PIBB 生物化学与生物物理进展
Progress in Biochemistry and Biophysics
2016, 43(10): 997~1003
www.pibb.ac.cn

Application of Wavelet Transform Algorithm and Rolling Ball Algorithm in Single Molecule Fluorescence Resonance Energy Transfer Images*

ZHANG Xiang[1,2)**,***], LI Min[2)**], JIA Ce[2)]

([1)] *College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan* 430074, China;
[2)] *Institute of Biophysics, Chinese Academy of Sciences, Beijing* 100101, China)

Abstract Single molecule fluorescence resonance energy transfer (smFRET) is a technique used to study the conformational change of a molecule by detecting the transfer efficiency of fluorescence energy between donor and acceptor fluorescence inside a single molecule. However, to obtain the information of these biological macromolecules, the statistical analysis of a large number of single molecular signals is required. Manual analysis is time-consuming and laborious and lacks objectivity and reproducibility. Therefore, the wavelet transform and rolling ball algorithms are applied to smFRET images to analyze single molecule signals. Based on the accurate detection of single molecule signals, we analyzed the linearity of the images processed by the rolling ball and wavelet transform algorithms. The results show that the two methods can not only remove the background noise of smFRET images but also maintain the linearity of single molecule fluorescence signals. At the end of this paper, we draw a statistics histogram of 15 bp DNA FRET efficiency and calculated the FRET efficiency value using rolling ball algorithm processing.

Key words single molecule fluorescence resonance energy transfer, wavelet transform algorithm, rolling ball algorithm
DOI: 10.16476/j.pibb.2016.0218

Single molecule fluorescence resonance energy transfer (smFRET) is a technique used to study the conformational change of a molecule by detecting the transfer efficiency of fluorescence energy between donor and acceptor fluorescence inside a single molecule. By marking the dye on the biological molecules to be studied and then fixing the biological molecules to a glass or quartz surface, the single molecule fluorescence image can be captured by total internal reflection fluorescence microscopy (TIRFM). Through the analysis of single-molecule fluorescence images, we can gather information about the folding and conformational changes of proteins[1], RNA folding and catalysis[2-3], the motion and signal transduction of membrane fusion proteins and actins[4-6], *etc*.

However, to obtain the information of these biological macromolecules, the statistical analysis of a large number of single molecular signals is required. Manual analysis is time-consuming and laborious and lacks objectivity and reproducibility. Therefore, it is necessary to apply automatic image processing algorithms to analyze the single molecule signals. We are apt to control the concentration of biological molecules in a sample to prevent its density from being too high so that many algorithms can detect these single-molecular signals accurately. However, single-molecule signal is comparatively weak and easily susceptible to background noise. Accordingly, what we need is an algorithm to remove background noise as well as maintain the linearity of the fluorescence intensity of a single-molecule signal. A traditional way used for background removal is to select a background region in the image and then use the region's average intensity values as background

*This work was supported by a grant from The National Natural Science Foundation of China (31300701).
**These authors contributed equally to this work.
***Corresponding author.
Tel: 86-10-64888523, E-mail: zhangxiang@ibp.ac.cn
Received: July 6, 2016 Accepted: September 30, 2016

value. The fluorescence intensity of a single molecule of an image is corrected by subtracting the background value. Although the method is simple, there are some problems: first, under the influence of subjective factors, the selection of different background regions leads to different results; secondly, the background value of the same region fluctuates at different times due to time series imaging; thirdly, it may result in incorrect analytical results on the condition that a signal occurred at some time in the selected background region. Therefore, we attempt to remove background noise by applying some image processing algorithms.

