Detection of melanocytes in skin histopathological images using radial line scanning

Cheng Lu [a,*], Muhammad Mahmood [b], Naresh Jha [c], Mrinal Mandal [a]

[a] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 2V4
[b] Division of Anatomical Pathology, Walter Mackenzie Health Sciences Centre, Edmonton, Alberta, Canada T6G 2B7
[c] Department of Oncology, University of Alberta, Cross Cancer Institute, Edmonton, Alberta, Canada T6G 1Z2

ABSTRACT

In the diagnosis of skin melanoma by analyzing histopathological images, the detection of the melanocytes in the epidermis area is an important step. However, the detection of the melanocytes from the epidermis area is difficult because other keratinocytes that are very similar to the melanocytes are also present. This paper proposes a novel computer-aided technique for detection of the melanocytes in the epidermis area of skin histopathological images. An adaptive threshold technique is first applied to segment all the keratinocytes in the image. In order to distinguish the melanocytes from other keratinocytes, a novel technique based on radial line scanning is proposed to estimate the halo region of the melanocytes. Based on the estimated halo region of all the nuclei, an area ratio of estimated halo region and the nuclei is used to detect the melanocytes from all the keratinocytes. Experimental results on 40 different histopathological images of skin tissue containing 341 melanocytes show that the proposed technique provides a superior performance.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Skin cancer is the most frequent and malignant types of cancer [14] and melanoma is the most aggressive type of skin cancer. According to a recent article, approximately 70,000 people are diagnosed with melanoma skin cancer, and about 9000 die from it in the United States every year [22]. The early detection of malignant melanoma will help to lower the mortality from this cancer. Approaches to melanoma diagnosis have dynamically evolved during the last 25 years [18]. Although there are many new emerging techniques, e.g., confocal microscopy [16], which could provide satisfactory performance, pathological examination remains the gold standard for diagnosis as the histopathology slides provide a cellular level view of the disease [10].

Traditionally, the histopathology slides are examined under a microscope by pathologists. The diagnosis decisions are then made based on their personal experience. However, this judgement is subjective and often leads to intra-observer and inter-observer variability [7,19,8,1,4]. For example, it has been reported that in the diagnosing of follicular variant of papillary carcinoma (FVPC), the inter-observer agreement on benign and malignant diagnoses is only 27% from 6 experts on 15 cases, and the intra-

observer agreement range from 17% to 100% [4]. Ruijter et al. [19] stated that at least 17% of all the grading errors result from the misinterpretation of the pathologists. To address this problem, automated computational tools are needed which can provide reliable and reproducible objective results.

In melanoma diagnosis, the segmentation and detection of the melanocytes in the epidermis area is an important step before the diagnosis is made. If the melanocytes can be found correctly, architectural and cellular features (e.g. size, distribution, location) can then be used to grade or determine the malignancy of the melanotic skin tissue.

In the epidermis area of skin, a normal melanocyte is typically a cell with a dark nuclei, lying singly in the basal of epidermis. However, in the melanoma or nevus, the melanocytes are abnormally growing, and can be distributed in the middle layer of epidermis. The digitized histopathological image we used in this work is stained with haematoxylin and eosin (H&E). Three examples of the skin epidermis image are shown in Fig. 1(a)–(c). The cell nuclei are observed as dark blue whereas the intra-cellular material and cytoplasm are observed as bright pink. Note that the bright seed points indicate the location of melanocytes whereas other nuclei are the other keratinocytes. It is observed that the differences between melanocytes and other keratinocytes are the surrounding region. In the case of melanocyte, it appears to lie in a brighter halo-like region and retracted from other cells, due to the shrinkage of cytoplasm [23]. One close up example of the melanocyte is shown in Fig. 1(d), where the outer dotted contour

