DOI:10. 19431/j. cnki. 1673-0062. 2018. 06. 014

基于小波新阈值函数的标记荧光蛋白信号去噪

曹　亮[1],朱卫华[1*],林晓明[2],陈　权[1]

(1.南华大学 电气工程学院,湖南 衡阳 421001;2.广东顺德创新设计研究院理学部,广东 佛山 528000)

摘　要:针对微流控电泳仪采集的微弱标记荧光蛋白信号易受到干扰,影响对样品生物蛋白定性定量分析.提出可调新阈值函数小波去噪,解决了软、硬阈值函数去噪出现的常见问题.通过 matlab 仿真对比,寻找出去噪质量最佳 db7 小波基,结合可调新阈值函数处理重叠峰严重的标记荧光蛋白信号,去噪后波形最光滑、信噪比最高、获得最小均方根误差、更能突出真实信号.结果表明,新的阈值函数更适用于重叠峰严重的标记荧光蛋白信号,为后续定性定量分析样品中的生物蛋白含量提供保障.

关键词:小波变换;荧光蛋白信号;小波阈值去噪;阈值函数

中图分类号:TN911.4　**文献标志码**:A　**文章编号**:1673-0062(2018)06-0085-06

De-Noising of Labeled Fluorescence Protein Signal Based on New Wavelet Threshold Function

CAO Liang[1], ZHU Weihua[1*], LIN Xiaoming[2], CHEN Quan[1]

(1.School of Electrical Engineering, University of South China, Hengyang, Hunan 421001, China; 2.Science Department, Guangdong Shunde Innovation Design Institute, Foshan, Guangdong 528000, China)

Abstract: The weakly labeled fluorescence protein signal collected by the microfluidic electrophoresis apparatus is susceptible to interference, which affects the qualitative and quantitative analysis of the biological protein in the sample. An adjustable new threshold function wavelet denoising was proposed to solve the common problems of soft and hard threshold function denoising. Through matlab simulation and comparison, the best db7 wavelet base for out-going noise quality was found, and the labeled new fluorescence protein signal was processed with the new adjustable threshold function. The denoised waveform was the smoothest, the signal-to-noise ratio was the highest, and the minimum root

收稿日期:2018-06-14

作者简介:曹　亮(1995-),男,硕士研究生,主要从事色谱信号采集处理方面的研究. E-mail:313397053@qq.com.
*通信作者:朱卫华(1964-),男,副教授,主要从事信号处理和嵌入式系统及应用方面的研究. E-mail:zhuweihua64@163.com

mean square error was obtained, which can better highlight the real signal. The results show that the new threshold function is more suitable for the labeling of fluorescence proteins with severe overlapping peaks, providing a guarantee for subsequent qualitative and quantitative analysis of biological protein content in samples.

key words: wavelet transform; fluorescence protein signal; wavelet threshold denoising; threshold function

0 引 言

微流控电泳仪采集的标记荧光信号非常微弱,易受噪声的污染,严重干扰对检测样品中所含物质成分进行定性定量分析.对检测到的标记荧光蛋白信号去噪,提高了分析样品成分的准确度.在信号处理中具有多分辨率分析的小波变换引起了学者们的关注,在文献[1]中就使用小波变换应用于矿物油荧光光谱去噪研究中.考虑微流控电泳仪检测信号为低信噪比以及去噪过程计算复杂度,本文采用小波阈值法对采集的数据进行去噪.

小波阈值去噪中应用得最为广泛的是硬阈值函数以及软阈值函数[2].硬阈值函数的计算简单,但是具有不连续性,导致在处理后的信号中可能会存在伪吉布斯现象.软阈值函数在时域内是连续的,但是经过其处理后的信号与未带噪前相比会出现偏差问题.为了解决硬、软阈值函数上述的问题,文献[3]中对阈值函数进行改进,克服了软、硬阈值函数已有的问题,但该文献提出的改进阈值函数适用于重叠峰很少的标记荧光蛋白信号.

为了处理重叠峰严重的标记荧光蛋白信号,提出了具有可调的新阈值函数.经过仿真实验,发现新可调阈值函数去噪质量优于传统的硬、软阈值函数,同时也证实具有可调特性的新阈值公式在处理重叠峰较严重的数据时,具有更好的去噪效果.

