



(12)发明专利申请

(10)申请公布号 CN 110236523 A
(43)申请公布日 2019.09.17

(21)申请号 201910521097.3

(22)申请日 2019.06.17

(71)申请人 杭州电子科技大学
地址 310018 浙江省杭州市下沙高教园区2号大街

(72)发明人 王建中 黄泽银 曹九稳

(74)专利代理机构 杭州君度专利代理事务所
(特殊普通合伙) 33240
代理人 朱月芬

(51) Int. Cl.
A61B 5/0402(2006.01)
A61B 5/0472(2006.01)
A61B 5/00(2006.01)

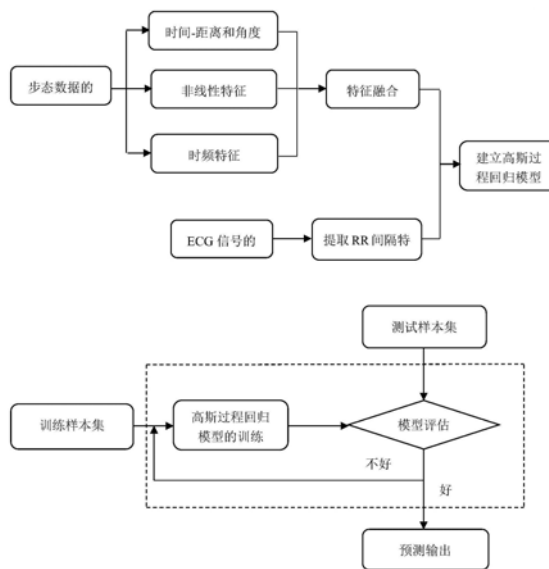
权利要求书3页 说明书12页 附图3页

(54)发明名称

基于高斯回归的步态-心电RR间隔关联方法

(57)摘要

本发明公开了一种基于高斯回归的步态-心电RR间隔关联方法。本发明包括以下步骤:步骤1.分别利用三维运动跟踪系统和心电采集设备,同步采集受试者运动的三维坐标数据及对应的ECG信号;步骤2.对三维坐标数据进行步态定量分析、非线性分析、时频分析,提取出17种特征,构成特征向量,提出融合方法;步骤3.对ECG信号进行预处理提取RR间隔;步骤4.构建步态特征与心电RR间隔之间的高斯过程回归(GPR)关联预测模型;步骤5.训练和测试高斯过程回归模型,进行关联预测分析。本发明自适应强、容易实现、超参数自适应获取等优势,具有较好的鲁棒性和泛化性能,能够更好地揭示步态与心电之间的关联信息。



1. 基于高斯回归的步态-心电RR间隔关联方法,其特征在于包括如下步骤:

步骤1. 获取人体下肢足部标记点的运动三维坐标数据及对应ECG信号;

采用三维运动跟踪系统和心电采集设备,同步捕捉受试者下肢足部标记点运动时的三维坐标数据以及对应的同步ECG信号;

步骤2. 对实时捕捉的运动三维坐标数据进行分析,提取时间-距离及角度参数、非线性特征、时频特征共17种步态特征,构成特征向量,提出融合方法;

步骤3. ECG信号的预处理及RR间隔提取;

步骤4. 构建步态特征与心电RR间隔之间的高斯过程回归(GPR)关联预测模型;

通过步骤3和步骤4所得的心电RR间隔和融合后的步态特征向量,将融合后的步态特征向量作为输入样本集,对应的心电RR间隔特征作为期望输出样本集,建立两者间的高斯过程回归模型;

步骤5. 训练和测试高斯过程回归模型,进行关联预测分析;

随机输入训练样本集和测试样本集占比为4:1,采用三种核函数:有理二次核函数(RQ)、平方指数核函数(SE)、Matern5/2核函数(MA),进行关联预测分析,将测试结果十次取其平均值。

2. 根据权利要求1所述的基于高斯回归的步态-心电RR间隔关联方法,其特征在于将步骤1所得去除噪声的三维动作数据,选用人体下肢部位的标记点进行分析;以脚跟前进方向的动作数据提取出每个步态周期,并以每十个步态周期为一个子集,将数据划分为多个子集;

提取出步态的时间-距离参数及角度参数,包括步态周期、步长、步高、步宽、步速、步速加速度、足偏角最大值及最小值,取每一子集的平均值;再采用非线性动力学分析的方法提取脚跟前向运动的三维动作数据每一子集的非线性特征,包括近似熵、样本熵、模糊熵、LZ复杂度、C0复杂度特征;最后采用时频分析的方法,通过快速傅里叶变换计算功率谱密度最大值,通过连续小波变换的C-morlet3-3、C-gaus1、mexh三种小波函数计算在尺度2的小波系数模值均值特征;将所提取的特征构成了特征向量,提出融合方法。

3. 根据权利要求2所述的基于高斯回归的步态-心电RR间隔关联方法,其特征在于步骤3具体实现如下:

3-1. 通过级联的低通和高通滤波器组成的带通滤波器进行处理,抑制基线漂移和T波造成的干扰;其二阶低通和高通滤波器的差分方程分别为:

$$y(nT) = 2y(nT-T) - y(nT-2T) + x(nT) - 2x(nT-6T) + x(nT-12T)$$

$$y(mT) = 32x(mT-16T) - [y(mT-T) + x(mT) - x(mT-32T)]$$

其中,T为采样周期,x(nT)表示离散时间系统心电信号时间序列的输入,y(nT)为低通滤波后的心电信号,其截止频率设为12Hz,增益为2,处理延迟约为6个采样周期;高通滤波器的截止频率约为5Hz,增益为32,16个采样周期的延迟;

3-2. 经过带通滤波器处理后的心电信号,再经过导数滤波器处理,提供QRS复斜率信息;其差分方程为:

$$y(mT) = (1/8T) [-x(mT-2T) - 2x(mT-T) + 2x(mT+T) + x(mT+2T)]$$

3-3. 经导数滤波器处理后的输出利用非线性平方函数达到非线性放大的效果,强调高频特性,非线性平方函数的差分方程为:

$$y(mT) = [x(mT)]^2$$

3-4. 通过移动窗口积分器 (MVI) 处理产生包络形状的特征波形, 来进一步增强特征与噪声的区分; 其差分方程为:

$$y(mT) = (1/N) [x(mT - (N-1)T) + x(mT - (N-2)T) + \dots + x(mT)]$$

其中, 若采集心电信号的采样频率 500Hz, 即每秒采集 500 个样本点, N 的样本点数为 75 个;

3-5. 使得阈值和其它参数进行周期性调整, 并检测 R 波或 QRS 波的位置; 最后对平滑信号中检测到的 R 波输出进行分析, 并对带通信号的输出进行两次检查, 最终找到真实 R 波的原始指标。

4. 根据权利要求 3 所述的基于高斯回归的步态-心电 RR 间隔关联方法, 其特征在于步骤 4 具体实现如下:

4.1 高斯过程 (GP) 原理

高斯过程是指一个正态的随机过程, 其任意维有限变量的联合分布服从高斯分布; 对于任意有限个 x , 即 $x_1, x_2, \dots, x_n \in N$, 其相对应的有限个随机变量 $F = (f(x_1), f(x_2), \dots, f(x_n))^T$, 记 $F \sim GP(m(x_i), K)$, 即

$$\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} \sim N \left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_n) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \dots & k(x_n, x_n) \end{bmatrix} \right)$$

