

A Center Location Algorithm Used in Cells Microscopic Image

Pu Hui

School of Software,
Northwestern Polytechnical
University,
Xi'an, P.R. China
huipulolita@mail.nwpu.edu.cn

Jinqiu Sun, Danna Xue

School of Astronomy,
Northwestern Polytechnical
University,
Xi'an, P.R. China
sunjinqiu@nwpu.edu.cn

Chen Ding, Yanning Zhang

School of Computer Science,
Northwestern Polytechnical
University,
Xi'an, P.R. China
ynzhang@nwpu.edu.cn

Abstract—Using microscopy technology to analysis human cells is an important method for pathological study, diagnostic etc. As a pioneering imaging technology, the microscopy has many advantages: reduced photo toxicity and photo bleaching as well as enhanced imaging penetration depth. There are many center location methods for cells microscopic image. However, location results of these methods cannot meet the requirements of precision analysis. In this paper, We present a novel algorithm to locate cell nucleus in microscopic image. Our method consists of the following components: Firstly, to improve robustness of the algorithm the initial segmented regions are obtained by using Otsu method. Secondly, in order to improve accuracy of location results, the random sample consensus (RANSAC) method is used to eliminate exterior point after find the minimum bounding circle of cells. Finally, an ellipse fitting method is introduced to get the location of cell nucleus. Experiments on challenging microscopic images show that the proposed algorithm performs favorably against several state-of-the-art methods.

Keywords—component; Center Location; microscopic image; Cell image; segmentation;

I. INTRODUCTION

Cell's microscopic images has significant meaning for diagnosis and treatment of diseases in medicine. Pathologists often make diagnostic decision by observing the specimen cells and the geometric parameters of the cell such as area, radius, central position and the circumference [1]. For example, in the early stages of the cancer, the shape of cancerous cell nucleus will be transform, and the physical characteristics of cancerous cells is different from normal cells. To get the parameters, accurate location and segmentation of cell images is required. Thus, locating cell nucleus is one of the most important tasks in analyzing and quantifying cell's microscopy images. There is no doubt that it will great assist to treat cancer if we can find these changes in a timely manner.

However, those various factors have limited the quality of captured microscopic image because of the variety and complexity of organism. On the one hand, in the process of captured microscopic image inevitable causes noise such as electric coupling, magnetic coupling, electric magnetic noise, background noise and so on. These noises influence detection and segmentation of small cells. On the other hand, signal to

noise ratio of microscopic images is affected by background and cells brightness, contrast ratio. Especially, in the situations of the brightness of cells is similar to background (too high or too low), it is hard to detect those cells in microscopic images. The above factors become the important problem how to accurate positioning and segmentation.

II. RELATED WORK

There are many approaches have been made for locate the center of object. Gander [2] presented an algorithm which compute the ellipse through calculating the minimal sum of the squares of the distances to the given points. Mai [3] raised a hierarchical approach for fast and robust ellipse extraction from images. Zhang [4] detected edge points by using Canny operator in the region of interest, then the accurate edge points are acquired by Gauss curve fitting based on the information of the ellipse and the gray-distribution characteristics of the edge points. Han [5] located the center of ellipse on account of the CACH and the least square method. However, those methods is not robust for microscopic images.

Xie [6] presented a new center location method based on fuzzy mathematics, which can be easily found the center of cell nucleus even when they have different size in one cell image. Liu [7] thought the direction of the gradients on edge pixels and geometry is applied to acquire the cell center. Hu [8] proposed a new method that can locate and detect nucleus effectively in critical noisy, fuzzy boundary and complex background. Those algorithms depend on the ability of edge detection.

The aim of this paper is to present a novel center location algorithm in cells microscopic image. This method is able to directly yield the accurate center of cell nucleus by binarization, RANSAC and ellipse fitting. Firstly, in order to get the initial segmented result, otsu method was used to detect cells which is more robust to background noise. Secondly, while achieving binary image, we need to extract the region of interest(ROI) so as to locate the center of nucleus, so we obtain the minimum bounding circle of cells. Thirdly, RANSAC (random sample consensus) method is employed to eliminate exterior point and improve accuracy of location results, fitting the ellipse by least square method in order to finish the center location. Experimental results show that the proposed algorithm

performs favorably against several state-of-the-art methods in cells microscopic image.

III. CENTER LOCATION ALGORITHM

In this section, the center location algorithm in cells microscopic image will be introduced in detail. It is inevitable that noise maybe existed in the Microscopic Image, it is essential preprocess the origin image. We use Gaussian filter to smooth and remove the noise.