* Corresponding author. Tel.: +1 780 492 0294; fax: + 1 780 492 1811.
E-mail addresses: lcheng4@ualberta.ca (C. Lu),
Nausherwan.Mahmood@albertahealthservices.ca (M. Mahmood),
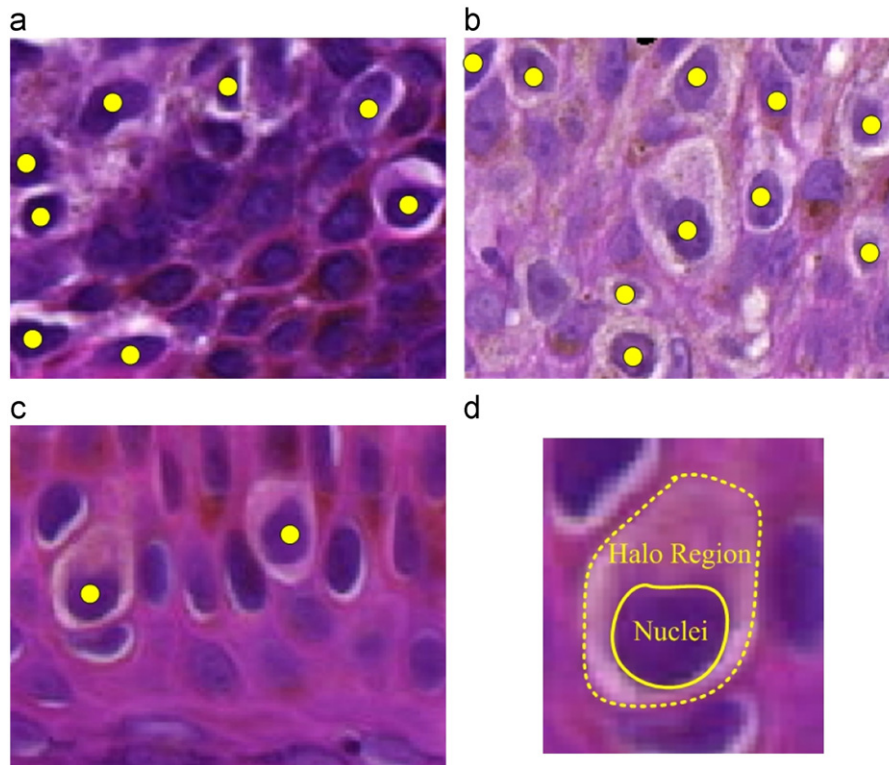Naresh.Jha@albertahealthservices.ca (N. Jha), mmandal@ualberta.ca (M. Mandal).

Fig. 1. Melanocytes in epidermis area from different skin tissues. Inter- and intra-image variations are observed in terms of the color. These images are sampled from the skin digitized slide under 30 × magnification. In (a)–(c), the bright seed points indicate the location of melanocytes whereas other nuclei are keratinocytes. (d) A close up image of a melanocyte. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

represents the halo region and the inner solid contour represents the nuclei. In contrast, the other keratinocytes are closely contact with the cytoplasm and it has no or little brighter area. The brighter halo-like region of the melanocyte is an important pattern for differentiation of the melanocytes and other keratinocytes. In this work, we refer this pattern as "MPattern" and the bright region around the melanocytes as *halo regions*.

Several works have been conducted on the segmentation or detection of various biological components in the histopathological image using image processing techniques such as thresholding [6,17,12] and watershed [3]. These techniques typically fail when considerable intensity variations are present in the images. By incorporating the image color, texture and shape information, Naik et al. [15] proposed to segment the nuclei using the Bayesian classifier. Sertel et al. [20] computed the probability map of karyorrhexis cells based on the estimated likelihood function, and the cell nuclei are then segmented using thresholding. Although these techniques have been reported to provide good performance, the performance is sensitive to the training samples.

In the melanocytes detection problem, the main difficulty is how to differentiate the melanocytes and other keratinocytes in the skin epidermis area. Similar problem has been addressed by Basavanhally et al. [2] in breast cancer diagnosis where the lymphocyte nuclei are differentiated from the cancer cell nuclei in H&E stained histopathological images. In their work, the basic assumption to differentiate two kinds of cells is based on the domain knowledge regarding the nuclei size, intensity of the nuclei and spatial proximity. However, in skin histopathological images, the size of melanocytes are very similar to that of other keratinocytes. In addition, the intensity value of the melanocytes and other keratinocytes are very close to each other. Therefore, the domain knowledge used in breast cancer [2] does not work well in the case of melanocytes detection.

There is another closely related work in the literature where the keratinocytes nuclei are segmented in the skin epidermis area [9]. In this work, a threshold is decided based on the assumption that cell nuclei covers approximately the darkest 20% pixels in the image. The pixels whose values are less than the threshold are labeled as nuclei regions. Morphological operations are then used to do the refinement. However, this global threshold based technique only works under the assumption that there are no intensity variations in the image and usually generate the under-segmentation results (many of the nuclei are grouped together). Also, there is no attempt to differentiate the melanocytes and other keratinocytes.

Template matching technique is a popular technique in computer vision for pattern detection. Naik et al. [15] have used four binary elliptical templates with different major and minor axes to detect the nuclei in breast cancer histopathological images. It is observed in Fig. 1 that the melanocytes typically have low intensity value while its spatial surrounding space presents a brighter intensity value. It may be possible to detect the melanocytes using a template matching technique with templates that have similar appearance of the melanocyte. However, several difficulties need to be addressed. First, the size of the template is hard to decide due to the size variations of the melanocytes even under the same magnification level. In the case of cancer skin, the melanocytes are abnormal and larger than that in the case of normal skin or nevus skin. Second, the intensity level of the template is hard to determine. Therefore, it is difficult to decide a 'good' template to match the melanocyte patterns.