1 小波阈值去噪

假设微流控电泳仪采集到的含噪标记荧光蛋白信号用式(1)模型表示:

$$G(t) = X(t) + N(t) \tag{1}$$

式(1)中:$G(t)$表示含噪标记荧光蛋白信号,$X(t)$代表纯净的标记荧光蛋白信号,噪声信号用$N(t)$表示.

小波阈值去噪过程如图 1 所示,主要为信号小波分解、阈值处理、小波重构三个步骤[4].

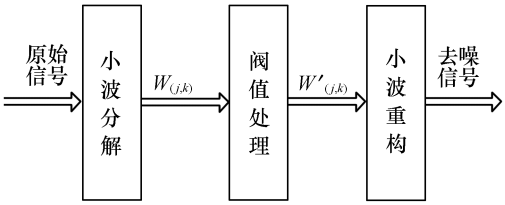

图 1 小波阈值去噪过程

Fig.1 Wavelet threshold denoising process

小波分解:根据 Mallat 鱼骨算法对数据进行解析,Mallat 的核心处理思想如式(2)所示:

$$A_j f = A_{j+1} f + D_{j+1} f \tag{2}$$

$$A_j f = \sum_k C_{(j,k)} \varphi_{(j,k)} \text{、} D_j f = \sum_k D_{(j,k)} \psi_{(j,k)}$$

在式(2)中:j 是小波分解层数,k 是小波分解系数.在任一分解层中,式(2)的 $A_j f$、$A_{j+1} f$、$D_{j+1} f$ 分别对应式(1)的 $G(t)$、$X(t)$、$N(t)$.多分辨分析思想是将带噪信号分开成详细分量 D_{j+1}(高频成分)和模糊分量 A_{j+1}(低频成分).

阈值处理:对分解出来的细节系数 $w_{(j,k)}$ 经过阈值量化分析后得到 $w'_{(j,k)}$.

小波重构:把上一步骤计算出的细节系数 $w'_{(j,k)}$ 和模糊系数结合,获得去噪后的信号.

1.1 阈值规则的选择

在小波阈值去噪中,结合待处理数据特点,选择最优的阈值判定方法和处理公式在整个去噪中非常重要的,有利于去噪后信号整体质量提升.阈值常用式(3)固定阈值规则进行判定,同时还有学者对固定阈值规则进行改进,如文献[5]中提出式(4)还有文献[6]中提出式(5)的改进阈值规则.

$$\lambda_j = \sigma_{dj} \sqrt{2\ln(N_j)} \tag{3}$$

$$\lambda_j = \frac{\sigma_{dj} \sqrt{2\ln(N_j)}}{\ln(j+1)} \tag{4}$$

$$\lambda_j = \frac{\sigma_{dj} \sqrt{2\ln(N_j)}}{\log_2(j+1)} \tag{5}$$

上式中:$\sigma_{dj}=\frac{\mathrm{median}(|w_{(j,k)}|)}{0.6745}$,$\sigma_{dj}$为第 j 层的噪声估计,N_j 是信号在 j 层的长度.该固定阈值规则是对每一层进行噪声估计,具有一定去噪效果.

式(4)和式(5)阈值规则都是根据小波分解层的不同,除以一个相应的调节系数,将会增长计算阈值的时间,从而延长了阈值处理的时间.本研究中的去噪阈值由是式(6)计算并确定:

$$\lambda_j=\sigma_{d1}\sqrt{2\ln(N_j)} \tag{6}$$

式(6)中:σ_{d1}为第 1 层的噪声估计,该阈值规则只需要对第一层的噪声进行估计,大大节省了阈值去噪过程中计算阈值的时间,同时只要挑拣相宜的阈值处理公式,该阈值选取方法的去噪质量将高于其它方法.

1.2　新阈值函数的提出

在阈值函数中,当分解后细节系数低于阈值时,系数被判定噪声而去除,当分解后细节系数高于阈值时,说明该系数很可能为有用信号,将对其进行保留或衰减处理[7].被广泛使用的阈值处理公式由 Donoho 提出的硬阈值式(7)和软阈值式(8):

$$w'_{(j,k)}=\begin{cases}w_{(j,k)} & |w_{(j,k)}|>\lambda\\ 0 & |w_{(j,k)}|\leqslant\lambda\end{cases} \tag{7}$$

$$w'_{(j,k)}=\begin{cases}\mathrm{sign}(w_{(j,k)})(|w_{(j,k)}|-\lambda) & |w_{(j,k)}|>\lambda\\ 0 & |w_{(j,k)}|\leqslant\lambda\end{cases} \tag{8}$$

在式(7)和式(8)中:$w_{(j,k)}$为待处理系数,$w'_{(j,k)}$为公式处理之后的系数.sign(·)为符号函数,λ 为阈值规则计算出来的阈值.

