



(12) 发明专利申请

(10) 申请公布号 CN 105403705 A

(43) 申请公布日 2016. 03. 16

(21) 申请号 201510889922. 7

(22) 申请日 2015. 12. 05

(71) 申请人 浙江大学

地址 310027 浙江省杭州市西湖区浙大路
38 号

(72) 发明人 赵春晖 宋广健

(74) 专利代理机构 杭州求是专利事务所有限公
司 33200

代理人 邱启旺

(51) Int. Cl.

G01N 33/66(2006. 01)

权利要求书3页 说明书8页 附图6页

(54) 发明名称

一种包含血糖分类功能故障检测模块的连续
血糖监测设备

(57) 摘要

本发明公开了一种包含血糖分类功能故障检测模块的连续血糖监测设备,该设备通过建立血糖 PCA 监测图对连续血糖监测仪进行故障检测,以便提供准确的血糖信号。由于外源输入的影响(如饮食、病人情感变化等),病人的血糖波动情况会发生改变,相应的血糖信号 SPE 值也有明显改变变化;此外,血糖信号间存在明显的非线性,因此无法建立起一个统一有效的控制限来检测血糖仪是否发生故障。本发明中将血糖信号分为若干类,保证类内信号局部线性化且血糖波动情况相近,从而保证每类内能更好的进行血糖仪故障检测,为后续的血糖预测、控制等提供更为准确可信的数据,具有重要的作用。

1. 一种包含血糖分类功能故障检测模块的连续血糖监测设备,其特征在于,该设备包括:用于采集人体血糖信息,输出血糖可用信号的传感器;用于对传感器的输出信号进行放大处理的信号放大器;用于对信号放大器输出的模拟信号进行数字转换的单片机;用于对单片机输出的数字信号进行数据处理的故障检测模块,故障检测模块可以集成在单片机中,也可以单独使用;用于将故障检测模块输出的血糖监测值 SPE 显示的显示器;用于数据存储的存储器;所述故障检测模块进行数据处理的过程包括以下步骤:

(1) 血糖数据预处理:将以一定采样周期 Δt 获得的单片机输出的连续血糖信号组合成一维时序数据 $x_{g_{1 \times 1}}$,其中 x_g 表示检测到的血糖信号,1 为采样个数;

(2) 对于连续血糖信号建立 PCA 模型:将连续 m 个血糖信号作为 PCA 模型多维输入,即 $x(k) = [x_g(k-m+1), x_g(k-m+2), \dots, x_g(k)]^T$

$$T = XP$$

$$E = X - XPP^T$$

其中 $X = [x(1), x(2), \dots, x(n)]^T$ 是由 $n \times m$ 维的血糖数据组成的训练矩阵; P 为 $m \times r$ 维负载矩阵, r 为保留的主元个数, $r \leq m$; T 为 $n \times r$ 维得分矩阵,为血糖数据矩阵 X 在主元空间的投影; $E = [e(1), e(2), \dots, e(n)]^T$ 为 $n \times m$ 维的残差矩阵。通过统计量 SPE 对血糖数据进行监测:

$$SPE(i) = e(i)^T e(i)$$

SPE 的控制限可以通过以下公式近似求取:

$$SPE \sim g\chi_{h,\alpha}^2$$

其中 $g = v/2m$, $h = 2m^2/v$, m 为全部 SPE 的均值, v 为全部 SPE 的方差, α 为控制线的置信度。

(3) 通过 PCA 迭代的方法将病人的血糖信号进行分类,并建立相应的 PCA 模型;

(4) 根据步骤 (3) 的分类结果,建立血糖仪在线故障检测的 PCA 监测图,具体包括以下子步骤:

(4.1) 将在线传入的血糖数据分别代入各类 PCA 模型,计算 SPE 值:

$$e_{test,i}^T = x_{test}^T - x_{test}^T P_i P_i^T$$

$$SPE_{new,i} = e_{test,i}^T e_{test,i}$$

其中 x_{test} 是当前时刻输入的 $m \times 1$ 维的血糖测试数据, P_i 是第 i 类 PCA 模型的负载矩阵,而 $e_{test,i}$ 是测试数据代入第 i 类 PCA 模型所得的残差向量, $SPE_{test,i}$ 是 $e_{test,i}$ 对应的 SPE 值;

(4.2) 确定检测结果:

$$result_{test,i} = SPE_{test,i} < SPE_Limit_i$$

$$result_{test} = result_{test,1} \vee result_{test,2} \vee \dots \vee result_{test,k}$$

其中 $result_{test,i}$ 是测试数据在第 i 类 PCA 模型下检测结果,其中 $result_{test,i} = 1$ 代表检测结果正常, $result_{test,i} = 0$ 代表检测结果异常; \vee 为取并集的逻辑运算; $result_{test}$ 为最终检测结果,为所有类 PCA 模型下检测结果的并集,当连续若干时刻的故障检测结果均为异常,则可说明血糖仪发生故障。

2. 根据权利要求 1 所述一种包含血糖分类功能故障检测模块的连续血糖监测设备,其

特征在于,所述步骤 (3) 具体包括以下子步骤:

(3.1) 对于某个特定病人选取足量血糖仪正常工作时的血糖数据作为训练数据,用于病人血糖分类;并将这些训练数据如步骤(2)铺展成n个采样点组成的n×m维数据矩阵X。假定全部训练数据属于同一类(即第一类),即:

$$X_1 = X = [x(1), x(2), \dots, x(n)]^T;$$

(3.2) 利用 PCA 迭代的方法将血糖分为若干类,具体步骤如下:

(3.2.1) 根据步骤(2)对第一类的数据矩阵建立 PCA 模型:

$$E_1 = X_1 - X_1 P_1 P_1^T$$

$$SPE_1(i) = e_1(i)^T e_1(i)$$

$$SPE_Limit_1 = g_1 \chi_{h,\alpha}^2$$

(3.2.2) 通过 PCA 模型控制限 SPE_Limit₁ 可以将数据 X₁ 分为两类,即 SPE 在控制限以下的 X_{1, <SPE_Limit₁} 和 SPE 在控制限以上的 X_{1, >SPE_Limit₁}。其中 X_{1, >SPE_Limit₁} 传入第二类,而 X_{1, <SPE_Limit₁} 作为新的第一类,即 X₁ = X_{1, <SPE_Limit₁};

(3.2.3) 重复步骤(3.2.1)和(3.2.2)直至符合第一类迭代终止条件;

(3.2.4) 根据步骤(3.2.1)-(3.2.3),可以获得第二类及之后各类,直至所有类都停止迭代且没有新的类分离出来,则判定完成分类;

(3.2.5) 利用各类的最终分类结果建立各类最终的 PCA 模型。

3. 根据权利要求2所述一种包含血糖分类功能故障检测模块的连续血糖监测设备,其特征在于,所述步骤(3.2.3)中,所述迭代终止条件判定包括线性判定和 SPE 离散程度判定。

所述线性判定是基于 PCA 模型,检验分类后血糖信号间是否为线性关系:

a) 提取第 i 类 X_i 的主元 T_i, 并进行 K 均值聚类将 X_i 分为 m 个区域;

b) 分别计算每个区域的相关系数矩阵,记为 R^(h), h = 1, 2, ..., m。根据每个变量的均值和方差可计算出 R^(h) 中每个元素的置信限,用矩阵表示为:

$$R^{(h)} = \begin{bmatrix} 1 & r_{12}^{(h)L} \leq r_{12}^{(h)} \leq r_{12}^{(h)U} & \cdots & r_{1N}^{(h)L} \leq r_{1N}^{(h)} \leq r_{1N}^{(h)U} \\ r_{21}^{(h)L} \leq r_{21}^{(h)} \leq r_{21}^{(h)U} & 1 & \cdots & r_{2N}^{(h)L} \leq r_{2N}^{(h)} \leq r_{2N}^{(h)U} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N1}^{(h)L} \leq r_{N1}^{(h)} \leq r_{N1}^{(h)U} & r_{N2}^{(h)L} \leq r_{N2}^{(h)} \leq r_{N2}^{(h)U} & \cdots & 1 \end{bmatrix}$$

