# A Computer Vision Approach to Micro-Nucleus Automatic Detection for Protozoan Parasites Segmentation

Hsiao-Yu Wang*, Hung-Yuan Chung

*Abstract*—**Till now, protozoan parasites cause many diseases, for examples, malaria, EHEC infection, shigellosis and *amoebiasis* etc. The kinds and growing stages of protozoan parasites would lead to different treatments. The most significant characteristic of different growing stages is the number of nucleuses, but partial nucleuses of a cell may be more unclear than the others causing the missing in nucleus detection. This paper presents a novel multiple nucleus detection schemes which are composed from the adaptive protozoan parasite erasure, gamma equalization, Fuzzy C-means clustering algorithm, modified connected component detection method, and circle mask scoring method. For each cell, the proposed scheme first detects the most significant nucleus, and then performs gamma equalization iteratively to extract a correct nucleus. In each iteration, only the remained region would be considered. Iterations are terminated when all parameters of gamma equalization are considered. The adaptive protozoan parasite erasure method is used to erase the boundary of a protozoan parasite by dynamic size mask. The modified connected component detection method labels each connected component more accurately than the traditional method. The iterative gamma equalization performs gamma equalization iteratively by different parameters to enhance the boundaries of nucleuses with different edge intensities. The circular mask scoring method can help estimate the circular degree of objects. The experiment shows that the proposed scheme can detect the nucleuses with indistinct boundaries effectively.**

***Index Terms* — Nucleus detection, FCM, boundary erasure, connected component.**

## I. Introduction

Most protozoans are constructed with only one cell (unicellular organisms) and they belong to eukaryotes. In the natural world, protozoans are classified into about 65,000 species and most protozoans exist in ocean, soil, water or rotten things, etc., without depending on any other organisms. Others belong to protozoan parasites and contain about 10,000 species.

Manuscript received January 19, 2010.

Hsiao-Yu Wang is with the No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan (R.O.C.) (phone: 0918190357;e-mail: 985201091@cc. ncu.edu.tw).

Hung-Yuan Chung is with the No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan (R.O.C.) (phone: (03)422-7151-34575; e-mail: hychung@cc.ncu.edu.tw).

They exist in the inside or outside of other organisms and about 40 species of them are found as human parasites [10].Most protozoan parasites exist on lumens, body fluid, tissues or cells of human and can be classified into pathogenic and nonpathogenic protozoan parasites [11].

The common intestinal infectious parasitic diseases include cholera, typhoid fever, Enterohemorrhagic Escherichia coli infection (EHEC) [7], shigellosis and amebic dysentery (or intestinal amebiasis) etc [5]. The cholera is caused by Vibrio cholerae serogroup O1 which can be divided into two types, cholera classical and E1 Tor; the typhoid fever is caused by Salmonella typhi and Salmonella paratyphi; the EHEC infection is caused by Escherichia coli; the shigellosis is caused by Shigella dysenteriae, Shigella flexneri, Shigella boydii and Shigella sonnei [8]; the amebic dysentery is caused by the Entamoeba histolytica [9], and it has two types, trophozoite and cyst in its life history [5, 9]. The medical imaging techniques are widely applied to assist doctors and pathologists to diagnose diseases. Nowadays, the general medical imaging techniques include: computerized tomography (CT) [21], magnetic resonance imaging (MRI) [22], ultrasound imaging [23] and electron microscopic imaging etc [24]. Also, the electron microscopic imaging techniques are used to observe protozoan parasites, blood cells, tumor cells and so on.

In recent years, digital medical image processing techniques and systems are gradually applied to assist doctors and pathologists to diagnose and analyze more efficiently and accurately by computers. The cell segmentation is a popular research topic of medical images processing. It is performed before the cell recognition to extract the sharps and regions of cells from a cell image which contains background or other objects. Hence, it is very important and affects the performance of a whole cell image recognition procedure. The most common cell segmentation research issues include red and white blood cells, tumor cells and other tissue cells of human. Most cell segmentation techniques are classified into texture-based and edge-based segmentation. The edge-based methods are set to detect the edges of cells by edge processing and analysis techniques. The common techniques include Sobel edge detection [15], Canny edge detection [13] and gradient vector flow for snake (GVF-snake) [14] etc. Edge-based methods are usually affected by noise and contrast more easily than texture-based methods. Texture-based methods classify texture features between target objects and others regions to obtain the cell regions in a cell image by classification

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

algorithms. Common texture-based classification methods include k-means clustering [1], support vector machine (SVM) [17], watersheds algorithm
[16], fuzzy c-means clustering [18] and artificial neural network (ANN) [19, 20] etc. The cell segmentation researches for protozoan parasite are few. Some cell segmentation researches for cancer cells and white blood cells are described as follows.