In this paper, we propose a novel technique to detect the melanocytes in the skin epidermis area. To our best knowledge, this is the first automated technique for the detection of the melanocytes in histopathological image of skin tissue. This technique operates on reliable quantitative measures and provides

objective and reproducible information complementary to that of a pathologist. Such quantitative analysis of melanocytes is important for clinical applications, as well as for research purpose.

The organization of this paper is as follows. The proposed technique is described in Section 2, followed by the performance evaluations in Section 3. The conclusion is presented in Section 4.

2. The proposed technique

The schematic of the proposed technique for melanocytes detection is shown in Fig. 2. It is observed that there are three modules. At first, the nuclei in the image are segmented using an adaptive thresholding technique which incorporates the domain knowledge. In the second module, the halo region is estimated for each pre-segmented nuclei region by using the radial line scanning (RLS) algorithm. Finally, the melanocytes are detected based on the estimated halo region and the pre-segmented nuclei region. Details of these three modules are presented in the following.

2.1. Segmentation of the nuclei

Before performing the melanocytes detection, we first segment all the nuclei in the image. Considering the intensity variations, an adaptive threshold technique is applied to segment the nuclei regions [13]. In order to reduce the influence from undesirable variations in the image and to make the nuclei region homogenous, a hybrid gray-scale morphological reconstruction method is first applied. The hybrid gray-scale morphological reconstruction method mainly consists of the Opening-by-Reconstruction and Closing-by-Reconstruction [5]. Next, a local region adaptive threshold selection method is used to achieve the segmentation. In this method, the local threshold is determined by minimizing a pre-defined cost function which incorporates two domain specific knowledge: the size and the shape of the nuclei. Fig. 3(a) shows an original image containing three nuclei, and the melanocyte is label with letter 'M'. In Fig. 3(b), the segmented nuclei regions are shown as the white regions in a binary image.

2.2. Estimation of the halo regions using RLS

After the segmentation of nuclei, we are now ready to detect the melanocytes from the pre-segmented nuclei. In order to capture the 'Mpattern', we propose the RLS algorithm, which makes use of the radial lines originating from the centroid of a nuclei. The RLS algorithm consists of four steps. The first step is to initialize the radial center and radial lines for each nuclei regions. These radial lines will be used for measuring the information from the image in order to estimate the halo region. The second step generates the gradient map of the image, which provides the edge information for the radial lines. In the third step, the halo region of each nuclei is estimated based on the gradient information measured by the radial lines. In the final step, the overlapping region problem is resolved. These four steps are described in the following.

2.2.1. Initialization of radial center and radial lines

In this step, we first calculate the centroid of each pre-segmented nuclei. Let the centroid coordinate be denoted by (C_x, C_y). Using the centroid as a radial center, a set of radial lines
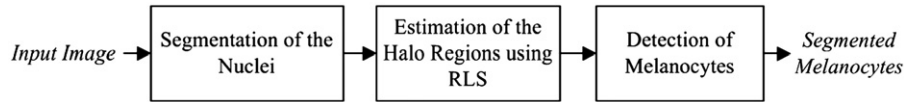

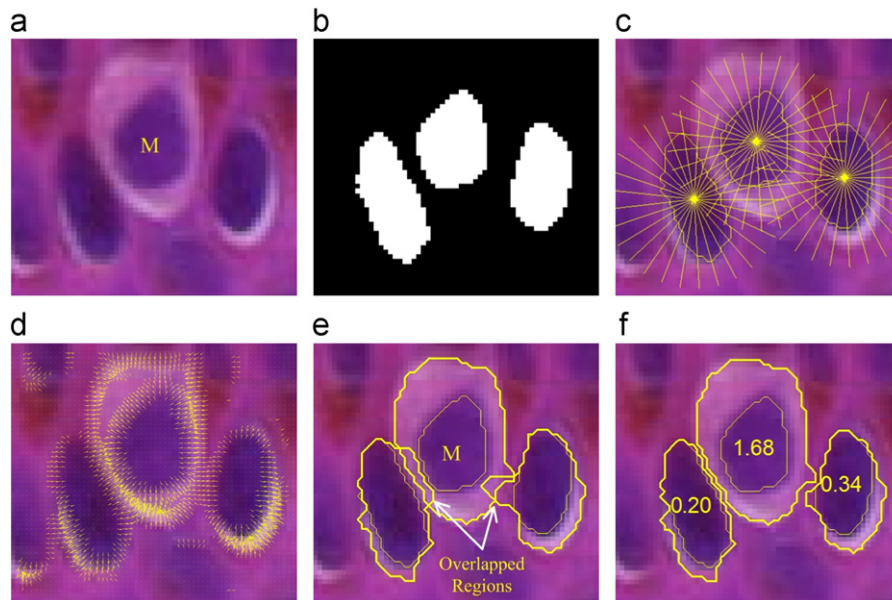**Fig. 2.** The schematic of the proposed melanocyte detection technique.