其中 $r_{ij}^{(h)L} = \frac{\exp(2(\zeta_{ij}^{(h)} - \varepsilon)) - 1}{\exp(2(\zeta_{ij}^{(h)} - \varepsilon)) + 1}$ 和 $r_{ij}^{(h)U} = \frac{\exp(2(\zeta_{ij}^{(h)} + \varepsilon)) - 1}{\exp(2(\zeta_{ij}^{(h)} + \varepsilon)) + 1}$ 分别为第 i 变量和第 j 变量

的相关系数 $r_{ij}^{(h)}$ 的下界和上界, $\zeta_{ij}^{(h)} = \frac{1}{2} \ln \frac{1 + r_{ij}^{(h)*}}{1 - r_{ij}^{(h)*}}$ 和 $\varepsilon = \frac{C_\alpha}{\sqrt{n^{(h)} - 3}}$, C_α 是标准正态分布检验水平为 α 时的临界值, n^(h) 是第 h 区域的样本数;

c) 利用粒子群优化算法获得相关系数矩阵的第 k 个特征值 λ_k 的上下限 λ_{k,max} 和 λ_{k,min}, 满足:

$$\lambda_{k_{\max}} = \arg \max_{\Delta R_{\max}} \lambda_k (R^{(h)} + \Delta R_{\max})$$

$$\lambda_{k_{\min}} = \arg \max_{\Delta R_{\min}} \lambda_k (R^{(h)} + \Delta R_{\min})$$

并计算残差矩阵的弗罗贝尼乌斯范数边界限： $\sigma_{\max} = \sum_{k=n+1}^N \lambda_{k_{\max}}$ 以及 $\sigma_{\min} = \sum_{k=n+1}^N \lambda_{k_{\min}}$ ，其中 n 为 PCA 模型保留主元个数， ΔR_{\max} 和 ΔR_{\min} 中的元素分别是 R 中非对角元素的波动，用于决定最大值 $\lambda_{k_{\max}}$ 和最小值 $\lambda_{k_{\min}}$ ；

d) 比较各区域的残差矩阵的弗罗贝尼乌斯范数 $\sigma = \frac{1}{n^{(h)} - 1} \sum_{i=1}^{n^{(h)}} \sum_{j=1}^N e_{ij}^2$ 是否落在边界内，若

所有范数均落在边界以内，则说明数据为线性数据；

所述 SPE 离散程度判定即相关特性判定：

若类内数据的 SPE 满足：

$$\text{var}\left(\frac{SPE_i}{m_i}\right) < 0.5$$

且只有少数采样点超出 SPE 控制限，即：

$$\frac{X_{i, >SPE \text{ Limit}_i}}{X_i} \leq \alpha$$

则说明类内数据的相关特性近似，即血糖波动情况相同。其中 m_i 为第 i 类 SPE 的均值， α 为控制线的置信度。

一种包含血糖分类功能故障检测模块的连续血糖监测设备

技术领域

[0001] 本发明属于血糖数据处理和分析的研究领域,特别是涉及一种包含血糖分类功能故障检测模块的连续血糖监测设备。

背景技术

[0002] 为了管理监测血糖水平,必须要对血糖水平进行测量,目前采用的是连续血糖监测设备。随着连续血糖监测 (CGM) 设备的快速发展,使得更高精度的控制变为了可能,实时的 CGM 系统对高 / 低血糖的提前检测起到了重要的作用。通过比较当前的测量值与高 / 低血糖阈值之间的关系就可以产生报警,而及时的高低血糖报警尤其是夜间的报警对于糖尿病患者尤为重要。但是通常来说,CGM 系统会发生各种故障,从而导致提供的血糖信息不再可靠,甚至会导致错误的高低血糖血糖警报和错误的血糖治疗决策。因此,CGM 血糖仪的在线故障显得尤为重要。

[0003] 本发明基于 PCA 监测图,通过分析血糖相关特性的变化情况进而检测 CGM 血糖仪是否发生故障。在实际应用中,血糖的相关特性受饮食等外源输入影响而发生变化。此外,血糖信号间存在明显的非线性,因此无法建立起一个统一有效的 PCA 模型进行故障检测。通过 PCA 迭代,可以将血糖数据按相关特性分类,同时实现局部线性化,从而提高 PCA 监测图的性能

发明内容

[0004] 本发明的目的在于针对传统 PCA 监测图的不足,提出一种包含血糖分类功能故障检测模块的连续血糖监测设备

[0005] 本发明的目的是通过以下技术方案来实现的:一种包含血糖分类功能故障检测模块的连续血糖监测设备,该设备包括:用于采集人体血糖信息,输出血糖可用信号的传感器;用于对传感器的输出信号进行放大处理的信号放大器;用于对信号放大器输出的模拟信号进行数字转换的单片机;用于对单片机输出的数字信号进行数据处理的故障检测模块,故障检测模块可以集成在单片机中,也可以单独使用;用于将故障检测模块输出的血糖监测值 SPE 显示的显示器;用于数据存储的存储器;所述故障检测模块进行数据处理的过程包括以下步骤:

[0006] (1) 血糖数据预处理:将以一定采样周期 Δt 获得的单片机输出的连续血糖信号组合成一维时序数据 $x_{g,1 \times 1}$,其中 x_g 表示检测到的血糖信号,1 为采样个数;

[0007] (2) 对于连续血糖信号建立 PCA 模型:将连续 m 个血糖信号作为 PCA 模型多维输入,即 $x(k) = [x_g(k-m+1), x_g(k-m+2), \dots, x_g(k)]^T$

[0008] $T = XP$

[0009] $E = X - XPP^T$

[0010] 其中 $X = [x(1), x(2), \dots, x(n)]^T$ 是由 $n \times m$ 维的血糖数据组成的训练矩阵; P 为 $m \times r$ 维负载矩阵, r 为保留的主元个数, $r \leq m$; T 为 $n \times r$ 维得分矩阵,为血糖数据矩阵 X 在

主元空间的投影； $E = [e(1), e(2), \dots, e(n)]^T$ 为 $n \times m$ 维的残差矩阵。通过统计量 SPE 对血糖数据进行监测：

$$[0011] \quad SPE(i) = e(i)^T e(i)$$

[0012] SPE 的控制限可以通过以下公式近似求取：

$$[0013] \quad SPE \sim g\chi_{h,\alpha}^2$$

[0014] 其中 $g = v/2m$, $h = 2m^2/v$, m 为全部 SPE 的均值, v 为全部 SPE 的方差, α 为控制线的置信度。

[0015] (3) 通过 PCA 迭代的方法将病人的血糖信号进行分类, 并建立相应的 PCA 模型；

[0016] (4) 根据步骤 (3) 的分类结果, 建立血糖仪在线故障检测的 PCA 监测图, 具体包括以下子步骤：

[0017] (4.1) 将在线传入的血糖数据分别代入各类 PCA 模型, 计算 SPE 值：

$$[0018] \quad e_{\text{test},i}^T = x_{\text{test}}^T - x_{\text{test}}^T P_i P_i^T$$

$$[0019] \quad SPE_{\text{new},i} = e_{\text{test},i}^T e_{\text{test},i}$$

[0020] 其中 x_{test} 是当前时刻输入的 $m \times 1$ 维的血糖测试数据, P_i 是第 i 类 PCA 模型的负载矩阵, 而 $e_{\text{test},i}$ 是测试数据代入第 i 类 PCA 模型所得的残差向量, $SPE_{\text{test},i}$ 是 $e_{\text{test},i}$ 对应的 SPE 值；

[0021] (4.2) 确定检测结果：

$$[0022] \quad \text{result}_{\text{test},i} = SPE_{\text{test},i} < SPE_Limit_i$$

$$[0023] \quad \text{result}_{\text{test}} = \text{result}_{\text{test},1} \vee \text{result}_{\text{test},2} \vee \dots \vee \text{result}_{\text{test},k}$$

