩阵 Σ_c^{t+1} :

$$[0043] \quad W_c^{t+1} = \frac{N_c^t}{N_{\beta}} \quad (7)$$

$$[0044] \quad \mathbf{u}_c^{t+1} = \frac{1}{N_c^t} \sum_{i=1}^{N_{\beta}} P_{ic}^t \mathbf{V}_i \quad (8)$$

$$[0045] \quad \Sigma_c^{t+1} = \frac{1}{N_c^t} \sum_{i=1}^{N_{\beta}} P_{ic}^t (\mathbf{V}_i - \mathbf{u}_c^{t+1})' (\mathbf{V}_i - \mathbf{u}_c^{t+1}) \quad (9)$$

[0046] 其中, P_{ic}^t 是所述特征向量 $\mathbf{V}_i$ 属于第c类的概率, N_c^t 是所述特征向量 $\mathbf{V}_i$ 属于第c类的总数,且 $N_c^t = \sum_{i=1}^{N_{\beta}} P_{ic}^t$;

[0047] S404:根据第t+1次迭代的权重 W_c^{t+1} 、聚类中心 $\mathbf{u}_c^{t+1}$ 和协方差矩阵 Σ_c^{t+1} ,通过公式(5)得到第t+1次迭代所述特征向量 $\mathbf{V}_i$ 对应的概率 $P(\mathbf{V}_i)^{t+1}$,再通过公式(10)计算得到第t+1次迭代的对数似然函数:

$$[0048] \quad L^{t+1} = \sum_{i=1}^{N_{\beta}} \log P(\mathbf{V}_i)^{t+1} \quad (10)$$

[0049] 其中, $P(\mathbf{V}_i)^{t+1}$ 为第t+1次迭代所述D维特征向量 $\mathbf{V}_i$ 对应的概率;

[0050] 当 $|L^{t+1} - L^t| < E$ 且 $t < T-1$,或者当 $t = T$ 时,结束所述基于期望最大化高斯混合模型的聚类算法,得到聚类结果;其中,E是设定的误差阈值。

[0051] 进一步地,在步骤S103中,使用箱形图得到中位数和四分位距两个指标,对不同类别的统计学特征进行分析,确定出每个类别对应的波形状态,从而检测出高频振荡节律。

[0052] 本发明提供的技术方案带来的有益效果是:提高癫痫脑电信号的高频振荡节律的

检测速度,帮助医生进行癫痫诊断和致痫灶切除。

附图说明

[0053] 下面将结合附图及实施例对本发明作进一步说明,附图中:

[0054] 图1是本发明实施例中基于模糊聚类优化高斯混合模型的高频振荡节律检测方法的流程图;

[0055] 图2是本发明实施例中高频振荡节律和非高频振荡节律两种状态下4个特征的曲线示意图;

[0056] 图3是本发明实施例中某病人的量化误差模型及参数化成本函数的示意图。

具体实施方式

[0057] 为了对本发明的技术特征、目的和效果有更加清楚的理解,现对照附图详细说明本发明的具体实施方式。

[0058] 本发明的实施例提供了基于模糊聚类优化高斯混合模型的高频振荡节律检测方法。

[0059] 请参考图1,图1是本发明实施例中基于模糊聚类优化高斯混合模型的高频振荡节律检测方法的流程图,具体包括以下步骤:

[0060] S101:获取癫痫脑电信号的四个特征:模糊熵、短时能量、功率比和频谱质心构成特征向量,在不同时刻得到多组所述四个特征,构成多个特征向量;其中,癫痫脑电信号的选取过程为:利用与高频振荡节律相关的特征,来区别高频振荡节律与其他脑电活动;选取模糊熵、短时能量、功率比和频谱质心作为癫痫脑电信号的特征,分别计算高频振荡节律和非高频振荡节律两种状态下的特征;所述模糊熵的具体计算步骤如下:

[0061] S201:获取癫痫脑电信号时间序列 $X = [x(1), x(2), \dots, x(N)]$;

[0062] S202:将所述癫痫脑电信号时间序列按顺序构造 m 维向量 $X_i^m = [x(i), x(i+1), \dots, x(i+m-1)] - u_0(i) \cdot \mathbf{I}$,其中, $i = 1, 2, \dots, N-m+1$, m 为正整数,且 $m \geq 1$, $\mathbf{I}$ 是元素全为1的 m 维行向量, $u_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} x(i+j)$;

[0063] S203:计算出向量 X_i^m 与 X_j^m 之间的距离 d_{ij}^m , $d_{ij}^m = \max |x(i+k) - u_0(i) - (x(j+k) - u_0(j))|$,其中, $i, j \in 1, 2, \dots, N-m+1, j \neq i, k = 0, 1, \dots, m-1$;

[0064] S204:根据所述距离 d_{ij}^m ,采用模糊隶属度函数 $\mu(d_{ij}^m, n, r)$,计算所述 X_i^m 与所述 X_j^m

之间的相似度 D_{ij}^m , $D_{ij}^m = \mu(d_{ij}^m, n, r) = \exp \left[- \left(\frac{d_{ij}^m}{r} \right)^n \right]$,其中, D_{ij}^m 为 X_i^m 与 X_j^m 之间的相似度, r

为相似容忍度;

[0065] S205:根据相似度 D_{ij}^m ,得到函数 $\Phi^m(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m-1} \sum_{j=1, j \neq i}^{N-m} D_{ij}^m \right)$,其中,

D_{ij}^m 为 X_i^m 与 X_j^m 之间的相似度, N 为所述癫痫脑电信号时间序列的长度, m 为所述向量 $\mathbf{I}$ 的维

数;

[0066] S206:将所述向量I的维数增加到m+1,得到函数 $\Phi^{m+1}(n,r)$;

[0067] S207:根据所述函数 $\Phi^m(n,r)$ 和所述函数 $\Phi^{m+1}(n,r)$,得到所述模糊熵:FuzzyEn(n,r,N)= $\ln \Phi^m(n,r) - \ln \Phi^{m+1}(n,r)$;

[0068] 采用均方法计算得到所述短时能量的估计值: $E^*(t) = \frac{1}{N} \sum_{k=t-N+1}^t x^2(k)$,其中, $E^*(t)$ 为所述短时能量的估计值,x(k)为脑电信号的幅值;

[0069] 所述功率比R的计算公式为: $R = \frac{P_{[80-200]}}{P_{[250-500]}}$,其中, $P_{[80-200]}$ 是疑似高频振荡节律在ripples带宽内的功率, $P_{[250-500]}$ 是疑似高频振荡节律在fast ripples带宽内的功率;

[0070] 所述频谱质心 f_c 的计算公式为: $f_c = \frac{\sum_{k=0}^{L/2} \frac{k}{LT} |M(k)|^2}{\sum_{k=0}^{L/2} |M(k)|^2}$,其中,T是采样周期,L是窗长,

M(k)是多窗口功率谱密度估计值,且 $M(k) = \sum_{l=0}^{L-1} w(l)x(l)e^{-j(2\pi/L)lk}$,w(1)是长度为L的汉明窗。

[0071] 四个特征分别在高频振荡节律和非高频振荡节律两种不同状态下的曲线图如图2所示,图2是本发明实施例中高频振荡节律和非高频振荡节律两种状态下4个特征的曲线示意图;图2(a)是模糊熵在高频振荡节律和非高频振荡节律两种状态下的曲线图,从中可以看出模糊熵在高频振荡节律状态下的值总体低于在非高频振荡节律状态下的值;图2(b)是短时能量在不同状态下的曲线图,从中可以看出非高频振荡节律状态下的短时能量很小,几乎为0,短时能量在高频振荡节律的值明显高于在非高频振荡节律的值;图2(c)是功率比在不同状态下的曲线图,功率比在高频振荡节律的值明显高于在非高频振荡节律的值;图2(d)是频谱质心在不同状态下的曲线图,从中可以看出频谱质心在高频振荡节律状态下的值总体高于在非高频振荡节律状态下的值。通过图2可知,非高频振荡节律状态下的短时能量和功率比很小;在高频振荡节律和非高频振荡节律两种状态下,模糊熵、短时能量、功率比和频谱质心有着显著不同,可以用于高频振荡节律分类。