其中, F 为联合分布; 均值函数和协方差函数 (核函数) 表示如下:

$$m(x) = E[f(x)]$$

$$K = E[(f(x) - m(x))(f(x') - m(x'))]$$

4.2 GPR 回归模型的建立

建立高斯回归模型时, 考虑目标值 y 含有噪声, 即定义为:

$$y = f(x) + \varepsilon; \text{ 其中: } f(x) \text{ 为高斯过程, } x \in N, \varepsilon \sim N(0, \sigma_n^2); \text{ 训练集为:}$$

$D: y_i = f(x_i) + \varepsilon_i (i = 1, 2, \dots, n)$, 在测试集: $D_* = \{(x_{i*}, y_{i*}) | i = n+1, n+2, \dots, n+n_*\}$ 预测对应的输出值 f_* 的多元高斯分布为:

$$\begin{bmatrix} y \\ f_* \end{bmatrix} \sim N \left(\mathbf{0}, \begin{bmatrix} K + \sigma_n^2 I & K_*^T \\ K_* & K_{**} \end{bmatrix} \right)$$

其中, $K_* = [k(x_*, x_1), k(x_*, x_2), \dots, k(x_*, x_n)]$, $K_{**} = k(x_*, x_*)$, σ_n 为噪声的标准差; 根据多元高斯分布的条件, 可得高斯过程回归模型预测方程:

$$(f_* | X, y, X_*) \sim GP(m(x_*), \text{cov}(f_*))$$

式中: 矩阵 X 由训练集的输入 x_i 的列向量组成; 矩阵 X_* 由测试集的输入 x_{i*} 的列向量组成;

$$m(x_*) = K_*(K + \sigma_n^2 I)^{-1} y$$

$$\text{cov}(f_*) = K_{**} - K_*(K + \sigma_n^2 I)^{-1} K_*^T$$

实测步态预测心电的 RR 间隔回归模型表示为:

$$f(x) = \varphi(x)^T \omega, \{x \in R^n, \varphi(x), \omega \in R^m, m > n\}$$

其中： $\omega \in \mathbb{R}^m$ 表示权值向量，均值函数 $m(x) = E[f(x)]$ ，核函数 $k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))]$ ；

4.3核函数的选取

采用了3种核函数：

二次有理核函数 (RQ)：

$$k(x_i, x_j) = \sigma_f^2 \exp\left[1 + \frac{(x_i - x_j)^T M (x_i - x_j)}{2\alpha}\right]^{-\alpha} + \sigma_n^2 \delta^{ij}$$

平方指数核函数 (SE)

$$k(x_i, x_j) = \sigma_f^2 \exp\left[-\frac{1}{2}(x_i - x_j)^T M (x_i - x_j)\right] + \sigma_n^2 \delta^{ij}$$

Matern核函数 (M)

$$k(x_i, x_j) = \sigma_f^2 \exp\left[1 + \sqrt{3M}(x_i - x_j)\right] \exp\left[-\sqrt{3M}(x_i - x_j)\right] + \sigma_n^2 \delta^{ij}$$

其中， $M = \text{diag}(l^{-2})$ ， l 为方差尺度， σ_f^2 为信号方差； l 为关联性测定超参数； δ^{ij} 为符号函数， i 和 j 相同时为1，不同时为0； α 为核函数的形状参数；

定义 $\theta = [\sigma_1, \sigma_f^2, \sigma_n^2]$ 为超参数向量，通过极大似然法自适应获得最优解；首先建立训练样本条件概率的负对数似然函数，并令其对超参数求偏导，然后采用共轭梯度法对该偏导数进行最小化，得到超参数最优解。

5. 根据权利要求4所述的基于高斯回归的步态-心电RR间隔关联方法，其特征在于步骤5具体实现如下：

随机输入训练样本集和测试样本集占比为4:1，采用三种核函数：有理二次核函数 (RQ)、平方指数核函数 (SE)、Matern5/2核函数 (MA)，进行关联预测分析，将测试结果十次取其平均值；

通过采用均方根误差 (RMSE) 和确定系数 (R^2) 对预测模型的准确性和有效性进行评价，其RMSE越小，预测结果越好；其表达式如下：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - t_i)^2} ;$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - t_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

其中， n 表示预测数据的数目， y_i 是实际预测数据值， t_i 代表期望输出值。

基于高斯回归的步态-心电RR间隔关联方法

技术领域

[0001] 本发明涉及一种基于高斯回归的步态-心电RR间隔关联方法,涉及数字信号处理与机器学习、医疗健康、人体运动分析等多种技术领域。

背景技术

[0002] 近年来,随着人类生活水平的提高,肥胖、压力等是导致心血管疾病的多重因素。目前,心血管疾病是威胁人类生命健康的重要原因之一,尤其是针对于心血管疾病患者在运动时,由运动过量造成身体不适甚至猝死等特别严重的后果。心律失常也是一种极其常见的电活动异常状态,严重时甚至可以导致猝死,在运动环境中也越来越多地出现在健康人身上。

[0003] 目前,步态分析包括定性分析和定量分析,定性分析通过观察法进行研究,但其具有一定的主观性,难以在多部位、多环节进行分析;步态的定量分析是通过器械和专门的设备获得客观数据对步态进行分析的方法,即可利用简单的测量工具来测量,也可通过复杂设备仪器对运动学参数、动力学参数、肌电活动参数及能量参数的获取,是目前使用广泛的方法。研究表明,人体下肢的运动呈周期性的变化。近年来,随着生物医学信号的研究兴起,非线性动力学和时频分析的方法被广泛应用,并取得了较好的成果。因非线性特征能体现其动态的本质特征,时频分析在时域和频域具有很好的局部性,分析结果较稳定,能更好地体现信号的本质特性。心电图是记录心脏的电活动状态,QRS波群代表心室肌除极的电位变化,是继P波之后出现的一个狭窄但振幅高的波群。它是由Q波、R波、S波组成,是心电图中最突出的部分。QRS波群的检测是ECG信号的关键,目前的方法主要包括神经网络和小波变换,但其初始阶段的训练时间较长,计算量较大,不适合实时检测。目前关于步态和心电的分析都是各子系统的单独研究,对于步态和心电的关联性研究的甚少。通过从各种步态信息中准确地预测心电RR间隔,对人体运动分析、运动康复治疗、患者的健康监护和智慧医疗具有十分重要的意义。

发明内容

[0004] 针对上述背景和存在的问题与不足,为了从各种步态信息中准确预测心电RR间隔,本发明提出了一种基于高斯回归的步态-心电RR间隔关联方法。本发明通过Codamotion红外三维实时运动捕捉系统和Trigno™无线系统,将实时捕捉的身体下肢部位的运动三维坐标数据及同步采集的ECG信号进行分析处理。首先,通过运动学分析、非线性动力学分析、时频分析提取了17种步态特征,包括步态周期、步长、步高、步宽、足偏角最大值及最小值、步速、步速加速度、近似熵、样本熵、模糊熵、LZ复杂度、C0复杂度以及三种小波(Cmorlet、Cgaus、mexh)变换的系数模值,构成特征向量,提出融合方法;然后,将ECG信号进行预处理,通过实时算法来检测QRS波群,提取与三维动作数据对应的心电RR间隔;建立步态特征向量与心电RR间隔之间的高斯过程回归关联模型;最后,进行性能评估。实验结果表明,本发明提取的步态特征及所提出的的高斯回归方法构建的步态与心电RR间隔预测的模型,具有较

高的预测精度。

[0005] 为实现上述目的,采用了以下技术方案,所述方法包括以下步骤:

[0006] 步骤1.获取人体下肢足部标记点的运动三维坐标数据及对应ECG信号;

[0007] 采用三维运动跟踪系统和心电采集设备,同步捕捉受试者下肢足部标记点运动时的三维坐标数据以及对应的同步ECG信号。

[0008] 步骤2.对实时捕捉的运动三维坐标数据进行分析,提取时间-距离及角度参数、非线性特征、时频特征共17种步态特征,构成特征向量,提出融合方法;

