



(12) 发明专利申请

(10) 申请公布号 CN 103488884 A

(43) 申请公布日 2014. 01. 01

(21) 申请号 201310415076. 6

(22) 申请日 2013. 09. 12

(71) 申请人 北京航空航天大学

地址 100191 北京市海淀区学院路 37 号

(72) 发明人 党香俊 范晔 孙富强 姜同敏

(74) 专利代理机构 北京永创新实专利事务所

11121

代理人 官汉增

(51) Int. Cl.

G06F 19/00(2011. 01)

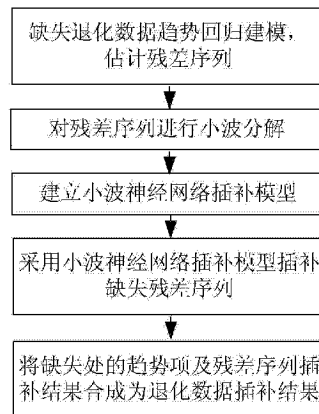
权利要求书2页 说明书4页 附图4页

(54) 发明名称

基于小波神经网络的退化数据缺失插补方法

(57) 摘要

本发明公开了一种基于小波神经网络的退化数据缺失插补方法,包括以下几个步骤:步骤一、缺失退化数据趋势回归建模,估计残差序列;步骤二、对残差序列进行小波分解;步骤三、分解子序列预处理;步骤四、采用小波神经网络插补模型插补缺失残差序列;步骤五、将缺失处的趋势项及残差序列插补结果合成为退化数据插补结果。本发明对退化数据的趋势进行回归建模,保证了缺失数据趋势与整体数据的统一。本发明对残差序列采用小波神经网络模型进行插补,避免了退化数据细节的丢失。本发明在插补过程中不存在对数据统计特性的假设,使得本方法对退化数据具有广泛的适用性。



1. 一种基于小波神经网络的退化数据缺失插补方法,其特征在于,包括以下几个步骤:

步骤一、缺失退化数据趋势回归建模,估计残差序列;

根据退化数据中已有数据的趋势特点,选择多项式函数、指数函数或幂律函数作为回归函数,得到退化趋势的函数表达式,并计算已有数据去除趋势项之后的残差序列;令  $E = (e_1, e_2, \dots, e_n)$  和  $E' = (e'_{1}, e'_{2}, \dots, e'_{m})$  分别表示缺失数据之前和之后的残差序列,而  $n$  和  $m$  表示残差序列的长度;

步骤二、对残差序列进行小波分解;

步骤三、小波分解子序列预处理;

步骤四、采用小波神经网络插补模型插补缺失残差序列;

以步骤三中预处理得到的各子序列训练输入矩阵和训练输出向量为基础,对每一级子序列构建一个  $q$  输入单输出小波神经网络,训练得到网络权值参数;将缺失节点前  $q$  个值作为网络插补的输入,若以  $G_L$  和  $G'_L$  为例,则其对应的  $L$  级高频插补的输入向量为  $\text{in}G_L G'_L = [g_{L, n-q+1} \ g_{L, n-q+2} \ \dots \ g_{L, n}]$ ,待插补序列  $G''_L$  的第一个节点为  $\hat{g}_{L, n+1}''$ ;在插补  $r$  个缺失节点的过程中,不断更新输入向量,将已经插补得到的节点值补充到输入向量中,舍弃与待插补点距离超过长度  $q$  的节点值;

最终得到各级插补子序列;

$$\begin{cases} \mathbf{H}_L'' = (\hat{h}_{L,1}'', \hat{h}_{L,2}'', \dots, \hat{h}_{L,r}'') \\ \mathbf{G}_l'' = (\hat{g}_{l,1}'', \hat{g}_{l,2}'', \dots, \hat{g}_{l,r}''), \quad l=1,2,\dots,L \end{cases} \quad (4)$$

步骤五、将缺失处的趋势项及残差序列插补结果合成为退化数据插补结果;

根据步骤一中得到的回归函数,计算缺失数据段的趋势序列  $\hat{\mathbf{X}} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_r)$ ;将  $\hat{\mathbf{X}}$  与由公式 (4) 中子序列重构得到的残差插补序列,合成最终插补序列  $\hat{\mathbf{Y}} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_r)$ ,完成退化缺失数据插补。