[0024] 其中 $\text{result}_{\text{test},i}$ 是测试数据在第 i 类 PCA 模型下检测结果, 其中 $\text{result}_{\text{test},i} = 1$ 代表检测结果正常, $\text{result}_{\text{test},i} = 0$ 代表检测结果异常； \vee 为取并集的逻辑运算； $\text{result}_{\text{test}}$ 为最终检测结果, 为所有类 PCA 模型下检测结果的并集, 当连续若干时刻的故障检测结果均为异常, 则可说明血糖仪发生故障。

[0025] 进一步地, 所述步骤 (3) 具体包括以下子步骤：

[0026] (3.1) 对于某个特定病人选取足量血糖仪正常工作时的血糖数据作为训练数据, 用于病人血糖分类；并将这些训练数据如步骤 (2) 铺展成 n 个采样点组成的 $n \times m$ 维数据矩阵 X 。假定全部训练数据属于同一类 (即第一类), 即：

$$[0027] \quad X_1 = X = [x(1), x(2), \dots, x(n)]^T;$$

[0028] (3.2) 利用 PCA 迭代的方法将血糖分为若干类, 具体步骤如下：

[0029] (3.2.1) 根据步骤 (2) 对第一类的数据矩阵建立 PCA 模型：

$$[0030] \quad E_1 = X_1 - X_1 P_1 P_1^T$$

$$[0031] \quad SPE_1(i) = e_1(i)^T e_1(i)$$

$$[0032] \quad SPE_Limit_1 = g_1 \chi_{h,\alpha}^2$$

[0033] (3.2.2) 通过 PCA 模型控制限 SPE_Limit_1 可以将数据 X_1 分为两类, 即 SPE 在控制限以下的 $X_{1, <SPE_Limit_1}$ 和 SPE 在控制限以上的 $X_{1, >SPE_Limit_1}$ 。其中 $X_{1, >SPE_Limit_1}$ 传入第二类, 而 $X_{1, <SPE_Limit_1}$ 作为新的第一类, 即 $X_1 = X_{1, <SPE_Limit_1}$ ；

[0034] (3.2.3) 重复步骤 (3.2.1) 和 (3.2.2) 直至符合第一类迭代终止条件；

[0035] (3.2.4) 根据步骤 (3.2.1)——(3.2.3), 可以获得第二类及之后各类, 直至所有类

都停止迭代且没有新的类分离出来,则判定完成分类;

[0036] (3.2.5) 利用各类的最终分类结果建立各类最终的 PCA 模型;

[0037] 进一步地,所述步骤 (3.2.3) 中,所述迭代终止条件判定包括线性判定和 SPE 离散程度判定。

[0038] 所述线性判定是基于 PCA 模型,检验分类后血糖信号间是否为线性关系:

[0039] a) 提取第 i 类 X_i 的主元 T_i , 并进行 K 均值聚类将 X_i 分为 m 个区域;

[0040] b) 分别计算每个区域的相关系数矩阵,记为 $R^{(h)}$, $h = 1, 2, \dots, m$ 。根据每个变量的均值和方差可计算出 $R^{(h)}$ 中每个元素的置信限,用矩阵表示为:

[0041]

$$R^{(h)} = \begin{bmatrix} 1 & r_{12}^{(h)L} \leq r_{12}^{(h)} \leq r_{12}^{(h)U} & \cdots & r_{1N}^{(h)L} \leq r_{1N}^{(h)} \leq r_{1N}^{(h)U} \\ r_{21}^{(h)L} \leq r_{21}^{(h)} \leq r_{21}^{(h)U} & 1 & \cdots & r_{2N}^{(h)L} \leq r_{2N}^{(h)} \leq r_{2N}^{(h)U} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N1}^{(h)L} \leq r_{N1}^{(h)} \leq r_{N1}^{(h)U} & r_{N2}^{(h)L} \leq r_{N2}^{(h)} \leq r_{N2}^{(h)U} & \cdots & 1 \end{bmatrix}$$

[0042] 其中 $r_{ij}^{(h)L} = \frac{\exp(2(\zeta_{ij}^{(h)} - \varepsilon)) - 1}{\exp(2(\zeta_{ij}^{(h)} - \varepsilon)) + 1}$ 和 $r_{ij}^{(h)U} = \frac{\exp(2(\zeta_{ij}^{(h)} + \varepsilon)) - 1}{\exp(2(\zeta_{ij}^{(h)} + \varepsilon)) + 1}$ 分别为第 i 变量和第 j 变

量的相关系数 $r_{ij}^{(h)}$ 的下界和上界, $\zeta_{ij}^{(h)} = \frac{1}{2} \ln \frac{1 + r_{ij}^{(h)*}}{1 - r_{ij}^{(h)*}}$ 和 $\varepsilon = \frac{C_\alpha}{\sqrt{n^{(h)} - 3}}$, C_α 是标准正态分布检验

水平为 α 时的临界值, $n^{(h)}$ 是第 h 区域的样本数;

[0043] c) 利用粒子群优化算法获得相关系数矩阵的第 k 个特征值 λ_k 的上下限 $\lambda_{k_{\max}}$ 和 $\lambda_{k_{\min}}$, 满足:

$$[0044] \quad \lambda_{k_{\max}} = \arg \max_{\Delta R_{\max}} \lambda_k (R^{(h)} + \Delta R_{\max})$$

$$[0045] \quad \lambda_{k_{\min}} = \arg \max_{\Delta R_{\min}} \lambda_k (R^{(h)} + \Delta R_{\min})$$

[0046] 并计算残差矩阵的弗罗贝尼乌斯范数边界限: $\sigma_{\max} = \sum_{k=n+1}^N \lambda_{k_{\max}}$ 以及

$\sigma_{\min} = \sum_{k=n+1}^N \lambda_{k_{\min}}$, 其中 n 为 PCA 模型保留主元个数, ΔR_{\max} 和 ΔR_{\min} 中的元素分别是 R 中非对

角元素的波动,用于决定最大值 $\lambda_{k_{\max}}$ 和最小值 $\lambda_{k_{\min}}$;

[0047] d) 比较各区域的残差矩阵的弗罗贝尼乌斯范数 $\sigma = \frac{1}{n^{(h)} - 1} \sum_{i=1}^{n^{(h)}} \sum_{j=1}^N e_{ij}^2$ 是否落在边界

内,若所有范数均落在边界以内,则说明数据为线性数据;

[0048] 所述 SPE 离散程度判定即相关特性判定:

[0049] 若类内数据的 SPE 满足:

$$[0050] \quad \text{var}\left(\frac{SPE_i}{m_i}\right) < 0.5$$

[0051] 且只有少数采样点超出 SPE 控制限,即:

$$[0052] \quad \frac{X_{i, >SPE_Limit}}{X_i} \leq \alpha$$

[0053] 则说明类内数据的相关特性近似,即血糖波动情况相同。其中 m_i 为第 i 类 SPE 的均值, α 为控制线的置信度。

[0054] 本发明的有益效果是:本发明所提出的含血糖分类功能故障检测模块的连续血糖监测设备能够根据病人血糖信号间不同的相关特性将血糖分为若干类,再分别进行故障检测,以便提高 PCA 监测图的检测精度,从而提升血糖信号的可信度。本发明易于实施,为血糖处理和分析的研究指明了新的方向。

附图说明

[0055] 图 1 是本发明连续血糖监测设备的结构框图;

[0056] 图 2 是本发明连续血糖监测设备中血糖分类的流程图;

[0057] 图 3 是本发明连续血糖监测设备中在线故障检测的实现流程图;

[0058] 图 4 是原始 PCA 的 SPE 监测图;

[0059] 图 5 是通过 PCA 迭代分类的 SPE 监测图,(a) 第一类 SPE 监测图 (b) 第二类 SPE 监测图 (c) 第三类 SPE 监测图;

[0060] 图 6 是 PCA 迭代分类的检测结果(正常情况),(a) 正常血糖信号 (b) 最终检测结果 (c) 第一类 SPE 监测图 (d) 第二类 SPE 监测图 (e) 第三类 SPE 监测图;

