



(12) 发明专利申请

(10) 申请公布号 CN 105138800 A

(43) 申请公布日 2015. 12. 09

(21) 申请号 201510610969. 5

(22) 申请日 2015. 09. 17

(71) 申请人 天津理工大学

地址 300384 天津市西青区宾水西道 391 号

(72) 发明人 孙杰 吴海强 郑龙洋

(51) Int. Cl.

G06F 17/50(2006. 01)

G01N 21/64(2006. 01)

权利要求书1页 说明书3页 附图3页

(54) 发明名称

基于分段拟合数据处理算法的荧光光谱数据  
噪声滤波方法

(57) 摘要

本发明公开一种基于分段拟合数据处理算法的荧光光谱数据噪声滤波方法，将分段拟合数据处理算法运用到荧光光谱的数据处理中，并对光谱曲线进行分段拟合时做多次迭代运算，滤掉光谱数据中的随机干扰，使最终的光谱曲线还原其曲线特性，包括以下过程：实测光谱数据；确立极值点；以单调区间划分拟合区域；分段拟合数据滤波；输出处理后的数据。采用本发明所述方法处理的光谱数据后可消除由于噪声影响而出现的异常数据，能有效提高光谱仪测量精度，为以光谱数据为依据的测量和检测提供保障。



1. 一种基于分段拟合数据处理算法的荧光光谱数据噪声滤波方法,其特征在于:将分段拟合数据处理算法运用到荧光光谱的数据处理中,并对光谱曲线进行分段拟合时做多次迭代运算,包括以下过程:实测光谱数据;确立极值点;以单调区间划分拟合区域;分段拟合数据滤波;输出处理后的数据;

具体步骤如下:

(1) 处理原始实测光谱数据的步骤:将采集到的光谱数据通过取点  $x_i$  附近等间距的  $n_l+n_r+1$  个点,用  $n_l$  表示  $x_i$  左边点的个数,用  $n_r$  表示  $x_i$  右边点的个数,  $x_{i+1}-x_i = \Delta x$  为均匀间距;

(2) 对生成的序列进行计算处理的步骤:通过对数据的求导得出极值点,划分出单调区间,用于分段拟合区域;

(3) 对处理后的序列处理的步骤:多项式在  $x_i$  的值,就给出了它的光滑数值  $g_i$  ( $g_i \in G$ ),因此

$$g_i = \sum_{k=0}^M b_k \left[ \frac{x - x_i}{\Delta x} \right] k$$

由于  $x_{i+1}-x_i = \Delta x$  的均匀间距,设实测数据为  $y_i$ ,  $P_i(x)$  表示相对于点  $x_i$  的一个  $M$  次多项式,用它在最小二乘意义下拟合这  $n_l+n_r+1$  个点,为了使用  $P_i(x)$  拟合测试数据,必须定义系数  $b_k$ ,使得下式达到最优,

$$\min \sum_{j=1}^{l+n_r} [p_i(x_j) - y_j]^2$$

$P_i(x)$  即为拟合后的光谱曲线。

2. 根据权利要求 1 所述的基于分段拟合数据处理算法的荧光光谱数据噪声滤波方法,其特征在于:其中迭代运算次数为 2-4 次。

3. 根据权利要求 1 所述的基于分段拟合数据处理算法的荧光光谱数据噪声滤波方法,其特征在于:分段拟合区域的处理数据区域段的选取以光谱曲线的单调区域为段选取依据。

## 基于分段拟合数据处理算法的荧光光谱数据噪声滤波方法

### 技术领域：

[0001] 本发明涉及一种荧光光谱数据的降噪处理算法，能有效滤除光谱数据中随机干扰，提高其仪器测量精度。

### 背景技术：

[0002] 光谱分析作为现代分析测试中最为常见的一种技术手段，能够以较小的样品损耗提供较全面的物质特征信息，因而在冶金矿业、地质地貌、工业机械、石油化工、医药安全、卫生质检等领域中一直发挥着重要作用。

[0003] 在光谱测量中，由于外界干扰和仪器本身噪声等因素的影响，测得的光谱数据中在不同程度上含有随机噪声，噪声会在引起谱图的抖动、毛刺的出现，将影响谱图的特征分析，增加了后续分析和检测的误差。

