

Research on the application of fluorescence signal peak detection algorithm

Changjun Li Wei Cai Chao Hu

General Technology Group Guoce Time Grid Technology Co., Ltd., Chongqing, 400039, China

Abstract

The principle of the currently widely used digital PCR fluorescence detection method is: to detect the fluorescence signal generated by the target substance, then process the collected fluorescence signal through an algorithm to determine the number of positive reaction units, and then calculate the initial concentration of the target substance by Poisson distribution. Among them, the detection and processing of fluorescence signals are the core links of the digital PCR fluorescence detection system. The rate of fluorescence data acquisition can directly reflect the throughput of the target substance, and the accuracy of the processing algorithm directly affects the detection accuracy. In this paper, a fluorescence signal peak detection method based on continuous wavelet transform is adopted, and the algorithm is implemented through MATLAB language. Several groups of actual collected fluorescence signals are detected and evaluated.

Keywords

Fluorescence signal; Wavelet transform; Peak detection

荧光信号峰值检测算法应用研究

李常军 蔡伟 胡超

通用技术集团国测时栅科技有限公司, 中国·重庆 400039

摘要

目前常用的数字PCR荧光检测方法的原理是: 对被检测物产生的荧光信号进行检测, 然后对采集到的荧光信号进行算法处理判定阳性反应单元个数, 再由泊松分布进行计算, 推断出目标检测物起始浓度。其中, 荧光信号的检测与处理是数字PCR荧光检测系统的核心环节。荧光数据采集的速率能直接反映被检测物的通量, 处理算法的精度直接影响检测精度。本文采用了一种基于连续小波变换的荧光信号峰值检测方法, 通过MATLAB语言实现该算法, 并对实际采集到的几组荧光信号进行检测并评估。

关键词

荧光信号; 小波变换; 峰值检测

1 引言

数字PCR技术可实现对检测样本的绝对定量, 其实现过程是: 首先对待检测DNA样本进行稀释和分液, 其目的是让每个待测微滴中的分子数少于或等于1个。然后让这些微滴在同样条件下进行PCR扩增。最后对样品进行检测计数, 其中, 在进行荧光信号读数时, 阳性反应单元记为1, 阴性反应单元记为0。

当含有荧光标记的分子的待测微滴经过PCR扩增后, 由激光激发产生的荧光信号将会很强, 以至于可采用荧光计数的方法进行绝对定量。虽然待测微滴样品中可能含有其他杂质分子, 但它们不会发生PCR扩增反应, 由其产生的干扰信号不会影响到最后的检测结果^[1]。

【作者简介】李常军(1993-), 男, 中国重庆人, 硕士, 工程师, 从事信号处理、FPGA设计研究。

为了统计出阳性反应单元个数以实现目标检测样本的绝对定量, 就要对采集到的荧光信号进行峰值检测以判断微滴的阴阳性。本文采用了基于连续小波变换进行荧光信号峰值检测的方法, 并通过波峰高度设定阈值找出合格的阳性反应单元。

2 峰值检测方法研究

小波变换是一种时频分析方法, 目前在信号分析与处理领域已得到广泛的应用。小波变换表征信号局部特征的原理是: 通过选择合适的尺度因子和平移因子可以得到一个变化的窗函数, 通过伸缩、平移运算可聚焦到信号的细节部分。小波变换分为连续小波变换(CWT)和离散小波变换(DWT)。连续小波变换可进行任意尺度的平移和缩放, 但离散小波变换只能进行特定尺度的平移和缩放^[2]。

小波是指满足一定条件的函数 $\psi(t)$ (称为基小波或母小波)通过伸缩和平移产生的一个函数族 $\psi_{a,b}(t)$:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad a, b \in \mathbb{R}; a \neq 0 \quad (1)$$

其中, a 是用于控制伸缩的变量, 称为尺度因子, b 是用于控制平移的变量, 称为平移因子。

给定一个基小波 $\psi(t)$, 对于一维连续信号 $f(t)$, 它的连续小波变换 (CWT) 定义为:

$$W_f(a, b) = \int_{-\infty}^{\infty} f(t) \psi_{a,b}^*(t) dt = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t) \psi^*\left(\frac{t-b}{a}\right) dt \quad (2)$$

任意函数 $f(t)$ 的小波变换都是一个二元函数, 这是小波变换和傅里叶变换很不相同的地方。函数的小波变换 $W_f(a, b)$ 数值的本质是原来的函数或者信号 $f(t)$ 在 $t=b$ 点附近按 $\psi_{a,b}(t)$ 进行加权的平均。

