Malaria Cell Counting Diagnosis within Large Field of View

Li-hui Zou, Jie Chen, and Juan Zhang  
Education Ministry Key Laboratory of Complex System Intelligent Control and Decision, School of Automation Beijing Institute of Technology Beijing, China  
{zoulihui, chenjie, zhjuan}@bit.edu.cn

Narciso García  
Grupo de Tratamiento de Imágenes, E.T.S.I. Telecomunicación Universidad Politécnica de Madrid Madrid, Spain  
narciso@gti.ssr.upm.es

Abstract—Malaria is one of the most serious parasitic infections of human. The accurate and timely diagnosis of malaria infection is essential to control and cure the disease. Some image processing algorithms to automate the diagnosis of malaria on thin blood smears are developed, but the percentage of parasitaemia is often not as precise as manual count. One reason resulting in this error is ignoring the cells at the borders of images. In order to solve this problem, a kind of diagnosis scheme within large field of view (FOV) is proposed. It includes three steps. The first step is image mosaicing to obtain large FOV based on space-time manifolds. The second step is the segmentation of erythrocytes where an improved Hough Transform is used. The third step is the detection of nucleated components. At last, it is concluded that the counting accuracy of malaria infection within large FOV is finer than several regular FOVs.

Keywords—malaria diagnosis; cell counting; image mosaicing; Circle Hough Transform

I. INTRODUCTION

Malaria is one of the most common infectious diseases and an enormous public health problem. There were an estimated 190–311 million clinical episodes of malaria, and 708,000–1,003,000 deaths in 2008 [1]. It becomes the 5th cause of death from infectious diseases worldwide in low-income countries.

Malaria is caused by protozoan parasites of the genus Plasmodium. Four species of the plasmodium parasite infect humans, which are Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale, and Plasmodium malariae. The most serious and virulent forms of the disease are caused by Plasmodium falciparum which contributes to the majority of deaths associated with the disease. Others only cause milder diseases in humans. Each malaria parasite life cycle can be divided into four stages which are merozoites, ring stage trophozoites, mature stage trophozoites, and then schizonts or gametocytes if sexual erythrocytic stages [2]. Rapid and accurate diagnosis facilitates prompt treatment, and it is an essential requirement to control malaria. The most economic, preferred, and reliable diagnosis is microscopic examination of blood films, especially the thin blood film examination which is the gold standard in diagnosis of malarial infection. Generally speaking, the diagnosis of malaria involves the following two or more tasks [3]:

• To determine the percentage of parasitaemia;

• To identify the species and the life-cycle stages of the detected parasites;

Usually, these jobs are taken by experienced microscopists manually. However, it is a time-consuming and tiring process that can be significantly affected by the expertise of the observer and, thus, have variable accuracy. Therefore, an automated image analysis system for accurate, fast and reliable determination of parasitemia is desirable.

Diagnosis of Malaria based on digital image analysis is totally a new research project started only in recent years [4]. Dealing with these two tasks above, blood smear image analysis has been tackled by using conventional image processing techniques like morphology, edge detection, region growing, thresholding, and pattern recognition techniques etc., which all have shown certain degrees of success with respect to the used sample images [5]. For example, a segmentation scheme for Red Blood Cells (RBCs) and parasites based on HSV color space was presented in [6] via detecting dominant hue range and calculating optimal saturation; in [7], the algorithm is based on edge detection and splitting large clumps made up from erythrocytes edge linking; in [8], it analyzes the infected blood cell images using morphological operators to segment cell images and to classify the parasites; F. B. Tek et al. proposed a novel binary parasite detection scheme that is based on a modified K nearest neighbor (KNN) classifier and compared three different classification schemes for the identification of the infecting species and life-cycle stages [9].

Here, in our research we mainly focus on the first task, determining the percentage of parasitaemia, because it is the most essential and time-consuming item in the diagnosis of malaria. The recommended procedure is by counting the numbers of parasitised red cells and all RBCs in a thin blood film and expressing the number of infected cells as a percentage of the RBCs. Some image processing algorithms to automate the diagnosis of this item on thin blood smears are developed, but the diagnosis results are sometimes higher than the manual counts as shown in the statistical results in [7]. One of the reasons is supposed to be that the cells at the borders of the images are often ignored. Besides, the statistic results in a single regular FOV can not reflect the global distribution of parasitaemia ratio due to the limitation of the amount of cells. In order to improve this statistical accuracy, a processing scheme within large FOV is proposed in this paper. The analysis process can be divided into three steps,
image mosaicing, segmentation of erythrocytes, and detection of nucleated components. Particularly, in the first step, an optimal image mosaicing method which is based on space-time manifolds is presented to obtain large FOV. In the second step, an improved Hough Transform is used to segment the erythrocytes to get the whole number of RBCs. The third step is to detect the nucleated components mainly based on the basic difference between the blue and green channels. The work flow of our processing is shown in Fig. 1.