N. Gao *et al*. [6] proposed a multispectral imaging technique for white blood cell segmentation. They obtained multiple marrow microscopic images with different bands of wavelengths for the same sample. The signal intensities for different wavelengths of each pixel are the features used to perform classification by support vector machine (SVM). The major classes include nucleus, erythrocytes, cytoplasm and background in their proposed scheme. In order to reduce the number of feature data in classification, the authors used two data reducing techniques and ignore background in their classification procedure. They found that only the signal intensities of background pixels are more than 200 when the wavelength is 530 nm. In addition, they use sequential minimal optimization (SMO) technique to reduce the number of features according to empirical risk. Their experiment represented that their segmentation correct rate is about 94% for certain classes of white blood cells.

Jiang *et al*. [12] proposed a cell segmentation scheme to segment white blood cells according to the information extracted from feature space. They transform a cell image into a histogram that contains feature information. First, the Gaussian filter is applied iteratively with larger variance to reduce the number of local minimums and obtain two valley points in the histogram. These two points are used to track the related gray levels in the original histogram by the fingerprint technique. The related two gray levels are used to cluster all pixels of the original image into three classes. The darkest class contains the nucleus of white blood cell pixels. After obtaining the nucleus regions, the cytoplasm regions are extracted by using Gaussian filter with a variance of 0.75 and a watershed clustering algorithm without considering the nucleus regions.

Wu *et al*. [25] presented an adaptive thresholding scheme to approximate and segment microscopic cell images. A parametric image was obtained by iterative minimizing its mean squared error with the original image and was used to perform thresholding scheme to extract cell regions. But, some cell priori-knowledge such as background, cell shape and size, are needed for the minimizing procedure.

Mat-Isa *et al*. [26] proposed a region growing feature extraction (RGFE) algorithm to extract four features from a region of interest (ROI) in a cervical cancer cell image for cell segmentation. These features are the size, gray level of nucleus, cytoplasm. However, the ROI and the threshold value must be manually selected by the users of the application system, which reduces the automatic ability.

Now, we focus on the protozoan parasite identification process. The first step of intestinal infectious parasitic disease diagnosis is usually to obtain fecal or tissue section samples, then pathologists use electron microscope to find protozoan parasites and to diagnose the species of the protozoan parasites by their biomorphic, such as, cell tissue distribution, nucleus numbers and smoothness of cell wall. Nowadays, most above works are still performed manually. But to manually diagnose the species of protozoan parasites exist some difficulties and problems. The species of protozoan parasites are many and protozoan parasites are tiny. In addition, the protozoan parasites have different forms according to their maturity. Hence it is difficult that the pathologists accurately and quickly diagnose the species of the protozoan parasite according to their hairline features, and it would make mistakes easily.

The protozoan parasite microscopic image processing exists some difficulties: (1) The illuminations, colors, noises and size of microscopic images are dependent on experiment devices and environments. (2) The contents of protozoan parasite microscopic images are complex when they are obtained from fecal or tissue section samples, especially feces contain many other unknown things. (3) The protozoan parasites have different forms according to their maturity. (4) The number of nucleuses is unfixed and the boundary contrast of nucleuses is variable.

In this paper, a novel protozoan parasite nucleus detection scheme is proposed to hurdle the above difficulties and to accurately detect the region of each protozoan parasite in color microscopic images. The proposed scheme is presented in Section 2. The proposed scheme including several images processing techniques are used to detect and segment all nucleuses of a protozoan parasite for protozoan parasites recognition. In the experiment, a protozoan parasite image which includes four nucleuses (one is nuclear) was used to test the performance of the proposed scheme. The experimental results show that the proposed scheme can detect all nucleuses even with different edge intensities. The experimental results are shown in Section 3. Finally, conclusions are given in Section 4.

## II. Feature Extraction Scheme

This paper proposes an effective scheme to segment protozoan parasites and their nucleuses from microscopic images. The proposed scheme is divided into two major stages: The first stage is to segment a whole protozoan parasite. The protozoan parasite microscopic images are photographed from human samples, hence the background of each image exists many unknown impurities to increase difficulties of the extraction of protozoan parasites. The second stage is to detect all nucleuses of each extracted protozoan parasite. At this stage several difficulties must be overcome. The first difficulty is that the number of nucleuses is various. The proposed scheme must to exert itself to detect all nucleuses but to ignore the cytoplasm in each protozoan parasite. The second difficulty is that the boundaries of some nucleuses are blurred or incomplete. This condition could affect the sizes,

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

shapes and positions of nucleuses and hinder detecting nucleuses correctly.