Fig. 3. Illustration of the RLS algorithm. (a) The original image with three nuclei regions where the melanocyte is indicated with letter 'M'. (b) The segmented nuclei regions. (c) The radial lines for each nuclei region (nuclei regions are shown with thick contours). (d) The suppressed smoothed gradient map. (e) The estimated halo regions without resolving the overlap. (f) The overlap resolved results of (e). The value of the parameter R_{HN} discussed in Section 2.3 is shown inside the nuclei region as well.

with different radial angles are calculated as follows:

$$\theta_i = \frac{2\pi}{k}(i-1), \quad i = 1 \ldots k, \tag{1}$$

where k indicates the number of radial lines. An illustration of the radial center (C_x, C_y) and its radial lines θ_i, $i = 1 \ldots k$ is shown in Fig. 4(a). Note that the nuclei contour is shown as the thick solid contour whereas the dotted circle indicates the maximum range of the radial line (denoted as R_{max}). The solid point in the center indicates the radial center (C_x, C_y) and a set of 16 radial lines are shown as the thick solid lines. In this paper, the radial lines are denoted by its radial angle θ. For example, the first and second radial lines θ_1 and θ_2 are shown in Fig. 4(a) and the radial angles for these two radial lines are 0 and $\pi/8$, respectively. Denote the nuclei boundary point at given radial angle θ_i as $b(\theta_i)$. In Fig. 4(a), $b(\theta_1)$ and $b(\theta_2)$ indicate the nuclei boundary point on given radial lines θ_1 and θ_2, respectively.

Note that R_{max} specifies the maximum range of the radial line and it is identical for all the radial lines, i.e., $R_{max}(\theta_i) = R_{max}$. In Fig. 4(a), the maximum range R_{max} of a radial line is labeled.

A radial line consists of R_{max} points and is index from 1 to R_{max}. We define a valid scanning range for each radial line in order to reduce the computational time. We define $R_{min}(\theta_i)$ as the distance from the radial center to the nuclei boundary point $b(\theta_i)$. Note that the minimum valid scanning range maybe different for each radial line. Now, let m be the index of the point on the radial line, and the valid scanning range of the index should range from $R_{min}(\theta_i)$ to R_{max}, i.e., $R_{min}(\theta_i) \leq m \leq R_{max}$. An illustration of the above mentioned notation is shown in Fig. 4(b). Note that one of the radial lines is shown as a thick solid line. The hollow point on the radial line represents the mth valid scanning point in the range of $[R_{min}(\theta_i), R_{max}]$. The shadow region shown in Fig. 4(b) indicates the valid scanning range. The radial centers and radial lines for Fig. 3(a) is shown in Fig. 3(c). These radial lines will be used later to measure the information from the image.

2.2.2. Gradient map generation

In the proposed technique, we estimate the halo region for each nuclei by analyzing the image features on the pre-initialized radial lines. The gradient is a very useful feature in the image that can help to estimate the halo region, and hence in this step, we generate the gradient map for the latter analysis. Denote the image as I and its 2-D coordinates as (x,y). In order to reduce the noise, a smooth image I_s is obtained by convolving the image with a Gaussian kernel $G_\sigma(x,y) = \sigma^{-1/2}e^{-(x^2+y^2)/4\sigma}$. In other words, $I_s(x,y) = G_\sigma(x,y) * I(x,y)$. The gradient magnitude map $\mathcal{G}_m$ and gradient direction map $\mathcal{G}_d$ are then generated using the following equations:

$$\mathcal{G}_m(x,y) = \|\nabla I_s(x,y)\| = \sqrt{\left(\frac{dI_s(x,y)}{dx}\right)^2 + \left(\frac{dI_s(x,y)}{dy}\right)^2}, \tag{2}$$

$$\mathcal{G}_d(x,y) = arctan\left(\frac{dI_s(x,y)}{dy} \middle/ \frac{dI_s(x,y)}{dx}\right). \tag{3}$$

To suppress the weak gradients, the following equations are used:

$$\mathcal{G}'_m(x,y) = \begin{cases} \mathcal{G}_m(x,y) & \text{if } \mathcal{G}_m(x,y) \geq \text{mean}(\mathcal{G}_m(x,y)), \\ 0 & \text{otherwise}, \end{cases} \tag{4}$$

$$\mathcal{G}'_d(x,y) = \begin{cases} \mathcal{G}_d(x,y) & \text{if } \mathcal{G}_m(x,y) \geq \text{mean}(\mathcal{G}_m(x,y)), \\ 0 & \text{otherwise}. \end{cases} \tag{5}$$

Eqs. (4) and (5) suppress the gradients whose magnitude are less than the mean gradient magnitude. The gradient map $\mathcal{G}'(x,y)$ is shown in Fig. 3(d) where the gradient (shown as arrows) is super imposed on the image.