分析式(7),硬阈值处理函数在 λ 处将会出现间断点,经过硬阈值函数处理后的信号在某些点处会产生震荡,造成信号不光滑.分析式(8),软阈值函数克服了式(7)中的不足,但是经过软阈值函数处理后的信号会留存一个固定的偏差.新的可调阈值函数的提出主要是围绕以下三点:

1)在$-\lambda$ 和 λ 处都不发生间断,即克服式(7)不连续性问题.

2)在小波分解出来的待处理系数越大于 λ 时,处理后的值越贴近原信号,即需要克服式(8)出现恒定偏差的现象.

3)对采集的数据使用新阈值公式方法去噪质量要高于其它阈值公式.

经过查阅大量国内外学者的阈值公式的研究[8-9],最后提出了新的可调阈值处理函数如式(9):

$$w'_{(j,k)}=\begin{cases}\mathrm{sign}(w_{(j,k)})\times\left(|w_{(j,k)}|-\dfrac{5\lambda}{4+\exp(2\alpha(|w_{(j,k)}|-\lambda)/\lambda}\right), & |w_{(j,k)}|>\lambda\\ 0, & |w_{(j,k)}|\leqslant\lambda\end{cases} \tag{9}$$

在式(9)中:

当 $w_{(j,k)}\to\lambda$ 时,$w'_{(j,k)}=0$;当 $w_{(j,k)}\to-\lambda$ 时,$w'_{(j,k)}=0$.即该阈值函数在任何点处连续,避免了去噪处理后信号产生伪吉布斯现象.

当 $w_{(j,k)}\to\infty$ 时,$w'_{(j,k)}\to w_{(j,k)}$;当 $w_{(j,k)}\to-\infty$ 时,$w'_{(j,k)}\to w_{(j,k)}$.即当待处理系数$|w_{(j,k)}|$越大于 λ 时,处理后的 $w'_{(j,k)}$越靠近 $w_{(j,k)}$.说明使用新阈值公式方法处理后的信号不会产生一个固定偏离.

式(9)波形对比如图 2 所示.该阈值函数具有适应性能力,可以根据待处理数据的不同,修改 α 调节因子,使去噪效果达到最优.

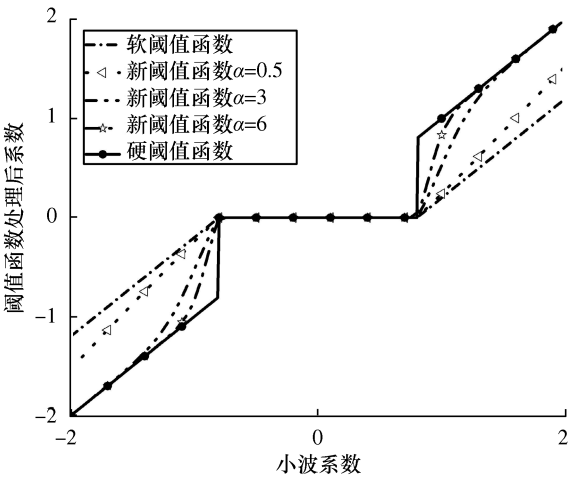

图 2　阈值函数波形对比图

Fig.2　Comparison of threshold function waveforms

1.3　去噪评价指标

为了方便确定最优的去噪方法,运用信噪比($R_{s,n}$)和均方根误差($E_{r,m,s}$)的阈值去噪评价指标[10],其式如下:

$$R_{s,n}=10\lg\left(\frac{\sum_{i=1}^{n}y_0^2(i)}{\sum_{i=1}^{n}(y_0(i)-y_1(i))^2}\right) \tag{10}$$

$$E_{r,m,s}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_0(i)-y_1(i))^2} \tag{11}$$

在式(10)~式(11)中:$y_0(i)$是不带噪信号,$y_1(i)$为处理后的数据,n 是处理信号的总长度.