First, the protozoan parasite segmentation is applied by our previous proposed protozoan parasite segmentation scheme. The scheme first transforms color microscopic images into gray-level ones. The gamma equalization is a novel non-linear based equalization scheme for equalizing color space of images. This method can be used to normalize the gray-level distribution of images. In order to erase noises and enhance the boundaries of protozoan parasites, two filtering schemes which are the median-mean filter and the two-class edge enhancement algorithm are used. Then, a two-means algorithm which is the well-known *k*-mean algorithm when *k*=2 is used to cluster the pixels of each image into two major clusters, the protozoan parasite and background. Some clustered pixels still are clustered into incorrect cluster. The largest connected object is considered as the protozoan parasite. Thus, we seek the largest connected object only for the pixels which are belonged to the cluster of protozoan parasite. Then a morphological closing operation is used to obtain a more complete protozoan parasite region. All pixels of the raw image which do not exist in the domain of protozoan parasite region are marked as background pixels. This segmentation scheme can gain 96.64% correct rate for 112 experimental protozoan parasite microscopic images and is reused in this paper for protozoan parasite segmentation. The details of gamma equalization and two-means algorithms are described in the following sub-sections.

Second, an iterative-based nucleus detection scheme is proposed to detect multiple nucleuses for each protozoan parasite. This nucleus detection scheme contains several stages with different methods which are: adaptive boundary erasure, iterative gamma equalization, two-means, modified connected component detection and circular mask scoring methods. This scheme is performed iteratively and first detects a most significant nucleus in a protozoan parasite. Then the height and the width of this nucleus can be measured and used as the conditions for further detecting other nucleuses in the following iterative procedure.

The details of each method are described in following sub-sections, respectively.

### A. Adaptive Boundary Erasure

The segmentation can effectively eliminate the background of a protozoan parasite image, but the unnecessary boundary pixels of the protozoan parasite are still preserved. This paper proposes an adaptive boundary erasure method as the pre-process of nucleus detection to eliminate the boundary region by an adaptive mask with various sizes. The size of a mask is determined by estimating the width of the boundary. In order to measure the width of the boundary, the propose method transforms the gray values into gradients *G* by using Sobel operator shown in Eq. (1) and defines *m* sample points on the most external boundary of protozoan parasite evenly. Each scan is performed on each sample point to find a local maximum of gradient from outside to inside in an interval $\delta$. Each scan can obtain a boundary width $w_i$ from calculating the distance of the local maximum and background. The boundary width for the protozoan parasite is considered as the mean of the boundary widths of all scans. Then, the size of adaptive mask *s* is obtained by Eq. (2). The boundary elimination is performed to mark boundary pixels as background pixels by an adaptive mask. This method scans all pixels except marked background pixels and checks neighboring pixels of each corresponding pixel by an adaptive mask, when existing one or more neighboring pixels are background pixels, the corresponding pixel also is marked as background pixel. All background pixels are ignored in the latter image processing.

$$G = \sqrt{G_X{}^2 + G_Y{}^2} \tag{1}$$

Where

$$G_X = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad , \quad G_Y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$s = 2w+1$$

Where

$$w = \frac{\sum_{i=1}^{m} w_i}{m} \tag{2}$$

### B. Iterative Gamma Equalization

In most cases, nucleuses and cytoplasm have different brightness in protozoan parasite images, and contours of nucleus regions usually are the most significant parts. But microscope images of our image database are obtained from different experiment conditions or environments. Hence each microscope image has different distributions of illuminations. Image equalization is a luminance transform method to equalize the illuminations of an image to distribute over whole gray-levels. Hence this method can be used to normalize the illuminations of our microscope images. The traditional equalization method [2, 3] is performed by Eq. (3), where *I*Gray (*i*, *j*) is the gray-level image, *I*Eq (*i*, *j*) is the image which performed equalization, *I*Max and *I*Min are the maximum and minimum of the pixel values of *I*Gray (*i*, *j*). The traditional equalization method has a major problem that the illumination contradistinction and distribution are changed linearly and are restricted by the few pixels located near by maximal and minimal pixel values of an image. Hence this method is not enough to enhance the contradistinction of nucleuses and cytoplasm when only using the traditional equalization method.