The rolling-ball algorithm, based on morphology analysis, to remove background is put forward [7]. It constructs a three-dimensional contour diagram with each image pixel coordinate and gray value. Then, a small ball is used to roll along the contour, on the condition that the radius of the ball is larger than that of the signal; in this way, the background can be removed as the ball rolls up and down, and the signal is simultaneously retained. Wavelet transform method, which is based on multi-resolution analysis, was used to remove background [8]. Through applying à trous wavelet transform to the image, the image was decomposed into wavelet coefficients of different frequency. Then, we calculated these wavelet coefficients to remove background noise and reconstruct the image with these wavelet coefficients. Together, the rolling ball and wavelet transform algorithms remove background related to morphology and multi-resolution analysis, respectively; these algorithms are used in many articles to remove the background noise and extract useful information from biological fluorescence images [9–12]. However, when processing smFRET images, we pay attention not only to the detection accuracy of single molecule signals but also to the linearity of the single molecule signals after image processing because there would be errors in smFRET image quantitative analysis if the linearity is damaged. In this paper, based on the accurate detection of single molecule signals, we analyzed the linearity of the images processed by the rolling ball and wavelet transform algorithms. The results show that the two methods can not only remove the background noise of smFRET images but also maintain the linearity of single molecule fluorescence signals. Therefore, both of these methods can be applied to single molecule signal detection and quantitative analysis of smFRET images.

1 Methods

1.1 Description of the rolling ball algorithm

The rolling ball algorithm can be achieved by the morphological-opening operator. Let $B1$, $B2$, …, Bk Bk (i, j) is the gray level of the pixel at coordinate location (i, j) in the kth neighborhood element, where i and j equal –1, 0, 1, and varies over a 3×3 window. The original image is denoted by $I(x, y)$. The image after processing is denoted by $F(x, y)$. Then, $F(x, y)$ can be calculated as follows[7]:

$$C(x, y)=\min_{i, j}[I(x-i, y-j)-Bk(-i, -j)] \quad (1)$$

$$F(x, y)=\max_{i, j}[C(x-i, y-j)+Bk(i, j)] \quad (2)$$

1.2 Description of the à trous wavelet transform algorithm

The original image is denoted by $I_0(x, y)$. We use the kernel [1/16, 1/4, 3/8, 1/4, 1/16] convolved with $I_0(x, y)$, first, row by row and then column by column. The convolution gives a smooth image $I_1(x, y)$. The first transform coefficient $W_1(x, y)$ can be obtained by the difference between $I_0(x, y)$ and $I_1(x, y)$: $W_1(x, y)=I_0(x, y)-I_1(x, y)$. For ith wavelet transform, $I_i(x, y)$ can be calculated by the convolution of $I_{i-1}(x, y)$ with the kernel, where $0<i\leqslant J$. At the scale i, the kernel can be obtained by inserting $2^{i-1}-1$ zeros between two gaps. At the final level J, we can obtain a set of $J+1$ images. They are W_1, W_2, … W_j and I_J. Thus, the wavelet transform can be calculated as follows[8]:

$$W_i(x, y)=I_{i-1}(x, y)-I_i(x, y),\ 0<j\leqslant J \quad (3)$$

$$F(x, y)=\sum_{i=2}^{J} W_i(x, y) \quad (4)$$

where $F(x, y)$ is the image after processing, and J is the decomposition level.

1.3 DNA preparation

We used dsDNA FRET standards labeled with a donor fluorophore, Sulfo-Cyanine3 (Cy3) NHS ester (Lumiprobe, America), and an acceptor fluorophore, Sulfo-Cyanine5 (Cy5) NHS ester (Lumiprobe, America). The oligonucleotides, 5′-(NH_2 C6) CATG-ACCATGACCAG -3′ and 5′-(NH_2 C6) CTGGTC-ATGGTCATG (Biotin)-3′ were synthesized by Takara Biotechnology (Dalian). Cy3- and Cy5-labeled ssDNAs were hybridized by heating at 75℃, followed by passive cooling to room temperature.