[0072] 为了验证从病人数据中提取的特征也可以用来区分高频振荡节律和非高频振荡节律,这里采用曼惠特尼U检验来分析从不同病人高频振荡节律和非高频振荡节律状态下提取的4个特征是否有显著性差异。曼惠特尼U检验得到的统计量绝对值|Z|如表1所示:

[0073] 表1:5个病人高频振荡节律和非高频振荡节律状态下曼惠特尼U检验

病人	模糊熵的 Z 值	短时能量的 Z 值	功率比的 Z 值	频谱质心的 Z 值
1	8.268	12.131	10.519	9.541
2	7.936	10.379	12.536	7.928
3	9.439	10.156	9.179	8.975
4	6.566	9.755	10.192	7.946
5	8.274	11.172	11.536	8.176

[0075] 从表1可以看出,统计量绝对值 $|Z|$ 都超过 $Z_{\alpha/2}$ (α 为显著性水平,设为0.05,此时 $Z_{\alpha/2}$ 值为1.96),落在否定域内,拒绝原假设,因此,5个病人高频振荡节律和非高频振荡节律状态下提取的4个特征有显著性差异,即这4个特征可以用来从癫痫脑电信号中区分出高频振荡节律。

[0076] S102:采用期望最大化高斯混合模型聚类算法对所述特征向量进行分类,得到聚类结果;用于提高运算速度和癫痫脑电信号的高频振荡节律的检测精度;

[0077] 分类过程中,采用模糊c-均值聚类算法对所述期望最大化高斯混合模型聚类算法的参数进行初始化,得到初始化参数;采用模糊c-均值聚类算法解决了期望最大化高斯混合模型聚类算法的运算复杂度较高的问题,提高癫痫脑电信号的高频振荡节律的检测速度,帮助医生进行癫痫诊断和致痫灶切除。

[0078] 基于期望最大化高斯混合模型的聚类算法对所述多个特征向量进行分类的具体步骤为:

[0079] 301:采用模糊c-均值聚类算法得到的参数初始值作为基于期望最大化高斯混合模型的聚类算法中的初始化参数;模糊c-均值聚类算法降低了期望最大化高斯混合模型聚类算法的运算复杂度;其中,采用期望最大化高斯混合模型聚类算法对所述多个特征向量进行分类时,根据模糊c-均值聚类算法得到所述期望最大化高斯混合模型聚类算法的初始化参数,具体步骤为:

[0080] 401:随机产生初始聚类中心 $v_c, c=1, 2, \dots, C, C$ 是聚类数目;

[0081] 本实施例中,可以采用传统算法对聚类数目进行提前设定,也可以建立的量化误差模型对聚类数目进行优化,得到最优的聚类数目,从而根据期望最大化高斯混合模型聚类算法得到聚类结果,为了提高癫痫脑电信号的高频振荡节律检测精度,本实施例采用建立量化误差模型对聚类数目进行优化,聚类数目的优化过程为:

[0082] 量化误差 J_c 曲线,被称为率失真(Rate-distortion, R-D)曲线,其形状依赖于数据点在D维空间的分布; J_c 通常是一个单调递减函数,根据量化误差 $J_c: J_c = \frac{Const}{C^{2/D}}$, D是所述D

维特征向量 V_i 的维数,建立量化误差模型: $\hat{J}_c = \frac{Const}{C^{2/a}}$,其中,a是模型参数,Const是一个常量;

