The local-threshold 2D-tophat Cell Segmentation for the Two-photon Confocal Microscope Image

Xiaoyang Yuan, Lei Gu, Lei Sun, and Takeshi Ikenaga
The Graduate School of Information,Production and Systems
Waseda University
2-7 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135
chobitsnerv@toki.waseda.jp

Abstract

The Two-photon Confocal Microscopy (TCM) is a new technology which is useful in nondestructive analysis of tissue. However its cell experiment images are excessive and difficult for machine processing since the low resolution and lots of noise. Thus nowadays such a work is mainly done manually and costs a lot of time and effort. This paper proposes a cell segmentation algorithm for this situation by the local-threshold 2D-tophat which is based on the tophat combined with the OTSU and Mathematic Morphology. The proposed method uses the cell region gradient information to do the cell segmentation which uses a local-dynamic threshold instead of a static one and uses the 2D information only to minimize the influences caused by the background and also the algorithm complexity. This method has been applied to experiment and is proved that the true-positive ratio (TPR) can be kept above 85% and false-positive ratio (FPR)+miss-positive ratio(MPR) can be kept under 20% when compared with the original tophat method that the TPR is about 80% and FPR+MPR above 20%.

1. Introduction

Automatic analysis of biomedical images becomes more and more important since the experiment data increase sharply. In TCM, a new technology which make it possible to observe the living cell tissue without destruction, the observed objects are labeled by fluorescent proteins and appear in the experiment image as bright circular areas, the radius of each is 20 pixels or so. Figure 1 is the experiment sample images of mouse cerebral cells. Image segmentation aims to partition the image into regions homogenous with respect to certain features and hopefully correspond to real objects in the actual scene [1]. Using machine to do the digital analysis can help the biological research support and substantiate its corresponding assumed theory. However the experiment result data gained by TCM is just grayscale image and the image quality is also not adequate because of the low resolution. At the same time, not like other equipment which just deal with the prepared slice of cell tissue sample, since the tissue is alive the thickness of observed sample cannot be too thin. Thus besides the observed cells there still exist lots of interference from background patterns. These influential background patterns may have the same gray value or similar shape which leads it impossible for traditional methods to do the cell segmentation correctly and efficiently. Thus nowadays such a kind of statics works is still done manually with some specialized tools such as ImageJ. Although the ImageJ is a powerful biological image processing tools, obtaining the whole cell information for once experiment maybe cost several days by several people since both of the cell amount in each image and image amount are not small. If the TPR of processing result obatined by compurter can reach up to 90% and the FPR+MPR can decrease down to 10%, it will improve the efficiency alomost 10 times than doing the work only by human according to the evaluation of biological researchers..

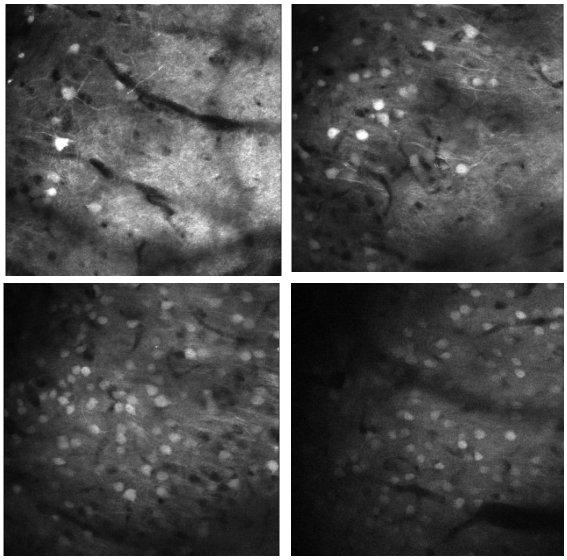

Figure 1. Sample Image of mouse cerebral cell labeled by fluorescent protein imaged by Two-photon Confocal Microscopy (TCM)

In this paper, the local-threshold 2D-tophat algorithm, based on tophat with OTSU and Mathematic Morphology and, is introduced to solve such a kind of troublesome problem. Firstly, three kinds of Mathematic Morphology operation, close operation, tophat operation and dilation operation, are applied to reduce the noise and enhance the cell regions. Secondly a local gradient threshold is calculated based on every line of pixels with OTSU threshold after the first step processing. Then the gradient threshold is applied into the segmentation method which will be introduced in detailed in the third section of this paper. After this step, a preparative segmentation result can be obtained. At last the preparative result will be processed in order to remove the superfluous results and gain the final segmentation result.