$$I_{eq}(i, j) = I_{Gray}(i, j) * \frac{I_{Max} - I_{Min}}{255} \tag{3}$$

Gamma equalization is proposed to overcome the

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

previous problem. In most natural images, the gray values of only few pixels are near 0 or 255. But these pixels could restrict the degree of image equalization. In order to avoid this restriction, the gamma equalization defines a threshold *t* used to decide how many pixels of an image should be ignored when they are near 0 or 255 in a gray-level histogram. For example, when *t* = 12, the 1/12 pixels nearest 0 and 1/12 pixels nearest 255 are ignored in gamma equalization and their pixel values all are set to 0 and 255, respectively. Figure 2 shows the actually performed and ignored ranges of the gamma equalization. The low-bound and up-bound are the smallest and largest gray values of the equalization range in a gray-level histogram.

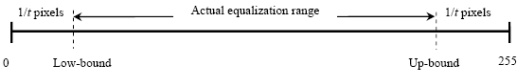


Fig. 1: A description of equalization range in the gamma equalization.

The gamma equalization reduces the effect when few pixels are near extreme values. This
method equalized all pixels which are located in the actual equalization range of the histogram nonlinearly using another parameter $\gamma$. The major equation of gamma equalization is shown in Eq. (4).

$$GE(i) = \frac{(i-L)^r}{R^r} * 255 \qquad r \in [0,1] \tag{4}$$

Where *i* is the *i*-th gray level, *L* is the low-bound, *R* is the actual equalization range and *GE* (*i*) is the result of the *i*-th gray level obtained from gamma equalization.

Larger *t* and smaller $\gamma$ would cause the processed image become darker with a less contrast that $\gamma$ is smaller than 1. When different *t* and $\gamma$ are used in gamma equalization, different enhancement effects are brought to the pixel values in different gray level nonlinearly. In this paper, the proposed scheme performs an iterative procedure with different values of the parameters *t* and $\gamma$ in gamma equalization to enhance nucleus with different brightness and sharpness. In each iteration, a nucleus is detected and is removed from the image in next iteration in order to further enhance the more unclear nucleuses.

### C. Fuzzy C-Means(FCM) Algorithm with Spatial Constraint

The purpose of this paper is to accurately extract the nucleuses from digital protozoan parasite microscopic images with complex cytoplasm background. When the images are processed by the above digital image process techniques, all protozoan parasite images are uniformly transformed into the equalized gray-level space, the contrast between nucleuses and cytoplasm are enhanced. This paper shows segmentation method which is based on a clustering algorithm called fuzzy c-means algorithm.

A modification of the fuzzy c-means (FCM) algorithm is used to initialize the level set segmentation refinement process. FCM minimizes the sum of similarity measures objective function J (U, V) given by

$$J(U,V) = \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^m \left\| x_j - v_i \right\|^2 \tag{5}$$

Where $X = \{x_1, x_2, \ldots x_N\}$ denotes the set of data (pixel feature vectors), $V = \{v_1, v_2, \ldots, v_C\}$ represents the prototypes, known as the clusters centers, $U = [u_{ij}]$ is the partition matrix which satisfies the condition, $\sum_i^C u_{ij} = 1 \ \forall j$, and *m* is a fuzzifier which indicates the fuzziness of membership for each point. The FCM algorithm is an iterative process for minimizing the membership distance between each point and the prototypes. However, the objective function Eq. (5) does not explicitly include any spatial information. Incorporating spatial information provides more robustness and efficiency to the fuzzy c-means algorithm by adding a second term to the FCM objective function,

$$J_M(U,V) = \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^m \left\| x_j - v_i \right\|^2 + \alpha \sum_{i=1}^{C}\sum_{j=1}^{N} u^m{}_{ij} e - \sum_{k\in\Omega} u_{ik}^m \tag{6}$$

Where $\Omega$ is a set of neighbors. The parameter $\alpha$ is a weight that controls the influence of the second term. The objective function (6) has two components. The first component is the same as FCM; the second is a penalty term. This component reaches a minimum when the membership value of neighbors in a particular cluster is large. The optimization of (6) with respect to U is solved by using Lagrange multipliers and the membership function update equation is,

$$u_{ij} = \frac{1}{\sum_{p=1}^{C} \left( \frac{\left\| x_j - v_i \right\|^2 + \alpha e^{-\sum_{k\in\Omega} u_{ik}^m}}{\left\| x_j - v_p \right\|^2 + \alpha e^{-\sum_{k\in\Omega} u_{pk}^m}} \right)^{\frac{1}{m-1}}} \tag{7}$$