[0061] 图 7 是 PCA 迭代分类的检测结果(故障情况),(a) 故障血糖信号 (b) 最终检测结果 (c) 第一类 SPE 监测图 (d) 第二类 SPE 监测图 (e) 第三类 SPE 监测图。

具体实施方式

[0062] 下面结合附图和具体实施例对本发明作进一步详细说明。

[0063] 如图 1 所示,本发明提供了一种包含血糖分类功能故障检测模块的连续血糖监测设备,该设备包括:用于采集人体血糖信息,输出血糖可用信号的传感器;用于对传感器的输出信号进行放大处理的信号放大器;用于对信号放大器输出的模拟信号进行数字转换的单片机;用于对单片机输出的数字信号进行数据处理的故障检测模块,故障检测模块可以集成在单片机中,也可以单独使用;用于将故障检测模块输出的血糖监测值 SPE 显示的显示器;用于数据存储的存储器;所述故障检测模块进行数据处理的过程包括以下步骤:

[0064] 步骤 1:血糖数据预处理

[0065] 将以采样周期为 5 分钟所获得的连续血糖监测信号组合成一维时序数据 $x_{g_1 \times 1}$, 其中, y 表示检测到的血糖信号, 1 为样个数:本实例中,共有三组对象的采样信号,采样周期为 5 分钟,第 1 组为青少年组,第 2 组为成人组,第 3 组为儿童组,每组 2 人,三组共 6 人,每个对象的采样信号包括五天的数据。分别添加单时刻数据异常的尖峰类 (spike) 故障以及持续一段时间数据异常的丧失敏感性类 (lose sensitivity) 故障;

[0066] 步骤 2:对于连续血糖信号建立 PCA 模型:将连续 m 个血糖信号作为 PCA 模型维多输入,即 $x(k) = [x_g(k-m+1), x_g(k-m+2), \dots, x_g(k)]^T$, 令 $m = 7$

[0067] $T = XP$

[0068] $E = X - XPP^T$

[0069] 其中 $X = [x(1), x(2), \dots, x(n)]^T$ 是由 $n \times m$ 维的血糖数据组成的训练矩阵; P 为 $m \times r$ 维负载矩阵, r 为保留的主元个数, $r \leq m$; T 为 $n \times r$ 维得分矩阵, 为血糖数据矩阵 X 在主元空间的投影; $E = [e(1), e(2), \dots, e(n)]^T$ 为 $n \times m$ 维的残差矩阵。通过统计量 SPE 对血糖数据进行监测:

$$[0070] \quad \text{SPE}(i) = e(i)^T e(i)$$

[0071] SPE 的控制限可以通过以下公式近似求取:

$$[0072] \quad \text{SPE} \sim g \chi_{h, \alpha}^2$$

[0073] 其中 $g = v/2m$, $h = 2m^2/v$, m 为全部 SPE 的均值, v 为全部 SPE 的方差, α 为控制线的置信度。

[0074] 步骤 3: 通过 PCA 迭代的方法将病人的血糖信号进行分类, 并建立相应的 PCA 模型;

[0075] (3.1) 对于某个特定病人选取一天血糖仪正常工作时的血糖数据作为训练数据, ($l = 288$) 用于病人血糖分类; 并将这些训练数据如步骤 (2) 铺展成 n 个采样点组成的 $n \times m$ 维数据矩阵 X 。假定全部训练数据属于同一类 (即第一类), 即:

$$[0076] \quad X_1 = X = [x(1), x(2), \dots, x(n)]^T;$$

[0077] (3.2) 利用 PCA 迭代的方法将血糖分为若干类, 具体步骤如下:

[0078] (3.2.1) 根据步骤 (2) 对第一类的数据矩阵建立 PCA 模型:

$$[0079] \quad E_1 = X_1 - X_1 P_1 P_1^T$$

$$[0080] \quad \text{SPE}_1(i) = e_1(i)^T e_1(i)$$

$$[0081] \quad \text{SPE_Limit}_1 = g_1 \chi_{h, \alpha}^2$$

[0082] (3.2.2) 通过 PCA 模型控制限 SPE_Limit_1 可以将数据 X_1 分为两类, 即 SPE 在控制限以下的 $X_{1, < \text{SPE_Limit}_1}$ 和 SPE 在控制限以上的 $X_{1, > \text{SPE_Limit}_1}$ 。其中 $X_{1, > \text{SPE_Limit}_1}$ 传入第二类, 而 $X_{1, < \text{SPE_Limit}_1}$ 作为新的第一类, 即 $X_1 = X_{1, < \text{SPE_Limit}_1}$;

[0083] (3.2.3) 重复步骤 (3.2.1) 和 (3.2.2) 直至符合第一类迭代终止条件 (详见步骤 3.3);

[0084] (3.2.4) 根据步骤 (3.2.1)——(3.2.3), 可以获得第二类及之后各类, 直至所有类都停止迭代且没有新的类分离出来, 则判定完成分类;

[0085] (3.2.5) 利用各类的最终分类结果建立各类最终的 PCA 模型;

[0086] (3.3) PCA 迭代终止条件判定包括线性判定和 SPE 离散程度判定。

[0087] (3.3.1) 线性判定是基于 PCA 模型, 检验分类后血糖信号间线性是否提升:

[0088] a) 提取第 i 类 X_i 的主元 T_i , 并进行 K 均值聚类将 X_i 分为 m 个区域;

[0089] b) 分别计算每个区域的相关系数矩阵, 记为 $R^{(h)}$, $h = 1, 2, \dots, m$ 。根据每个变量的均值和方差可计算出 $R^{(h)}$ 中每个元素的置信限, 用矩阵表示为:

[0090]

$$R^{(h)} = \begin{bmatrix} 1 & r_{12}^{(h) L} \leq r_{12}^{(h)} \leq r_{12}^{(h) U} & \cdots & r_{1N}^{(h) L} \leq r_{1N}^{(h)} \leq r_{1N}^{(h) U} \\ r_{21}^{(h) L} \leq r_{21}^{(h)} \leq r_{21}^{(h) U} & 1 & \cdots & r_{2N}^{(h) L} \leq r_{2N}^{(h)} \leq r_{2N}^{(h) U} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N1}^{(h) L} \leq r_{N1}^{(h)} \leq r_{N1}^{(h) U} & r_{N2}^{(h) L} \leq r_{N2}^{(h)} \leq r_{N2}^{(h) U} & \cdots & 1 \end{bmatrix}$$

[0091] 其中 $r_{ij}^{(h) L} = \frac{\exp(2(\zeta_{ij}^{(h)} - \varepsilon)) - 1}{\exp(2(\zeta_{ij}^{(h)} - \varepsilon)) + 1}$ 和 $r_{ij}^{(h) U} = \frac{\exp(2(\zeta_{ij}^{(h)} + \varepsilon)) - 1}{\exp(2(\zeta_{ij}^{(h)} + \varepsilon)) + 1}$ 分别为第 i 变量和第 j 变

量的相关系数 $r_{ij}^{(h)}$ 的下界和上界, $\zeta_{ij}^{(h)} = \frac{1}{2} \ln \frac{1 + r_{ij}^{(h)*}}{1 - r_{ij}^{(h)*}}$ 和 $\varepsilon = \frac{C_\alpha}{\sqrt{n^{(h)} - 3}}$, C_α 是标准正态分布检验

水平为 α 时的临界值, $n^{(h)}$ 是第 h 区域的样本数;

[0092] c) 利用粒子群优化算法获得相关系数矩阵的第 k 个特征值 λ_k 的上下限 $\hat{\lambda}_{k_{\max}}$ 和 $\hat{\lambda}_{k_{\min}}$, 满足:

$$[0093] \quad \hat{\lambda}_{k_{\max}} = \arg \max_{\Delta R_{\max}} \lambda_k(R^{(h)} + \Delta R_{\max})$$

$$[0094] \quad \hat{\lambda}_{k_{\min}} = \arg \max_{\Delta R_{\min}} \lambda_k(R^{(h)} + \Delta R_{\min})$$