2. Related Research Work

The cell segmentation is not an old problem since the development of computer vision makes it possible to observe the live cell quantitatively, which paves the way toward mathematical modeling of protein kinetics and biochemical signaling networks [1].

Such a problem can be tackled by some existing algorithms. For example, Voting Based Technique (VBT) [2], Watershed Algorithms (WA) [3], Tophat filter (TH) [4], [5], Grayscale Opening Tophat Filter (GOTH) [6], [7] and so on. All existing method can be divided into two categories. The first kind of category can be called as machine learning based. Another kind can be called as processing and threshold based. VBT and WA are treated as the first kind. TH and GOTH are treated as the second type. For the first kind of segmentation method, it suffers from lots of problem such as unreliable segmentation result and high demand for preprocessing. The VBT may gain an unreliable result because when a property receives votes during the early rounds then more votes are tend to be received in the following steps. Although WA doesn't meet such a kind of problem and show a good segmentation result, this method is deeply influenced by smoothing operation in the initialization stage. For the second kind, the cell segmentation workflow is reducing the noise at first and then enhancing the useful signals finally using a threshold to obtain the final result. However most such a kind of methods use a static threshold to do screening which leads to miss in the segmentation result. For other methods which are also using dynamic threshold such TH, the interference background patterns occurred in current situation influence the segmentation result seriously because these interference background patterns cannot be decreased by Gaussian filter or other smoothing filter easily. Thus such a kind of situation is not easy for machine to do the cell segmentation.

For other related works, the status of observed object is almost the same or better but the background is much simpler. This makes the segmentation is not necessary to concern a lot about interference from background and this is also the reason why the final TPR can reach up to 95% or much higher but under situation discussed in this paper the TPR cannot reach 80% and the FPR also above 20%.

3. Local-threshold 2D-tophat Cell Segmentation Algorithms

Since the gray gradient of cell region for each region in the image is different, it cannot gain a good cell segmentation result with a uniform threshold. The application of dynamic local threshold, which is calculated based on the OTSU threshold and gray value of each line of pixel in the image, is applied to solve such a kind of problem. At the same time using 2D information, which means just using the gray information of the x-axis and y-axis, decreases the segmentation algorithm complexity and influence from the background. The proposed cell segmentation algorithm is based on the tophat algorithm but just using the 2D information benefits a lot for this. Also during the preprocessing step, a

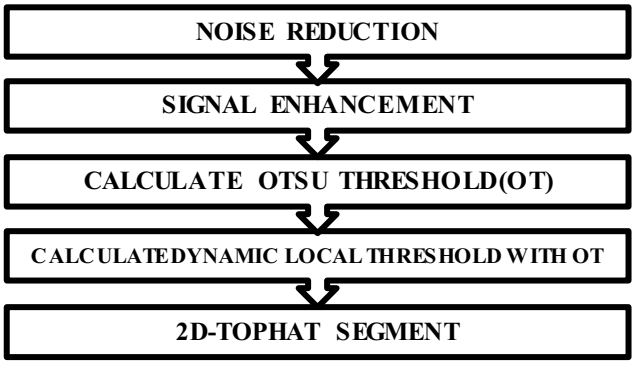

Figure 2. The operation flow of local-threshold 2D-tophat

mathematic morphology (MM) tophat operation is applied to reconstruct the image in order to decrease influence of the background. At the same time, in order to make the image become better to segment the choosing of MM open and close operation elements are also targeted. The flow of proposed method, local-threshold 2D-tophat algorithm, is shown as the figure 2.