[0009] 通过步骤1所得去除噪声的三维动作数据,选用人体下肢部位的标记点进行分析。我们以脚跟前进方向的动作数据提取出每个步态周期,并以每十个步态周期为一个子集,将数据划分为多个子集。

[0010] 通过对下肢足部三维动作数据进行定量分析,提取出步态的时间-距离参数及角度参数,包括步态周期、步长、步高、步宽、步速、步速加速度、足偏角最大值及最小值,取每一子集的平均值;再采用非线性动力学分析的方法提取脚跟前向运动的三维动作数据每一子集的非线性特征,包括近似熵、样本熵、模糊熵、LZ复杂度、C0复杂度特征;最后采用时频分析的方法,通过快速傅里叶变换计算功率谱密度最大值,通过连续小波变换的C-morlet3-3、C-gaus1、mexh三种小波函数计算在尺度2的小波系数模值均值特征。将所提取的特征构成了特征向量,提出融合方法。

[0011] 步骤3.ECG信号的预处理及RR间隔提取;

[0012] 3-1.通过级联的低通和高通滤波器组成的带通滤波器进行处理,抑制基线漂移和T波造成的干扰;其二阶低通和高通滤波器的差分方程分别为:

$$[0013] \quad y(nT) = 2y(nT-T) - y(nT-2T) + x(nT) - 2x(nT-6T) + x(nT-12T)$$

$$[0014] \quad y(mT) = 32x(mT-16T) - [y(mT-T) + x(mT) - x(mT-32T)]$$

[0015] 其中,T为采样周期,x(nT)表示离散时间系统心电信号时间序列的输入,y(nT)为低通滤波后的心电信号,其截止频率设为12Hz,增益为2,处理延迟约为6个采样周期。高通滤波器的截止频率约为5Hz,增益为32,16个采样周期的延迟。

[0016] 3-2.经过带通滤波器处理后的心电信号,再经过导数滤波器处理,提供QRS复斜率信息;其差分方程为:

$$[0017] \quad y(mT) = (1/8T) [-x(mT-2T) - 2x(mT-T) + 2x(mT+T) + x(mT+2T)]$$

[0018] 3-3.经导数滤波器处理后的输出利用非线性平方函数达到非线性放大的效果,强调高频特性,非线性平方函数的差分方程为:

$$[0019] \quad y(mT) = [x(mT)]^2$$

[0020] 3-4.通过移动窗口积分器(MVI)处理产生包络形状的特征波形,来进一步增强特征与噪声的区分。其差分方程为:

$$[0021] \quad y(mT) = (1/N) [x(mT-(N-1)T) + x(mT-(N-2)T) + \dots + x(mT)]$$

[0022] 其中,移动积分窗口的宽度应该与尽可能宽的QRS复合体大致相同,与原始心电信号的采样周期有关联。若采集心电信号的采样频率500Hz,即每秒采集500个样本点,因此一般来说N的样本点数约为75个(即对应的窗口的实时宽度周期为0.15秒)。

[0023] 3-5.使得阈值和其它参数进行周期性调整,并检测R波或QRS波的位置。最后对平滑信号中检测到的R波输出进行分析,并对带通信号的输出进行了两次检查,提高了检测精

度,最终找到真实R波的原始指标。

[0024] 步骤4. 构建步态特征与心电RR间隔之间的高斯过程回归(GPR)关联预测模型;

[0025] 通过步骤3和步骤4所得的心电RR间隔和融合后的步态特征向量,将融合后的步态特征向量作为输入样本集,对应的心电RR间隔特征作为期望输出样本集,建立两者间的高斯过程回归模型。

[0026] 4.1高斯过程(GP)原理

[0027] 高斯过程是指一个正态的随机过程,其任意维有限变量的联合分布服从高斯分布。对于任意有限个 x ,即 $x_1, x_2, \dots, x_n \in N$,其相对应的有限个随机变量 $F = (f(x_1), f(x_2), \dots, f(x_n))^T$,记 $F \sim GP(m(x_i), K)$,即

$$[0028] \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} \sim N \left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_n) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \cdots & k(x_n, x_n) \end{bmatrix} \right)$$

[0029] 其中, F 为联合分布;均值函数和协方差函数(核函数)表示如下:

$$[0030] m(x) = E[f(x)]$$

$$[0031] K = E[(f(x) - m(x))(f(x') - m(x')))]$$

[0032] 4.2GPR回归模型的建立

[0033] 建立高斯回归模型时,考虑目标值 y 含有噪声,即定义为:

$$[0034] y = f(x) + \varepsilon. \text{其中: } f(x) \text{ 为高斯过程, } x \in N, \varepsilon \sim N(0, \sigma_n^2). \text{训练集为:}$$

[0035] $D: y_i = f(x_i) + \varepsilon_i (i = 1, 2, \dots, n)$, 在测试集: $D_* = \{(x_{i*}, y_{i*}) | i = n+1, n+2, \dots, n+n_*\}$ 预测对应的输出值 f_* 的多元高斯分布为:

$$[0036] \begin{bmatrix} y \\ f_* \end{bmatrix} \sim N \left(\mathbf{0}, \begin{bmatrix} K + \sigma_n^2 I & K_*^T \\ K_* & K_{**} \end{bmatrix} \right)$$

[0037] 其中, $K_* = [k(x_*, x_1), k(x_*, x_2), \dots, k(x_*, x_n)]$, $K_{**} = k(x_*, x_*)$, σ_n 为噪声的标准差。根据多元高斯分布的条件,可得高斯过程回归模型预测方程:

$$[0038] (f_* | X, y, X_*) \sim GP(m(x_*), \text{cov}(f_*))$$

[0039] 式中:矩阵 X 由训练集的输入 x_i 的列向量组成;矩阵 X_* 由测试集的输入 x_{i*} 的列向量组成。

$$[0040] m(x_*) = K_*(K + \sigma_n^2 I)^{-1} y$$

$$[0041] \text{cov}(f_*) = K_{**} - K_*(K + \sigma_n^2 I)^{-1} K_*^T$$

[0042] 实测步态预测心电的RR间隔回归模型表示为:

$$[0043] f(x) = \varphi(x)^T \omega, \{x \in R^n, \varphi(x), \omega \in R^m, m > n\}$$

[0044] 其中: $\omega \in R^m$ 表示权值向量,均值函数 $m(x) = E[f(x)]$,核函数 $k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))]$ 。

[0045] 4.3核函数的选取

[0046] 核函数的选取对于高斯过程至关重要,在高斯过程回归模型中协方差函数(即核函数)必须满足Mercer条件。本发明采用了3种核函数:

[0047] 二次有理核函数 (RQ) :

$$[0048] \quad k(x_i, x_j) = \sigma_f^2 \exp\left[1 + \frac{(x_i - x_j)^T M (x_i - x_j)}{2\alpha}\right]^{-\alpha} + \sigma_n^2 \delta^{ij}$$

[0049] 平方指数核函数 (SE)

$$[0050] \quad k(x_i, x_j) = \sigma_f^2 \exp\left[-\frac{1}{2}(x_i - x_j)^T M (x_i - x_j)\right] + \sigma_n^2 \delta^{ij}$$

[0051] Matern核函数 (M)

$$[0052] \quad k(x_i, x_j) = \sigma_f^2 \exp\left[1 + \sqrt{3M}(x_i - x_j)\right] \exp\left[-\sqrt{3M}(x_i - x_j)\right] + \sigma_n^2 \delta^{ij}$$