[0095] 并计算残差矩阵的弗罗贝尼乌斯范数边界限: $\sigma_{\max} = \sum_{k=n+1}^N \hat{\lambda}_{k_{\max}}$ 以及 $\sigma_{\min} = \sum_{k=n+1}^N \hat{\lambda}_{k_{\min}}$,

其中 n 为 PCA 模型保留主元个数, ΔR_{\max} 和 ΔR_{\min} 中的元素分别是 R 中非对角元素的波动, 用于决定最大值 $\hat{\lambda}_{k_{\max}}$ 和最小值 $\hat{\lambda}_{k_{\min}}$;

[0096] d) 比较各区域的残差矩阵的弗罗贝尼乌斯范数 $\sigma = \frac{1}{n^{(h)} - 1} \sum_{i=1}^{n^{(h)}} \sum_{j=1}^N e_{ij}^2$ 是否落在边界

内, 若所有范数均落在边界以内, 则说明数据为线性数据;

[0097] 所述 SPE 离散程度判定即相关特性判定:

[0098] 若类内数据的 SPE 满足:

$$[0099] \quad \text{var}\left(\frac{SPE_i}{m_i}\right) < 0.5$$

[0100] 且只有少数采样点超出 SPE 控制限, 即:

$$[0101] \quad \frac{X_{i, >SPE_Limit_i}}{X_i} \leq \alpha$$

[0102] 则说明类内数据的相关特性近似, 即血糖波动情况相同。其中 m_i 为第 i 类 SPE 的均值, α 为控制线的置信度。

[0103] (3.3.2) SPE 离散程度判定即相关特性判定: 若类内数据的 SPE 满足:

$$[0104] \quad \text{var}\left(\frac{SPE_i}{m_i}\right) < 0.5$$

[0105] 且只有少数采样点超出 SPE 控制限, 即:

[0106]
$$\frac{X_{i, >SPE_Limit_i}}{X_i} \leq \alpha$$

[0107] 则说明类内数据的相关特性近似,即血糖波动情况相同。其中 m_i 为第 i 类 SPE 的均值。

[0108] (4) 根据步骤 (3) 的分类结果,建立血糖仪在线故障检测的 PCA 监测图,具体包括以下子步骤:

[0109] (4.1) 将在线传入的血糖数据分别代入各类 PCA 模型,计算 SPE 值:

[0110]
$$e_{test,i}^T = x_{test}^T - x_{test}^T P_i P_i^T$$

[0111]
$$SPE_{new,i} = e_{test,i}^T e_{test,i}$$

[0112] 其中 x_{test} 是当前时刻输入的 $m \times 1$ 维的血糖测试数据, P_i 是第 i 类 PCA 模型的负载矩阵,而 $e_{test,i}$ 是测试数据代入第 i 类 PCA 模型所得的残差向量, $SPE_{test,i}$ 是 $e_{test,i}$ 对应的 SPE 值;

[0113] (4.2) 确定测试数据的检测结果:

[0114]
$$result_{test,i} = SPE_{test,i} < SPE_Limit_i$$

[0115]
$$result_{test} = result_{test,1} \vee result_{test,2} \vee \dots \vee result_{test,k}$$

[0116] 其中 $result_{test,i}$ 是测试数据在第 i 类 PCA 模型下检测结果,其中 $result_{test,i} = 1$ 代表检测结果正常, $result_{test,i} = 0$ 代表检测结果异常; \vee 为取并集的逻辑运算; $result_{test}$ 为最终检测结果,为所有类 PCA 模型下检测结果的并集,当连续若干时刻的故障检测结果均为异常,则可说明血糖仪发生故障。

[0117] (4.3) 采用以下两个指标作为检测性能的评价指标:

[0118] (4.3.1) 误报率

[0119]

$$误报率 = \frac{误报样本点}{全部正常样本点} \times 100\%$$

[0120] 误报样本点是本为正常样本点但发生故障警报;误报率越低则说明检测性能越好;

[0121] (4.3.2) 漏报率

[0122]

$$漏报率 = \frac{漏报故障样本点}{全部故障样本点} \times 100\%$$

[0123] 漏报样本点是本为故障样本点但没发生故障警报;漏报率越低则说明检测性能越好;

[0124] 从表 1 表 2 中可以看出,基于 PCA 迭代分类的 PCA 监测图要明显好于原始 PCA 监测图,误报率、漏报率都有明显的下降。

[0125] 表 1 针对 3 组(青少年组、成人组和儿童组)共 6 个对象的血糖采样数据分别采用分类 PCA 和原始 PCA 的故障检测的误报率结果对比(结果用均值 \pm 标准差表示)

[0126]

对象 #	分类 PCA	原始 PCA

1	1.51	5.41
2	1.42	3.37
3	5.14	5.05
4	1.33	5.14
5	0.27	4.61
6	4.52	6.56
均值 ± 标准差	2.36 ± 1.8	5.02 ± 0.95

[0127] 表 2 针对 3 组（青少年组、成人组和儿童组）共 6 个对象的血糖采样数据分别采用分类 PCA 和原始 PCA 的故障检测的漏报率结果对比（结果用均值 ± 标准差表示）

[0128]

故障大小	SPIKE		LOSE SENSITIVITY	
	分类 PCA	原始 PCA	分类 PCA	原始 PCA
12mg/dl	14.01 ± 19.06	26.75 ± 21.62	9.66 ± 7.19	9.1 ± 8.86
15mg/dl	6.72 ± 15.03	13.45 ± 18.72	3.08 ± 3.24	3.64 ± 6.7
18mg/dl	1.4 ± 3.13	7.56 ± 16.91	0.84 ± 1.53	1.12 ± 2.51
20mg/dl	0 ± 0	4.06 ± 9.08	0.14 ± 0.31	0.14 ± 0.31