[0083] 根据所述量化误差模型,得到对数线性模型: $\lg(\hat{J}_c) + \frac{2}{a} \cdot \lg(C) + b = 0$,其中,根据对

数线性回归的系数,计算模型参数a和模型参数b: $(a, b) = \arg \min \left\{ \sum_{c=2}^{C_{\max}} \left(\lg(\hat{J}_c) + \frac{2}{a} \cdot \lg(C) + b \right)^2 \right\}$,

即所述模型参数a和所述模型参数b满足公式 $\arg \min \left\{ \sum_{c=2}^{C_{\max}} \left(\lg(\hat{J}_c) + \frac{2}{a} \cdot \lg(C) + b \right)^2 \right\}$ 取值最小,此时得到所述模型参数a和所述模型参数b;在后续分析中只用到参数a,模型参数a为:

$$a = 2 \cdot \frac{M \sum_{C=2}^{C_{\max}} \lg^2(C) - (\sum_{C=2}^{C_{\max}} \lg(C))^2}{M \sum_{C=2}^{C_{\max}} \lg(C) \cdot \lg(J_C) - \sum_{C=2}^{C_{\max}} \lg(C) \cdot \sum_{C=2}^{C_{\max}} \lg(J_C)}, \text{其中}, C \text{是聚类数目}, M = C_{\max} - 1, C_{\max} \text{是聚类数}$$

目的最大值,通过聚类误差和聚类数目形成的R-D曲线确定模型参数a,作为特征向量的维数;

[0084] 通过公式 $I(C) = \hat{J}_C \cdot C^{2/a}$,得到量化误差模型的一个常量,基于这个常量的不变性,构造参数化成本函数PCF: $\text{PCF}(C) = J_C \cdot C^{2/a}$,所述参数化成本函数PCF最小时,对应的聚类数目C*最优: $C^* = \argmin \{\text{PCF}(C)\}$ 。

[0085]

[0086] 402:根据聚类中心 v_c ,计算第i个特征向量属于第c类的隶属度值 μ_{ic} :

$$u_{ic} = \left(\sum_{k=1}^C \left(\frac{\|V_i - v_c\|}{\|V_i - v_k\|} \right)^{\frac{2}{p-1}} \right)^{-1}, \text{其中}, \mu_{ic} \text{满足条件: } \mu_{ic} \in [0,1], \forall i, c, \sum_{c=1}^C \mu_{ic} = 1, \forall i, 0 < \sum_{i=1}^{N_{fl}} \mu_{ic} < N_{fl}, \forall c,$$

C是聚类数目,p是加权模糊参数, V_i 是所述D维特征向量,且由提取的所述模糊熵、所述短时能量、所述功率比和所述频谱质心构成;

$$[0087] \quad 403: \text{根据公式 } v_c = \frac{\sum_{i=1}^{N_{fl}} \mu_{ic}^m V_i}{\sum_{i=1}^{N_{fl}} \mu_{ic}^m}, \text{更新聚类中心 } v_c; \text{其中}, N_{fl} \text{是特征向量的长度}, V_i \text{是所}$$

述D维特征向量,由所述模糊熵、所述短时能量、所述功率比和所述频谱质心构成, μ_{ic} 是第i个特征向量属于第c类的隶属度值;

[0088] 404:当满足 $\max_{1 \leq c \leq C} \|v_{c, \text{new}} - v_{c, \text{old}}\| < \varepsilon$ 时,所述模糊c-均值聚类算法结束;其中, $v_{c, \text{new}}$ 是更新后的聚类中心, $v_{c, \text{old}}$ 是更新前的聚类中心, ε 是设定的精度误差;