$R_{s,n}$值越高体现了去噪后的纯净程度越好,掺杂的噪声越少;$E_{r,m,s}$值越低代表了处理后数据越贴近真实值.

1.4 验证新阈值函数和阈值规则

为了验证改进的新阈值函数和使用的阈值规则方法处理后信号最优.用 matlab 模拟一组高斯峰数据,同时为了消除小波基在小波阈值去噪中的影响,统一使用 db7 小波基进行带噪数据小波分解,寻找出各阈值公式下对应的最好评价指标($R_{s,n}$、$E_{r,m,s}$)、最佳分解层还有最佳 α 调节因子.

各阈值函数方法处理后信号如图 3 所示,使用新调节阈值处理公式方法去噪后波形 f 明显比 b、e 要好.去噪后波形 c 在信号开始具有抖动现象,而 f 的波形一直都光滑.分析表 1:带噪信号经过新阈值函数处理后,$R_{s,n}$提升 2.09 倍,$E_{r,m,s}$由未处理时 31.471 6 减少到 5.283 7.结合图 3 和表 1 可以得出使用新的可调节阈值函数去噪质量优于图 3 中的其它方法.

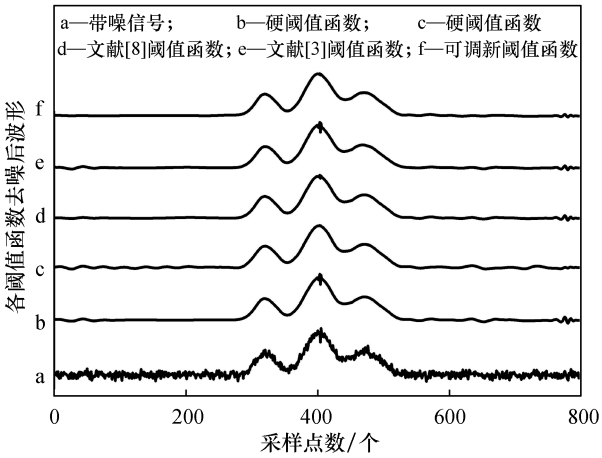

图 3 模拟验证信号去噪

Fig.3 Simulated verification signal denoising

表 1 阈值函数去噪对比

Table 1 Threshold function denoising contrast

阈值函数	$R_{s,n}$/dB	$E_{r,m,s}$	分解层	最佳 α
含噪信号	14.313 0	31.471 6	-	-
硬阈值	25.141 8	9.046 5	7	-
软阈值	25.489 7	8.691 3	4	-
文献[8]	28.160 1	6.391 0	6	0
文献[3]	26.242 7	7.969 5	6	0.2
新阈值	29.812 6	5.283 7	7	2.7

为了验证引用的阈值规则去噪效果,使用 db7 小波基和提出的新调节阈值函数进行去噪.去噪后评价指标如表 2 所示:可以得出本文引用的阈值规则是最优的.

表 2 阈值规则去噪对比

Table 2 Threshold rules denoising contrast

阈值规则	$R_{s,n}$/dB	$E_{r,m,s}$	分解层	最佳 α
含噪信号	14.313	31.471 6	-	-
固定阈值	25.473 4	8.707 6	4	-
文献[5]	27.779 8	6.677	6	1.2
文献[6]	26.017 9	8.178 5	6	0.3
本文	29.812 6	5.283 7	7	2.7

1.5 小波基和分解层的确定

根据小波变换的原理,为了获得最优去噪质量,选取小波基函数时考虑其对称特点、紧支撑能力、正交性等原则[11],对比 dbN、symN 小波基,使用新阈值函数和引用的最优阈值规则,在不同分解层下,对模拟的高斯峰数据进行去噪,根据去噪后信噪比的评价指标,选择最佳的小波基和最合适的分解层数.

不同小波基以及分解层去噪后 $R_{s,n}$对比如图 4 所示.在 db 小波族中,db7 去噪后的 $R_{s,n}$相比其它 dbN 小波基高;当小波基达到最佳分解层后,若继续对信号进行分解,分解后的信噪比会下降,说明分解层不是多越好.在 sym 小波族中,sym9 去噪后信号的 $R_{s,n}$明显高于 symN 的其它小波基.根据表 3 统计 db7 和 sym9 部分分解层去噪后 $R_{s,n}$指标,使用小波基 db7 对波峰属于高斯峰的带噪信号 7 层分解处理后的质量最好.