对尺度因子 a 和平移因子 b 进行离散化, 其中令 $a = a_0^j, b = ka_0^j b_0, j \in \mathbb{Z}$ 可得离散小波基函数为:

$$\psi_{j,k}(t) = a_0^{-j/2} \psi\left(\frac{t - ka_0^j b_0}{a_0^j}\right) = a_0^{-j/2} \psi(a_0^{-j} t - kb_0) \quad (3)$$

则离散小波变换定义为:

$$W_{j,k}(t) = \int_{-\infty}^{\infty} f(t) \psi_{j,k}^*(t) dt \quad (4)$$

3 连续小波变换在峰值检测中的应用

传统的峰值检测算法一般分为 3 个步骤: 谱线平滑、基线校准与峰值识别 [30]。在峰值识别方面, 基于局部极大值、信噪比、峰面积等方法都有报道, 但这些算法的成效均在很大程度上依赖于前期的谱线平滑和基线校准 [3-7]。随着小波理论的发展, 运用小波方法进行信号分析与处理成为主流。基于连续小波变换的峰值检测方法有良好的抗噪性能, 相比于传统的峰值检测算法有更好的综合性能。本文根据 *haar* 小波的特点, 提出一种基于 *haar* 小波的荧光信号峰值检测的方法。运用连续小波变换检测峰值的原理是根据小波变换后的各尺度小波系数与原始信号的关系识别出原始信号的峰

值, 该方法能方便地得到原始信号和小波系数的关系。

实际应用中, 所分析的信号通常是离散的采样数据, 所以该情况下的连续小波变换一般使用式 (2) 的离散形式:

$$Wf(a, iT_s) = T_s \frac{1}{\sqrt{a}} \sum_n f(nT_s) \psi\left(\frac{(n-i)T_s}{a}\right) \quad (5)$$

其中, $1/T_s$ 是采样频率, $\psi(t)$ 由小波母函数决定, 其中 *haar* 小波母函数为:

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \\ 0 & \text{其他} \end{cases} \quad (6)$$

可以看出, *haar* 小波函数是一个梯形函数。因此在计算时, 信号与 *haar* 小波的卷积, 经过一定处理后, 能够等效于原始信号的近似导数, 从而很容易找到原始信号与小波系数的关系。这种方法不仅能够快速地进行计算, 而且具有良好的局部化特性, 可进行任意小的尺度分解。

本文采用的峰值检测方法是: 利用 *haar* 小波与原始荧光数据进行连续小波变换, 根据小波变换后得到的小波系数与原始荧光数据的关系检测峰值。这里选择的小波母函数为 *haar* 小波, 采用的尺度系数为 16。具体步骤如下: 原始荧光数据通过小波变换后, 得到对应尺度的小波系数。设 $\text{coef}(i)$ 表示小波系数向量中的第 i 个数, 则小波系数向量中单调递减且过零点的系数应满足以下准则:

$$\begin{aligned} \text{coef}(i-1) &\geq 0 \\ \text{coef}(i+1) &\leq 0 \\ \text{coef}(i) &\leq \text{coef}(i-1) \\ \text{coef}(i) &\geq \text{coef}(i+1) \end{aligned} \quad (7)$$

小波系数向量中单调递减且过零点附近的系数近似对应于原始荧光数据的极大值, 也即这里的小波系数相当于原始荧光数据的近似导数, 能够反映原始荧光数据的变化特征。图 1 是根据小波系数与原始荧光数据的关系获取的可疑峰值点。