[0089] 根据所述模糊c-均值聚类算法,通过公式(1)、公式(2)和公式(3)分别得到所述基于期望最大化高斯混合模型的聚类算法的初始化参数:权重 W_c 、聚类中心 u_c 和协方差矩阵 Σ_c :

$$[0090] \quad W_c = \frac{N_c}{N_{fl}} \quad (1)$$

$$[0091] \quad u_c = v_c \quad (2)$$

$$[0092] \quad \Sigma_c = \text{cov}(V_i^c) \quad (3)$$

[0093] 其中, N_c 是第c类特征向量的数目之和, V_i^c 是属于第c类的特征向量构成的矩阵;

[0094] 302:根据所述初始化参数,通过公式(4)计算D维特征向量 V_i 属于第c类的概率 P_{ic}^t :

$$[0095] \quad P_{ic}^t = \frac{W_c^t P_c(V_i | u_c^t, \Sigma_c^t)}{P(V_i)^t}, t = 1, 2, \dots, T \quad (4)$$

[0096] 其中,t是迭代次数(number of the iteration), $T \geq 1$, W_c^t 、 u_c^t 和 Σ_c^t 分别是第t次迭代的权重、聚类中心和协方差矩阵, $P(V_i)^t$ 是第t次迭代时所述D维特征向量 V_i 对应的概

率；

[0097] 所述D维特征向量 V_i 对应的概率分布由C个高斯成分构成,计算所述D维特征向量 V_i 对应的概率的公式是:

$$[0098] \quad P(V_i) = \sum_{c=1}^C W_c P_c(V_i | \mathbf{u}_c, \Sigma_c), i=1,2,\dots,N_{fl} \quad (5)$$

[0099] 其中,C是聚类数目, W_c 是第c个高斯成分的权重,并且满足 $\sum_{c=1}^C W_c=1, c=1,2,\dots, C, \mathbf{u}_c$ 为聚类中心, $\mathbf{u}_c \in \mathbb{R}^D, \Sigma_c$ 为协方差矩阵, $\Sigma_c \in \mathbb{R}^{D \times D}, D$ 为所述D维特征向量的维数;

[0100] 303:根据计算出的所述D维特征向量 V_i 属于第c类的概率,通过公式(6)、公式(7)和公式(8),更新高斯混合模型的参数,分别得到第t+1次迭代的权重 W_c^{t+1} 、聚类中心 $\mathbf{u}_c^{t+1}$ 和协方差矩阵 Σ_c^{t+1} :

$$[0101] \quad W_c^{t+1} = \frac{N_c^t}{N_{fl}} \quad (6)$$

$$[0102] \quad \mathbf{u}_c^{t+1} = \frac{1}{N_c^t} \sum_{i=1}^{N_{fl}} P_{ic}^t V_i \quad (7)$$

$$[0103] \quad \Sigma_c^{t+1} = \frac{1}{N_c^t} \sum_{i=1}^{N_{fl}} P_{ic}^t (V_i - \mathbf{u}_c^{t+1})' (V_i - \mathbf{u}_c^{t+1}) \quad (8)$$

[0104] 其中, P_{ic}^t 是所述D维特征向量 V_i 属于第c类的概率, N_c^t 是所述D维特征向量 V_i 属于第c类的总数,且 $N_c^t = \sum_{i=1}^{N_{fl}} P_{ic}^t$;

[0105] 304:根据第t+1次迭代的权重 W_c^{t+1} 、聚类中心 $\mathbf{u}_c^{t+1}$ 和协方差矩阵 Σ_c^{t+1} ,通过公式(5)得到第t+1次迭代所述D维特征向量 V_i 对应的概率 $P(V_i)^{t+1}$,再通过公式(9)计算得到第t+1次迭代的对数似然函数:

$$[0106] \quad L^{t+1} = \sum_{i=1}^{N_{fl}} \log P(V_i)^{t+1} \quad (9)$$

[0107] 其中, $P(V_i)^{t+1}$ 为第t+1次迭代所述D维特征向量 V_i 对应的概率;