2. 根据权利要求 1 所述的一种基于小波神经网络的退化数据缺失插补方法,其特征在于:步骤二中选择同一种小波函数,分别对残差序列  $E$  和  $E'$  中数据进行  $L$  级小波分解;

分解后得到残差序列  $E$  的子序列为:

$$\begin{cases} \mathbf{H}_L = (h_{L,1}, h_{L,2}, \dots, h_{L,n}) \\ \mathbf{G}_l = (g_{l,1}, g_{l,2}, \dots, g_{l,n}), \quad l=1,2,\dots,L \end{cases} \quad (1)$$

采用同样的方法,得到残差序列  $E'$  的小波分解子序列:

$$\begin{cases} \mathbf{H}'_L = (h'_{L,1}, h'_{L,2}, \dots, h'_{L,m}) \\ \mathbf{G}'_l = (g'_{l,1}, g'_{l,2}, \dots, g'_{l,m}), \quad l=1,2,\dots,L \end{cases} \quad (2)$$

3. 根据权利要求 1 所述的一种基于小波神经网络的退化数据缺失插补方法,其特征在于:步骤三具体为:小波分解子序列只具有一个维度,若以其作为小波神经网络的输入,则输入节点数为 1,只能够利用待插值节点前 1 位的数据;为了充分利用待插值节点前更多节点的信息,需要对子序列做插值前的预处理,调整数据维度;以  $E$  和  $E'$  的高频分解子序列  $G_L$  和  $G'_L$  为例,构建  $L$  级高频的插值训练输入矩阵  $\text{trin}G_L G'_L$  和训练输出向量

$\text{troutG}_L G'_L$ , 如公式 (3) 所示:

$$\text{trinG}_L G'_L = \begin{bmatrix} g_{L,1} & g_{L,2} & \cdots & g_{L,q} \\ g_{L,2} & g_{L,3} & \cdots & g_{L,q+1} \\ \vdots & \vdots & \ddots & \vdots \\ g_{L,n-q} & g_{L,n-q+1} & \cdots & g_{L,n-1} \\ g'_{L,1} & g'_{L,2} & \vdots & g'_{L,q} \\ g'_{L,2} & g'_{L,3} & \vdots & g'_{L,q+1} \\ \vdots & \vdots & \ddots & \vdots \\ g'_{L,m-q} & g'_{L,m-q+1} & \cdots & g'_{L,m-1} \end{bmatrix}, \quad \text{troutG}_L G'_L = \begin{bmatrix} g_{L,q+1} \\ g_{L,q+2} \\ \vdots \\ g_{L,n} \\ g'_{L,q+1} \\ g'_{L,q+2} \\ \vdots \\ g'_{L,m} \end{bmatrix} \quad (3)$$

其中  $q$  为小波神经网络输入节点数;

$\text{trinG}_L G'_L$  中的每一行作为小波神经网络的一组输入向量, 而对应训练输出值则为  $\text{troutG}_L G'_L$  中对应行的元素; 其他高频及低子序列的训练输入矩阵和训练输出向量的构建参照公式 (3), 替换对应位置的元素即可。

## 基于小波神经网络的退化数据缺失插补方法

### 技术领域

[0001] 本发明涉及一种基于小波神经网络的退化数据缺失插补方法,属于加速退化试验技术领域。

### 背景技术

[0002] 由于在传输过程中发生数据丢失,或者在人工记录时遗漏部分数据,以及监测设备中途发生故障等原因,性能退化数据经常会出现有缺失的情况。数据缺失给性能退化数据的处理带来了很大的困难:常用的性能退化数据处理方法都是针对完整数据研究出来的,而对存在部分缺失的数据进行统计分析,则存在适用性上的问题,使得处理结果的准确性难以满足工程需求。

[0003] 面对这种问题,存在两种可能的解决思路:其中一种是改进现有的退化数据处理方法,使之能够适应有缺失的性能退化数据,另一种则是针对有缺失的数据进行处理,使之转化为符合下一阶段处理要求的完整数据。对于前一种方案,在实现上有难度,针对大量现有成熟处理方法的改进是一件庞大的工程;后一种方案的在实施上则更切合实际,并且将有缺失的数据转化为完整数据后,可以应用到更多的现有数据处理方法中,对下一阶段数据处理的通用性更强。