1.4 smFRET data acquisition

The sample coverslip was coated by APTES (Sigma, America) and then the amine-modified

surfaces were coated with biotin-PEG (Laysan Bio, America) and mPEG (Laysan Bio, America). dsDNA (15 bp) with one strand modified with biotin at one end was immobilized on the coverslips through biotin-streptavidin (Invitrogen, America) interaction. PEG-coated coverslips were used within a few days. smFRET images were acquired using a TIRFM imaging system that was constructed based on the prismless and through-the-lens configuration as previously described [13]. The TIRFM imaging system was equipped with a 150 × 1.45 NA oil-immersion objective (Olympus, Japan), an EMCCD camera (Andor iXon 897, England), and lasers (Coherent, America). The EMCCD camera was driven by Andor Solis (Andor, England). The 532 nm laser was used to excite fluorescence during the experiment.

2 Results

An example smFRET fluorescence image is shown in Figure 1. The smFRET image consists of two channels: one is the donor channel, and the other is the acceptor channel. In our smFRET experiment, the Cy5-Cy3 fluorescence dye was used to mark DNA; Cy5 is the donor channel, and Cy3 is the acceptor channel. The shape of a single molecule signal in the image is a diffraction spot. The full width at half maximum of the spot is about hundreds of nanometer, which is due to the diffraction limit of the optical imaging system.

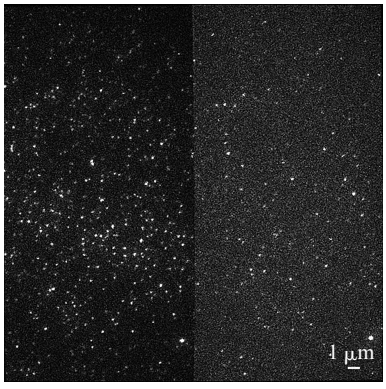

Fig. 1 An example smFRET image of DNA labeled with Cy5-Cy3 dye

In the figure, the left half part is donor channel, and the right half part is acceptor channel.

The single molecule signal was located by Gaussian fitting of the diffraction spot. Meanwhile, the single molecule signal was influenced by various noise, including background noises, thermal noise, quantization noise and shot noise, *etc*. Therefore, we needed algorithms to process the image and remove the noise, while maintaining the linearity of the fluorescence signal.

The rolling ball algorithm based on morphology and the wavelet transform algorithm based on multi-resolution analysis were applied to the signal detection in the smFRET images. Figure 2a shows the image processing by the wavelet transform algorithm, and Figure 2b shows the image processing by the rolling ball algorithm. From Figure 2, we can see that the background noise in the original image was removed cleanly and that the single molecule signal was highlighted, which provided convenience for the following threshold segmentation. The single molecule signal can be detected easily by an adaptive threshold algorithm[14]. If the two single molecule signals are too close to be distinguished, a watershed algorithm[15] can be used to separate them. Finally, Gaussian fitting was used to locate the center of the single molecule signal. The whole flow chart of image processing is shown in Figure 3.

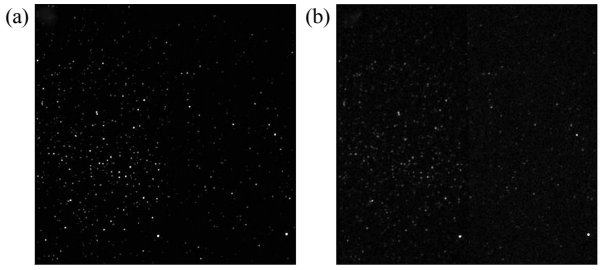

Fig. 2 Denoising images processing by wavelet transform algorithm and rolling ball algorithm

(a) Denoising image processing by the wavelet transform algorithm; (b) Denoising image processing by the rolling ball algorithm.