表 3 db7 和 sym9 部分分解层处理后的信噪比

Table 3 $R_{s,n}$ after processing of db7 and sym9 partial decomposition layers

小波基	6	7	8	9	10
db7	29.714	29.813	29.707	29.604	29.592
sym9	29.127	28.345	28.650	28.637	28.661

2 仿真实验与分析

微流控电泳仪检测出标记荧光蛋白信号满足高斯分布,检测出信号的波峰属于高斯峰,用 matlab 软件模拟出高斯峰进行分析.微流控电泳仪检测标记荧光蛋白信号时,需要进行样品进样、

样品分离、样品检测等操作.在这一系列操作过程中,微流控电泳仪信号采集系统将在检测点处采集大量信号,为了提高系统资源利用率,将在标记荧光蛋白信号的前后分别加上起始峰和结束峰,系统检测到起始峰后,对信号进行进一步处理,直到遇到结束峰时,标志荧光蛋白信号采集完毕.

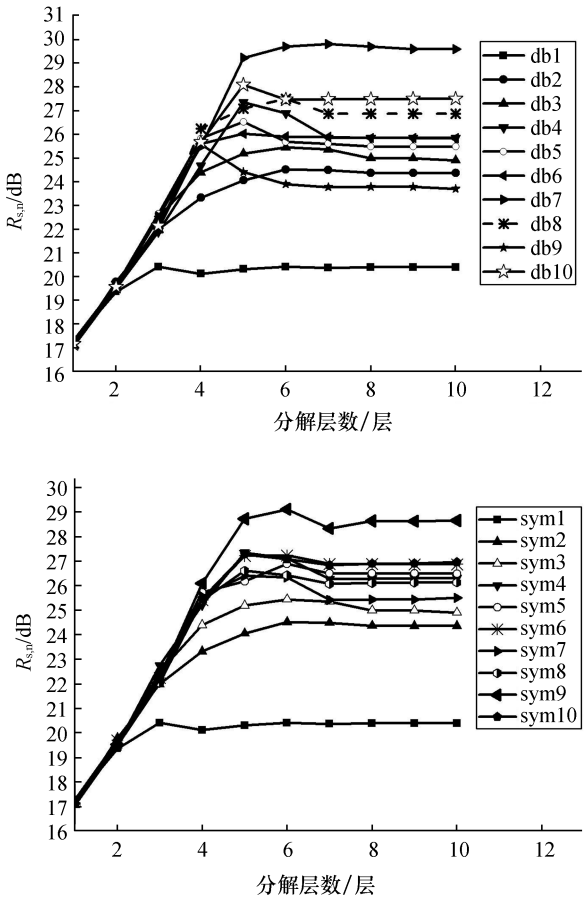

图 4　不同小波基处理后的信噪比

Fig.4　$R_{s,n}$ ratio after different wavelet bases processing

使用 matlab 仿真出重叠峰严重的数据,数据中有 7 个高斯峰,如图 5 所示.第 1 个单独的高而窄的峰为起始峰,第 2、3、4、5、6 个峰为模拟的荧光蛋白信号高斯峰,其中第 2、6 峰比较微弱,第 3、4、5 峰重叠比较多,第 7 个单独矮而宽的峰为结束峰.图 5 中 a 为添加−30 dB 高斯白噪声信号,由于噪声干扰,在带噪信号中只能明显看到 5 个峰.从图 5 中可以看出,去噪后 b、d、e 的波形在某些点处不光滑.对比 f 和 c 方法处理后的波形,可以发现 c 的信号波峰没有 f 高,说明经过软阈值公式处理造成信号发生了偏差.波形 f 下的虚线为信号的基底,经过和基底信号对比发现,新阈值去噪后信号的 7 个高斯峰都还存在.在表 4 的去噪评价指标中,本文提出的可调新阈值函数处理后信号的 $R_{s,n}$是带噪信号的 1.88 倍,$E_{r,m,s}$是处理前的 0.178 倍,经过可调新阈值函数处理后信号的 $R_{s,n}$相比文献[3]提升了 1.956 dB.经过图 5 和表 4 分析得出,新阈值函数在处理重叠高斯峰信号时去噪效果最佳.