II. PROPOSED SCHEME

A. Acquisition of Large Field of View

A general image acquisition mode for obtaining large FOV is shown in Fig. 2. There are parts of horizontal and vertical overlaps between neighboring images. Then a high-resolution mosaic with large FOV is obtained by image mosaicing technique. Considering the convenience for the operators, we recommend to shoot a video sequence with high-overlapped ratio between neighbor frames, which is also beneficial to create smooth mosaics. The original video sequence can be obtained through the trinocular biological microscope which can mount a video camera equipment to get digital images by moving the sample film smoothly along the travelling stage. The image acquisition system can be constructed under the simplest observation conditions without knowing explicitly the full motion or the calibration of the camera.

B. Image Mosaicing

Image mosaicing is an essential step for improving the counting precision in the proposed scheme. It can solve the limitation of the FOV. The performance of image mosaicing algorithm will affect the following steps directly. In order to avoid additional introduced artifacts as much as possible, a smooth mosaic is required, not only smooth in color intensities, but also in geometric structure.

The research achievements in image mosaicing are fruitful during recent years [10]. Taking account of dense image capturing and the motion mode of the sample film in our application, the manifold-based image mosaicing technique [11] is more appropriate to solve the problems for video sequences under general motion conditions. A specific flow chart for image mosaicing is proposed in our scheme shown as Fig. 3. Since seldom motion and calibration information is known from an external device, we compute the motion parameters between neighboring frames and align the sequence by using phase correlation method. If great brightness differences caused by the auto-exposure or other natural factors exist among the neighboring images, they are eliminated by integrating with the motion estimation results to guarantee the smoothness in color intensities. And then the whole sequence whose brightness is already uniform is stitched into a high resolution image by using the space-time manifold-based mosaicing technique which utilizes the motion information among the frames more efficiently and ensures the smoothness in geometric structure of the synthetic mosaic.

- Motion estimation based on phase correlation

The efficient phase correlation method is used to determine the displacement \((u, v)\) between two overlapped images. It is based on the translation property of the Fourier transform, that is, shifting a signal in time domain introduces linear phase difference in frequency domain.

Assume that there is a displacement \((u, v)\) between image \(I_1(x, y)\) and image \(I_2(x, y)\), and \(I_2(x, y) = I_1(x-u, y-v)\). The Fourier transforms of these two images, \(G_1(\xi, \eta)\) and \(G_2(\xi, \eta)\), satisfy the following relationship:

\[
G_2(\xi, \eta) = e^{-j2\pi(\xi u + \eta v)} G_1(\xi, \eta)
\]  

Hence the normalized cross-power spectrum of two images \(I_1\) and \(I_2\) with Fourier transform \(G_1\) and \(G_2\) is given by:

\[
\frac{G_1(\xi, \eta) \times G_2^*(\xi, \eta)}{|G_1(\xi, \eta) \times G_2(\xi, \eta)|} = e^{-j2\pi(\xi u + \eta v)}
\]
where ‘*’ is the symbol of conjugation [12]. The translation property ensures the phase of the cross-power spectrum is equivalent to the phase difference between the images. The inverse Fourier transform of both sides of (2) can generate an impulse function δ at the point \((u, v)\). Therefore, the displacement of these two images, \((u, v)\), can be obtained by locating the peak of the inverse Fourier transform result. Additionally, if needed, the extended phase correlation method [13] can handle more general case, not only translation but also rotation and scale between \(I_1(x,y)\) and \(I_2(x,y)\).

One thing should be noted that other fast motion computation algorithms can be utilized too, as long as they can satisfy the demand of the processing efficiency, such as iterating Lucas-Kanade optical flow under Gaussian pyramid strategy for the motion estimation [14].