[0053] 其中, $M = \text{diag}(1^{-2})$, l 为方差尺度, σ_f^2 为信号方差, l 为关联性测定超参数; δ^{ij} 为符号函数, i 和 j 相同时为 1, 不同时为 0; α 为核函数的形状参数。

[0054] 定义 $\theta = [\sigma_f^2, \sigma_n^2]$ 为超参数向量, 通过极大似然法自适应获得最优解。首先建立训练样本条件概率的负对数似然函数, 并令其对超参数求偏导, 然后采用共轭梯度法对该偏导数进行最小化, 得到超参数最优解。

[0055] 步骤 5. 训练和测试高斯过程回归模型, 进行关联预测分析;

[0056] 随机输入训练样本集和测试样本集占比为 4:1, 采用三种核函数: 有理二次核函数 (RQ)、平方指数核函数 (SE)、Matern5/2 核函数 (MA), 进行关联预测分析, 将测试结果十次取其平均值。

[0057] 通过采用均方根误差 (RMSE) 和确定系数 (R^2) 对预测模型的准确性和有效性进行评价, 其 RMSE 越小, 预测结果越好。其表达式如下:

$$[0058] \quad RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - t_i)^2};$$

$$[0059] \quad R^2 = 1 - \frac{\sum_{i=1}^n (y_i - t_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

[0060] 其中, n 表示预测数据的数目, y_i 是实际预测数据值, t_i 代表期望输出值。

[0061] 本发明有益效果如下:

[0062] 1、本发明通过 Codamotion 三维运动捕捉系统和 TrignoTM 无线系统对人体运动的三维坐标数据和同步的 ECG 信号进行采集, 相比目前普遍使用的便携式采集设备, 更具精准性; 本发明通过提出一种基于高斯回归的步态-心电 RR 间隔关联方法, 该方法具有自适应强、容易实现、超参数自适应获取等优势, 具有较好的鲁棒性和泛化性能, 能够更好地揭示步态与心电之间的关联信息。

[0063] 2、本发明通过将步态的时间-距离及角度参数、非线性特征和时频特征提出融合方法, 建立步态与心电 RR 间隔的高斯回归关联预测模型, 相比单个特征的 RR 间隔预测有着较好的预测效果。揭示了人在运动时, 速度越快, RR 间隔的预测越准确的规律。通过非线性动力学分析和时频分析方法提取步态特征, 非线性特征能较好的体现人体步态在运动时的动态变化, 时频特征在时域和频域有着很好的局部性, 分析结果较稳定, 能够很好地反映信号本身。

附图说明

[0064] 图1为本发明体流程图；

[0065] 图2为本发明实施例的单个步态特征及融合特征的实验预测结果表；

[0066] 图3为本发明实施例的五种运动模式的实验预测结果表；

具体实施方式

[0067] 为了使本发明的技术方案更加清晰，下面结合附图和实施例，对本发明作进一步的详细说明。

[0068] 参数说明：

[0069] GC—步态周期；ApEn—近似熵；SL—步长SampEn—样本熵；SH—步高；FuzzyEn—模糊熵；SW—步宽 C_{LZ} —LZ复杂度；SV—步速； C_{C0} —C0复杂度；SA—步速加速度PSD—功率谱密度； FA_{max} —足偏角最大值； $CWT_{Cmorlet}$ —复morlet小波变换系数模值； FA_{min} —足偏角最小值； CWT_{Cgaus} —复Cgaus小波变换系数模值；SV—步速； CWT_{mexh} —墨西哥帽小波变换的系数模值；SA—步速加速度；F—融和特征向量集；GPR-RQ—二次有理核函数的高斯过程回归；GPR-SE—平方指数核函数的高斯过程回归；GPR-M—Matern5/2核函数的高斯过程回归；BP—前馈神经网络；BT—提升树；RF—随机森林。

[0070] 如图1所示，为本发明实施例的高斯回归的步态-心电RR间隔关联方法整体流程图，包括如下步骤：

[0071] 步骤1. 获取人体下肢足部标记点的运动三维坐标数据及对应ECG信号；

[0072] 运动三维坐标数据的获取通过Codamotion三维实时运动捕捉系统进行跟踪，通过WiFi与计算机连接，使用步态分析软件设置100Hz的采样频率下记录身体下肢运动位置。在被测者足部的左、右跟肘水平，左、右后跟低水平，左肘和右肘，第五左、右趾位置放置8个标记点。通过放置在实验室工作空间中的Coda CX1单元检测以覆盖运行范围。ECG信号的获取通过TrignoTM无线系统记录心电图（ECG）活动，采集1条通道的心电图，其采样频率为1000Hz，并与100Hz的身体部位运动学坐标数据同步。

[0073] 五种运动模式方案包括：(1) 舒适行走：参与者提高跑步机速度，直到达到舒适行走速度，并以这个速度步行2min；(2) 步行：参与者增加或降低跑步机速度，直到4km/h，并以这个速度行走2min；(3) 最大速度行走：参与者提高速度，直到它的最大步行速度，并以这个速度行走1min；(4) 行走和跑步：休息后，参与者将跑步机的速度提高到最大步行速度，步行30s，然后以同样的速度跑1min；(5) 最大速度跑：参与者提高了速度，直到他跑步的极限，并以该速度进行2min。17名健康志愿者参与数据记录，所有参与者都没有被诊断出任何疾病，并为正常体型，且都爱好运动。所有受试者都采用同样的设备，相同的标记集和同等的实验条件。

[0074] 步骤2. 对实时捕捉的运动三维坐标数据进行分析，提取步态特征；

[0075] 通过定量分析、非线性动力学分析、时频分析方法，提取时间-距离及角度参数、非线性特征、时频特征共17种步态特征，构成特征向量，提出融合方法；具体包括以下方法：

[0076] (1) 首先，通过步骤1所得去除噪声的三维动作数据，选用人体的下肢部位的标记点进行分析。我们以脚跟前进方向的步态数据提取出每个步态周期，并以每十个步态周期为一个子集，将数据划分为多个子集。

[0077] (2) 定量分析常利用相关仪器来获得一组步态参数,如时间-空间距离参数,包括与步态和速度相关参数的测量,运动学涉及运动几何相关的参数,比如关节角度,进而完成对步态的分析。我们将下肢三维动作数据进行定量分析,提取出步态周期(GC)、步长(SL)、步高(SH)、步宽(SW)、步速(SV)、步速加速度(SA)、足偏角最大值(FA_{max})及最小值(FA_{min}),再取每一子集的均值为特征。特征表示具体如下:

[0078] 步态周期(GC)描述了从支撑脚跟在地面的初始位置到同一脚跟第二次接触地面的运动过程。步长(SL)是行走时左右足跟先后着地时连续接触点之间的纵向直线距离。每步行周期过程中足跟抬起的最大高度来表示步幅高度(SH)。步幅宽度则是指人行走时,两脚跟之间的横向距离(SW)。足角指贯穿一侧足底的中心线与前进方向所成的夹角。步速(SV)指人体步行的平均速度,步速加速度(SA)是单位时间内足跟前进方向速度的平均变化量。

[0079] (3) 再采用非线性动力学分析的方法提取脚跟前向运动的三维动作数据每一子集的近似熵(ApEn)、样本熵(SampEn)、模糊熵(FuzzyEn)、LZ复杂度(C_{LZ})、C0复杂度特征(C_{C0});方法具体如下:

[0080] 近似熵(ApEn)是一种用于量化时间序列波动的规律性和不可预测性的非线性动力学参数,反映了时间序列中新信息发生的可能性,熵值越大越复杂。则近似熵的值可通过以下步骤求得:

[0081] 1) 将步态时间序列 $\{x_i\}$ 按顺序组成 m 维矢量,即

[0082] $X(i) = [x(i), x(i+1), \dots, x(i+m-1)]$,其中, $i=1, 2, \dots, N-m+1$

[0083] 2) 定义 $X(i)$ 与 $X(j)$ 间的距离 $d[X(i), X(j)]$ 为两者对应元素中差值最大的一个,即

[0084] $d[X(i), X(j)] = \max_k |x(i+k) - x(j+k)|$, $j=1, 2, \dots, N-m+1$

[0085] 3) 按给定的阈值 r ($r > 0$),对每一个 i 值统计 $d[X(i), X(j)] < r$ 的数目与总的矢量个数 $N-m+1$ 的比值,记为 $C_i^m(r)$:

[0086] $C_i^m(r) = \frac{1}{N-m+1} \text{num}\{d[X(i), X(j)] < r\}, i=1, 2, \dots, N-m+1$

[0087] 4) 定义函数 $\Phi^m(i)$

[0088] $\Phi^m(i) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(r)$

[0089] 5) 定义算法相关参数 $m=2, r=0.2*SD$,SD为标准差。计算该序列的近似熵为:

[0090] $ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r)$

[0091] 样本熵(SampEn)是一种新的时间序列复杂性测度方法。样本熵在算法上相对于近似熵算法的改进,其 j 的取值范围为 $[1, N-m+1]$,但 $j \neq i$ 。

[0092] $B^m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} B_i^m(r)$

[0093] 估计样本熵为: $SampEn(m, r, N) = -\ln[B^{m+1}(r)/B^m(r)]$

[0094] 我们在该方法中选取实际的计算中,取 $m=2, r=0.2*SD$

[0095] 模糊熵(FuzzyEn)是一种衡量时间序列规律性的方法,在FuzzyEn中模糊地定义了向量的相似性。FuzzyEn拥有SampEn熵的优良特性,且独立的、同分布的随机数和周期性正

弦信号表明,FuzzyEn能够更有效地测量时间序列的规律性。其算法描述如下:

[0096] 1) 对于给定的N点样本时间序列 $[x(1), x(2), \dots, x(N)]$

[0097] 2) 定义相空间维数 $m(m \leq N-2)$ 和相似容限度 r ,重构相空间如下: $X(i) = [x(i), x(i+1), \dots, x(i+m-1)] - x_0(i), i=1, 2, \dots, N-m+1$

[0098] 其中, $x_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} x(i+j)$,

[0099] 3) 模糊函数 A_{ij}^m 是指数函数,计算公式如下:

[0100] $A_{ij}^m = \exp[-\ln(2) \cdot (d_{ij}^m / r)^2], j=1, 2, \dots, N-m+1, j \neq i$

[0101] 对于向量 $X(i)$,将 $X(i)$ 与 $X(j)$ 之间的最大绝对距离 d_{ij}^m 表示如下:

[0102] $d_{ij}^m = d[X(i), X(j)] = \max_{p=1, 2, \dots, m} (|x(i+p-1) - x_0(i)| - |x(j+p-1) - x_0(j)|)$

[0103] 4) 对每个 i 的 A_{ij}^m ,求取平均值,得到

[0104] $C_i^m(r) = \frac{1}{N-m} \sum_{j=1, j \neq i}^{N-m+1} A_{ij}^m$

[0105] 5) 定义函数 $\Phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} C_i^m(r)$

[0106] 6) 因此,FuzzyEn表示为:FuzzyEn(m, r, N) = $\ln \Phi^m(r) - \ln \Phi^{m+1}(r)$

[0107] Lempel-Ziv复杂度(CLz)反映了一个时间序列随其长度的增长出现新模式的速度。复杂度值越大,说明数据在窗口长度时期内随时间出现的新变化越多,发生新变化的速率越快,表明这一时期的数据变化是无序而复杂的;反之,复杂度越小,则说明发生新变化的速率越慢,数据变化是规则的,周期性越强。具体算法如下:

[0108] 1) 采用二值化方法对序列 $X = \{X_1, X_2, \dots, X_n\}$ 进行粗粒化处理,形成“0-1”序列 $P = \{S_1, S_2, \dots, S_n\}$ 。

[0109] 2) 对上面所得到的“0-1”序列,依次检索其中的新字符串。新字符串需满足唯一性和连续性,并且用一个“·”将前后分开,其中新字符串的检索过程详细如下:

[0110] 3) 用 S 表示序列 P 的部分字符串,即 $S = \{S_1, S_2, \dots, S_r\} (r \leq n)$;用 Q 表示 S 的子串,即 $Q = \{Q_1, Q_2, \dots, Q_m\}$;用 SQ 表示 S, Q 的串联组合,即 $SQ = \{S, Q\}$

[0111] 用 SQ^π 表示删去 SQ 的最后一个字符所得的字符串;用 $V(SQ^\pi)$ 表示 SQ^π 中所得的子串集合;用 $d(n)$ 表示不同子串的数目。对于给定序列, $P = \{S_1, S_2, \dots, S_n\}$,开始时,取 $S = S_1, Q = S_2, SQ^\pi = S_1, d(n) = 1$ 。一般情况下,

[0112] 令 $S = \{S_1, S_2, \dots, S_r\} (r = 2, 3, \dots, n-1)$;那么 $SQ^\pi = \{S_1, S_2, \dots, S_r\} (r = 2, 3, \dots, n-1)$,

[0113] $Q = S_{r+1}$,判断 Q 是否为 S 的一个子串,如果 Q 属于 $V(SQ^\pi)$,则 Q 是 SQ^π

[0114] 的一个子串而非新子串, S 保持不变,令 $Q = S_{r+1}S_{r+2}$,只到 Q 不属于 $V(SQ^\pi)$ 为止,即有 $Q = \{S_{r+1}, S_{r+2}, \dots, S_{r+i}\}$ 不属于 $SQ^\pi = \{S_1, S_2, \dots, S_{r+i-1}\}$,此时 $d(n)$ 加1,一次重复上述过程直到最后一个字符。计算得到一个用“·”分成段的字符串的数目,定义复杂度 $d(n)$ 。

[0115] 4) 对复杂度进行归一化;为了获得不依赖于序列长度的复杂度,需要将 $d(n)$ 归一

化。发现所有的符号序列复杂度 $d(n)$ 都趋于一个确定的值,即:

$$[0116] \quad \lim_{n \rightarrow \infty} d(n) = b(n) = \frac{n}{\log a(n)}$$

[0117] 其中 n 为序列长度, a 为字符串中不同字符的个数(“0-1”序列 $a=2$)

[0118] 将 $b(n)$ 对 $d(n)$ 进行归一化,

[0119] 5) 得出归一化的LZ复杂度如下所示:

$$[0120] \quad D(n) = \frac{d(n)}{b(n)} = \frac{d(n) \cdot \log_a(n)}{n}$$

[0121] C0复杂度(C_0)算法思想主要是将要分析的步态时间参数序列进行分解,分为随机序列的部分以及规则序列部分,对序列的复杂度预测就被定义为做随机运动部分的步态时间参数序列长度和全部原始步态时间参数序列长度分别与时间轴之间围成的面积之比。该方法具体如下:

[0122] 1) 对于给定一个长度为 N 的步态时间参数序列 $x(t)$,其中 $t \in 0, 1, 2, \dots, N-1$;对其进行离散傅里叶变换:

$$[0123] \quad f(k) = \frac{1}{N} \sum_{t=0}^{N-1} x(t) e^{-2\pi i \frac{tk}{N}} = \frac{1}{N} \sum_{t=0}^{N-1} x(t) W_N^{-tk}$$