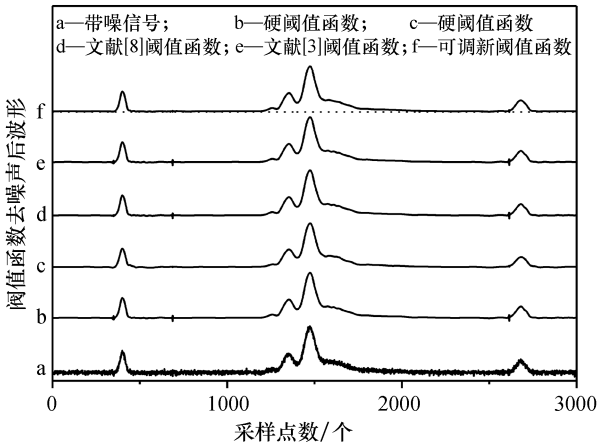

图 5　模拟数据添加−30 dB 噪声的最佳去噪对比

Fig.5　The best de-noising comparison for adding −30 dB noise to analog data

表 4　模拟加噪数据各方法最佳评判指标

Table 4　The best judgement index of each method of simulated noise data

阈值函数	$R_{s,n}$/dB	$E_{r,m,s}$
噪声信号	17.088	31.765
硬阈值	29.965	7.213
软阈值	26.773	10.417
文献[8]	30.445	6.825
文献[3]	30.159	7.053
新阈值	32.115	5.631

3　结　论

对微流控电泳仪采集到标记荧光蛋白信号进行去噪,有利于对检测样品中蛋白定性定量的进一步分析.为了提高对重叠峰严重的标记荧光蛋白信号去噪质量,提出具备可调因子的新阈值函数,弥补了硬、软阈值函数本身的不足.选择 db7 小波基,经过 matlab 实验仿真,在处理重叠严重的模拟高斯峰信号时,新阈值函数去噪后信噪比和均方根误差指标值均高于文献[3]中阈值处理函数和其它阈值处理函数.证实了可调新阈值函

数在处理带有重叠严重的高斯峰信号时,去噪质量最好.

参考文献:

[1] 王玉田,杨哲,侯培国,等.基于提升小波变换的矿物油荧光光谱去噪研究[J].光谱学与光谱分析,2016,36(7):2144-2147.

[2] ZHANG Q,WANG L,GAO P Y,et al.An innovative wavelet threshold denoising method for environmental drift of fiber optic gyro[J].Mathematical problems in engineering,2016,2016(3):1-8.

[3] XIE X Y,LIU Y,LIN X M,et al.The application of adaptive step threshold method in fluorescence signal detection[C]//IEEE International Conference on Computational Intelligence and Applications.Beijing,China:IEEE,2017:411-416.

[4] 周超,刘宁武,何天博,等.小波阈值去噪技术在呼出气体激光光谱诊断中的应用研究[J].中国激光,2017,44(11):263-271.

[5] 郑敏敏,高小榕,谢海鹤.心电信号小波去噪的改进算法研究[J].中国生物医学工程学报,2017,36(1):114-118.

[6] LU J Y,HONG L,DONG Y,et al.A new wavelet threshold function and denoising application[J].Mathematical problems in engineering,2016,2016(3):1-8.

[7] XING G,ZHANG Y,YAN Y,et al.An exponential threshold function wavelet denoising method based on gaussian distribution[C]//International conference on natural computation,fuzzy systems and knowledge discovery.Changsha,China:IEEE,2016:408-413.

[8] 江虹,苏阳.一种改进的小波阈值函数去噪方法[J].激光与红外,2016,46(1):119-122.

[9] ZHANG J,LIN J L,LI X L,et al.ECG signals denoising method based on improved wavelet threshold algorithm [C]//Advanced Information Management,Communicates,Electronic and Automation Control Conference.Xi'an,China:IEEE,2017:1779-1784.

[10] 陶珂,朱建军.小波去噪质量评价方法的对比研究[J].大地测量与地球动力学,2012,32(2):128-133.

[11] 吴莉莉,贾树恒,邢玉清,等.基于小波变换的昆虫刺吸电位(EPG)信号去噪研究[J].传感技术学报,2017,30(12):1895-1899.

(责任编辑:周泉)