2.2.3. Estimation of the halo regions

The melanocytes can be distinguished from other keratinocytes based on their bright halo-like regions. In this step, we use the radial lines and the pre-computed gradient map to detect the boundary points for the halo regions estimation. On each radial line with radial angle θ_i, there are two kinds of boundary point: the inner boundary point and outer boundary point. We assume that the inner boundary points are located at the boundary of the pre-segmented nuclei. Therefore, in this section, we aim to detect the outer boundary point, denoted as $BP(\theta_i)$, for the halo region. All the outer boundary points $\{BP(\theta_i), i = 1, \ldots, k\}$ will then form the outer halo contour which can constrain a halo region. In Fig. 1(d), the dotted contour corresponds to the outer boundary contour that consists of outer boundary points whereas the solid contour corresponds to the inner boundary contour that consists of inner boundary points. Ideally, for the melanocytes, the halo-like regions surrounding the nuclei are segmented by the detected outer boundary points; whereas for the other keratinocytes, the detected outer boundary points only encompass the nuclei themselves, i.e., the outer boundary points are equal to the inner boundary points since there is no halo region.

Before introducing how to estimate the boundary points of the halo region, let us first analyze the gradient information around the halo region of a typical melanocyte (see Fig. 5). It is observed that the gradient map (shown as small arrows) is superimposed onto the image where the gradient magnitude is represented by the length of the arrow and the gradient direction is represented by the direction of the arrow. Because the intensity of the halo region is higher than the neighboring regions, the gradient directions point towards the halo region both from the nuclei and from the outer region. The dotted contour indicates the ground truth halo region's outer contour which consists of several outer boundary points. It is noted that the outer boundary points should have gradient direction that opposite to the radial line. For
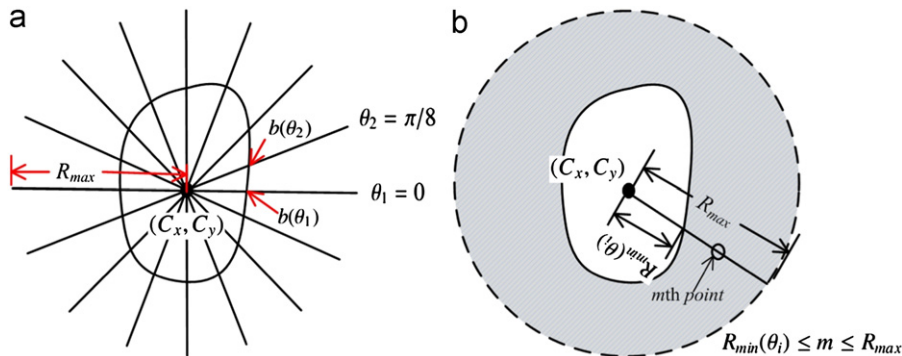

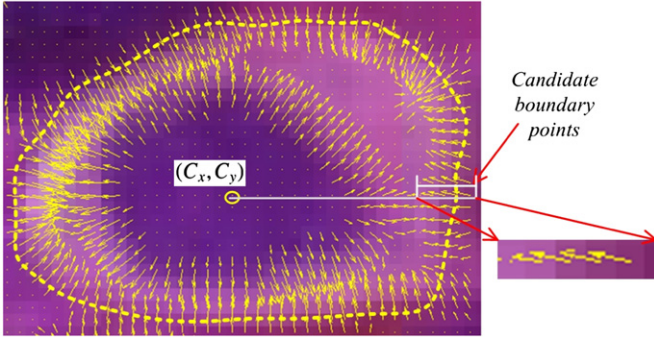**Fig. 4.** Illustration of the notation of radial center and radial line.

Fig. 5. A close up example of a melanocyte with gradient map. The gradient map (shown as small arrows) is superimposed onto the image. The dotted contour indicates the ground truth halo region's outer contour. An enlarged version of the boundary is shown in the small image on the right.