A. Image Segmentation

Before locating the center of cells, we should achieve the area of cell image, that is, the reduction of a gray level image to a binary image. We found that Otsu method can have a high segmentation correct rate, so we use Otsu to isolate background and objects. The algorithm assumes that the image contains two classes of pixels following bi-modal histogram (foreground pixels and background pixels), it then calculates the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal, or equivalently (because the sum of pairwise squared distances is constant), so that their inter-class variance is maximal. The Otsu methods is works as follows:

$$\sigma_{\omega}^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t) \quad (1)$$

Where, $\omega_0(t)$, $\omega_1(t)$ are the probabilities of the two classes separated by a threshold t ; $\sigma_1(t)$, $\sigma_2(t)$ are variances of these two classes. As you known, Otsu shows that minimizing the intra-class variance is the same as maximizing inter-class variance [9]:

$$\sigma_b^2(t) = \sigma^2 - \sigma_{\omega}^2(t) = \omega_0(t)\omega_1(t)[\mu_0(t) - \mu_1(t)]^2 \quad (2)$$

which is expressed in terms of class probabilities ω and class means μ . The detail of Otsu methods as follows.

- Compute histogram and probabilities of each intensity level;
- Set up initial $\omega_i(0)$ and $\mu_i(0)$;
- Step through all possible thresholds $t=1, \dots$, maximum intensity:
Update ω_i and μ_i ; Compute $\sigma_b^2(t)$
- Desired threshold corresponds to the maximum $\sigma_b^2(t)$.

The segmentation result is shown in Fig. 1(b). After Otsu, the image is divided to 2 parts: nucleus and cell intercellular substance. And the gray level of image is two. It provides a well the basis for accurate segmentation.

B. Extract Effective Regions

While achieving binary image, we need to extract the region of interest(ROI) so as to locate the center of nucleus. Extracting connected region too large or too small will have an disadvantageous impact on the results of the fitting process and positioning. If the selected area is too large, it will introduce more background pixels, so reducing fitting accuracy. On the contrary, if too small, it caused to misse the effective pixels so that the fitting is not accurate. At this point a number of discrete points to obtain a value of 1 or 0, the proper way is to find the minimum bounding circle of cells of them. we observe the use the following aspects to acquire the connected region:

- Traverse all all the discrete points, and calculate the distance between current point and other, return the the maximum distance of two points;
- Set length of the two points(LineA) is value of diameter;
- Calculate all other points distance to the center , if less than equal to the radius of the circle , the circle is returned;
- If the distance of some point to the center is greater than the radius, then set the largest of which is PntC , so we can get the circle of three points which from two points of LineA and PntC;
- Repeat c).

C. Center Location by Ellipse Fitting and modified RANSAC

Up to now we have grouped arc segments that presumably belong to the diffireent ellipse. An improved least square approach is applied to realize ellipse fitting. The ellipse-constraint algebraic fitting always provides an elliptical solution, but the bias is inevitably added to result because the sample data may including some biased data. Based on this situation, the method of RANSAC used to reject outliers iteratively at the process of each cell fitting.

a) The basic idea of Centering algorithm

As we know , most of cells usually have elliptically shaped boundaries.so we should extract the ellipse image accurately and fit it by implicit algebraic curve in order to finish the center location. According to the geometric properties of an ellipse, the ellipse can be determined by the following five parameters,:

$$\frac{((x-x_0)\cos\theta + (y-y_0)\sin\theta)^2}{a^2} + \frac{((y-y_0)\cos\theta + (x-x_0)\sin\theta)^2}{b^2} = 1 \quad (3)$$

Once detemining the parameter of equation(3), the center of cell can be accurately ensured. The equation can be denoted as:

$$f(x, y) = ax^2 + 2bxy + cy^2 + 2ex + f = X^T PX = 0 \quad (4)$$

That is:

$$f(x, y) = [x^2 \ 2xy \ y^2 \ 2x \ 2y \ 1][a \ b \ c \ d \ e \ f]^T = 0 \quad (5)$$

$$X = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, P = \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix} \quad (6)$$

where the rank of X is 3, and X represents the homogeneous coordinate of points, and P represents the coefficient matrix of ellipse.

For the ellipse fitting, only three sample points are required and they satisfy the equation :

$$\begin{bmatrix} x_1^2 & 2x_1y_1 & y_1^2 & 2x_1 & 2y_1 & 1 \\ x_2^2 & 2x_2y_2 & y_2^2 & 2x_2 & 2y_2 & 1 \\ x_3^2 & 2x_3y_3 & y_3^2 & 2x_3 & 2y_3 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = Ap = 0 \quad (7)$$

in which the rank of A is 3.