[0124] 式中 i 为虚数单位, $W_N = e^{\frac{2\pi i}{N}}$, $k = 0, 1, 2, \dots, N-1$;

[0125] 2) 求出 $f(k)$ 的均方值:

$$[0126] \quad G_N = \frac{1}{N} \sum_{t=0}^{N-1} |f(k)|^2$$

[0127] 3) 定义一个大于1的正常数 δ ,对 $f(k)$ 序列中超过均方值 δ 倍的频谱予以保留,反之则将其置零处理,得到转换后的序列:

$$[0128] \quad \tilde{f}(k) = \begin{cases} f(k), & |f(k)|^2 > rG_N \\ 0, & |f(k)|^2 \leq rG_N \end{cases}$$

[0129] 4) 对上式进行傅里叶逆变换可得:

$$[0130] \quad \tilde{x}(t) = \sum_{k=0}^{N-1} \tilde{f}(k) W_N^{tk}$$

[0131] 其中 $W_N = e^{\frac{2\pi i}{N}}$;根据C0复杂度算法定义,便可得到C0复杂度计算值:

$$[0132] \quad C_0 = \frac{\sum_0^{N-1} |x(t) - \tilde{x}(t)|^2}{\sum_0^{N-1} |x(t)|^2}$$

[0133] (4) 最后采用时频分析的方法,通过快速傅里叶变换计算每一子集的功率谱密度最大值,通过连续小波变换的三种小波函数(Cmorlet3-3、Cgaus1、mexh)计算在尺度2的三种小波系数模值均值特征(CWT_{cmorlet}、CWT_{cgaus}、CWT_{mexh})。时频分析方法具体如下:

[0134] 快速傅里叶变换:通过计算每一子集的傅里叶变换所得幅值进行功率谱密度

(PSD)的计算,以功率谱密度最大值作为一个特征,形成特征向量。快速傅里叶变换是计算离散傅里叶变换的快速算法。DFT的定义式为:

$$[0135] \quad X(k) = \sum_{j=0}^{N-1} x(j) \omega_N^{(j-1)(k-1)}, k=1,2,\dots,N$$

[0136] 在所有负指数 $\omega_N = e^{(-2\pi i)/N}$ 的值全部已算好的情况下,要计算一个 $X(k)$ 需要 N 次复数乘法和 $N-1$ 次复数加法。算出全部 N 点 $X(k)$ 共需 N^2 次复数乘法和 $N(N-1)$ 次复数加法。

[0137] 功率谱密度 (PSD) 计算公式为: $P = y \cdot \text{conj}(y) / N$ 。

[0138] 连续小波变换:通过对小波基函数的伸缩和平移来实现对信号的多尺度细化分析,其特点是能充分突出数据在局部方面的统计特征。由于小波变换具有良好的性质,如线性可加性、平移不变性、伸缩共变性、自相似性,利用这些性质可以更细化特征。 $f(t)$ 表示一维连续信号,则连续小波变换的基本公式为:

$$[0139] \quad CWT(f, a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt$$

[0140] 其中, $CWT(f, a, b)$ 为信号函数 $f(t)$ 在尺度 a 、位置 b 的小波变换系数, ψ^* 为满足一定条件的小波函数 ψ 的共轭函数。通过采用 `morlet3-3`、`C-gaus1`、`mexh` 三种小波函数的小波变换计算每一子集在尺度 2 的小波系数模值 (CWT_{morlet} 、 CWT_{cgaus} 、 CWT_{mexh})。

[0141] (5) 最后我们将提取的 17 种特征融合,形成特征向量 F 。

[0142] 步骤 3. ECG 信号的预处理及 RR 间隔提取;

[0143] (1) 通过级联的低通和高通滤波器组成的带通滤波器进行处理,抑制基线漂移和 T 波造成的干扰;其二阶低通和高通滤波器的差分方程分别为:

$$[0144] \quad y(nT) = 2y(nT-T) - y(nT-2T) + x(nT) - 2x(nT-6T) + x(nT-12T)$$

$$[0145] \quad y(mT) = 32x(mT-16T) - [y(mT-T) + x(mT) - x(mT-32T)]$$

[0146] 其中, T 为采样周期, $x(nT)$ 表示离散时间系统心电信号时间序列的输入, $y(nT)$ 为低通滤波后的心电信号,其截止频率设为 12Hz,增益为 2,处理延迟约为 6 个采样周期。高通滤波器的截止频率约为 5Hz,增益为 32,16 个采样周期的延迟。

[0147] (2) 经过带通滤波器处理后的心电信号,再经过导数滤波器处理,提供 QRS 复斜率信息;其差分方程为:

$$[0148] \quad y(mT) = (1/8T) [-x(mT-2T) - 2x(mT-T) + 2x(mT+T) + x(mT+2T)]$$

[0149] (3) 经导数滤波器处理后的输出利用非线性平方函数达到非线性放大的效果,强调高频特性,非线性平方函数的差分方程为:

$$[0150] \quad y(mT) = [x(mT)]^2$$

[0151] (4) 通过移动窗口积分器 (MVI) 处理产生包络形状的特征波形,来进一步增强特征与噪声的区分。其差分方程为:

$$[0152] \quad y(mT) = (1/N) [x(mT-(N-1)T) + x(mT-(N-2)T) + \dots + x(mT)]$$

[0153] 其中,移动积分窗口的宽度应该与尽可能宽的 QRS 复合体大致相同,与原始心电信号的采样周期有关联。在本研究中,采集心电信号的采样频率 1000Hz,即每秒采集 1000 个样本点,因此一般来说 N 的样本点数约为 150 个(即对应的窗口的实时宽度周期为 0.15s)。

[0154] (5) 使得阈值和其它参数进行周期性调整,并检测 R 波或 QRS 波的位置。最后对平滑信号中检测到的 R 波输出进行分析,并对带通滤波器的输出进行了两次检查,提高了检测精

度,最终找到真实R波的原始指标。

[0155] 设经过积分处理后的ECG信号为 $Y(n)$,经过下述公式对ECG信号进行训练,得到初始阈值,即:

$$[0156] \quad THR_{sig} = \frac{1}{3} \max_n X(n)$$

$$[0157] \quad THR_{noise} = \frac{1}{2} \left[\frac{1}{n} \sum_{i=1}^n X(i) \right]$$

[0158] 其中,以 $2s$ 作为初始化阈值的训练期,则 $n=2fs$, f_s 为ECG信号的采样频率, THR_{sig} 信号幅值初始阈值, THR_{noise} 为干扰初始阈值。为了使得初始阈值更具合理性,取信号 $X(n)$ 的前 $20s$ 的样本数据,并分为 10 组进行初始化阈值操作,根据上述公式计算出 10 个阈值,再去除 10 组中的最大值和最小值,以免造成可能出现的尖峰噪声干扰等情况导致阈值过大或过小引起的误差。最后再对剩下的 8 个值求平均值得到 THR_{sig} 和 THR_{noise} 分别作为信号幅值初始阈值和干扰初始阈值。设检测到的信号峰值为 $PEAK$,若 $PEAK > THR_{sig}$,则该峰值所对应的位置作为预选的QRS波,并得到一个估计信号电平值 LEV_{sig} 。

[0159] 更新的估计信号电平 LEV_{sig} 和估计干扰电 LEV_{noise} ,对阈值进行调整,公式如下:

$$[0160] \quad THR_{sig} = \alpha \cdot LEV_{sig} + \beta \cdot LEV_{noise}$$

$$[0161] \quad THR_{noise} = \frac{1}{2} THR_{sig}$$

[0162] 其中, α 、 β 为阈值调整贡献的加权因子,通过多次实验,取 $\alpha=0.25$, $\beta=0.75$ 。在QRS波群中检测到R波,并计算当前相邻两个R波峰峰值的时间为RR间隔,取平均值。