[0004] 对缺失数据进行插补是应用最广泛的缺失数据处理思路。目前在插补领域,已经基于统计理论,研究出了均匀插补、回归插补和经验最大化插补等统计插补方法。虽然可以将统计学中对缺失数据的处理方法借鉴到性能退化数据的缺失处理中,但性能退化数据的一些特点则可能对这些统计方法的应用带来挑战。通常情况下,退化数据具有一定的趋势性,而要求数据具有平稳性的插补方法则不再适用。另外,现有统计插补方法都是以假设待插补数据或残差服从某种统计分布为前提。然而从设备中直接采集到的退化数据则带有明显的系统特征,任意时刻前后的数据都具有难以忽略的相关性。若忽略这种相关性的存在,对退化数据进行统计分布假设,则存在明显的偏差。因此,对有缺失的性能退化数据的处理,所采用的插补方法是否正确有效,是影响最终结论可信性的关键因素之一。

### 发明内容

[0005] 本发明的目的是为了解决有缺失退化数据的插补问题,提出一种基于小波神经网络的退化数据缺失插补方法,采用的技术方案如下所述:

[0006] (1) 缺失退化数据趋势回归建模,估计残差序列;

[0007] (2) 对残差序列进行小波分解,得到小波分解子序列;

[0008] (3) 对小波分解子序列进行预处理;

[0009] (4) 采用小波神经网络插补模型插补缺失残差序列;

[0010] (5) 将缺失处的趋势项及残差序列插补结果合成为退化数据插补结果。

[0011] 本发明的优点在于:

[0012] (1) 对退化数据的趋势进行回归建模,保证了缺失数据趋势与整体数据的统一;

- [0013] (2) 对残差序列采用小波神经网络模型进行插补,避免了退化数据细节的丢失;  
 [0014] (3) 在插补过程中不存在对数据统计特性的假设,使得本方法对退化数据具有广泛的适用性。

### 附图说明

- [0015] 图 1 是本发明的方法流程图;  
 [0016] 图 2 是本发明残差序列 E 小波分解子序列(虚线框内);  
 [0017] 图 3 是本发明小波神经网络插补模型插补缺失残差序列;  
 [0018] 图 4 是本发明实施例含缺失数据的退化曲线;  
 [0019] 图 5 是本发明实施例残差序列;  
 [0020] 图 6 是本发明实施例前段残差序列 E 的小波分解;  
 [0021] 图 7 是本发明实施例各级插补子序列结果;  
 [0022] 图 8 是本发明实施例第 401 ~ 700min 缺失段的最终插补结果。

### 具体实施方式

[0023] 下面将结合附图和实施例对本发明作进一步的详细说明。

[0024] 本发明是一种基于小波神经网络的退化数据缺失插补方法,如图 1 所示流程,具体实施方法如下:

[0025] 步骤一、缺失退化数据趋势回归建模,估计残差序列;

[0026] 根据退化数据中已有数据的趋势特点,选择多项式函数、指数函数或幂律函数等常见函数作为回归函数,得到退化趋势的函数表达式,并计算已有数据去除趋势项之后的残差序列。令  $E = (e_1, e_2, \dots, e_n)$  和  $E' = (e'_{1}, e'_{2}, \dots, e'_{m})$  分别表示缺失数据之前和之后的残差序列,而 n 和 m 表示序列的长度。

[0027] 步骤二、对残差序列进行小波分解;

[0028] 选择同一种小波函数,分别对残差序列 E 和 E' 中数据进行 L 级小波分解。

[0029] 以 E 为例的小波分解示意图如附图 2。分解后得到的 L 阶低频和高频子序列为:

$$[0030] \begin{cases} \mathbf{H}_L = (h_{L,1}, h_{L,2}, \dots, h_{L,n}) \\ \mathbf{G}_l = (g_{l,1}, g_{l,2}, \dots, g_{l,n}), \quad l=1, 2, \dots, L \end{cases} \quad (1)$$

[0031] 采用同样的方法,得到 E' 的小波分解子序列:

$$[0032] \begin{cases} \mathbf{H}'_L = (h'_{L,1}, h'_{L,2}, \dots, h'_{L,m}) \\ \mathbf{G}'_l = (g'_{l,1}, g'_{l,2}, \dots, g'_{l,m}), \quad l=1, 2, \dots, L \end{cases} \quad (2)$$