The neighboring membership values ($u_{pk}$) influence $u_{ij}$ to follow the neighborhood behavior. For instance if a given point has a high membership value to a particular cluster and its spatial neighbors have a small membership values to this cluster, the penalty term plays the role to force the point to belong to the same cluster as its neighbors. The weight $\alpha$ controls the importance of the regularization term. The prototype update equation is the same as standard FCM. The spatial constraint FCM (SCFCM) algorithm consists in the same steps as the original fuzzy c-means algorithm.

### D. Modified Connected Component Detection.

Although the nucleuses and their boundaries are usually darker than the cytoplasm in a protozoan parasite, but some pixels of the cytoplasm also are darker and classified into darker class. This paper considers each independent black region (pixel values are 0) as a possible nucleus object, and the actual nucleuses are usually larger

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

than the black regions belonged to the cytoplasm. However, the number of nucleuses of protozoan parasites is not known. In order to accurately extract all nucleuses, the largest black region is extracted to be the first nucleus and the standard for detecting other nucleuses. During each iteration, the largest connected component is detected to be a candidate nucleus.

The connected component detection checks each group connected black pixels to construct an independent component and mark a unique index. A binary component labeling method [4] is used to find all connected components in the binary image obtained from two-means. Each pixel is scanned from bottom-up to left-right to refer its left and below indexed neighboring pixel to determine its index. The referred neighboring pixels and scanning directions are shown as Figure 4(a), where the black and white grids are the corresponding pixel and it's referred pixels. The indexing procedure is performed only when the corresponding pixel is belonged darker class, and its index is determined by follow three conditions: (a) If the left and below neighbor both are black pixels, the corresponding pixel is marked as the smallest index of the two neighbors. (b) If only left or below neighbor is black, the corresponding pixel is marked as the index of the left or below neighbor. (c) If both the left and below neighbor are not black pixels, the corresponding pixel is marked a new index (larger than the current largest index). The number of pixels for each connected component is accumulated as its size during the indexing procedure.

This indexing method has a drawback that this method is only performed well for indexing completed rectangular components and may mark more than one index for an actual connected component with different shape. Most nucleuses of protozoan parasites are oval and the two-means may make it include some white holds. Figure 2(a) shows an example of the largest independent component obtained by above method for an original protozoan parasite and it is apparent that the detected component is not the whole region of the protozoan parasite. In order to improve this drawback, our previous research presents a modified indexing method to obtain the largest complete component. This method adds three scanning models with different scanning orders and referred pixels to find largest connected components, and they are shown in Figure 4(b)-(d). In this modified approach, 4 largest connected components are obtained from different scanning models (see Figure 5(b)-(d)) and they are combined by OR operations to be the final object. This modified approach can improve the previous problem, but 4 scans and 3 OR operations must be performed and cause increasing computing time. This paper proposes a novel solution to reduce computing time. First, only one scanning model is used to identify all connected components, and then another scanning is performed to check indexes of each black pixel and its black neighbors, when existing different indexes, all pixels with the larger index are remarked as the smaller index. This modified approach can improve the traditional problem and save more computing time (only scanning twice for each image).

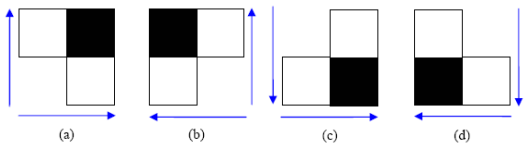


Fig. 2: The scanning directions and referred neighbors of different scanning models: (a)left-right and bottom-up; (b) right-left and bottom-up; (c) left-right and top-down; and (d)right-left and top-down.