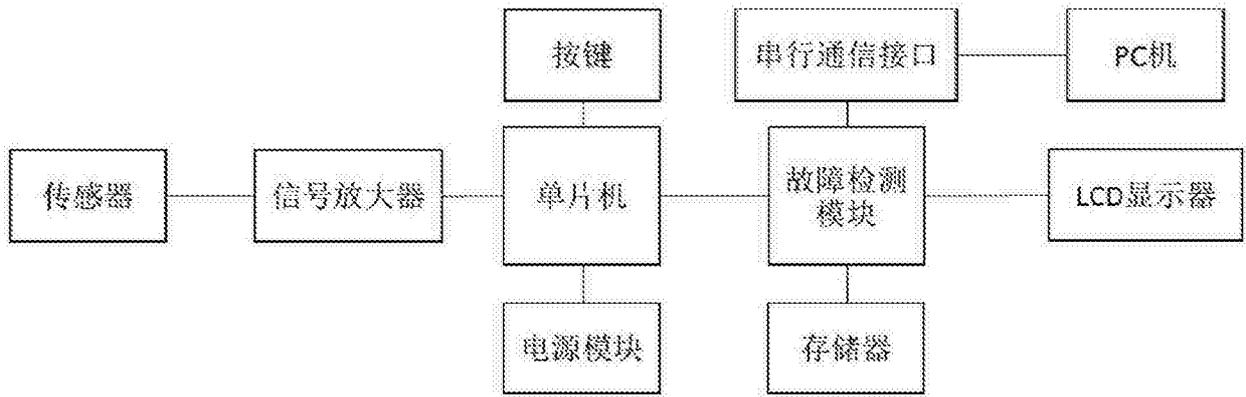


图 1

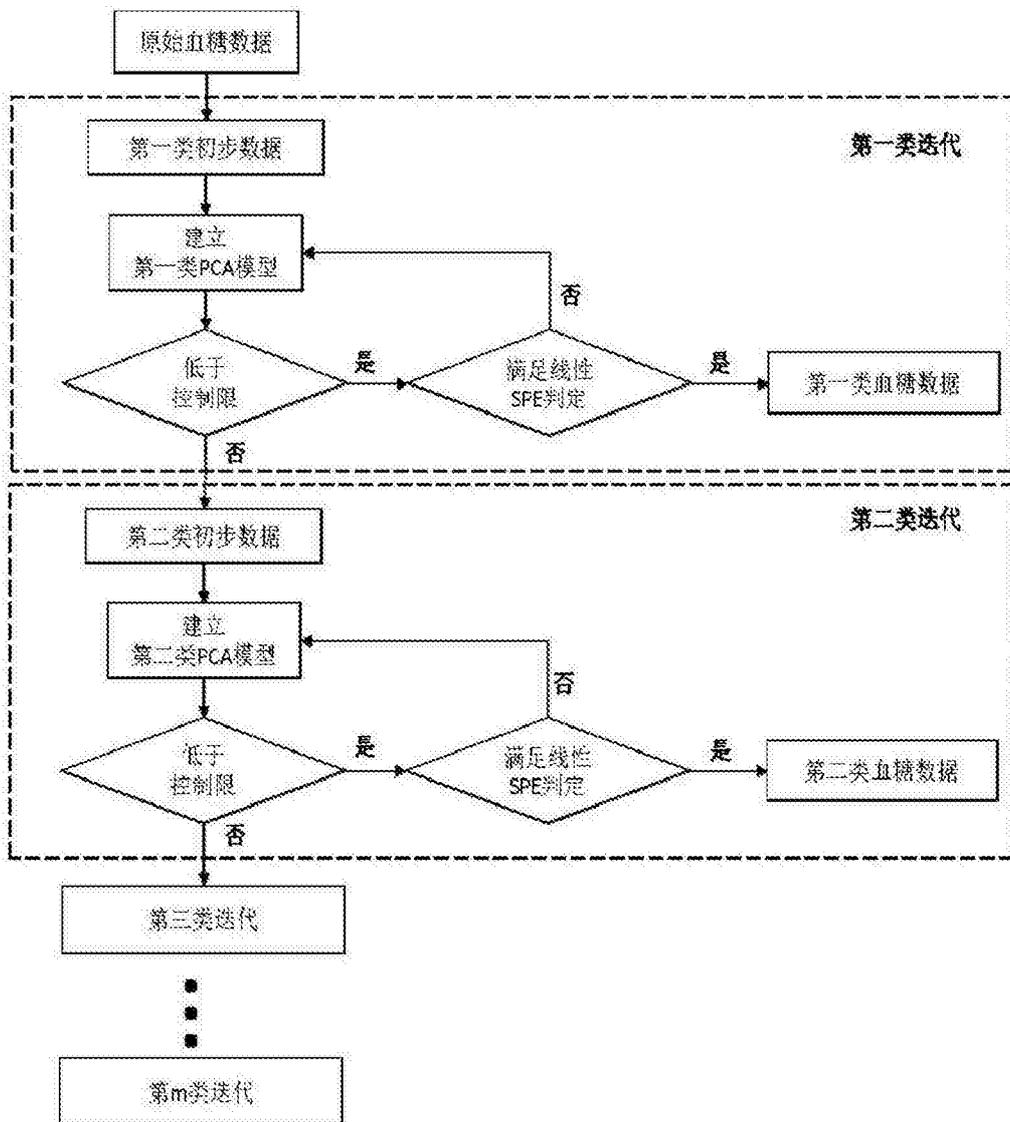


图 2

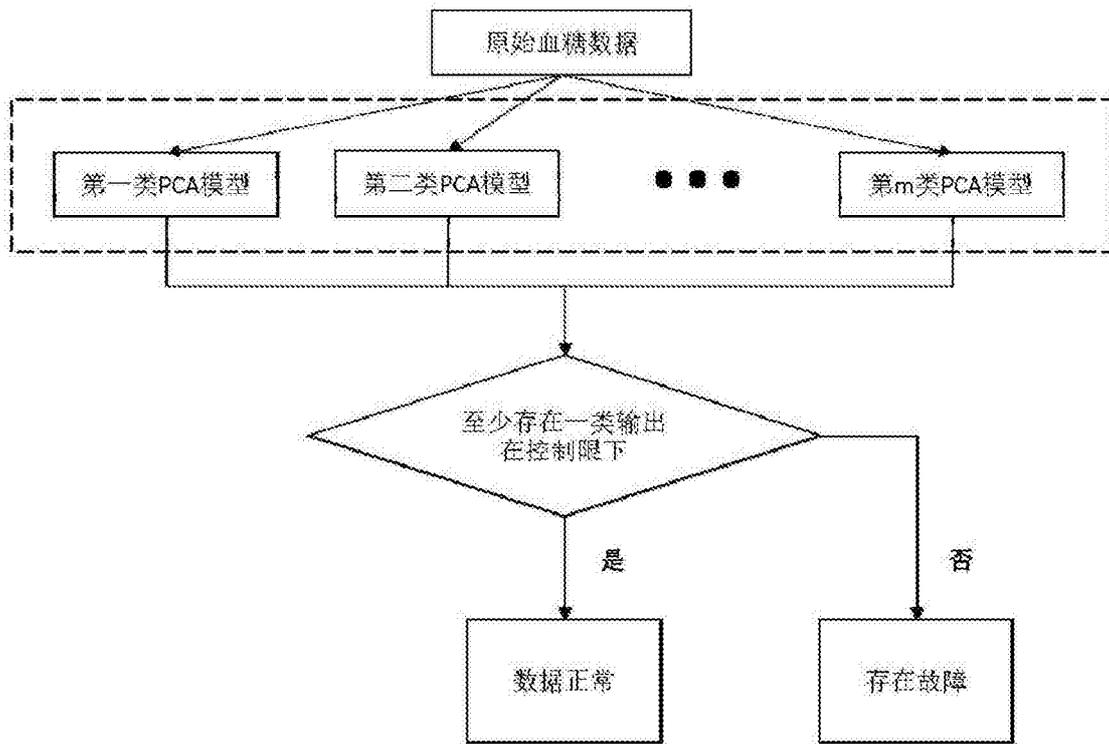


图 3

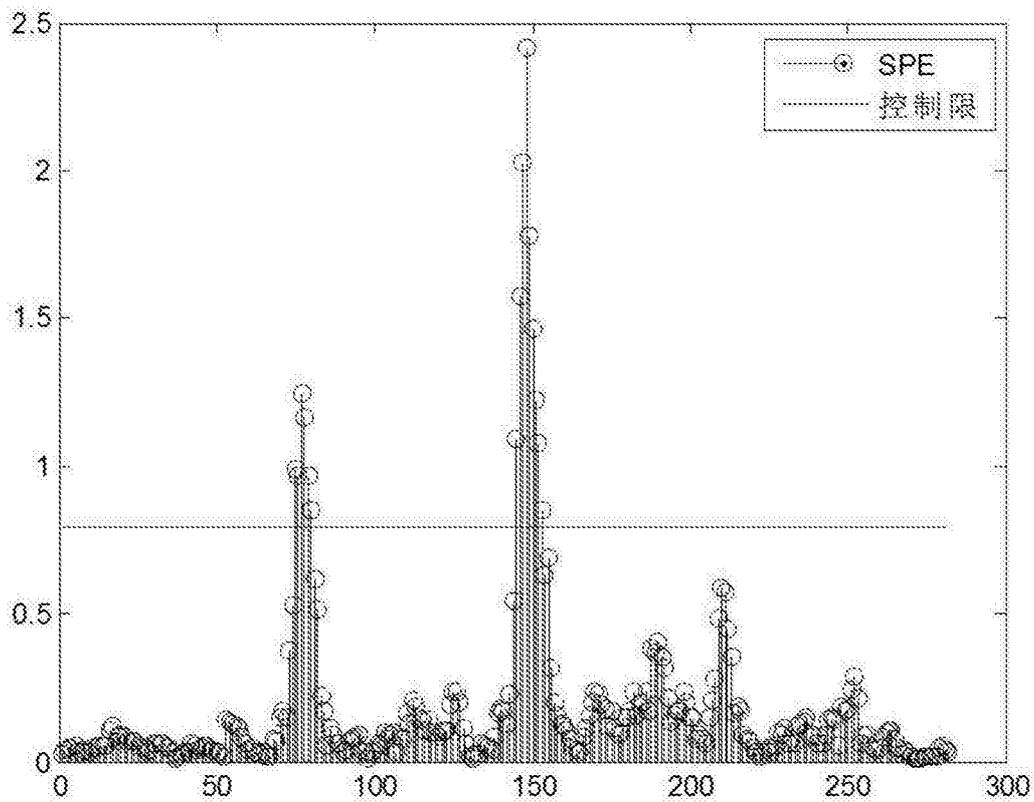
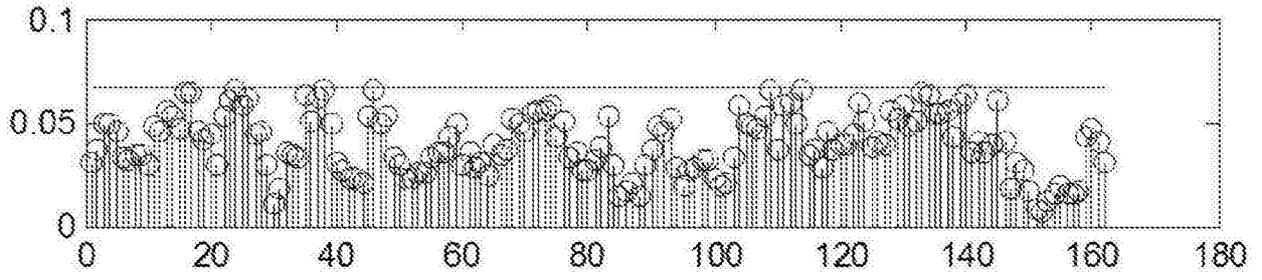
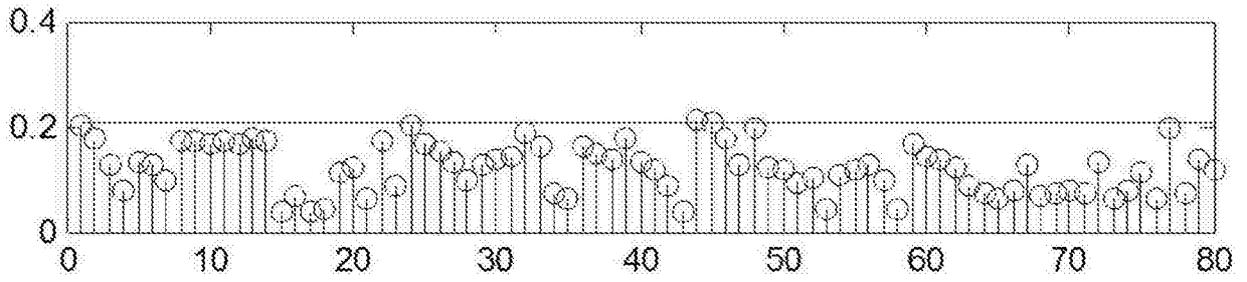