[0004] 通过对测量得到的图谱数据进行数字处理可降低噪声、提高信噪比，是光谱分析中常用的预处理方法。为了消除或减弱干扰的影响，提高曲线的平滑度，需对采样数据进行平滑处理。

### 发明内容：

[0005] 本发明的目的是通过设定相关的参数，以处理效果稳定性和误差大小为依据，给出了滤波的最佳模板，处理后避免出现锯齿状波动等异常数据，能有效提高光谱数据的计算精度。

[0006] 为了达到上述目的，本发明采用的技术方案是：

[0007] 一种基于分段拟合数据处理算法的荧光光谱数据噪声滤波方法，将分段拟合 数据处理算法运用到荧光光谱的数据处理中，并对光谱曲线进行分段拟合时做多次迭代运算，包括以下过程：实测光谱数据；确立极值点；以单调区间划分拟合区域；分段拟合数据滤波；输出处理后的数据；

[0008] 具体步骤如下：

[0009] (1) 处理原始实测光谱数据的步骤：将采集到的光谱数据通过取点  $x_i$  附近等间距的  $n_l+n_r+1$  个点，用  $n_l$  表示  $x_i$  左边点的个数，用  $n_r$  表示  $x_i$  右边点的个数， $x_{i+1}-x_i = \Delta x$  为均匀间距；

[0010] (2) 对生成的序列进行计算处理的步骤：通过对数据的求导得出极值点，划分出单调区间，用于分段拟合区域；

[0011] (3) 对处理后的序列处理的步骤：多项式在  $x_i$  的值，就给出了它的光滑数值  $g_i$  ( $g_i \in G$ )，因此

$$[0012] g_i = \sum_{k=0}^M b_k \left[ \frac{x - x_i}{\Delta x} \right]^k$$

[0013] 由于  $x_{i+1}-x_i = \Delta x$  的均匀间距，设实测数据为  $y_i$ ， $P_i(x)$  表示相对于点  $x_i$  的一个  $M$

次多项式,用它在最小二乘意义下拟合这  $n_l+n_r+1$  个点,为了使用  $P_i(x)$  拟合测试数据,必须定义系数  $b_k$ ,使得下式达到最优,

$$[0014] \quad \min \sum_{i=1}^{i+n_r} [P_i(x_j) - y_j]^2$$

[0015]  $P_i(x)$  即为拟合后的光谱曲线。

[0016] 进一步的,本发明的迭代运算次数为 2-4 次。

[0017] 进一步的,本发明分段拟合区域的处理数据区域段的选取以光谱曲线的单调区域为段选取依据。

[0018] 本发明的有益效果是:本发明所述方法通过分段拟合的光谱数据可过滤掉随机干扰数据,保证输出光谱曲线光滑,原有曲线特性不变。通过设定相关的参数,以处理效果稳定性和误差大小为依据,给出了滤波的最佳模板,处理后避免出现锯齿状波动等异常数据,能有效提高光谱数据的测量精度。

#### 附图说明:

[0019] 图 1:本发明数据处理方法流程图;

[0020] 图 2:模拟光谱数据与分段拟合处理对比图;

[0021] 图 3:理想信号与分段拟合处理对比图;

[0022] 图 4:分段拟合处理误差图。

#### 具体实施方式:

[0023] 下面结合附图举例对本发明做更详细地描述。

[0024] 本实施例是一种光谱数据过程平滑处理方法。所述方法的思路如下:

[0025] 将采集到的原始光谱数据假定为离散的随机信号,由于随机干扰的存在,使得随机信号绘成的过程多呈锯齿状,表明采样数据具有非平稳随机过程特性。为了消除或减弱干扰的影响,需对原光谱数据进行平滑处理。平滑的原则是既要消除数据中的干扰成分,又要保持原有曲线特性不变。

[0026] 参见附图 1,本实施例所述方法的具体步骤如下:

[0027] 处理原始实测光谱数据的步骤:将采集到的光谱数据通过取点  $x_i$  附近等间距的  $n_l+n_r+1$  个点,用  $n_l$  表示  $x_i$  左边点的个数,用  $n_r$  表示  $x_i$  右边点的个数,  $x_{i+1}-x_i = \Delta x$  为均匀间距。