- Eliminating brightness differences

It is a preprocessing for the purpose of smoothing the color intensities for the mosaics. Firstly, adjacent frames are aligned by the motion estimation results. It is deemed that a step change in brightness is occurred around the current frame \(I_t\) if the average brightness difference \(\Delta_{bt}\) of the overlapped region between the current frame \(I_t\) and the previous frame \(I_{t-1}\) exceeds a threshold \(T\). Then the current frame \(I_t\) is taken as a brightness reference image \(I_{ref}\). The rule can be described as follows:

\[
\text{if } |\Delta_{bt}| \geq T, \text{ then } I_{ref} = I_t \tag{3}
\]

\[
\Delta_{bt} = \frac{\text{mean}}{\forall t, x, y \in K} \left( I_t(x_{t-1} + u_t, y_{t-1} + v_t) - I_{t-1}(x_{t-1}, y_{t-1}) \right) \tag{4}
\]

The color compensation can be conducted by adjusting (increasing or decreasing) all the previous or the following frames (depending on the position of the reference image in the whole sequence, if \(t < N/2\), the previous, otherwise, the following) with the difference value \(\Delta_{bt}\), i.e.,

\[
I' = I + \Delta_{bt} \begin{cases} 
\text{if } t < N/2, & i = (1, \ldots, t) \\
\text{if } t \geq N/2, & i = (t, \ldots, N) 
\end{cases} \tag{5}
\]

It allows a brightness discontinuity between the adjusted image \(I'\) and the reference image \(I_{ref}\) no more than \(T\). If multiple brightness reference images are involved, the compensation can be handled iteratively until the global brightness difference under the threshold \(T\).

- Space-time manifold-based mosaicing

The sprite of manifold-based mosaicing technique is inspired from line scanning using a 1D linear CCD camera whose imaging process can be modeled by a multi-perspective projection. They differ on the width and the shape of the scanning line. The “scanning line” in manifold-based mosaicing is not just one single line, but several lines which are referred to as a strip. The shape of the strip is determined by the relative motion between the camera and the scanning scenes. In order to receive maximum information of the scanning scenes without missing or duplicating, the stripe requires being perpendicular to the optical flow and its width should be proportional to the motion.

Optical flow can be interpreted as the motion between two neighboring frames taken at times \(t\) and \(t + \delta t\) at every pixel position. Due to the high-overlapped ratio among original images and the characteristic that seldom considerable depth differences are introduced, we can approximately take the motion estimation result above as the optical flow vector for every single point. Moreover, the relationship between neighboring images can be described by affine transform and expressed as:

\[
\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x_t - x_{t-1} \\ y_t - y_{t-1} \end{pmatrix} = \begin{pmatrix} a + bx + cy \\ d + ex + fy \end{pmatrix} \tag{6}
\]

where \((x_{t-1}, y_{t-1})\) and \((x_t, y_t)\) are corresponding points in image \(I_{t-1}\) and image \(I_t\), \((u, v)\) is the optical flow vector as a function of position \((x_t, y_t)\), and \((a, b, c, d, e, f)\) are the parameters of the affine transform \(A\). A cut line \(F(x,y)=0\) which must be orthogonal to the optical flow \((u, v)\) is calculated first. The normal line to \(F(x,y)=0\) whose direction is \((\partial F/\partial x, \partial F/\partial y)\), the same direction as \((u, v)\), fulfills the request. This constraint can be expressed by:

\[
\frac{\partial F}{\partial x} = k \begin{pmatrix} u \\ v \end{pmatrix} = k \begin{pmatrix} a + bx + cy \\ d + ex + fy \end{pmatrix} \tag{7}
\]

The equation of the cut line is obtained if \(e = c\):

\[
0 = F(x,y) = ax + dy + \frac{b}{2} x^2 + \frac{c + e}{2} xy + \frac{f}{2} y^2 + M \tag{8}
\]

It is a family of lines that are all perpendicular to the optical flow, and \(M\) is used to select a specific line [15]. In our case, the relative movement between the sample film and the camera is smooth, only side-way translation or small vertical motion, the affine transformation \(A\) can take the form \(A = (a, 0, 0, d, 0, 0)\), thus the cut line becomes \(0 = F(x,y)=ax+dy+M\) whose slope is \(-a/d\), and \(M\) is set to be the value for which the line contains maximum number of pixels in the center of the image where the distortions are minimum.