Fig. 6. Illustration of the outer boundary point detection on a given radial line.

a given radial line with radial angle θ_i, we denote the gradient direction of a outer boundary point for the halo region as $BP_{\mathcal{G}_d'}(\theta_i)$. Based on the characteristic of the gradient direction, in this paper, we assume that the angle difference between $BP_{\mathcal{G}_d'}(\theta_i)$ and the radial angle of a radial line should satisfy the following condition:

$$\frac{\pi}{2} < |\theta_i - BP_{\mathcal{G}_d'}(\theta_i)| < \frac{3\pi}{2}. \tag{6}$$

In an ideal case, the gradient direction and the radial angle are strictly opposite, i.e., $|\theta_i - BP_{\mathcal{G}_d'}(\theta_i)| = \pi$. In order to allow the variations, we assume that $BP_{\mathcal{G}_d'}(\theta_i)$ to be $\theta_i \pm \pi/2$, or satisfies Eq. (6). In Fig. 5, a radial line with radial angle $\theta_1 = 0$ is shown. On this radial line, the candidate outer boundary points are labeled.

Let $\mathcal{G}_m'(\theta_i, m)$ and $\mathcal{G}_d'(\theta_i, m)$ represent, respectively, the gradient magnitude and direction on the mth point of the radial line θ_i. Based on the above analysis, an outer boundary point is expected to maximize the following cost function:

$$\mathcal{C}(\theta_i, m) = \cos[\alpha(\theta_i, m)] + \mathcal{G}_m'(\theta_i, m), R_{min}(\theta_i) \leq m \leq R_{max}, \tag{7}$$

where $\alpha(\theta_i, m)$ is the angle difference between gradient direction and the direction of the radial line, which is defined as follows:

$$\alpha(\theta_i, m) = |\theta_i - \mathcal{G}_d'(\theta_i, m)|. \tag{8}$$

For each radial line, the outer boundary point for the potential halo region is calculated using the following equation:

$$BP(\theta_i) = \begin{cases} \arg_m \max\{\mathcal{C}(\theta_i, m)\} & \text{if } \exists m, \ \frac{\pi}{2} < \alpha(\theta_i, m) < \frac{3\pi}{2}, \\ R_{min}(\theta_i) & \text{otherwise}. \end{cases} \tag{9}$$

Eq. (9) examining the gradient information (the direction and the magnitude) of each point on the radial line. If there exist one or more points satisfying $\pi/2 < |\alpha(\theta_i, m)| < 3\pi/2$, the one with maximum cost function value $\mathcal{C}(\theta_i, m)$ is estimated as the outer boundary point for the current radial line. If there is no point (on the radial line) satisfying $\pi/2 < |\alpha(\theta_i, m)| < 3\pi/2$, we choose the point corresponding to the minimum valid range as the estimated outer boundary point.

The idea of the outer boundary point estimation using Eq. (9) is further explained in Fig. 6. In Fig. 6(a), six blocks are shown which represent six pixels/points that comprise the valid range of a radial line θ_i (in this case $\theta_i = \pi/2$). We denote each block, from bottom to the top, as the first to sixth pixels on the radial line. On each pixel, we have corresponding gradient information which are represented by the arrows. Note that the gradient direction is indicated by the direction of the arrow whereas the gradient magnitude is indicated by the length of the arrow. It is observed that the third, fifth, and sixth pixel satisfies the condition $\pi/2 < |\alpha(\theta_i, m)| < 3\pi/2$. These pixels are the candidate outer boundary points on the current radial
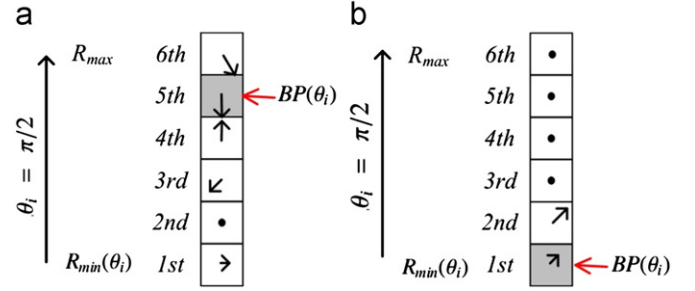
line. In order to determine the final outer boundary point, the cost function value $\mathcal{C}(\theta_i, m)$ is calculated (as shown in Eq. (7)). The pixel corresponding to the maximum cost function value will be chosen as the outer boundary point. In Fig. 6(a), the outer boundary point $BP(\theta_i)$ is the fifth pixel. Fig. 6(b) shows another case, where there is no pixel whose gradient satisfy the condition $\pi/2 < |\alpha(\theta_i, m)| < 3\pi/2$ in Eq. (9). In this case, the outer boundary point $BP(\theta_i)$ is the first pixel, i.e., the pixel corresponding to the minimum valid range $R_{min}(\theta_i)$.