[0033] 步骤三、小波分解子序列预处理;

[0034] 序列分解后得到的子序列只具有一个维度,若以其作为小波神经网络的输入,则输入节点数为 1,只能够利用待插值节点前 1 位的数据。为了充分利用待插值节点前更多节点的信息,需要对子序列做插值前的预处理,调整数据维度。以 E 和 E' 的高频分解子序列  $G_L$  和  $G'_L$  为例,构建 L 级高频的插值训练输入矩阵  $\text{trin}G_L G'_L$  和训练输出向量  $\text{trout}G_L G'_L$ ,如公式 (3) 所示。

[0035]

$$\text{trin}G_L G'_L = \begin{bmatrix} g_{L,1} & g_{L,2} & \cdots & g_{L,q} \\ g_{L,2} & g_{L,3} & \cdots & g_{L,q+1} \\ \vdots & \vdots & \ddots & \vdots \\ g_{L,n-q} & g_{L,n-q+1} & \cdots & g_{L,n-1} \\ g'_{L,1} & g'_{L,2} & \vdots & g'_{L,q} \\ g'_{L,2} & g'_{L,3} & \vdots & g'_{L,q+1} \\ \vdots & \vdots & \ddots & \vdots \\ g'_{L,m-q} & g'_{L,m-q+1} & \cdots & g'_{L,m-1} \end{bmatrix}, \quad \text{trout}G_L G'_L = \begin{bmatrix} g_{L,q+1} \\ g_{L,q+2} \\ \vdots \\ g_{L,n} \\ g'_{L,q+1} \\ g'_{L,q+2} \\ \vdots \\ g'_{L,m} \end{bmatrix} \quad (3)$$

[0036] 其中  $q$  为小波神经网络输入节点数。

[0037]  $\text{trin}G_L G'_L$  中的每一行作为小波神经网络的一组输入向量, 而对应训练输出值则为  $\text{trout}G_L G'_L$  中对应的元素。其他高频及低频子序列的训练输入矩阵和训练输出向量的构建可参照公式 (3), 简单替换对应位置的元素即可。

[0038] 步骤四、采用小波神经网络插补模型插补缺失残差序列;

[0039] 以步骤三中预处理得到的各子序列训练输入矩阵和训练输出向量为基础, 对每一级子序列构建一个  $q$  输入单输出小波神经网络, 训练得到网络权值参数。将缺失节点前  $q$  个值作为网络插补的输入, 若以  $G_L$  和  $G'_L$  为例, 则其对应的  $L$  级高频插补的输入向量为  $\text{in}G_L G'_L = [g_{L,n-q+1} \ g_{L,n-q+2} \ \cdots \ g_{L,n}]$ , 待插补序列  $G''_L$  的第一个节点为  $\hat{g}''_{L,n+1}$ 。在插补  $r$  个缺失节点的过程中, 不断更新输入向量, 将已经插补得到的节点值补充到输入向量中, 舍弃与待插补点距离超过长度  $q$  的节点值。此过程如图 3 所示。

[0040] 按照同样方法, 最终得到各级插补子序列:

$$\begin{cases} \mathbf{H}''_L = (\hat{h}''_{L,1}, \hat{h}''_{L,2}, \cdots, \hat{h}''_{L,r}) \\ \mathbf{G}''_l = (\hat{g}''_{l,1}, \hat{g}''_{l,2}, \cdots, \hat{g}''_{l,r}), \quad l=1, 2, \cdots, L \end{cases} \quad (4)$$

[0042] 步骤五、将缺失处的趋势项及残差序列插补结果合成为退化数据插补结果;

[0043] 根据步骤一中得到的回归函数, 计算缺失数据段的趋势序列  $\hat{\mathbf{X}} = (\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_r)$ 。将  $\hat{\mathbf{X}}$  与由公式 (4) 中子序列重构得到的残差插补序列, 合成最终插补序列  $\hat{\mathbf{Y}} = (\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_r)$ , 完成退化缺失数据插补。