### E. Circular Mask

Generally speaking, the size of the nucleuses is similar in a protozoan parasite, and the shape of most nucleuses is approximately circular. In the first iteration, a most observable nucleus is obtained from the largest connected component detection, and its width $w_S$ and height $h_S$ be looked as the standard for detecting other nucleus in the same protozoan parasite. However, the sharpness and brightness of nucleuses are slightly different to cause the size of the nucleuses different when they are obtained in different iteration. A tolerance difference value $\delta_s$ is defined to tolerate the size difference between the first nucleus and the following candidate nucleuses. Additionally, most nucleuses are approximately circular, hence, the shape of candidate nucleuses is also considered to determine whether they are actual nucleuses. Another tolerance difference value $\delta_R$ is defined to tolerate the proportion of the width and height. The perfect proportion is 1 when the object is circular and the tolerance range is $1\pm\delta_R$. However, the width and height of each object is measured by finding the smallest and fittest rectangle which can contain whole object. Hence, each object which fits the size condition is still not sure that its shape is circular. In order to further measure the shape of each object, this paper proposes a circular mask scoring method to measure the similarity of the shape of each object and a circle by a circular mask. The radius of the circular mask can be obtained by the average of width and height, and the circular mask can be divided into two regions to be an outer and an inner circle. The range of outer circle is determined by a parameter $r_O$ ( $0<r<r_O$ ) and the radius of inner circle is $r - r_O$. The scoring method accumulates the number of object pixels (pixel value = 0) in outer and inner circle ( $S_O$ and $S_I$ ) then calculates the circular score *CS* to be the degree of circularity for each object by Eq.(8). This method defines a threshold $T_{CS}$ to determine whether the circularity of the object is enough or not. Note that only objects fitting a preset size are processed to determine they are nucleus or not.

$$CS = \frac{S_O - S_I + C_I}{C_O + C_I} \times 100, \qquad (8)$$

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

where $C_O$ and $C_I$ are the total number of pixels in the outer and inner circle.

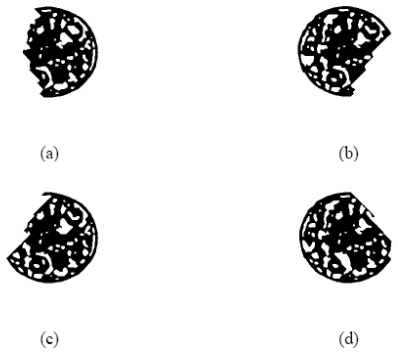


Fig. 3: The largest independent component from different scanning models in indexing method: (a) the original scanning model; (b) right-left and bottom-up; (c) left-right and up-bottom; and (d) right-left and up-bottom.

## III. Experimental Results

In our experiment, a protozoan parasite image is first used to show the result of each stage of the proposed scheme. This experiment shows an example image in Figure 4(a) and the segmented image obtained from the previous protozoan nucleus image segmentation scheme is shown in Figure 4(b).

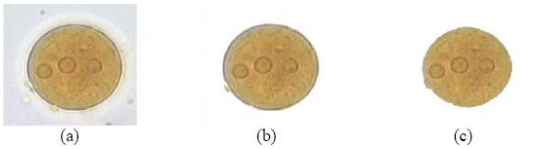


Fig. 4: (a) The original image; (b) the segmented image; (c) the result obtained from adaptive boundary erasure using 21 × 21pixel masks.

In Figure 4(b), the boundary region of the protozoan parasite is still preserved and similar as the boundaries of nucleuses. This paper uses the adaptive boundary erasure to remove most boundary pixels of the protozoan parasite. The size adaptive mask is obtained automatically and is 21× 21 for Figure 4(b). The erasure result is shown as Figure 4(c), and most boundary pixels are removed. In the following processing, only the preservative protozoan parasite region is considered in each image.

In order to enhance the boundary of each nucleus with different edge intensities, the iterative gamma equalization is performed iteratively with different parameters *t* and *γ*. Some results of gamma equalization with different parameters are shown in Figure 5, and it can be found that the boundaries with larger edge intensities can be preserved with larger parameter *t*; hence the parameter *t* is changed by a decreasing order in each iteration (the order in this experiment is from 35 to 1).

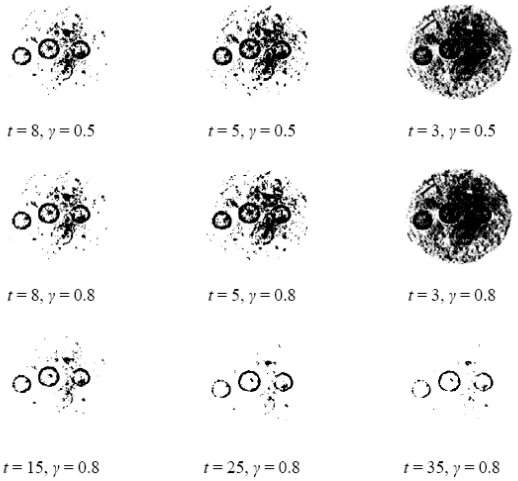


Fig. 5: The results of gamma equalization with different parameters *t* and *γ*.