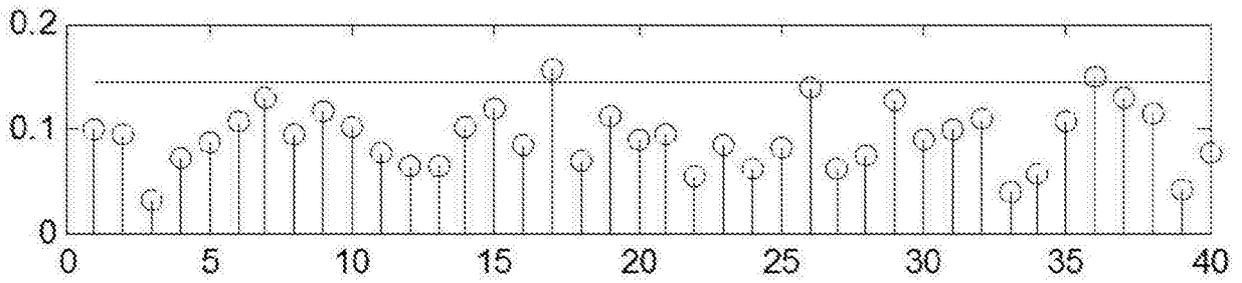
图 4



(a)

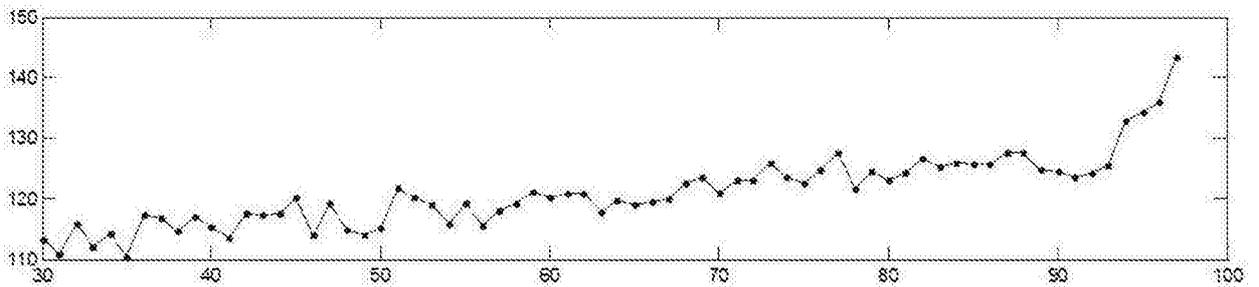


(b)

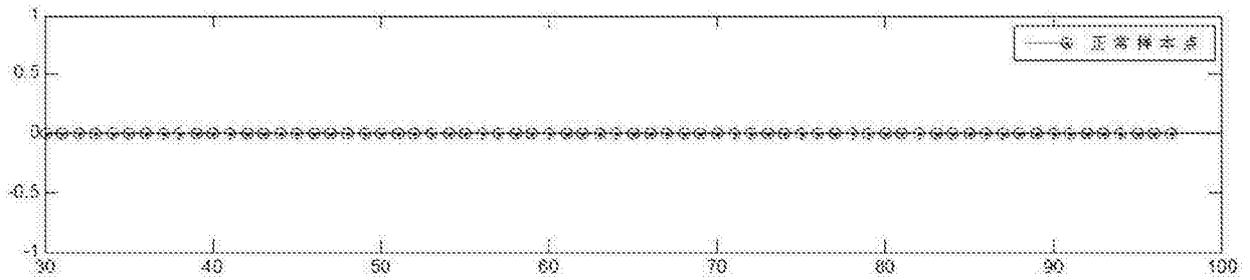


(c)

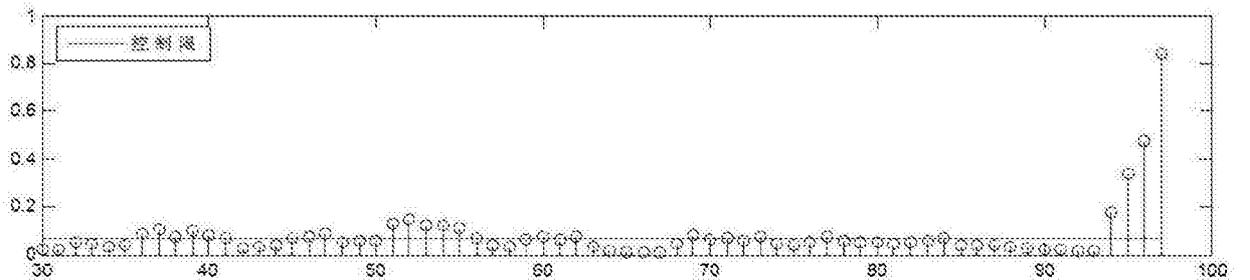
图 5



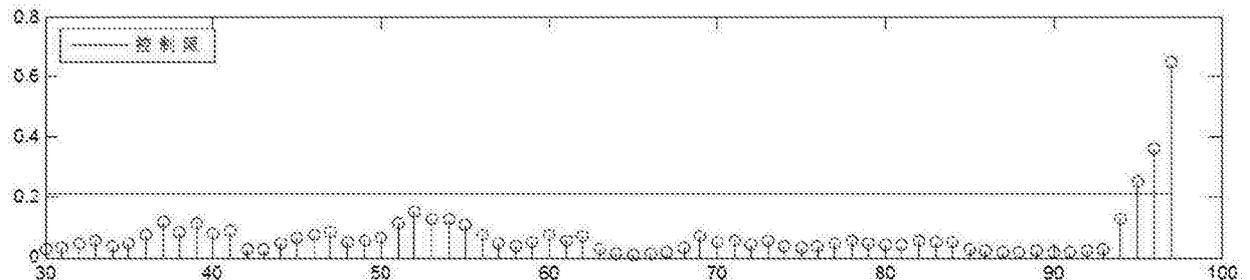
(a)