[0044] 实施例 1: 以某型号 SLD 退化数据为基础, 仿真得到含有缺失数据的性能退化曲线, 如图 4 所示。在不影响方法验证的情况下, 参数的量级已经做相关处理, 单位已略去。缺失数据段为第 401 ~ 700min, 相应的性能参数值为 0。

[0045] 步骤一、缺失退化数据趋势回归建模, 估计残差序列;

[0046] 回归得到退化数据的趋势项函数为  $x=0.0196t+0.1390$ , 计算已有数据去除趋势项之后的前后两段残差序列  $E$  和  $E'$ , 如图 5 所示。

[0047] 步骤二、对残差序列进行小波分解;

[0048] 选择 Daubechies5 小波对残差序列  $E$  和  $E'$  进行 4 级小波分解, 残差序列  $E$  的分解结果如图 6 所示。

[0049] 步骤三、分解子序列预处理;

- [0050] 按照上文公式 (3) 对分解后的子序列进行预处理, 其中  $q$  取为 10。
- [0051] 步骤四、采用小波神经网络插补模型插补缺失残差序列;
- [0052] 对每一级子序列构建一个 10 输入单输出小波神经网络, 完成网络训练。按照图 3 所示流程插补残差序列。最终得到各级插补子序列如图 7 所示。
- [0053] 步骤五、将缺失处的趋势项及残差序列插补结果合成为退化数据插补结果;
- [0054] 根据步骤一中得到的回归函数  $x=0.0196t+0.1390$ , 计算缺失数据段的趋势序列  $\hat{X}$ 。将  $\hat{X}$  与由图 7 各子序列重构得到的残差插补序列, 合成最终插补序列  $\hat{Y}$ , 完成退化数据缺失插补。最终插补结果如图 8 所示。