[0108] 当 $|L^{t+1}-L^t| < E$ 且 $t < T-1$,或者当 $t=T$ 时,结束所述基于期望最大化高斯混合模型的聚类算法,得到聚类结果;其中,E是设定的误差阈值。

[0109] S103:根据所述聚类结果,采用箱形图的方法对不同类别特征进行统计学分析,检测出癫痫脑电信号的高频振荡节律;使用箱形图得到中位数和四分位距两个指标,对不同类别的统计学特征进行分析,确定出每个类别对应的波形状态,从而检测出高频振荡节律。

[0110] 某病人的R-D曲线、量化误差模型及参数化成本函数如图3所示,图3是本发明实施例中某病人的量化误差模型及参数化成本函数的示意图;图3(a)表示,建立的量化误差模型基本能够描述R-D曲线的特征;图3(b)表示,当参数化成本函数最小时,聚类数目为4,此时的聚类数目是最优的,因此,期望最大化高斯混合模型聚类算法的聚类数目C为4。

[0111] 采用期望最大化高斯混合模型聚类算法对从病人癫痫脑电信号中提取的特征向量进行聚类,得到不同类别;为了对每个类别给出可描述的统计量,根据所述期望最大化高

斯混合模型聚类算得到的聚类结果,最后使用箱形图的方法得到中位数和四分位距两个指标,对不同类别特征进行统计学分析,检测出癫痫脑电信号的高频振荡节律。不同类别指标如表2所示:

[0112] 表2不同类别指标

[0113]

特征	类别 1		类别 2		类别 3		类别 4	
	中位数	四分位距	中位数	四分位距	中位数	四分位距	中位数	四分位距
模糊熵	0.636	0.131	1.07	0.333	0.687	0.108	0.86	0.233
短时能量	18.2	17.9	5.23	2.34	6.73	6.10	6.35	6.84
功率比	462	630	2.40	2.96	90.1	114	12.9	20.3
频谱质心	140	30.7	254	29.5	162	23.9	205	54.6

[0114] 从表2可以看出:类别1模糊熵中位数最小,短时能量和功率比中位数最大,且频谱质心的中位数为140Hz,因此类别1是ripples;相比于类别4,类别1、2和3的频谱质心更为集中,类别4是伪迹;类别2频谱质心最大,带宽比最小,且只有类别2的频谱质心在FRs频率范围内;类别3的功率比和频谱质心的中位数介于类别1和类别3之间,因此类别3是一种混合波形(ripples+FRs);其中,Ripples、FRs和混合波形(ripples+FRs)均为高频振荡节律。