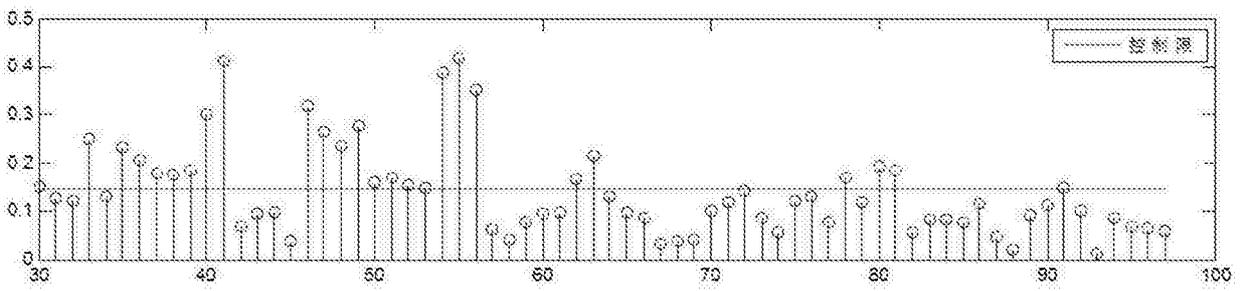
(b)



(c)

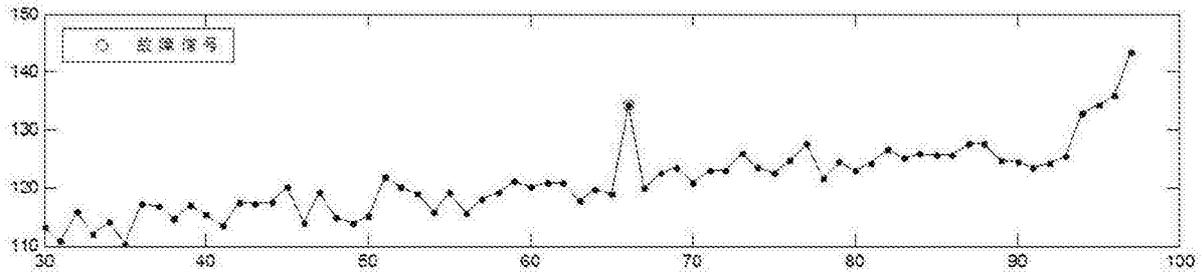


(d)

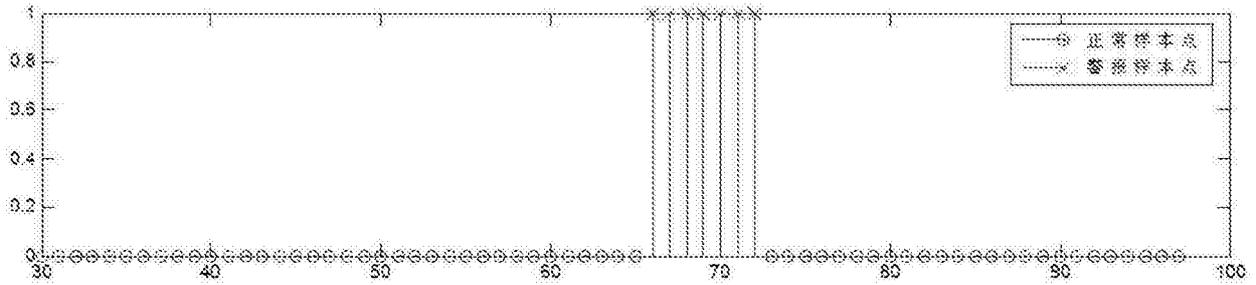


(e)

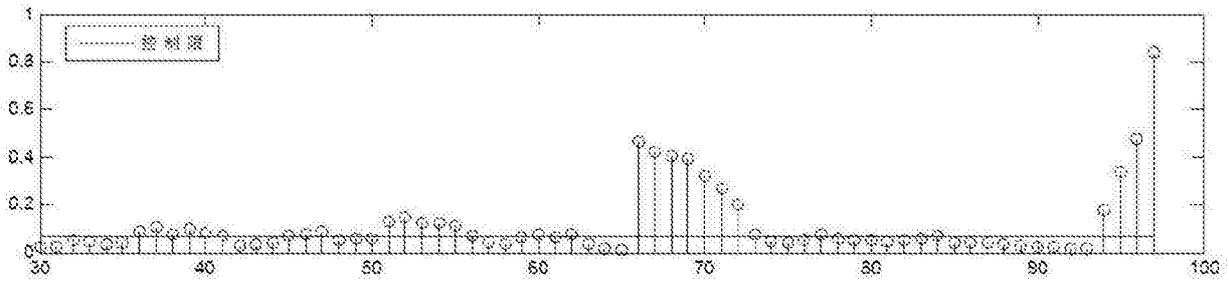
图 6



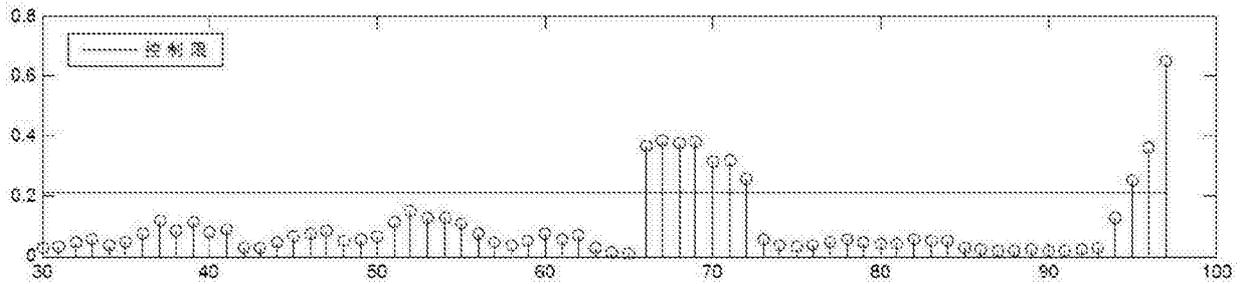
(a)



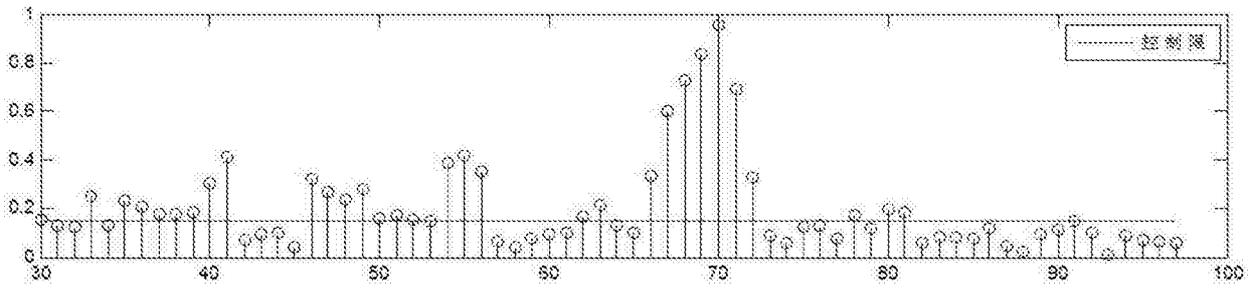
(b)



(c)



(d)



(e)

图 7