Narrow strips are selected from each image in the time domain to form the final space-time manifold. The schematic diagram of manifold-based image mosaicing approach is shown in Fig.4 [11]. Given the affine transformations \(A_t\) and \(A_{t+1}\), two cut lines \(F_t(x_t,y_t) = 0\) and \(F_{t+1}(x_{t+1},y_{t+1}) = 0\) can be selected respectively. They are taken as the left sides of strips \(S_t\) and \(S_{t+1}\). Furthermore, they determine the right sides \(F_t(x_{t-1},y_{t-1}) = 0\) and \(F_{t+1}(x_{t-1},y_{t-1}) = 0\) of their previous strips.

For example, the right side \(F_{t}(x_{t-1},y_{t-1}) = 0\) of strip \(S_t\) is computed by projecting the line \(F_t(x_t,y_t) = 0\) onto the image \(I_t\) using the transformation \(A_t\). It makes the right side of strip \(S_t\) has the same properties as the left side of strip \(S_t\) in the next image. The same corresponding between the right
side $F^{x}_{y}(x, y) = 0$ of strip $S$ in image $I,$ and the left side $F^{t}_{x}(x, y) = 0$ of strip $S$ in image $I_{t+1}$ is associated by the transformation $A_{t+1},$ etc. This selection of strips can keep the orthogonality to the optical flow. And then these narrow strips from multiple images are cut out and warped together. The warping method is keeping the left side of each strip unchanged and warping the right side to match the left side of the next strip. It makes sure that each strip can be stitched seamlessly with the next strip. The width of each strip is proportional to the motion between the neighbor images. As a result, we can finally get a smooth and continuous mosaic after pasting all the warped cut strips.

C. Segmentation of Erythrocytes

The purpose of this step is to determine the denominator of the parasitaemia percentage, the whole number of RBCs in large FOV. An improved Circle Hough Transform (CHT) is used for detecting RBCs in consideration of the feature that blood cells belong to circular pattern. Compared with other segmentation algorithms for blood cells, CHT is one of the optimal selections for erythrocyte detection due to its efficiency and robustness. It can also handle the segmentation for the highly overlapping and oval cells [16]. Usually, the stained sample of thin blood film represents different performance in RGB color space. The RBCs are much more prominent in green channel than others. Therefore, we chose the green channel of the whole mosaics to segment the RBCs. The pre-processing, such as the enhancement of image contrast via adaptive histogram equalization, is necessary before CHT to reduce the effect of noises.

The CHT aims to find circular formations of a given radius $R$ within an image by mapping the geometric space of images to the parameter space. A circle can be described by (9):

$$(x - a_x)^2 + (y - a_y)^2 = R^2 \tag{9}$$

where $(a_x, a_y)$ are the coordinates of the circle center, and $R$ is the radius of the circle. An example of CHT is shown in Fig. 5 [17]. A set of edge points is indicated by the bold circle. Each edge point in the original image contributes a circle of radius $R$ to the output parameter space, and this contribution is shown by these thin circles. These contributed circles in parameter space intersect at the $(a_x, a_y)$ as a peak that is the center of the original circle in geometric space.

The improved CHT for erythrocyte detection proceeds in the following steps [18]:

**Step 1:** Compute the gradient field of image intensity in green channel using the following equation:

$$\nabla I_b(i, j) = (V_x, V_y)_{(i, j)} = (I_b(i, j) - I_b(i, j - 1), I_b(i, j) - I_b(i - 1, j)) \tag{10}$$

where $\nabla I_b(i, j)$ is the gradient vector at pixel $(i, j).$ And a thresholding on the gradient magnitude is performed before the voting process of CHT to remove the ‘uniform intensity’ image background.

**Step 2:** Convert the gradient field to an accumulation array where the pixel intensity corresponds to the probability of that pixel being the center of a RBC. The construction of the accumulation array which has same dimensions as the gradient field is done by a voting process in which the magnitude of $\nabla I_b(i, j)$ is used as a weight value with a possible radius range of the RBCs.

**Step 3:** Detect the peaks in the accumulation array by a Laplacian of Gaussian (LoG) filter. The LoG filter can fill the transition areas between local peaks with negative values or zeros, which makes it easier to separate the peaks.