[0028] 对生成的序列进行计算处理的步骤:通过对数据的求导得出极值点,划分出单调区间,用于分段拟合区域。

[0029] 对处理后的序列处理的步骤:多项式在  $x_i$  的值,就给出了它的光滑数值  $g_i$  ( $g_i \in G$ )。因此

$$[0030] \quad g_i = \sum_{k=0}^M b_k \left[ \frac{x - x_i}{\Delta x} \right]^k$$

[0031] 由于  $x_{i+1}-x_i = \Delta x$  的均匀间距,设实测数据为  $y_i$ ,  $P_i(x)$  表示相对于点  $x_i$  的一个  $M$  次多项式,用它在最小二乘意义下拟合这  $n_l+n_r+1$  个点。为了使用  $P_i(x)$  拟合测试数据,必

须定义系数  $b_k$ ,使得下式达到最优。

$$[0032] \quad \min \sum_{i=1}^{l+n_r} [p_i(x_j) - y_j]^2$$

[0033]  $P_i(x)$  即为拟合后的光谱曲线。并且我们可以对光谱曲线进行多次平滑,分段拟合次数为 2-4 次。可根据实际资料及具体需求调试确定,其原则是在保持原有基本形态不变的基础上,尽量消除过程中的干扰成分。可以取多次进行比较,以获得最佳的平滑效果。

[0034] 表 1 为采用基于分段拟合数据处理算法处理后的光谱数据的对比。

[0035] 表 1 滤波数据采样对比

[0036]

采样点 x	0.000 0	0.571 2	1.205 9	1.840 5	2.475 2	3.109 9	3.744 5	4.379 2	5.013 9	5.648 5	6.283 2
理想信号	0.000 0	0.540 6	0.934 1	0.963 8	0.618 2	0.031 7	-0.56 71	-0.94 50	-0.95 49	-0.59 29	0.000 0
加入噪声信号	0.086 6	0.543 6	0.855 2	0.816 7	0.612 7	0.128 3	-0.59 60	-0.86 24	-0.86 66	-0.57 97	0.069 2
处理后信号	0.074 0	0.537 0	0.973 7	0.955 8	0.611 2	0.017 8	-0.53 84	-0.95 79	-0.94 16	-0.56 58	0.061 8
误差	0.074 0	0.003 6	0.039 6	0.008 0	0.007 0	0.013 9	0.028 7	0.012 9	0.013 3	0.027 1	0.061 8

[0037] 由表 1 我们可以计算出分段拟合数据处理算法处理后的光谱数据的平均误差  $\bar{x}=0.0264$ ,均方根误差  $\delta = 0.0345$ 。

[0038] 本发明运用了 MATLAB 进行仿真实验,以  $\sin(x)$  函数模拟无噪声理想收入光谱曲线,在理想信号中加入高斯白噪声,运用分段拟合算法进行滤波处理。

[0039] 仿真结果如图 2,3,4 所示:

[0040] 图 2 为模拟光谱数据与分段拟合曲线的对比图,通过对比可以发现分段拟合算法可以很好的对光谱数据进行拟合。

[0041] 图 3 为理想信号与分段拟合处理对比图,通过对比可以发现分段拟合算法拟合的曲线与理想曲线大体契合,能很好的对理想光谱数据进行还原。

[0042] 图 4 为分段拟合处理误差图,通过此图可以很直观的看出分段拟合处理与理想数据之间的误差。

[0043] 应当明确的是,上述对实施例的描述是为便于该技术领域的普通技术人员能理解和应用本发明。熟悉本领域技术的人员显然可以容易的对这些实施例做出各种修改,并把在此说明的一般原理应用到其他实施例中而不必经过创造性劳动。因此,本发明不限于这里的实施例,本领域技术人员根据本发明的揭示,对于本发明做出的改进和修饰都应该在本发明的保护范围之内。