3.1. Noise Reduction and Signal Enhancement

Just like others related works, the first step for cell segmentation is also noise reducing. Here MM close operation is applied to remove some macula in the inner of a cell region which may influence the cell segmentation result in the following segmentation steps greatly. The parameter for MM close operation is a 3*3 pixels cross structure element which is applied 3 times since the macula usually appears as vertical or horizontal line and size is around 3 pixels. After this Noise reduction, MM tophat is used to reconstruct the cell image in order to minimize the interference from background as much as possible because the cell regions are reconstructed fast than the background. Here, we use a 10*10 pixels ellipse structure element to do 2 times operation, which is the average cell radius and is initialized artificially for image samples of different cells. After this step, although the cell region can be reconstructed, it still shrinks. Thus a MM dilation operation is applied to recover the cell size. At the same time, some too small cells are also enlarged in order to detect them out. The parameter for MM dilation is 5*5 pixels eclipse structure element. Dilation operation is just done once and then the image under being segmented is obtained.

3.2. Calculate local gradient threshold of each pixel line with OTSU

OTSU is proposed in [9] which is based on discriminant analysis. With this algorithm the pixel of image is divided into two classes $C_0 = \{0,1,\dots,\mathrm{p}\}$ and $C_1 = \{p, p+1,\dots,l-1\}$ at l levels grayscale. According to the [9], σ_B^2 denotes the between-class variance. Then with

$$\sigma_B^2 = \omega_0\omega_1(\mu_1\mu_0)^2$$

an optimal threshold can be calculated by maximizing this formula. G_i denotes the probability of occurrence of grey-level i and is defined as:

$$G_i = \frac{n_i}{n} \quad n = \sum_{i=0}^{l-1} n_i$$

n is the total pixel numbers and n_i is the number of pixel in grey-level i. Also other parameters are defined as followed:

$$\mu_T = \sum_{i=0}^{l-1} iG_i$$

$$\omega_0 = \sum_{i=0}^{p} G_i \quad \omega_1 = 1 - \omega_0$$

$$\mu_1 = \frac{\mu_T - \mu_t}{1 - \mu_0} \quad \mu_0 = \frac{\mu_1}{\omega_0} \quad \mu_t = \sum_{i=0}^{t} iG_i$$

When σ_B^2 increase to the maximum, the corresponding t becomes the threshold. Then the gray value t_m, which is the average value of those that are larger than the t during the current pixel line, is used to calculate the gradient g with t. g is defined as:

$$g = \frac{tm - t}{10}$$

The 10 is the average radius which can be set manually according to the different types of cell samples.

Since the tm for different pixel line is different, variational g is applied. Thus the threshold can be treated as dynamic and local. This avoids the case that some cell region cannot be detected by just using the static threshold. After this step the gradient of cell existing areas is obtained.

3.3. Local-threshold 2D-tophat algorithm

After a series of operation explained in 3.1 and 3.2 section, the image under being segmented is obtained. Then the following algorithm is applied to segment the image in order to detect the cell exist regions.

Firstly, a median filter is used to process the image gray data matrix for x-axis and y-axis direction separately. That is to say two processed images are obtained which contain the corresponding direction processing information.

Secondly, calculate the difference between each pair of neighbor pixels for horizontal-information image and vertical-information image. Using P_{mn} denotes the grayscale value in corresponding position (m, n) of the image. This step is expressed as followed:

$$\mathbb{D}_x = \begin{pmatrix} P_{21} - P_{11} & \cdots & P_{m1} - P_{(m-1)1} \\ \vdots & \ddots & \vdots \\ P_{2n} - P_{1n} & \cdots & P_{mn} - P_{(m-1)n} \end{pmatrix}$$

$$\mathbb{D}_y = \begin{pmatrix} P'_{12} - P'_{11} & \cdots & P'_{m2} - P'_{m1} \\ \vdots & \ddots & \vdots \\ P'_{1n} - P'_{1(n-1)} & \cdots & P'_{mn} - P'_{m(n-1)} \end{pmatrix}$$

Then such two matrixes can be got which stand for the difference of x-axis and y-axis directions. After this a filter is used to remove the values which suddenly become below zero when comparing with their neighbors, which can be expressed as (D_x denotes one difference value in simple line arrange):

$$IF\ D_{x-1} > 0, D_{x+1} > 0, D_x < 0$$
$$THEN\ D_x = \frac{D_{x-1} + D_{x+1}}{2}$$

For the above matrix, the values of the $\mathbb{D}_x$ is just processed horizontally and $\mathbb{D}_y$ is the same, corresponding y-axis direction values.