[0115] 本发明的有益效果是:提高癫痫脑电信号的高频振荡节律的检测速度,帮助医生进行癫痫诊断和致痫灶切除。

[0116] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

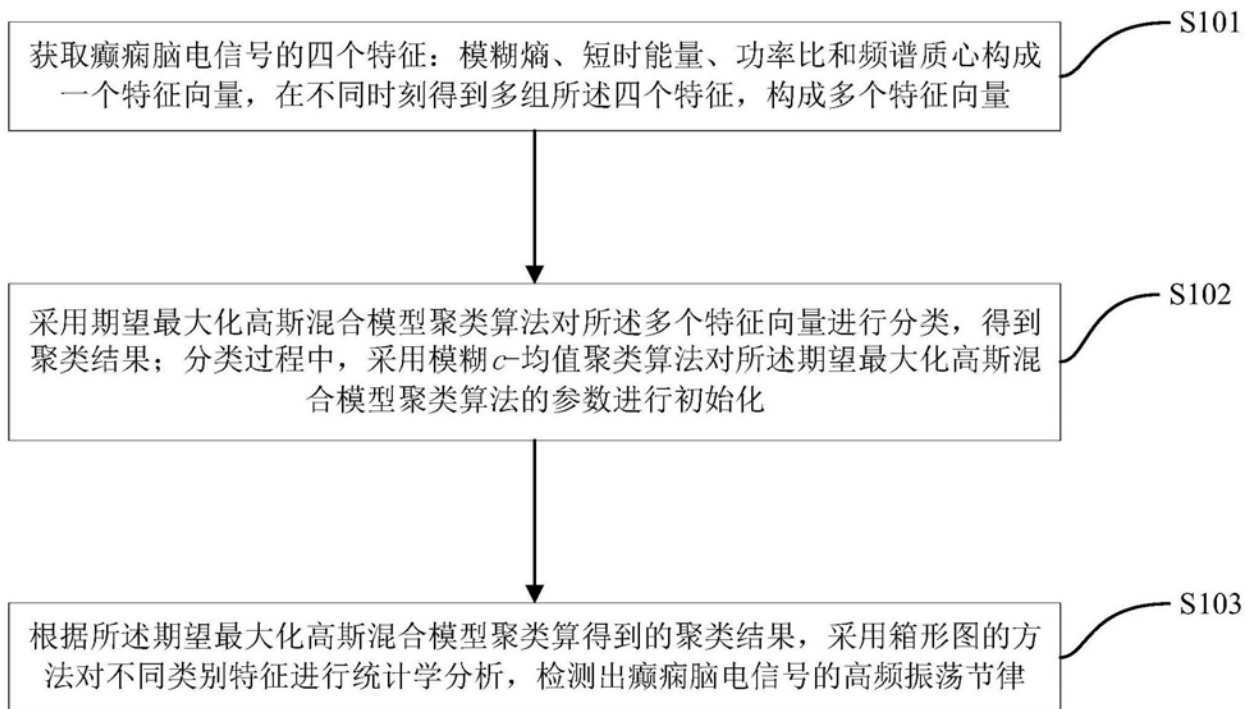


图1

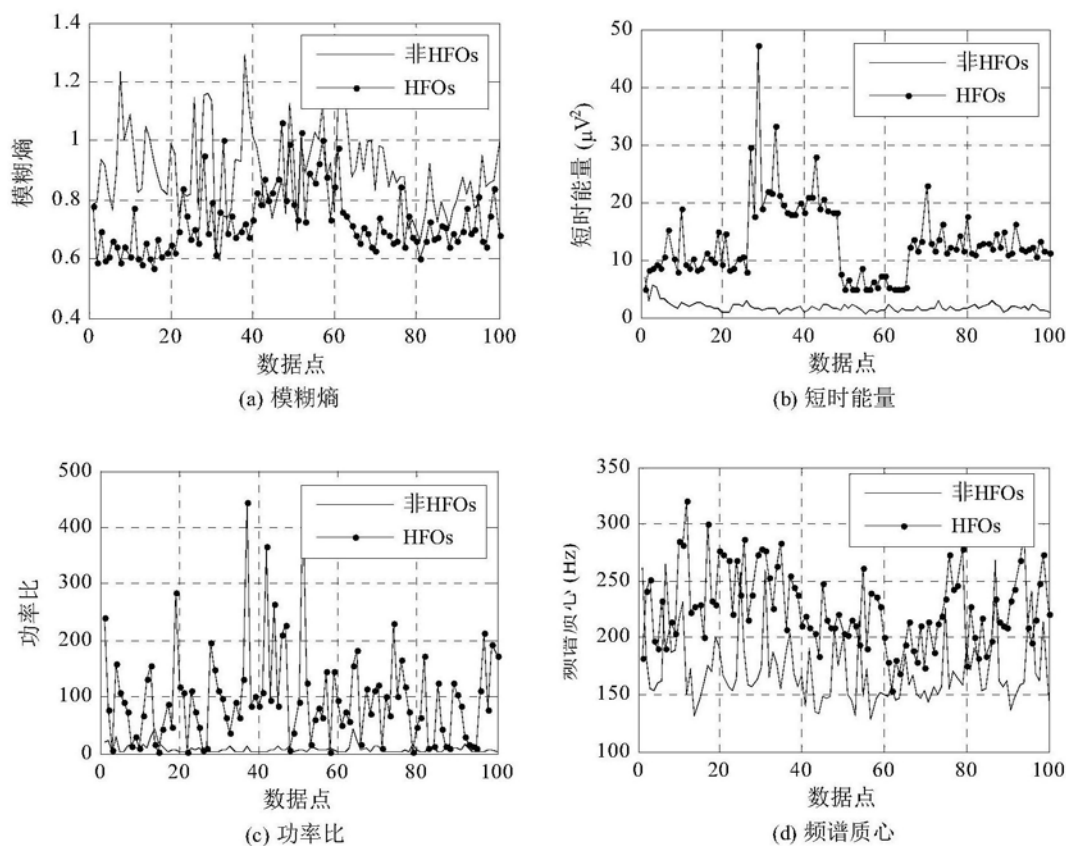


图2

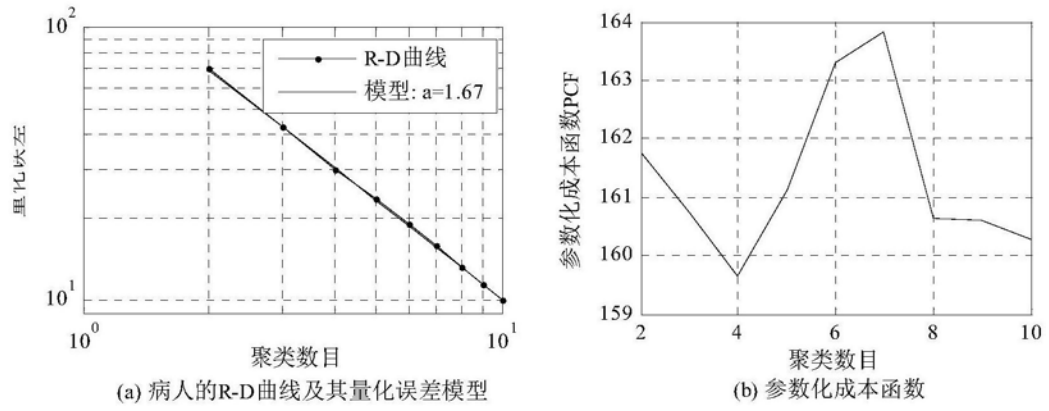


图3

专利名称(译)	基于模糊聚类优化高斯混合模型的高频振荡节律检测方法		
公开(公告)号	CN108875604A	公开(公告)日	2018-11-23
申请号	CN201810552305.1	申请日	2018-05-31
[标]申请(专利权)人(译)	中国地质大学(武汉)		
申请(专利权)人(译)	中国地质大学(武汉)		
当前申请(专利权)人(译)	中国地质大学(武汉)		
[标]发明人	吴敏 万雄波 方泽林 万婷 杜玉晓		
发明人	吴敏 万雄波 方泽林 万婷 杜玉晓		
IPC分类号	G06K9/00 A61B5/0476 A61B5/00		
CPC分类号	G06K9/00523 A61B5/0476 A61B5/4094 G06K9/00536		
外部链接	Espacenet SIPO		

摘要(译)

本发明提供了基于模糊聚类优化高斯混合模型的高频振荡节律检测方法，基于聚类分析的方法对高频振荡节律进行检测，选取模糊熵、短时能量、功率比和频谱质心为癫痫脑电信号的特征，将其构成特征向量作为聚类算法的输入，采用期望最大化高斯混合模型聚类算法对该特征向量进行分类，并采用模糊c-均值聚类算法得到期望最大化高斯混合模型聚类算法的初始化参数；选取中位数和四分位距分析每类的统计学特征，检测出高频振荡节律。本发明的有益效果：提高癫痫脑电信号的高频振荡节律的检测速度，帮助医生进行癫痫诊断和致痫灶切除。