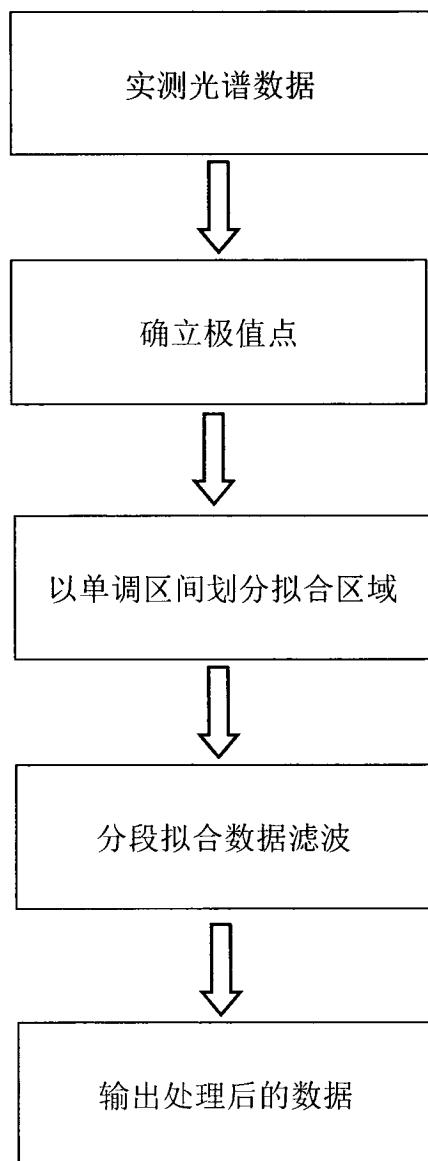


图 1

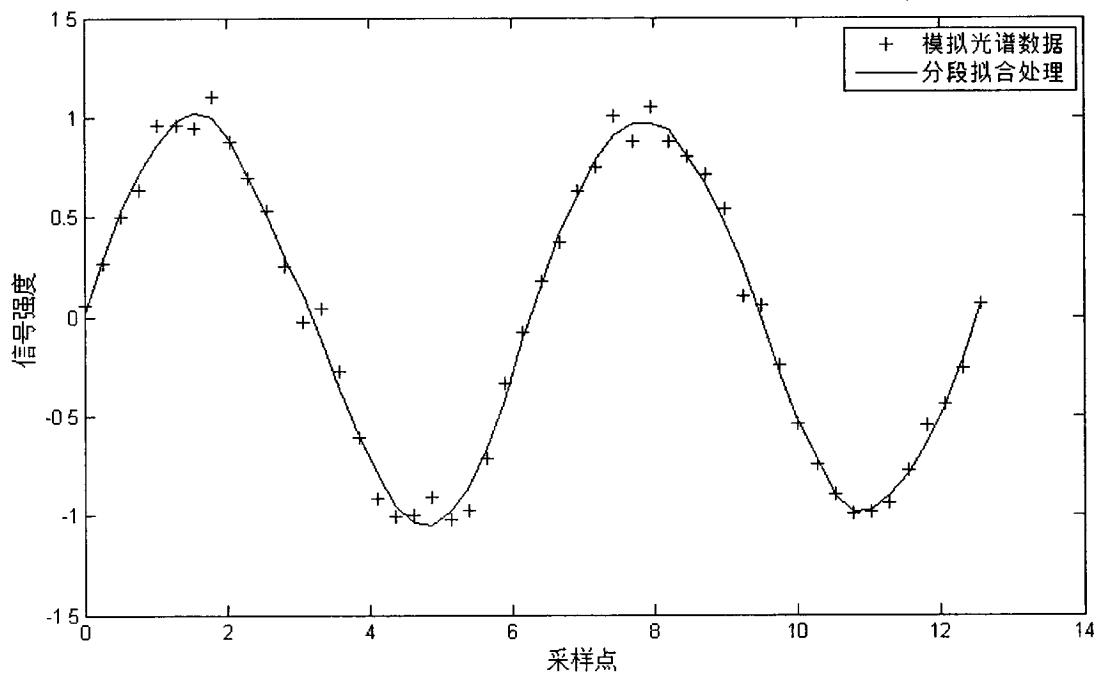


图 2