The fitting of a general conic may be approached by minimizing the sum of squared algebraic distances

$$\min \|Da\|^2, s.t. a^T Ca = 1 \quad (8)$$

where D is the data set, and C is a constraint matrix.

Introducing the Lagrange multiplier λ and differentiating, we arrive at the system of simultaneous equations

$$\begin{aligned} 2D^T Da - 2\lambda Ca &= 0 \\ a^T Ca &= 1 \end{aligned} \quad (9)$$

This may be rewritten as the system

$$Sa = \lambda Ca \quad (10)$$

$$a^T Ca = 1 \quad (11)$$

where S is the scatter matrix $D^T D$. This system is readily solved by considering the generalized eigenvectors of (10). If (λ_i, u_i) solves (10), then so does $(\lambda_i, \mu u_i)$ for any μ and from (11) we can find the value of μ_i as $\mu_i^2 u_i^T C u_i = 1$, giving

$$\mu_i = \sqrt{\frac{1}{u_i^T C u_i}} = \sqrt{\frac{1}{u_i^T S u_i}} \quad (12)$$

Finally, setting $\hat{a}_i = \mu_i u_i$, and then continuously take eigenvectors u_i corresponding to $\lambda_i > 0$, which can be seen the solutions of curve fitting equation.

b) The step of location algorithm

In this paper, every time fitting of ellipse by RANSAC for the segmented area is as follows:

- Start define iteration number $k=1$;
- Choose three points randomly from the observed points to fit the ellipse model;
- Calculate the distances from the points to the fitting model, then discriminate the distances whether in the threshold or not, if less than the threshold the points will be labeled as inliers. Define the set of inliers as consistent set;
- Compute the number of inliers set. If the number is less than the defined threshold, go to b), if greater, end the loop;
- Choose the consistence set with the largest number of inliers;
- If the iteration number exceeds the maximum number, stop the loop;
- Re-estimate the model using the inliers with the largestest number.

After this, the location of each cell can be determined through the experiment, we can get the best fitting of each input data (for every cell). Each ellipse obtain a set of parameters a, b, c, d, e, f , according equation (7) we can ensure the ellipse. The parameters expressed by center coordinate (x_0, y_0) , the major and minor axis a, b , the angle of major axis and x axis θ . And the final determination is (x_0, y_0) .

By the equation (3) and (4), we can infer:

$$\begin{cases} x_0 = \frac{be - 2cd}{4ac - b^2} \\ y_0 = \frac{bd - 2ae}{4ac - b^2} \end{cases} \quad (13)$$

The location result is shown in fig.1 (d). From the results we can find the location is accurate, almost all the nucleuses have been found out.

IV. EXPERIMENTAL RESULTS

We implemented our center location algorithm in Matlab language with Matlab R2013a in Windows7 environment and tested on a computer based on Intel Pentium processor which was sufficient for our purposes. We select 5 groups different cells microscopic images as the experiment data, as shown throughout the paper. These microscopic images are with complex background and contrast ratio in different levels. Fig.1 show the results of our method applied to microscopic image data.

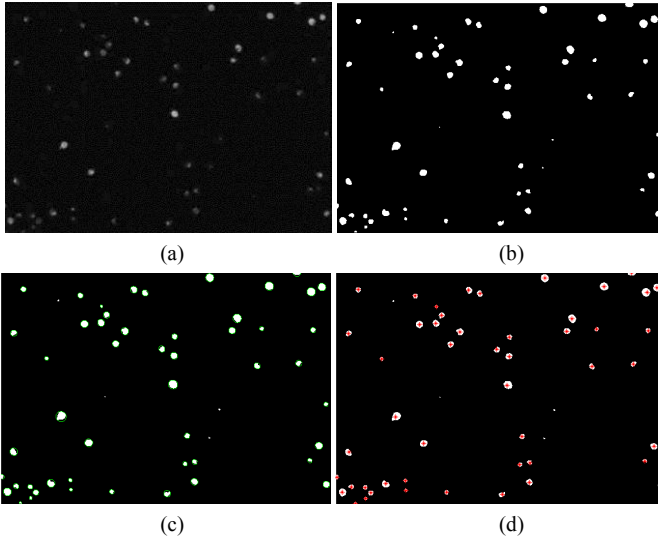


Fig. 1. The process of center an image using this paper's method. (a) Original image; (b) Image that has been segmented; (c) Result of connected region; (d) Result of center location

Fig. 1(a) shows the original image. The segmentation process is as follows: Use Otsu method to distinguish foreground pixels and background pixels. The result of segmented is shown as Fig. 1 (b), we can see that almost cells have been segment including the bright cells and the darker cells. However, binary image at the boundaries usually is not very smooth and background area also have a few small discrete noise, small hole or tiny spot. Taking advantage of morphology operation such as after many of corrode and dilate with the expansion of the same number can effectively improve those situation.