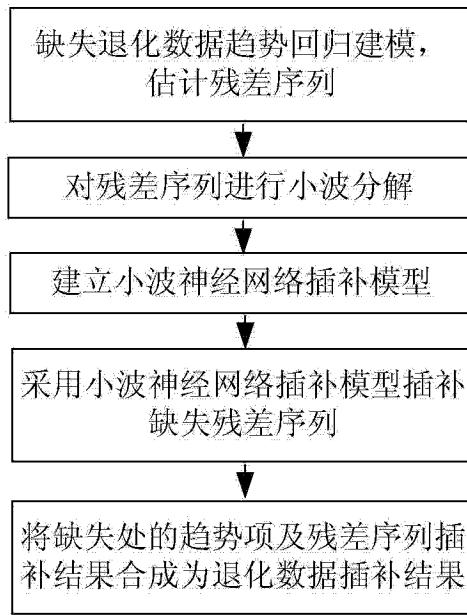


图 1

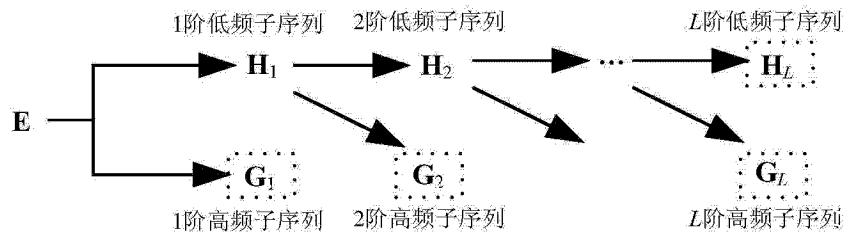


图 2

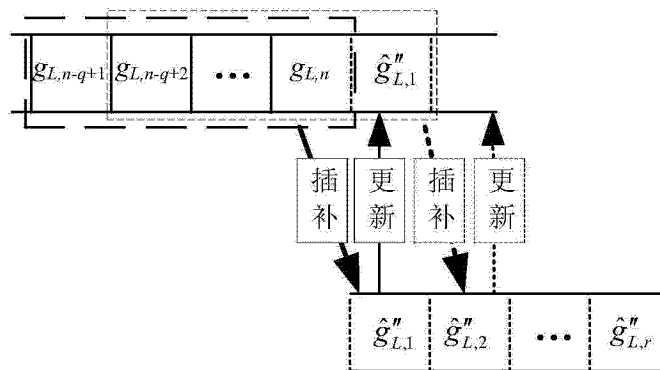