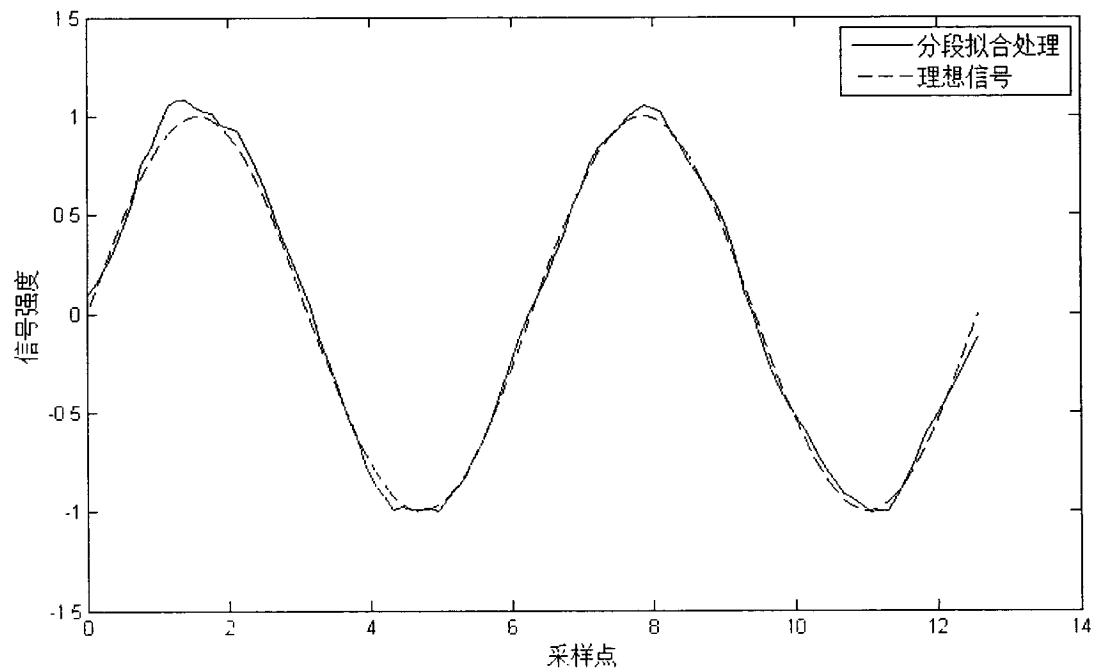


图 3

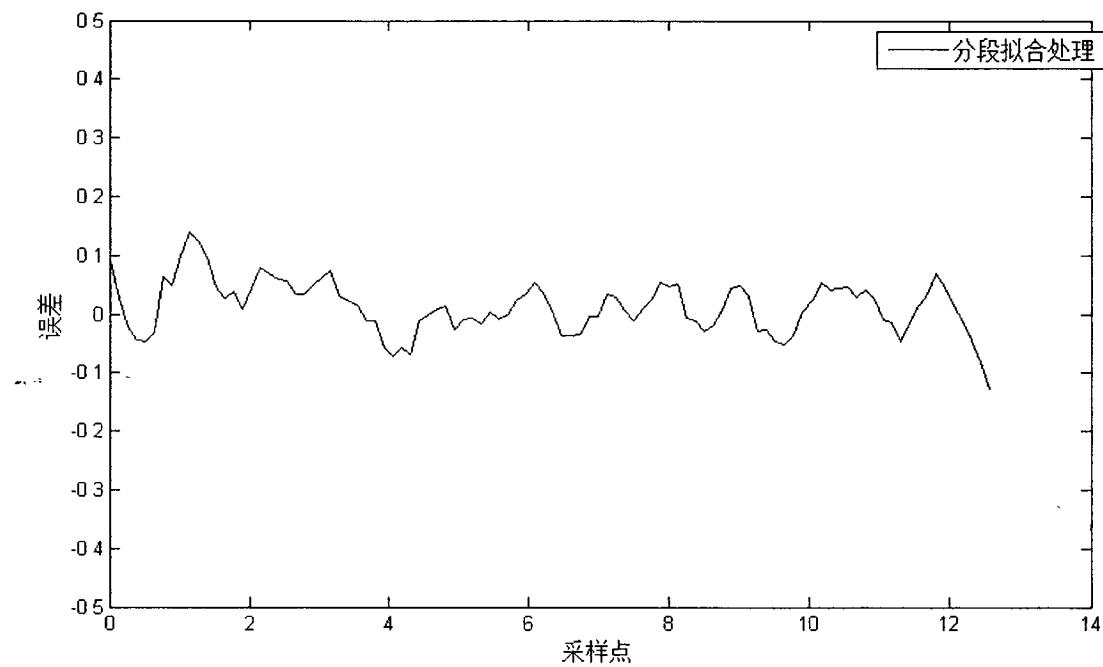


图 4