By using Eq. (9), we can estimate accurate halo regions for the melanocytes while reducing the influence of noise. After estimating the outer boundary point on each radial line using Eq. (9), for each nuclei region, the potential halo contour (HC) is estimated by the boundary points as follows:

$$HC = \{(x_{BP(\theta_i)}, y_{BP(\theta_i)})\}, \quad i = 1, \ldots, k, \tag{10}$$

$$x_{BP(\theta_i)} = C_x + BP(\theta_i) \times \cos(\theta_i), \tag{11}$$

$$y_{BP(\theta_i)} = C_y + BP(\theta_i) \times \sin(\theta_i). \tag{12}$$

The estimated halo contours (shown as thick contours) for Fig. 3(a) is shown in Fig. 3(e) where the contours are superimposed onto the original image. Note that the region within the halo contours excluding the nuclei regions are the halo regions we are looking for. Two more examples of the halo region estimation are shown in Fig. 7(a) and (c), where the melanocytes are label with letter 'M' in the figure.

2.2.4. Resolving the overlapped regions

In the skin epidermis, the melanocytes and other kind of keratinocytes are usually clustered with each others. It is observed that in Fig. 3(a), two other keratinocytes and one melanocyte (the middle one) are very close to each other. Such proximity will lead to the overlapping of the estimated halo regions (as shown in Fig. 3(e)). Such overlapping will result in erroneous detection in the latter processing. Note that the overlapping is due to the spatial closeness of the other keratinocytes and the halo region. The gradient information of the halo region is falsely captured by the neighboring keratinocytes which are not melanocytes. Generally, the estimated halo regions of the melanocytes have smooth boundaries and larger area compared to that of the other keratinocytes whereas the estimated halo regions of the other keratinocytes have non-smooth boundaries (see Figs. 3(e) and 7(a) and (c)). Given an estimated halo contour HC, denote A(HC) and S(HC) as the area inside HC and the smoothness measure of the HC. The smoothness of the HC is calculated as

$$S(HC) = \frac{1}{k} \sum_{i=1}^{k} \left| \frac{dBP(\theta_i)}{d\theta_i} \right|. \tag{13}$$
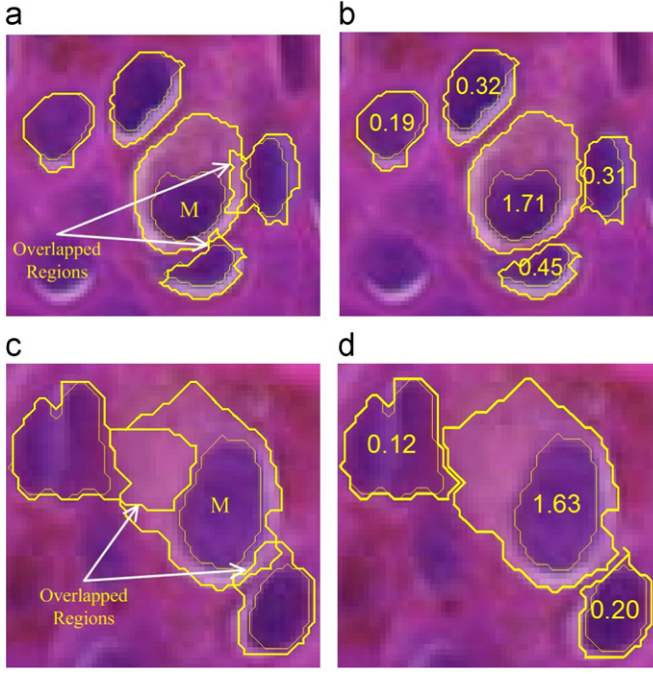
Fig. 7. Two examples of halo-like regions estimation. (a) and (c) show the estimated halo regions without resolving the overlap. The nuclei of melanocytes are labeled with letter 'M'. (b) and (d) are the overlap resolved results of (a) and (c), respectively. The value of the parameter R_{HN} discussed in Section 2.3 is shown inside the nuclei region as well.

Therefore, the ratio of the boundary smoothness and the area (RSA), i.e., $RSA = S(HC)/A(HC)$, is used as a criterion to resolve the overlapped regions. Specifically, for all the estimated halo regions, we first check if there are overlapped. Assume that two estimated halo regions are overlapped and denote these two estimated halo regions as HC_a and HC_b, respectively. The overlapped region is denoted as $HC_{a\cap b} = HC_a \cap HC_b$. We resolve the overlap by using the equation as follows:

$$HC_a = \begin{cases} HC_a & \text{if } \dfrac{S(HC_a)}{A(HC_a)} > \dfrac{S(HC_b)}{A(HC_b)}, \\ HC_a - HC_{a\cap b} & \text{otherwise,} \end{cases} \quad (14)$$

$$HC_b = \begin{cases} HC_a - HC_{a\cap b} & \text{if } \dfrac{S(HC_a)}{A(HC_a)} > \dfrac{S(HC_b)}{A(HC_b)}, \\ HC_b & \text{otherwise.} \end{cases} \quad (15)$$

Intuitively, for the two halo regions that are overlapped, the one with a higher RSA value will take over the overlapped region whereas the other will lose the overlapped region. The result of the overlapped region resolving for Figs. 3(e) and 7(a) and (c) are shown in Figs. 3(f) and 7(b) and (d), respectively. It is observed that the melanocytes correctly take over the overlapped regions which reflect the effectiveness of the proposed method.