In each iteration, a largest connected component is detected but it is not sure a validated nucleus before checking by the width, the height, the proportion of width and height and the circular mask scoring procedure. In circular mask scoring procedure, the radius *r* of circular mask is set as the average of the width and height of the first extracted nucleus. The most important parameter $r_O$ can affect the score. In order to compare the difference of different $r_O$, an experiment is performed to show the score for each pixel in the equalized example image using different $r_O$. The results are shown in Figure 6 and all score are normalized into 0 to255 and shown as gray-level images. It can be found by Figure 6 that the score is dispersive with too large or small $r_O$, when the $r_O$ is set around 10-5% of *r* can obtain the more concentrative score around the center of each nucleus.

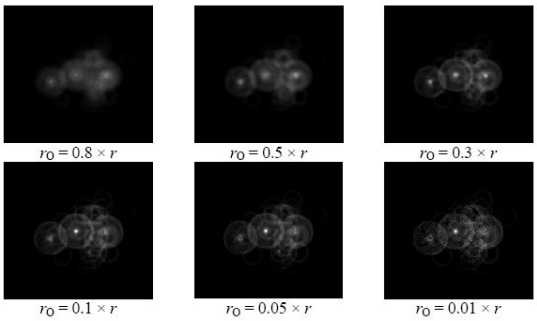


Fig. 6: The score images of the circular masks with different $r_O$.

In our experiment, the $r_O$ is set to 10% of *r*, and an object is determined to be a nucleus when its score is larger than 75. Note that, in order to further enhance the boundaries with less contrast, the extracted nucleus which obtained from each iteration is not considered in next iteration. The detection result of the example image is shown in Figure 7. This result shows that the proposed scheme can detect all nucleuses of the example image even the bottom unclear nucleus is also detected correctly

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

and that the iterative gamma equalization can be used to enhance dynamically the edge with different edge intensities in an image.

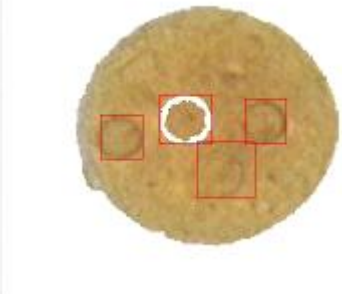

Fig. 7: The final detection result, the red rectangles are the regions of detected nucleuses and the white region is the first detected nucleus.

## IV. Conclusion

This paper presents a novel multiple nucleus detection schemes which include the protozoan parasite erasure, gamma equalization, and fuzzy C-means clustering algorithm; modified connected component detection method and circle mask scoring method. The proposed adaptive protozoan parasite erasure method can erase the boundary of a protozoan parasite by a mask. The mask size can be estimate by measuring the width of boundary automatically. Hence it can erase boundary by a precise and adaptive way. This paper modified the traditional connected component detection method to solve a drawback which labels an actual connected component with different indexes and can gain a better performance than our previous method, because it saves one image scan and four OR operations for each pixel. The iterative gamma equalization can perform gamma equalization iteratively by different parameters to enhance the boundary of nucleuses with different edge intensities. The circular mask scoring method can help estimate the circular degree of objects. The experiment shows that the proposed scheme can detect the nucleuses with indistinct boundaries effectively.

## References

[1] J. B. MacQueen, Some Methods for classification and Analysis of Multivariate Observations, *Proc. 5th Berkeley Symp. Math. Stati. Prob.*, University of California Press 1, pp.281-297, 1967.

[2] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd Edition, Addison-Wesley, 1992.

[3] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing using MATLAB*, Prentice Hall, 2004.

[4] E. Pietka, S. Popiech, A. Gertych, F. Cao, H. K. Huang, and V. Gilsanz, Computer Automated Approach to the Extraction of Epiphyseal Regions in Hand Radiographs, *J. Digit. Imaging*, vol.14, pp.165-172, 2001

[5] Center of Disease Control, R.O.C., Legal Infectious Diseases - Understanding Diseases, August 24, 2007, website available at http://www.cdc.gov.tw/.

[6] N. Guo, L. Zeng, and Q. Wu, A Method Based on Multispectral Imaging Technique for White Blood Cell Segmentation, *Comput. Biol. Med.*, vol.37, pp.70-76, 2006.

[7] T. V. Nguyen, V. P. Le, C. L. Huy, K. N. Gia, and A. Weintraub, Detection and Characterization of Diarrheagenic Escherichia Coli from Young Children in Hanoi, Vietnam, *J. Clin. Microbiol.*, vol.43, pp.755-760, 2005.