图 3



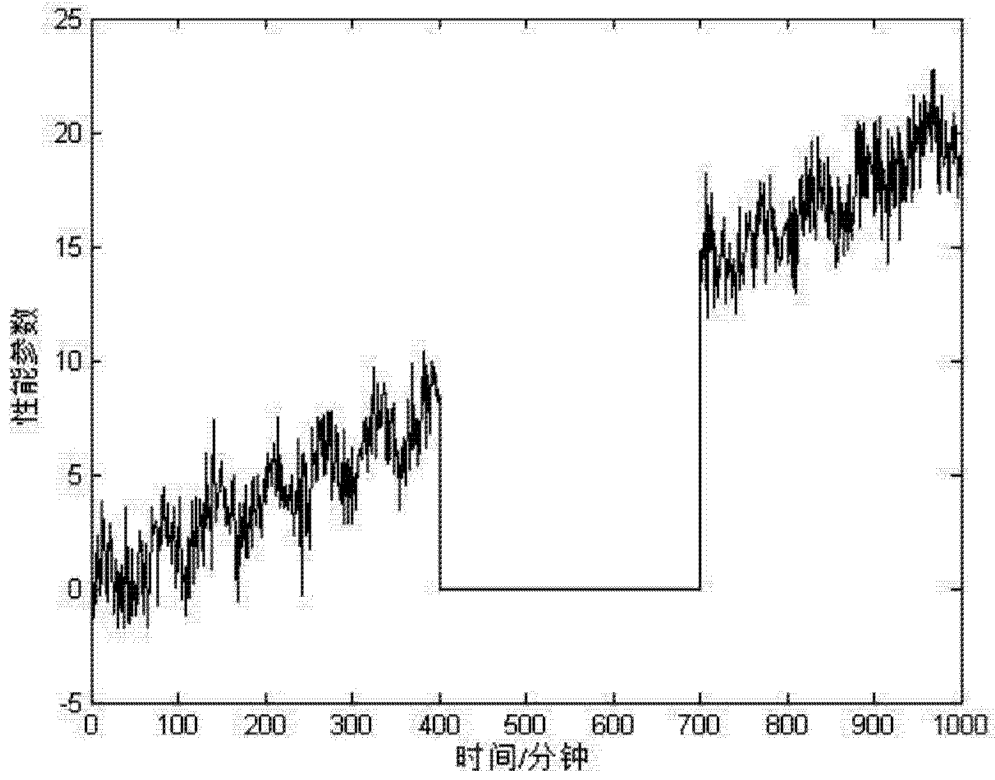


图 4

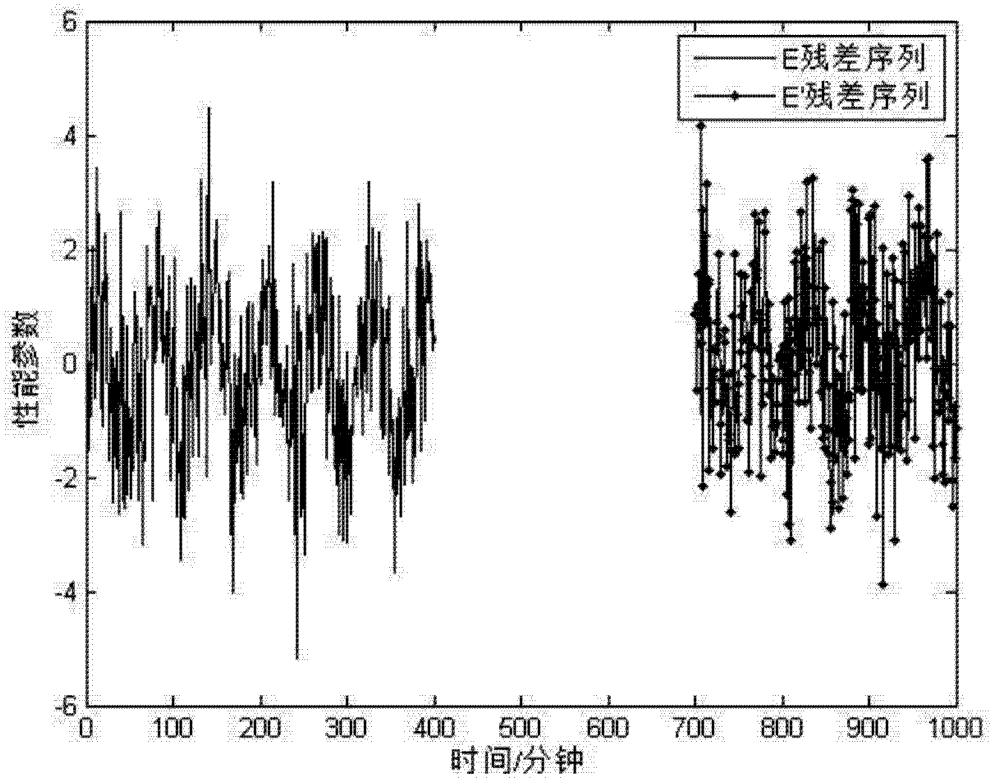


图 5