After all the above processing, the difference matrixes are used to do the segment from the x-axis and y-axis direction separately. First the values of matrix are scanned by each line separately. When continuous five plus values are found the fifth value is recorded. When the next minus value is met, the gradient is calculated. If it is larger than the local-threshold g, this is recorded as a strong rising edge. The start value point and high-value point are also recorded. Then the corresponding falling edge is checked out if existed, which is constituted by at least fine continuous five minus values and the gradient with high-value point is also larger than g. These are expressed simply as followed:

$$IF\ D_x, \cdots, D_{x+4} > 0\ then\ record\ the\ D_{x+4}$$
$$D_x, \cdots, D_{x+m} > 0; D_{x+m+1} < 0 \longrightarrow \mathcal{G}_{up} = \frac{P_{x+m+1} - P_x}{m+1}$$
$$IF\ \mathcal{G}_{up} > g \Rightarrow Strong\ rising\ edge$$
$$mirrore\ dge\ of\ strong\ rising\ edge$$
$$\Leftrightarrow Strong\ falling\ edge$$

After the corresponding strong rising and falling edge are decided, the width of interval between them is calculated. If the width is larger than 20pixel, this interval information, the coordinate values of start, high and end points, are recorded. If the area is two large, it is divided averagely by each part that is at least large than 20 pixels. Using the operation flow to scan the difference data matrixes, the cell region data set is obtained.

3.4. Merge and Screen the Segmentation result

After operation described in section 3.3, the cell regions can be obtained. Then the data is labels with white color in a black image which is in the same size with the original cell image. Some lines without polymerizing area are ignored. For the lines which polymerize into

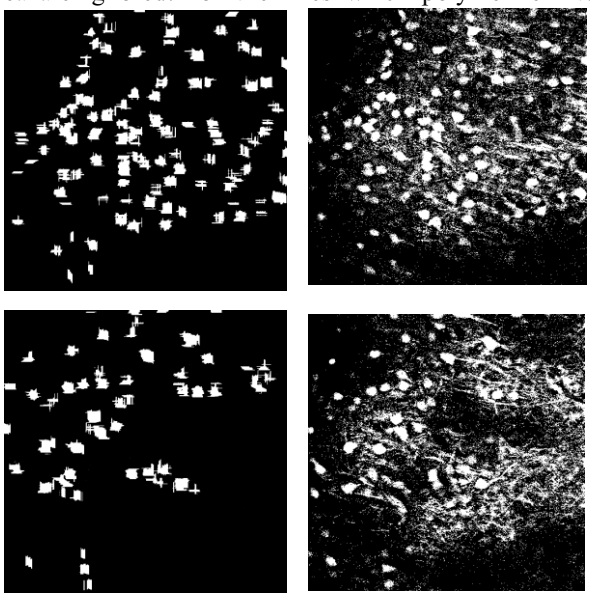

Figure 3. Segmentation simulation results comparison. The left two images are obtained by the 2D-tophat methods. The right two images are obtained by the original tophat method in [8].

areas, the gaps of area are filled. Then the centroid of each area is calculated. Then final results is obtained.

4. Experiment Result

This algorithms is simulated with the sample images which are displayed in the figure 1 of the section one and the background truth is obtained by the biology experts. The figure 3 shows some comparison samples of the final segmentation result with the tophat algorithms described in [8]. According to the images, it's found that the tophat algorithms cannot deal with the background noise effectively, which leads to influence the simulation results greatly. Although the tophat bases on the same algorithms principle as the proposed method of this paper and in [8] is also given a simulation results, but since the observed objects status of this paper isn't good enough like the one in [8] that the objects are the spots. In this paper, the observed objects are approximate circular area. At the same time, the background in [8] is not very complex. Thus this algorithm doesn't work very well under the image status of this paper which is also found in the table 1.1 and 1.2, the simulation results of 2D-tophat and original tophat method. The table 2 is corresponding true-positive ratio (TPR), false-positive ratio (FPR) and miss-positive ratio comparison between the 2D-tophat and original tophat method .