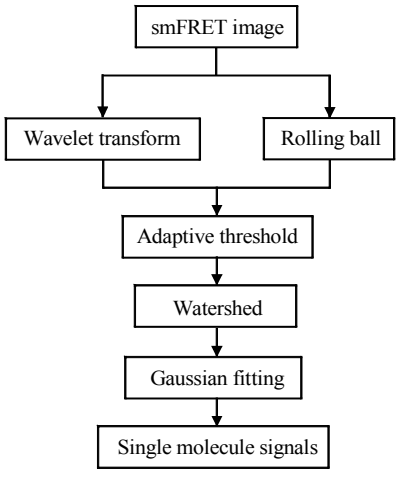

Fig. 3 Flow chart of smFRET image processing

We wanted to perform quantitative analysis of the smFRET signal, but we were concerned that the single molecule signal may not maintain linearity. For this reason, we constructed a series of images to simulate smFRET images. The simulated images of 512 pixels× 512 pixels size have two channels with different levels of background in each channel. 100 different intensities of single molecule signals were added to each channel. The intensity of the single molecule signal was stimulated with Gaussian distribution.

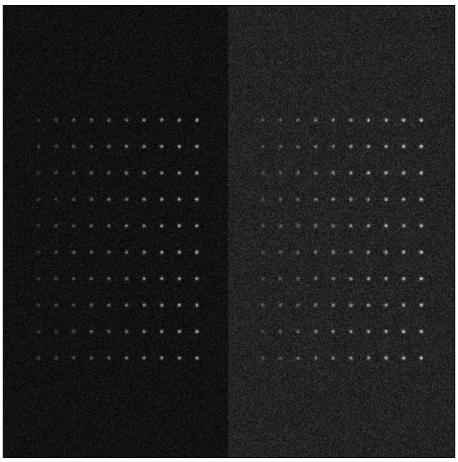

Fig. 4 An example of simulated image with Gaussian noise and Poisson noise

Meanwhile, we added different levels of Gaussian noise and Poisson noise to the images to evaluate the linearity of the rolling ball and wavelet transform algorithms. The simulated image is shown in Figure 4.

We used DIPimage to generate the Gaussian noise and Poisson noise. DIPimage is a free image processing and analysis toolbox, which can be download from the website http://www.diplib.org/. Gaussian noise and Poisson noise can be generated by input parameters to the noise function. The parameter of Gaussian noise is denoted as P_g. The greater the Pg is, the greater the noise level is. The parameter of the Poisson noise is denoted as P_p. The smaller the P_p is, the greater the noise level is.

As shown in Figure 5, we analyzed the linearity of single molecule signals in donor and acceptor channels after processing by the wavelet transform algorithm, where the horizontal coordinate is real signal intensity and the vertical coordinate is the signal intensity after processing. The data of both coordinates have been normalized and were carried out with linear fitting. Similarly, linearity analysis of the single molecule signal after processing by the rolling ball algorithm is shown in Figure 6.