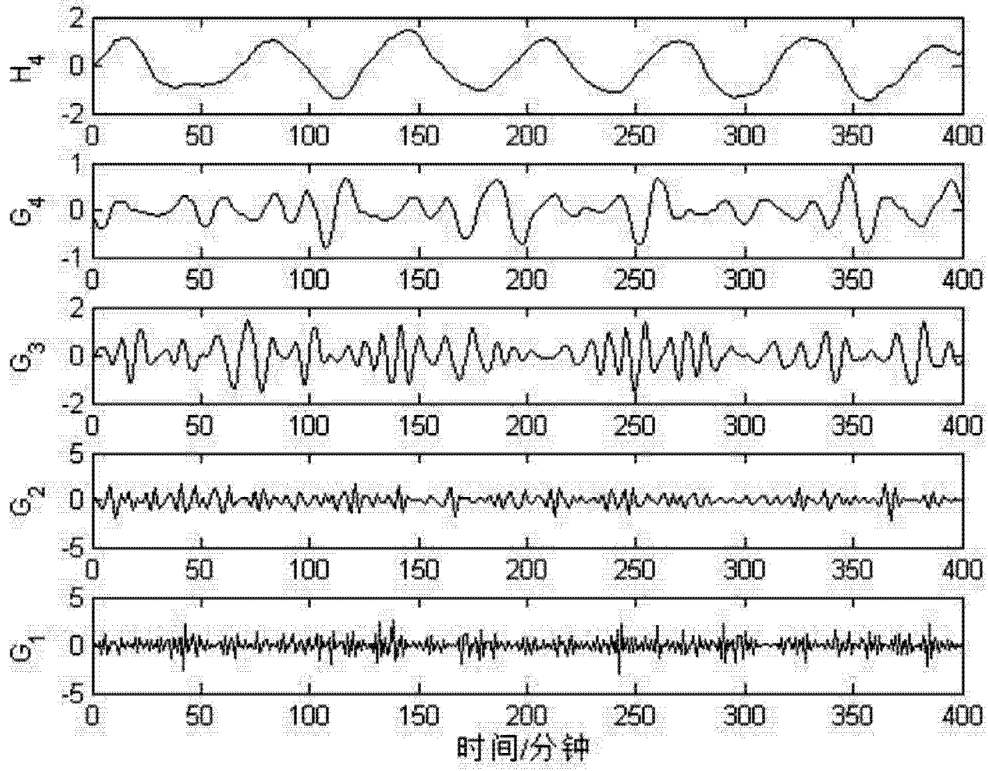


图 6

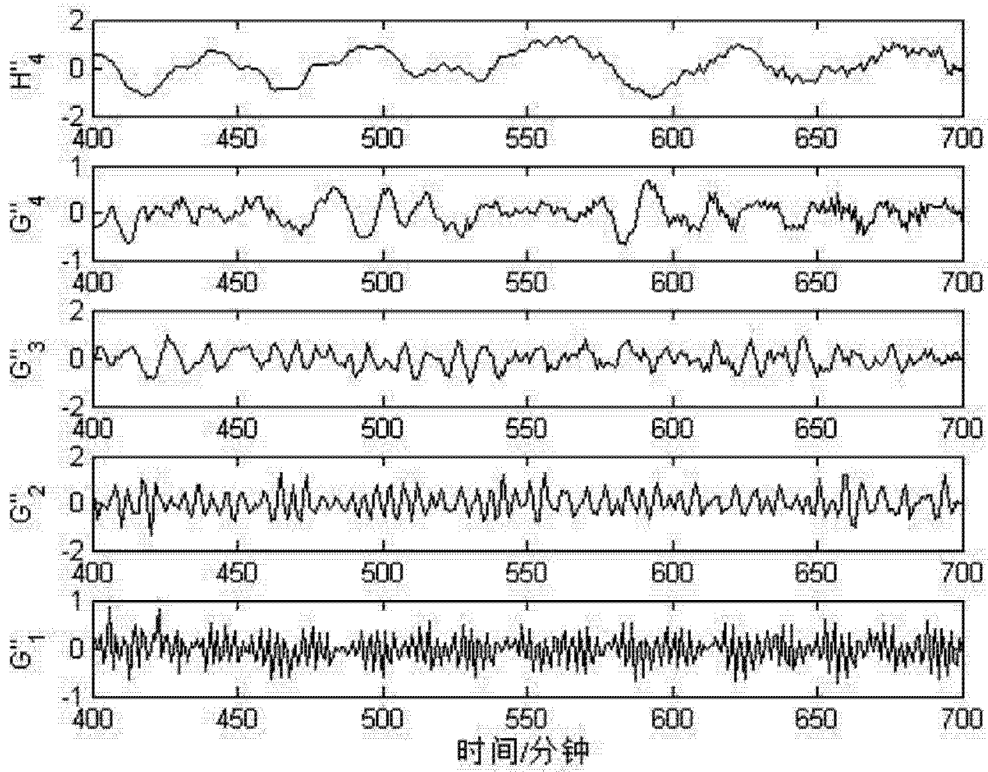


图 7

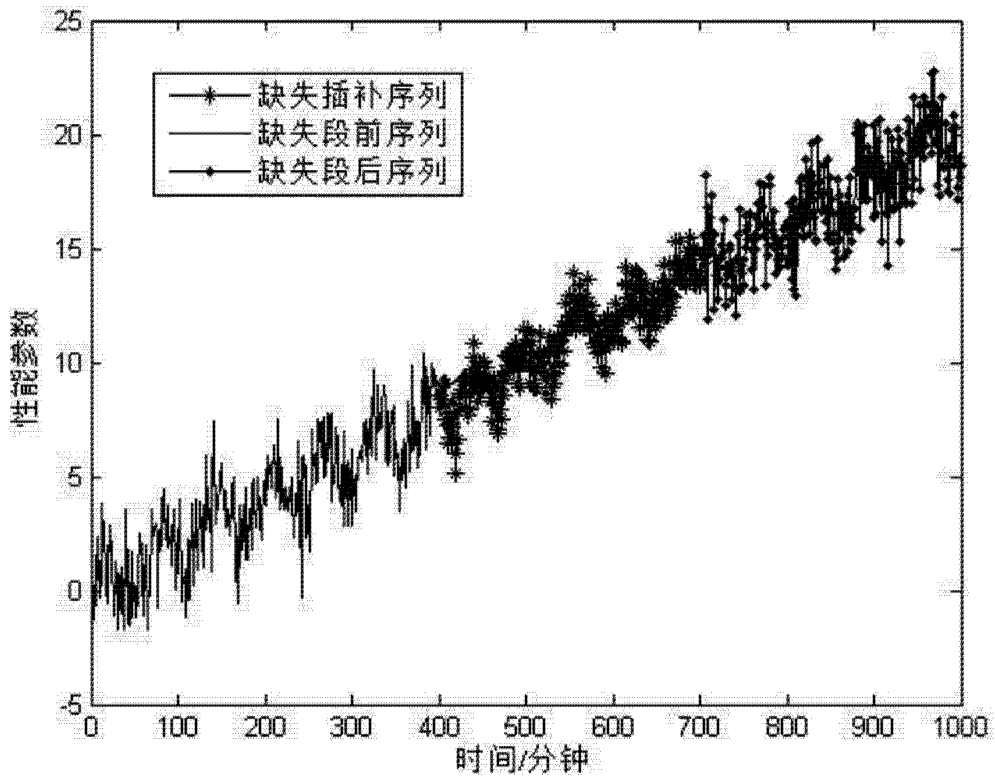


图 8