Then we desire to extract the connected regions through finding the minimum bounding circle of cells. Fig. 1(c) shows the extracted edge of nucleus, the green circle indicates the minimum circumscribed circle of the extracted cells. Those green circle surround all cell nucleus, this is the premise of locate the center of cell nucleus.

Assuming the outline of the cell nucleus are circle or ellipse, we formulate the location of cell nucleus as an ellipse fitting problem. Since the point of focus obtained by extraction point may not meet elliptic equations, so outliers should be screened and rejected. So the effective points of ellipse is obtained using the random sample consensus (RANSAC) paradigm.

Finally, the method of least squares was used to fit ellipse cells, at last we get the accurate location of all cells. The result is shown in Fig.1 (d), most of the nucleus has been found. The red point in the image is the accurate center of nucleus that has been located.

In this paper, we use this method to center 5 pieces of microscopic cell images; the statistics is shown in TABLE I. where the accuracy and loss rate are defined as:

$$\text{accuracy} = \text{correct} / (\text{correct} + \text{wrong}) \times 100\% \quad (14)$$

TABLE I. THE RESULTS FOR EXPERIMENT

No	Ground truth	correct	error	missing	accuracy	loss rate
1	59	52	3	4	94.5%	6.78%
2	25	22	1	2	88.0%	8%
3	37	37	0	0	100%	0%
4	51	44	4	1	86.3%	1.96%
5	184	164	12	8	89.1%	4.34%

$$\text{loss rate} = \text{miss} / \text{total} \times 100\% \quad (15)$$

From Table 1 we can infer that our method calculate location of high veracious. Compared with the result by Xie [6], which accuracy is 85.0%,88.9%,85.0%, it is obvious that our results are more precise. The experimental results shown that, nucleus can be located accurately.

V. CONCLUSION

In this paper, an effective novel center location algorithm in cells microscopic image is proposed. According image segmentation we get the initial segmented result. Then, RANSAC method is employed to eliminate exterior point and improve accuracy of location results, after extract the minimum bounding circle of cells. At last, we fit the ellipse image by least square method in order to finish the center location. Experiments on challenging datasets show robustness and effectiveness of the proposed algorithm.

ACKNOWLEDGMENT

The work is supported by grants NSF of China (61231016, 61301192, 61301193, 61303123), Natural Science Basis research Plan in Shaanxi Province of China (No. 2013JQ8032).

REFERENCES

- [1] G. Idoia, Perezn, "Applying deformable templates for cell mage segmentation," *Pattern Recognition*, vol. 33, no.5, pp. 821-832, May. 2000.
- [2] W. Gander, G. H. Golub, R. Strebler, "Least-squares fitting of circles and ellipses," *BIT Numerical Mathematics*, vol. 34, no. 4, 1994, pp. 558-578.
- [3] F. Mai, Y. S. Hung, H. Zhong. "A hierarchical approach for fast and robust ellipse extraction," *Pattern Recognition*, 2008, vol. 41, no. 8, pp. 2512-2524.
- [4] H. Zhang, F. P. Da, D. K. Xing. "Algorithm of centre location of ellipse in optical measurement," *Journal of Applied Optics*, vol. 29, no. 6, pp.905-911, 2008.
- [5] B. Han, J.T.Shangguan, Y. L. Li. "A Ellipse Center Positioning Method Based on The Least Square," *Journal of Changzhi University*, vol. 26, no. 5, 2009.
- [6] Z. M. Xie, G. N. Chen, J. X. Chen. "A New Center Location Algorithm Used in Cell Segmentation," *Image and Signal Processing*, 2008. CISP 08. Congress on. 2008, pp. 610-614.
- [7] D. S. Liu, X. J. Shen. "Automatic Segmentation of Cell Image," *Journal of DaLian Railway Institute*, vol.25, no.4, pp. 59-62, 2004
- [8] H. Y. Hu, G. N. Chen, P. Wang, Y. Liu. "Center Location and Segmentation of Cell Two-photon Microscopic Images Based on Random Sample Consensus," *Proceedings of the 2013 6th International Congress on Image and Signal Processing (CISP 2013)*, Hangzhou, 2013, v2, pp. 588-592.
- [9] N. Otsu "A threshold selection method from gray-level histograms," *IEEE Trans. Sys., Man., Cyber.* vol. 9, no. 1, pp. 62-66.