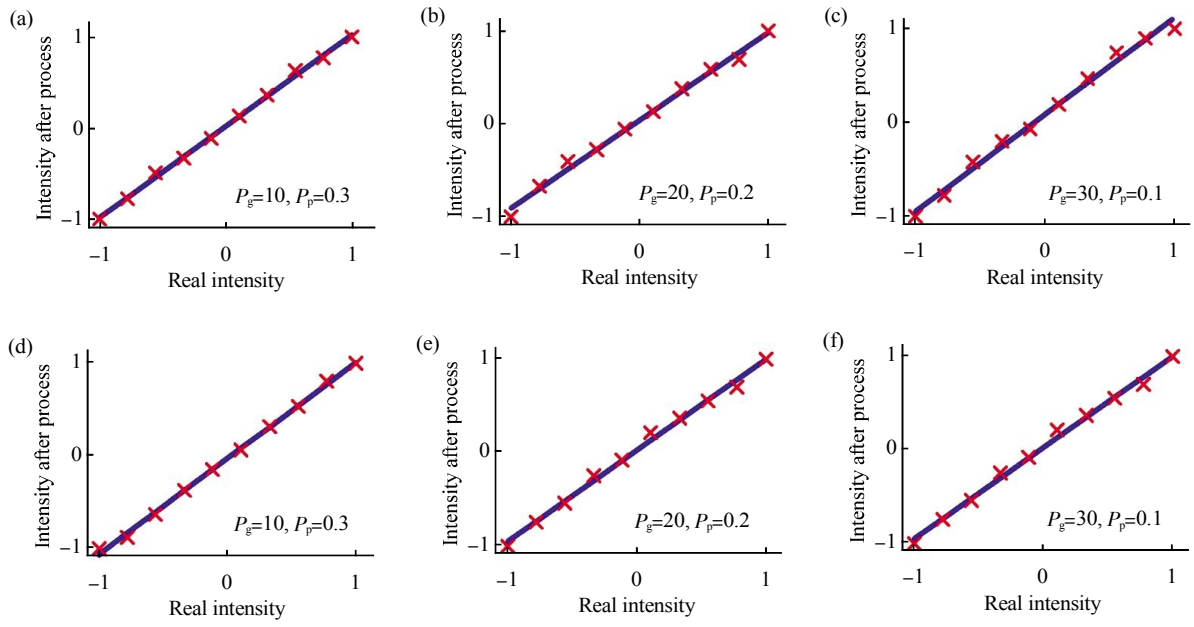

Fig. 5 Linearity analysis of single molecule signals after wavelet transform processing

(a–c) Linearity analysis of donor channel. The noise level increases successively from (a) to (c). (d–f) Linearity analysis of acceptor channel. The noise level increases successively from (d) to (f). P_g is the parameter of Gaussian noise, and P_p is the parameter of Poisson noise.

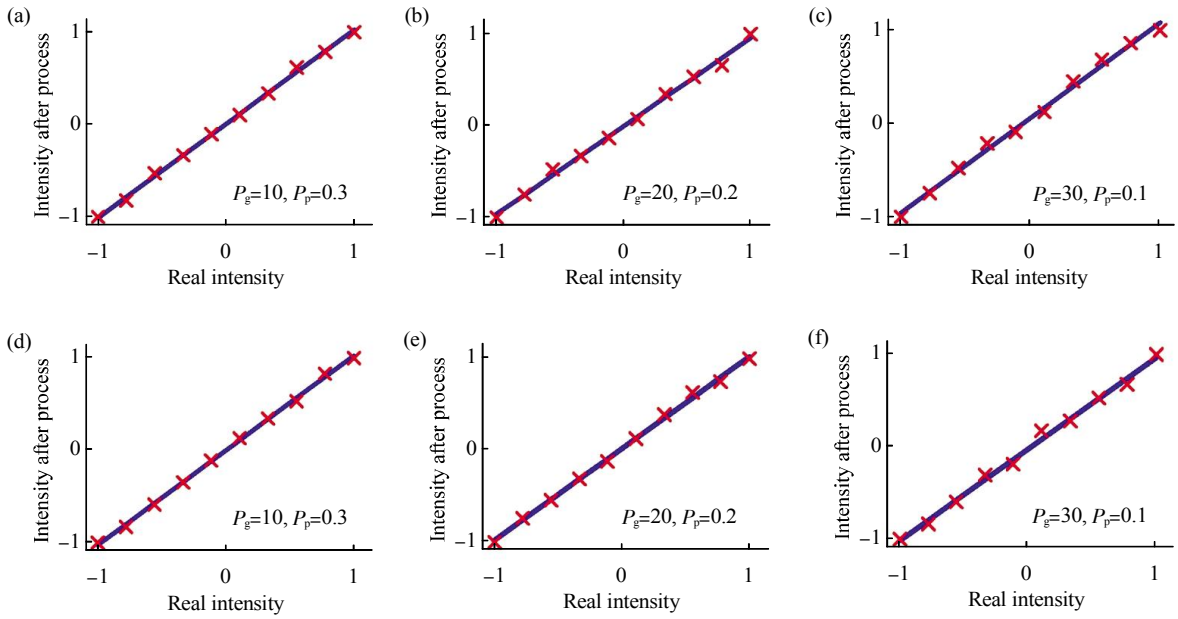

Fig. 6　Linearity analysis of single molecule signals after rolling ball processing

(a–c) Linearity analysis of donor channel. The noise level increases successively from (a) to (c). (d–f) Linearity analysis of acceptor channel. The noise level increases successively from (d) to (f). P_g is the parameter of Gaussian noise, and P_p is the parameter of Poisson noise.