Table 1. 1 Local-threshold 2D-tophat simulation result

Sample No.	Total	Right	Wrong	Miss
1	11	11	0	0
2	45	40	4	5
3	103	92	13	9
4	78	67	8	6

Table 1. 2 The original tophat simulation result

Sample No.	Total	Right	Wrong	Miss
1	11	11	0	0
2	45	40	12	4
3	103	82	18	21
4	78	60	8	10

(Total denotes the cell number according to the background truth. Right denote the right cell number. Wrong denote the cell wrong number. Miss denote the miss cell number)

Table 2 The comparison of TPR, FPR and MPR between local-threshold 2D-tophat and Original tophat

Sample No.	LT2D			Original tophat		
	TPR	FPR	MPR	TPR	FPR	MPR
1	1	0	0	1	0	0
2	0.88	0.08	0.11	0,88	0.26	0.08
3	0.89	0.12	0.08	0.80	0.17	0.20
4	0.86	0.1	0.08	0.77	0.10	0.13

TPR, FPR and MPR are defined as:

$$TPR = Right/Total$$
$$FPR = Wrong/Total$$
$$MPR = Miss/Total$$

According to the table 2, it's observed that the local-threshold 2D-tophat method keeps the TPR above 85% and the FPR+MPR below 20%. The TPR of original tophat algorithm is just around 80% and FPR+MPR keeps above 20%. And for the four kinds of image samples, which are one image contained a few of cells, one image contained normal amount of cells, one image contains lot of cells and one image contained many blurry cell. For these representative situations, local-threshold 2D-tophat works as stably as expected

5. Conclusion

This paper shows a cell segmentation algorithm for the TCM Image. As the proposed method in this paper is based on the tophat algorithm, it can deals with the situation that the background isn't able to remove by several steps thoroughly. Since the cell local information, that the gradient of cell region is larger than others, is used to do the cell segmentation, the proposed method is also robust under different situations. In the simulation the proposed method shows above 85% TPR. At the same time the MPR + FPR is kept under 20%, which is better than the situation of original tophat method that TPR is below 80% and MPR + FPR is above 20%. Although it's not achieve the desired result (above 90% TPR, below 10%FPR+MPR), it's more practical than the other methods under the TCM image situation.

6. Acknowledgement

This work is supported by KAKENHI (23300018)

References

[1] R.Elis and C.Athale, "Computational imaging in cell biology", JCB, Vol 161,No 3, May, 2003, 477-481.

[2] Yang, Q. et al ,"Perceptual Organization of Radial Symmetries", Proceedings of CVPR, 2004

[3] Viincen t, L. and Soille, P. Vincen t, L. and Soille, P. Watershed in Digital Spaces: "An Ef_cient algorithm Based on Immersion Simulations. IEEE Transactions on pattern analysis and machine intelligence", Vol. 13, No. 6, June 1991.

[4] D. S. Bright and E. B. Steel, "Two-dimensional top hat filter for extracting spots and spheres from digital images," J. Microsc., vol. 146, no. 2, pp. 191–200, 1987

[5] E. J. Breen, G. H. Joss, and K. L.Williams, "Locating objects of interest within biological images: The top hat box filter," J. Comput.-Assist. Microsc., vol. 3, no. 2, pp. 97–102, 1991.

[6] M. Sonka, V. Hlavac, and R. Boyle, "Image Processing, Analysis, and Machine Vision", 2nd ed. Pacific Grove, CA: PWS, 1999.

[7] P. Soille, "Morphological Image Analysis: Principles and Applications. Berlin", Germany: Springer-Verlag, 2003.

[8] Smal, I. Dept. of Med. Inf., Erasmus MC, Rotterdam, Netherlands Loog, M. ; Niessen, W. ; Meijering, E.,: "Quantitative Comparison of Spot Detection Methods in Fluorescence Microscopy," Medical Imaging, vol. 29 , Issue: 2, pp. 282 – 301, Feb. 2010

[9] N. Otsu, "A threshold selection method from grey-level histogram," IEEE Trans. Syst., Man, Cybern., vol. SMC-8, 1978, pp.62-66.