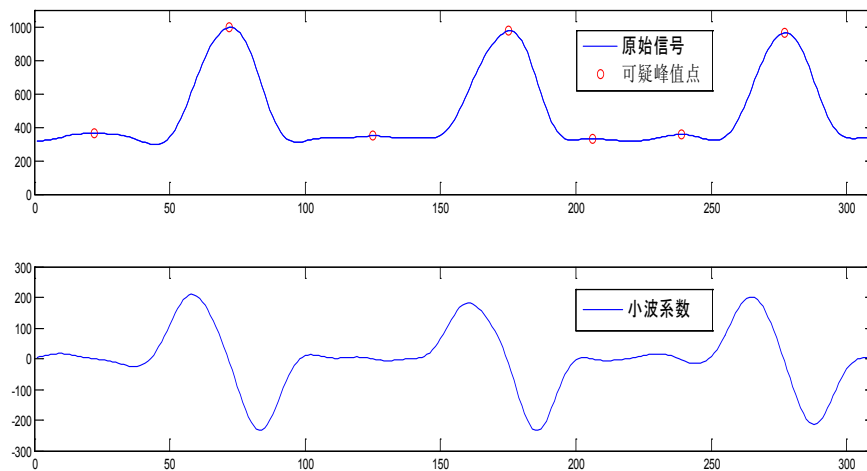


图 1 小波变换获取的可疑峰值点

由于这步求出的只是原始荧光数据的所有可疑局部极大值点，还应当设定阈值从而准确地获取真正的荧光信号的波峰值。本文中是将截取的荧光数据均值作为有效峰值判断的阈值，舍弃掉可疑峰值点从而准确获取有效峰值^[7]。图 2 是将获取的可疑峰值点经过阈值判断得到的峰值点。

4 结果与讨论

本文通过 MATLAB 语言实现了该荧光信号峰值检测方法，并对实际采集到的几组荧光信号进行了检测并评估。表

1 列出了该方法通过对数字 PCR 荧光检测系统采集到的荧光信号进行处理从而得到的阳性微滴的检测率。从表 1 可以看出，该方法具有很好的抗噪能力以及很高的正确检测率，阳性微滴的检测准确率达到 99.4%。该方法不仅简单并且易于实现，对荧光信号峰值特征的定位准确，迅速。另外，对误检的产生需要说明的是，除了外在物理条件造成某段信号的信噪比过低而使该算法产生误检，经过对错检处进一步的研究发现，还存在微流控芯片通道内微滴间距太小造成漏检以及微滴破损而产生误检的情况。

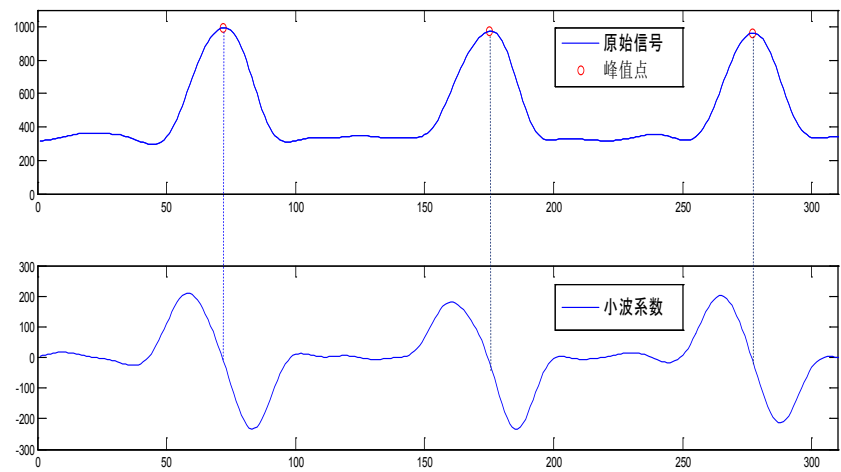


图 2 经过阈值判断获取的峰值点

表 1 荧光信号峰值检测结果

序号	阳性微滴数	误检数	检测准确率
第一组	267	0	100%
第二组	278	2	99.2%
第三组	274	3	98.9%
第四组	270	0	100%
第五组	264	2	99.2%
总计	1353	7	99.4%

所以，为了提高数字 PCR 荧光检测与处理系统的检测精度，不仅需要在信号采集与检测算法上下功夫，生成较好的微滴也有利于提高检测的精度。

参考文献

[1] 李鑫鑫, 邬少飞. 基于数字PCR荧光检测仪的软件系统设计模型. 科技视界, 2016(27):187-188.
[2] 候正信, 王国安. 数字信号处理基础. 第一版. 北京: 电子工业

出版社, 2003, 525-554.
[3] 周玲丽, 赖剑煌. 基于尺度不变特征变换优化算法的带遮挡人脸识别. 计算机应用, 2011, 31(a01): 64-66.
[4] Li C, Zheng C, Tai C. Detection of ECG characteristic points using wavelet transforms. IEEE Transactions on Biomedical Engineering, 1995, 42(1): 21.
[5] Du P, Kibbe W A, Lin S M. Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. Bioinformatics, 2006, 22(17): 2059-2065.
[6] Shao X, Ma C. A general approach to derivative calculation using wavelet transform. Chemometrics & Intelligent Laboratory Systems, 2003, 69(1): 157-165.
[7] 蔡涛, 王先培, 杜双育, 阳婕. 基于多尺度小波变换的红外光谱谱峰识别算法. 分析化学, 2011, 39(06): 911-914.