From Figure 5 and Figure 6, we can see that both algorithms have good linearity under different levels of noise. Meanwhile, we calculated their linear correlation coefficient of before and after processing, as shown in Figure 7.

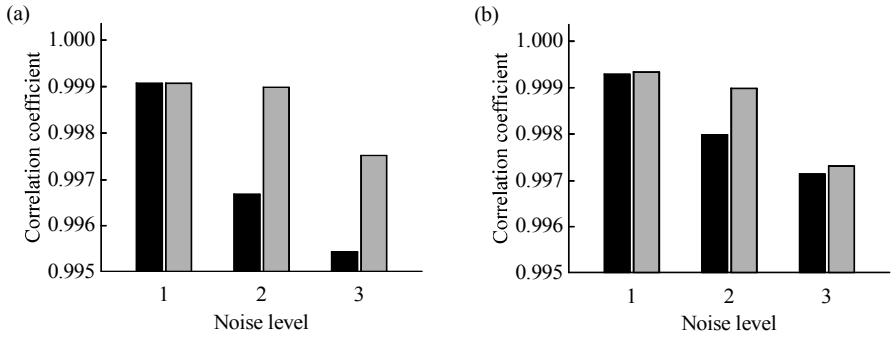

Fig. 7　Linear correlation coefficient of before and after processing

(a) Linear correlation coefficient of before and after wavelet transform algorithm processing. (b) Linear correlation coefficient of before and after rolling ball algorithm processing. Horizontal coordinates of 1, 2, 3 represents the gradual increase in the level of noise, the corresponding parameters are as follows: P_g=10, P_p=0.3; P_g=20, P_p=0.2; P_g=30, P_p=0.1. ■: Donor; □: Acceptor.

From Figure 7, we can see that the linearity of both algorithms reached more than 0.995. The rolling ball algorithm is slightly better than the wavelet transform algorithm. When the linear coefficient is equal to 1, it means the linear correlation is complete. The closer the linear coefficient is to 1, the better the linear correlation is. Therefore, the wavelet transform algorithm and the rolling ball algorithm have good linearity.

We used the rolling ball algorithm to remove the background noise and analyzed the DNA molecule labeled with Cy5-Cy3 fluorescence dye as shown in Figure 8.

FRET efficiency can be calculated with the following formula[16]:

$$E=\frac{I_A}{(I_D+\gamma\times I_A)} \tag{6}$$

where γ is denoted as the relative detection

efficiency and quantum yield of two dyes, which can be determined by photo bleaching. We used Cy3- Cy5 to obtain approximately 1.

As shown in Figure 8, we calculated the statistical histogram of 15 bp DNA FRET efficiency. By Gaussian fitting, we obtained the result that 15 bp DNA FRET efficiency was approximately 0.634. The distance of donor and acceptor can be calculated by $R=R_0\sqrt[6]{1/E-1}$, where R_0 was Foster radius [17], which is about 6 nm for Cy3 and Cy5 [18]. The calculated distance between these dyes was about 5.47 nm. The length of 15 bp DNA was about 5.1 nm. The discrepancy was 0.37 nm, which may be due to the orientation of the labeled dyes[17].