[0163] 步骤4. 构建步态特征与心电RR间隔之间的高斯过程回归(GPR)关联预测模型;

[0164] 通过步骤3和步骤4所得的心电RR间隔和融合后的步态特征向量,将融合后的步态特征向量作为输入样本集,对应的心电RR间隔特征作为期望输出样本集,建立两者间的高斯过程回归模型。

[0165] 4.1高斯过程(GP)原理

[0166] 高斯过程是指一个正态的随机过程,其任意维有限变量的联合分布服从高斯分布。对于任意有限个 x ,即 $x_1, x_2, \dots, x_n \in N$,其相对应的有限个随机变量 $F = (f(x_1), f(x_2), \dots, f(x_n))^T$,记 $F \sim GP(m(x_i), K)$,即

$$[0167] \quad \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} \sim N \left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_n) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \cdots & k(x_n, x_n) \end{bmatrix} \right)$$

[0168] 其中, F 为联合分布;均值函数和协方差函数(核函数)表示如下:

$$[0169] \quad m(x) = E[f(x)]$$

$$[0170] \quad K = E[(f(x) - m(x))(f(x') - m(x'))]$$

[0171] 4.2GPR回归模型的建立

[0172] 建立高斯回归模型时,考虑目标值 y 含有噪声,即定义为:

$$[0173] \quad y = f(x) + \varepsilon. \text{其中: } f(x) \text{ 为高斯过程, } x \in N, \varepsilon \sim N(0, \sigma_n^2). \text{训练集为:}$$

$$[0174] \quad D: y_i = f(x_i) + \varepsilon_i (i=1, 2, \dots, n)$$

[0175] 需要在测试集: $D_* = \{(x_{i*}, y_{i*}) | i = n+1, n+2, \dots, n+n_*\}$ 预测对应的输出值 f_* 的多元高斯分布为:

$$[0176] \quad \begin{bmatrix} y \\ f_* \end{bmatrix} \sim N \left(0, \begin{bmatrix} K + \sigma_n^2 I & K_*^T \\ K_* & K_{**} \end{bmatrix} \right)$$

[0177] 其中, $K_* = [k(x_*, x_1), k(x_*, x_2), \dots, k(x_*, x_n)]$, $K_{**} = k(x_*, x_*)$, σ_n 为噪声的标准差。

[0178] 根据多元高斯分布的条件, 可得高斯过程回归模型预测方程:

$$[0179] \quad (f_* | X, y, X_*) \sim GP(m(x_*), \text{cov}(f_*))$$

[0180] 式中: 矩阵 X 由训练集的输入 x_i 的列向量组成; 矩阵 X_* 由测试集的输入 x_{i*} 的列向量组成。

$$[0181] \quad m(x_*) = K_*(K + \sigma_n^2 I)^{-1} y$$

$$[0182] \quad \text{cov}(f_*) = K_{**} - K_*(K + \sigma_n^2 I)^{-1} K_*^T$$

[0183] 实测步态预测心电的RR间隔回归模型表示为:

$$[0184] \quad f(x) = \varphi(x)^T \omega, \{x \in R^n, \varphi(x), \omega \in R^m, m > n\}$$

[0185] 其中: $\omega \in R^m$ 表示权值向量, 均值函数 $m(x) = E[f(x)]$, 核函数 $k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))]$ 。

[0186] 4.3 核函数的选取

[0187] 核函数的选取对于高斯过程至关重要, 在高斯过程回归模型中协方差函数 (即核函数) 必须满足 Mercer 条件。本发明采用了 3 种核函数:

[0188] 二次有理核函数 (RQ):

$$[0189] \quad k(x_i, x_j) = \sigma_f^2 \exp \left[1 + \frac{(x_i - x_j)^T M (x_i - x_j)}{2\alpha} \right]^{-\alpha} + \sigma_n^2 \delta^{ij}$$

[0190] 平方指数核函数 (SE)

$$[0191] \quad k(x_i, x_j) = \sigma_f^2 \exp \left[-\frac{1}{2} (x_i - x_j)^T M (x_i - x_j) \right] + \sigma_n^2 \delta^{ij}$$

[0192] Matern 核函数 (M)

$$[0193] \quad k(x_i, x_j) = \sigma_f^2 \exp \left[1 + \sqrt{3M} (x_i - x_j) \right] \exp \left[-\sqrt{3M} (x_i - x_j) \right] + \sigma_n^2 \delta^{ij}$$

[0194] 其中, $M = \text{diag}(l^{-2})$, l 为方差尺度, σ_f^2 为信号方差, l 为关联性测定超参数; δ^{ij} 为符号函数, i 和 j 相同时为 1, 不同时为 0; α 为核函数的形状参数。

[0195] 定义 $\theta = [\sigma_1, \sigma_f^2, \sigma_n^2]$ 为超参数向量, 通过极大似然法自适应获得最优解。首先建立训练样本条件概率的负对数似然函数, 并令其对超参数求偏导, 然后采用共轭梯度法对该偏导数进行最小化, 得到超参数最优解。

[0196] 步骤 5. 训练和测试高斯过程回归模型, 进行关联预测分析;

[0197] 随机输入训练样本集和测试样本集占比为 4:1, 采用高斯回归的三种核函数: 有理二次核函数 (GPR-RQ)、平方指数核函数 (GPR-SE)、Matern 5/2 核函数 (GPR-M), 对五种运动场

景进行关联预测分析,将测试结果十次取其平均值。

[0198] 通过采用均方根误差 (RMSE) 对预测模型的准确性和有效性进行评价,其RMSE越小,预测结果越好。其表达式如下:

$$[0199] \quad RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - t_i)^2}$$

[0200] 其中, n 表示预测数据的数目, y_i 是实际预测数据值, t_i 代表期望输出值, \bar{t}_i 代表期望输出均值。

[0201] 如图2所示为本发明实施例的单个特征及融合特征的实验预测结果表,由实验结果可知:在三个核函数的GPR-RQ、GPR-SE、GPR-M高斯回归预测中,Matern核函数的高斯过程回归模型 (GPR-M) 对17种步态特征单个测试效果较好。其融合特征向量F相比单个特征的预测效果更好,具有更低的均方根误差。其中,定量分析提取的步态参数有步速 (SV)、步速加速度 (SA)、步态周期 (GC) 具有较好的预测效果;非线性特征的模糊熵 (FuzzyEn) 和近似熵 (ApEn) 以及时频特征的墨西哥帽小波变换系数模值 (CWT_{mexh}) 都具有较好的预测效果,能够较好的反映步态信号本身。而通过实验结果可以看出相比较于传统的预测方法,高斯回归方法的预测精度更高。如图3所示为本发明实施例的五种运动模式的实验预测结果表。对五种运动场景的预测中,也表现了GPR-M方法相比前馈BP神经网络、提升树 (BT) 和随机森林 (RF) 三种算法对舒适行走、步行、最大速度走、行走和跑、跑步五种运动模式的预测结果较好。而且随着行走速度的增加,表现其预测的结果越准确,从而揭示了步态与心电RR间隔的强关联性。由上述实验结果可看出高斯回归方法的泛化性能更好,预测精度更高,能够更好地揭示步态与心电之间的关联信息。