D. Nucleated Components Detection

This step is to determine the numerator of the parasitaemia percentage. Another observation among the color channels can be seen that in particular stained nucleated components result in distinctively high intensity values in the blue channel, while the same nucleated components in the green channel exhibit very low intensity values. Therefore, the differences between the blue $I_b$ and the green $I_g$ color intensity channels are utilized to stretch the contrast for visual perception and emphasize nucleated objects [5]:

$$\Delta_{bg} = I_b - I_g - \min_{\forall i,j \in \text{RBCs}} \{I_b - I_g\} \tag{11}$$

If $\Delta_{bg}$ exceeds certain threshold, then mask it as a possible nucleated component.

In order to identify the infected RBCs, the binary masks of nucleated component and the RBCs are used as follows. Since we can get every RBC mask which is separated with each other via the improved CHT, erode the RBC mask by applying a disk-shaped structural element of radius two to
obtain the inner region of the RBC where the nucleated components are possibly located. Then, the estimation of parasitemia for the separately occurring RBC can be accomplished straightforwardly by overlaying the binary masks of nucleated components and the RBC. Additionally, a cell is deemed to be infected if the area of its nucleated component masks exceeds a predefined threshold for removing the false positives.

III. EXPERIMENT RESULTS AND DISCUSSIONS

In our experiments, to be concise, instead of the manipulation on a real image acquisition system, the video sequences were gathered from a simulated experiment platform shown in Fig. 6. It uses a projection screen to simulate the large FOV and a camera with a small FOV to get video sequences by moving along the screen. The camera path is parallel to the screen to simulate the motion mode of the real microscopic image acquisition system.

Two examples of malaria cell counting within the simulated large FOV are given to prove the validity of our philosophy. Fig. 7 (a) shows several original images from the video sequence, (b) is the image mosaicing result, (c) is the accumulation array in pseudo-color after CHT, (d) is the detection result of nucleated components, and (e) is the final detection result where the infected cells were marked with red ‘+’ at their center positions. Another example is shown in Fig. 8. And the counting results compared with regular FOVs are shown in Table 1. The statistic results of regular FOVs were computed by counting several original images with minimum overlapped areas separately and adding them together. It showed that the mosaicing results were satisfied and under the proposed scheme the counting accuracy within large FOV was finer than regular FOVs. The counting results of RBCs in large FOV were close to the true numbers, which shows that the improved CHT is an optimal algorithm for segmentation of erythrocytes. However, the method for detecting nucleated components was not so effective, see for example Fig. 8 where the detected infected cells were less than the true number due to the limitations of the acquisition process. In this kind of simulated experiment environment, the image quality decreased greatly due to the indirect imaging, especially affected by the uneven illumination as seen in Fig. 7(a) and Fig. 8(a) that the center areas of source images were brighter than the borders, which also made the brightness of generated mosaic images non-uniform. It brought difficulties for the analysis of the final mosaic image as the original features in different color channels were changed, and led to the failure of nucleated component detection sometimes. Nevertheless, under the same detecting situation for nucleated components, it still can be seen that the statistical results of parasitemia percentage in large FOV were much more precise than those of several regular FOVs together.

As a matter of fact, there are various challenges to image analysis process in the practical microscopic malaria diagnosis system. Many inherent factors during the image acquisition process, such as different conditions of microscope components, different camera settings and so on, will influence the image quality. In addition, the differences in specimen preparation in terms of smearing, staining and
cleanness will cause variations as often as the imaging conditions. All these factors can bring technical challenges to every step in the process scheme. Exposure difference and motion blur are main problems to image mosaicing. Dealing with the situation of cells overlap and cluster is the biggest challenge to the segmentation of erythrocytes, and the appearance of nucleated components can be affected by both imaging conditions and specimen preparations. Addressing these variations can contribute to the robustness of the system.

IV. CONCLUSIONS

There are many advantages of digital image analysis for helping the diagnosis of malaria. It can greatly liberate the microscopists from the manual time-consuming and tiring work. In our research work, one of the most important contributions must be highlighted is that we propose the scheme of malaria diagnose within large FOV. It can reduce the error of the total number of RBCs, which is caused by ignoring the cells at the borders of images, and improve the statistic accuracy for the percentage of parasitaemia. Imagine that the statistical cells are 10 or 100 times larger, the improved accuracy is appreciable. It makes image mosaicing become a meaningful work in the diagnosis of malaria.

ACKNOWLEDGMENT

This work is supported by Chinese National Science Fund for Distinguished Young Scholars, No. 60925011. Also, Li-hui Zou wishes to thank the Sino-Spanish Student Exchange Program of the Universidad Politécnica de Madrid for a personal research grant.

REFERENCES