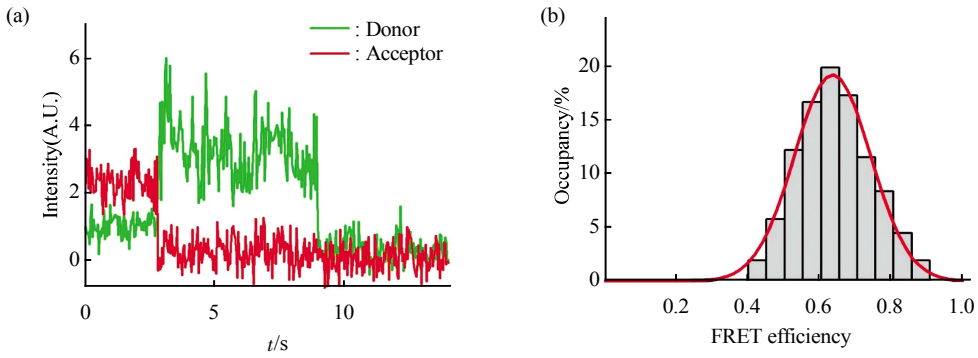

Fig. 8 FRET signal statistical analysis of DNA molecule labeled with Cy5-Cy3 fluorescence dye

(a) Intensity *vs*. time curve of donor and acceptor fluorescence signal. (b) FRET efficiency statistical histogram of 15 bp DNA. Red curve is the Gaussian fitting for the histogram.

3 Discussion

The application of total internal reflection fluorescence microscopy to study smFRET has made great progress in the biological field, which improves the understanding of some biological processes. However, manual analysis is time-consuming and laborious and lacks objectivity and reproducibility. In this paper, we presented a method to automatically analyze smFRET images. In this method, the wavelet transform and rolling algorithms are applied in smFRET image analysis to remove background noise.

To verify the effectiveness of the wavelet transform and rolling ball algorithms in smFRET image processing, we constructed a series of smFRET images with different levels of noise and added different intensities of single molecule signals to the image. Through analysis of the linearity of single molecule signals after wavelet transform and rolling ball algorithm processing, we found that both algorithms have good linearity under different levels of noise. Additionally, we calculated the linear correlation coefficient of before and after processing. The linearity of both algorithms reached more than 0.995, with the rolling ball algorithm being slightly better than the wavelet transform algorithm. Because the concentration of the sample can be controlled to make the single molecule signals non-overlapping in the experiment, the detection accuracy of these algorithms can achieve 100%. The computation time for processing one image by wavelet transform algorithm and rolling ball algorithm were about 4.52 s and 4.66 s separately. Moreover, we calculated the statistical histogram of 15 bp DNA FRET efficiency. By Gaussian fitting, we obtained the result that 15 bp DNA FRET efficiency was approximately 0.634 and the corresponding distance was about 5.47 nm. The discrepancy between the real distance and calculated distance was 0.37 nm, which may be due to the orientation of the labeled dyes. In general, these two algorithms can be applied to the single molecule FRET images to improve processing efficiency.

smFRET technology gives people new insight into biological problems, including the folding and conformational changes of proteins, RNA folding and catalysis, the motion and signal transduction of membrane fusion proteins and actins, *etc*. With the development of the fluorescence probes and data analysis algorithms, this field will develop by leaps and bounds.

References

[1] Haas E. The study of protein folding and dynamics by determination of intramolecular distance distributions and their fluctuations using ensemble and single-molecule FRET measurements. Chem Phys Chem, 2005, **6**(5): 858−870

[2] Xie Z, Srividya N, Sosnick T R, *et al*. Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. Proc Natl Acad Sci USA, 2004, **101**(2): 534−539

[3] Gell C, Sabir T, Westwood J, *et al*. Single-molecule fluorescence resonance energy transfer assays reveal heterogeneous folding ensembles in a simple RNA stem-loop. Journal of Molecular Biology, 2008, **384**(1): 264−278

[4] Sivaramakrishnan S, Ashley E A, Leinwand L A, *et al*. Insights into human beta-cardiac myosin function from single molecule and single cell studies. Journal of Cardiovascular Translational Research, 2009, **2**(4): 426−440