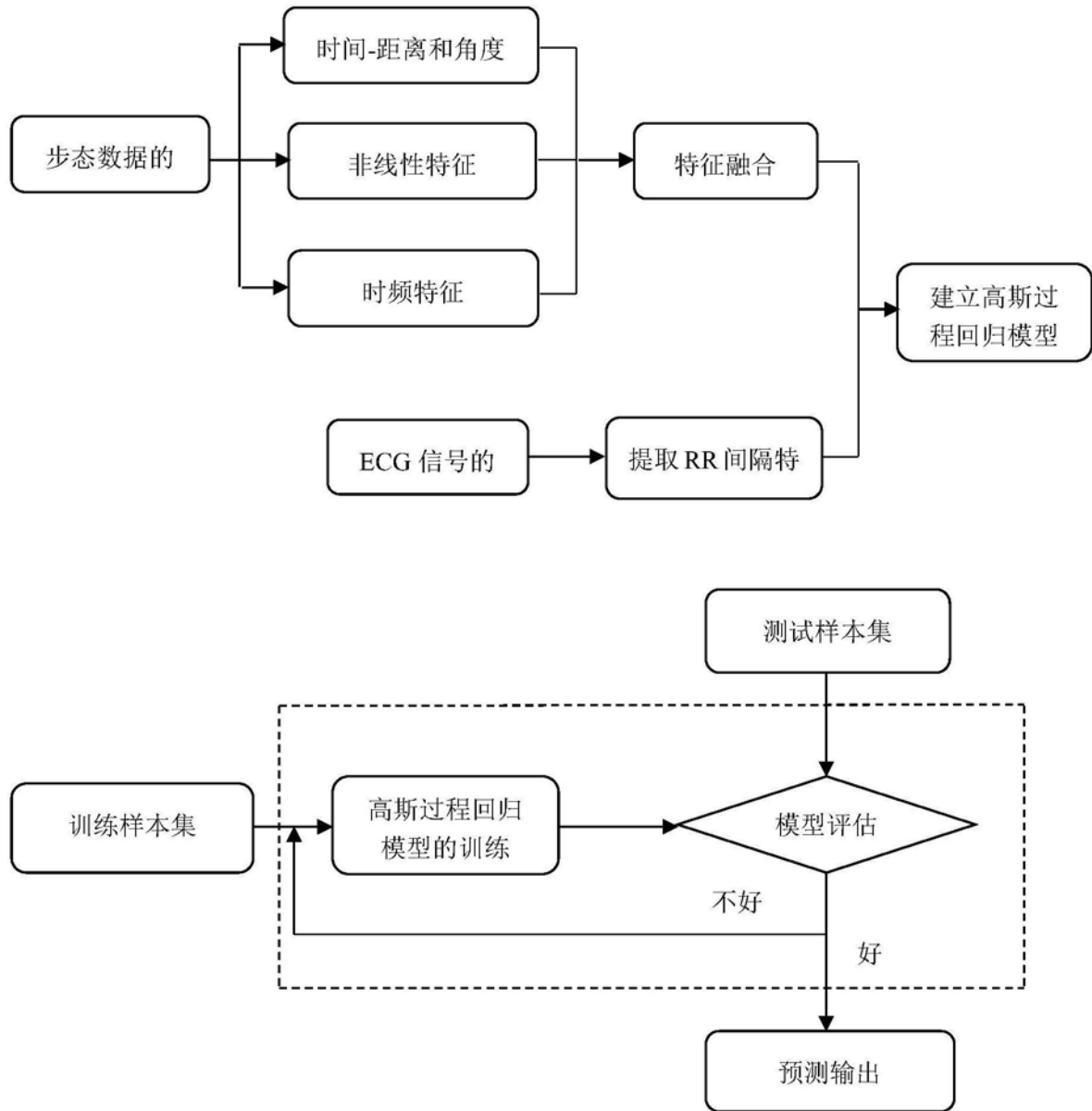


图1

特征	RMSE					
	<u>GPR-RQ</u>	<u>GPR-SE</u>	<u>GPR-M</u>	BP	BT	RF
GC	0.0518	0.0575	0.0532	0.0630	0.0579	0.0525
SL	0.1211	0.1206	0.1206	0.1340	0.1238	0.1298
SH	0.0930	0.0944	0.0938	0.0956	0.0962	0.0986
SW	0.1333	0.1334	0.1331	0.1434	0.1366	0.1440
<u>FA_{max}</u>	0.1347	0.1348	0.1345	0.1594	0.1373	0.1455
<u>FA_{min}</u>	0.1003	0.1018	0.1018	0.0980	0.1038	0.1062
SV	0.0680	0.0691	0.0687	0.0790	0.0732	0.0711
SA	0.0471	0.0486	0.0477	0.0601	0.0532	0.0497
<u>ApEn</u>	0.0617	0.0616	0.0612	0.0665	0.0672	0.0672
<u>SampEn</u>	0.0670	0.0670	0.0668	0.0898	0.0720	0.0713
<u>FuzzyEn</u>	0.0570	0.0592	0.0571	0.0657	0.0624	0.0614
<u>C₁₂</u>	0.0861	0.0867	0.0865	0.0708	0.0890	0.0899
<u>C₀₀</u>	0.1252	0.1246	0.1218	0.1218	0.1192	0.1235
PSD	0.0965	0.0906	0.0901	0.0878	0.0941	0.0955
<u>CWT_{C_{max}}</u>	0.0781	0.0779	0.0778	0.0765	0.0816	0.0833
<u>CWT_{C_{min}}</u>	0.0648	0.0622	0.0621	0.0721	0.0697	0.0679
<u>CWT_{avg}</u>	0.0594	0.0528	0.0520	0.0566	0.0583	0.0555
F	0.0201	0.0209	0.0200	0.0477	0.0343	0.0208

图2

特征	RMSE					
	<u>GPR-RQ</u>	<u>GPR-SE</u>	<u>GPR-M</u>	BP	BT	RF
舒适行走	0.0518	0.0575	0.0532	0.0630	0.0579	0.0525
步行	0.1211	0.1206	0.1206	0.1340	0.1238	0.1298
最大速度走	0.0648	0.0622	0.0621	0.0721	0.0697	0.0679
行走和跑	0.0594	0.0528	0.0520	0.0566	0.0583	0.0555
跑步	0.0201	0.0209	0.0200	0.0477	0.0343	0.0208

图3

专利名称(译)	基于高斯回归的步态-心电RR间隔关联方法		
公开(公告)号	CN110236523A	公开(公告)日	2019-09-17
申请号	CN201910521097.3	申请日	2019-06-17
[标]申请(专利权)人(译)	杭州电子科技大学		
申请(专利权)人(译)	杭州电子科技大学		
当前申请(专利权)人(译)	杭州电子科技大学		
[标]发明人	王建中 黄泽银 曹九稳		
发明人	王建中 黄泽银 曹九稳		
IPC分类号	A61B5/0402 A61B5/0472 A61B5/00		
CPC分类号	A61B5/0402 A61B5/0472 A61B5/7203 A61B5/725 A61B5/7267		
代理人(译)	朱月芬		
外部链接	Espacenet SIPO		

摘要(译)

本发明公开了一种基于高斯回归的步态-心电RR间隔关联方法。本发明包括以下步骤：步骤1.分别利用三维运动跟踪系统和心电采集设备，同步采集受试者运动的三维坐标数据及对应的ECG信号；步骤2.对三维坐标数据进行步态定量分析、非线性分析、时频分析，提取出17种特征，构成特征向量，提出融合方法；步骤3.对ECG信号进行预处理提取RR间隔；步骤4.构建步态特征与心电RR间隔之间的高斯过程回归(GPR)关联预测模型；步骤5.训练和测试高斯过程回归模型，进行关联预测分析。本发明自适应强、容易实现、超参数自适应获取等优势，具有较好的鲁棒性和泛化性能，能够更好地揭示步态与心电之间的关联信息。