[8] B. R Thapa, K. Ventkateswarlu, A. K. Malik and D. Panigrahi, Shigellosis in Children from North India: A Clinicopathological Study, *J. Trop. Pediatr.*, vol.41, pp.303-307, 1995.

[9] N. P. Sherwood, Further Studies on the Antigenic Properties of Pathogenic and Free Living Amebas. II. Complement Fixation in Amebic Dysentery, *Am. J. Epidemiol.*, vol.16, pp.124-136, 1932.

[10] N. D. Levine Chairman, J. O. Corliss, F. E. G. Cox, G. Deroux, J. Grain, B. M. Honigberg, G. F. Leedaile, A. R. Loeblich 3rd, J. Lom, D. Lynn, E. G. Merinfeld, F. C. Page, G. Poljansky, V. Sprague, J. Vavra, and F. G. Wallace, A Newly Revised Classification of the Protozoa, *J. Eukaryot. Microbiol.*, vol.27, pp.37-58, 1980.

[11] E. K. Markell, D. T. John, and W. A. Krotoski, *Medical Parasitology*, eighth edition, W.B. Saunders Company, 1999.

[12] K. Jiang, Q. M. Liao, and Y. Xiong, A Novel White Blood Cell Segmentation Scheme Based on Feature Space Clustering, *Soft Comput.*, vol.10, pp.12-19, 2006.

[13] J. Canny, A Computational Approach to Edge Detection, *IEEE Trans. Pattern Anal.*, vol.8, pp.679-698, 1986.

[14] C. Xu and J. L. Prince, Snakes, Shapes, and Gradient Vector Flow, *IEEE Trans. Image Process.*, Vol.7, pp.359-369, 1998.

[15] E. Davies, Machine Vision: Theory, Algorithms and Practicalities, Academic Press, 1990.

[16] L. Vincent and P. Soille, Watersheds in Digital Spaces: an Efficient Algorithm Based on Immersion Simulations, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.13, pp.583-598, 1991.

[17] V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, 1995.

[18] J. C. Bezdek, *Fuzzy* Mathematics *in Pattern Classification*, New York: Cornell University, 1973.

[19] T. W. Nattkemper, H. Wersing, W. Schubert, and H. Ritter, A Neural Network Architecture for Automatic Segmentation of Fluorescence Micrographs, *Neurocomputing*, vol.48, pp.357-367, 2002.

[20] O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*, 2nd Edition, Richard John Wiley & Sons, 2001.

[21] S. Arivazhagan, S. Deivalakshmi, K. Kannan, B. N. Gajbhiye, C. Muralidhar, S. N. Lukose, and M. P. Subramanian, Multi-resolution System for Artifact Removal and Edge Enhancement in Computerized Tomography Images, *Pattern Recogn. Lett.*, vol.28, pp.1769-1780, 2007.

[22] E. M. Haacke, M. Ayaz, A. Khan, E. S. Manova, B. Krishnamurthy, L. Gollapalli, C. Ciulla, I. Kim, F. Petersen, and W. Kirsch, Establishing a Baseline Phase Behavior in Magnetic Resonance Imaging to Determine Normal vs. Abnormal Iron Content in the Brain, *J. Magn. Reson. Imaging*, vol.26, pp.265-273, 2007.

[23] S. J. Lord, W. Lei, P. Craft, J. N. Cawson, I. Morris, S. Walleser, A. Griffiths, S. Parker, and N. Houssami, A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer, *Eur. J. Cancer*, vol.43, pp.1905-1917, 2007.

[24] N. Guo, L. Zeng, and Q. Wu, A Method Based on Multispectral Imaging Technique for White Blood Cell Segmentation, *Comput. Biol. Med.*, vol.37, pp.70-76, 2006.

[25] H. S. Wu, J. Gil, and J. Barba, Optimal Segmentation of Cell Images, *IEE Proc. Vis. Image Signal Processing*, vol.145, pp.50-56, 1998.

[26] N.A.Mat-Isa, M.Y. Mashor, Seeded Region Growing Features Extraction Algorithm; Its Potential Use in Improving Screening for Cervical Cancer, *Int. J. Comput. Internet Manage.*, vol.13, pp.61-70, 2005.

[27] H.Y. Wang, S.S. Yu, An Iterative Based Novel Multi-Nucleus Detection Scheme For Protozoan Parasite Microscopic Images. *ICIC Computing*, Information and Control vol.5, 2009.

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)