[5] Choi U B, Strop P, Vrljic M, *et al*. Single-molecule FRET-derived model of the synaptotagmin 1-SNARE fusion complex. Nature Structural & Molecular Biology, 2010, **17**(3): 318−324

[6] Xu X B J, Jin T. Monitoring dynamic GPCR signaling events using fluorescence microscopy, FRET imaging, and single-molecule imaging. Methods Mol Biol, 2009, **571**: 371−383

[7] Sternberg. Biomedical image processing. IEEE Computer, 1983, **16**(1): 22−34

[8] Olivomarin J. Extraction of spots in biological images using multiscale products. Pattern Recognition, 2002, **35**(9):1989−1996

[9] Brandenburg B, Zhuang X. Virus trafficking - learning from single-virus tracking. Nature Reviews Microbiology, 2007, **5** (3): 197−208

[10] Loerke D, Mettlen M, Yarar D, *et al*. Cargo and dynamin regulate clathrin-coated Pit maturation. PLOS Biology, 2009, **7** (3): 628−639

[11] Arhel N, Genovesio A, Kim K, *et al*. Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. Nature Methods, 2006, **3**(10): 817−824

[12] Wienisch M, Klingauf J. Vesicular proteins exocytosed and subsequently retrieved by compensatory endocytosis are nonidentical. Nature Neuroscience, 2016, **9**(8): 1019−1027

[13] Li C, Bai L, Li D, *et al*. Dynamic tracking and mobility analysis of single GLUT4 storage vesicle in live 3T3-L1 cells. Cell Research, 2004, **14**(6): 480−486

[14] Otsu N. A threshold selection method from gray level histograms. 1979, **9**(1): 62−66

[15] Meyer F. Topographic distance and watershed lines. Signal Processing, 1994, **38**(1): 113−125

[16] Ha T, Ting A Y, Liang J, *et al*. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. Proc Natl Acad Sci USA, 1999, **96**(3): 893−898

[17] Zhang H, Browne M. Construction of a laser combiner for dual fluorescent single molecule imaging of pRNA of phi29 DNA packaging motor. Biomedical Microdevices, 2010, **12**(1): 97−105

[18] Joo C, Ha T. Preparing sample chambers for single-molecule FRET. CSH Protocols, 2012, **10**: 1104−1108

小波变换及滚球算法在单分子荧光能量共振转移图像分析中的应用 *

张 翔[1,2)**,***] 李 敏[2)**] 贾 策[2)]

([1)] 华中科技大学生命科学与技术学院，武汉 430074；[2)] 中国科学院生物物理研究所，北京 100101)

摘要 单分子荧光共振能量转移技术是通过检测单个分子内的荧光供体及受体间荧光能量转移的效率来研究分子构象的变化．要得到这些生物大分子的信息就需要对大量的单分子信号进行统计分析，人工分析这些信息，既费时费力又不具备客观性和可重复性，因此本文将小波变换及滚球算法应用到单分子荧光能量共振转移图像中对单分子信号进行统计分析．在保证准确检测到单分子信号的前提下，文章对滚球算法和小波变换算法处理图像后的线性进行了分析，结果表明，滚球算法和小波变换算法不但能够很好地去除单分子 FRET 图像的背景噪声，同时还能很好地保持单分子荧光信号的线性．最后本文还利用滚球算法处理单分子 FRET 图像及统计 15 bp DNA 的 FRET 效率的直方图，通过计算得到了 15 bp DNA 的 FRET 效率值．

关键词 单分子荧光能量共振转移，小波变换算法，滚球算法

学科分类号 Q334，Q336 **DOI**: 10.16476/j.pibb.2016.0218

* 国家自然科学基金资助项目(31300701).

** 共同第一作者. *** 通讯联系人.

Tel: 010-64888523, E-mail: zhangxiang@ibp.ac.cn

收稿日期：2016-07-06，接受日期：2016-09-30