2.3. Detection of the melanocytes

As described in Section 1, the melanocytes generally have associate large halo regions whereas the other keratinocytes have no such large halo regions. Some halo region estimation results obtained by the RLS algorithm can be observed in Figs. 3(f) and 7(b) and (d). Based on this observation, we can now use the ratio of the estimated halo region area and the nuclei area, denote it as R_{HN}, as a parameter to differential the melanocytes and other

keratinocytes. The ratio R_{HN} is calculated as follows:

$$R_{HN} = \frac{A_{HR}}{A_{NR}}, \quad (16)$$

where A_{HR} and A_{NR} are the areas of the halo and nuclei regions, respectively. For the melanocyte detection, we select the nuclei regions as the melanocytes if its R_{HN} value is greater than a predefined threshold $T_{R_{HN}}$. In Figs. 3(f) and 7(b) and (d), the corresponding parameter R_{HN} for each nuclei region is show inside the nuclei region. It is observed that the true melanocytes regions have larger R_{HN} values compare to the other keratinocytes regions.

3. Performance evaluations

We have evaluated the proposed technique on 40 different skin histopathology images of epidermis which contains 341 melanocytes. These images are captured under $30\times$ magnification on Carl Zeiss MIRAX MIDI Scanning system.

3.1. Evaluation metrics

3.1.1. Evaluation metrics for melanocyte detection

In the melanocytes evaluation, the melanocytes manually identified by an expert are treated as the ground truths positives for the performance evaluation. The other keratinocytes identified by a pathologist are treated as the ground truths negatives. The number of positives and negatives are denoted as N_P and N_N, respectively. We also denote N_{DO} as the total number of detected objects, N_{TP} as the number of true-positives, N_{FP} as the number of false-positives, N_{TN} as the number of true-negative, N_{FN} as the number of false-negative. The true positive rate (TPR, also known as recall rate) and false positive rate (FPR) are defined as follows:

$$TPR = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100\%, \quad (17)$$

$$FPR = \frac{N_{FP}}{N_{FP} + N_{TN}} \times 100\%. \quad (18)$$

The precision rate (PRE) is defined as follows:

$$PRE = \frac{N_{TP}}{N_{DO}} \times 100\%. \quad (19)$$

3.1.2. Evaluation metrics for halo region estimation

For the halo region estimation evaluation, we randomly selected a set of 50 melanocytes and 50 other keratinocytes as the test set. We then manually labeled the boundaries of the halo region of these cells. These manually labeled boundaries are treated as the ground truth and will be compared with the estimation of the halo regions obtained by the proposed RLS technique.

We define GT as the ground truth halo region for each nuclei, EST as the estimated halo region obtained by the proposed RLS technique. Two area-based evaluation metrics: sensitivity ($\mathcal{A}_{SEN}$), precision ($\mathcal{A}_{PRE}$) are defined as follows:

$$\mathcal{A}_{SEN} = \frac{|GT \cap EST|}{|GT|} \times 100\%, \quad (20)$$

$$\mathcal{A}_{PRE} = \frac{|GT \cap EST|}{|EST|} \times 100\%, \quad (21)$$

where $|\cdot|$ is the cardinality operator. We also use another boundary-based evaluation metric: mean absolute distance (MAD) which is defined as follows:

$$MAD(C^{GT}, C^{EST}) = \frac{1}{M} \sum_{i=1}^{M} [\min \|C_i^{GT} - C_j^{EST}\|], \quad (22)$$

where C^{GT} and C^{EST} are the contour/boundary of the ground truth region and estimated region, respectively. C_i^{GT} and C_j^{EST} represent the ith pixel and the jth pixel on the contour of the ground truth region and estimated region, respectively. M is the total number of pixels on the contour of the ground truth region.

3.2. Parameters selection

To obtain good performance, two parameters that will be used in the proposed technique need to be selected appropriately. These parameters are R_{max} (the maximum range of the radial line) and $T_{R_{HT}}$ (the predefined threshold for discrimination of the melanocytes and other nuclei). For the parameter R_{max}, we chose a set of 20 manually labeled melanocytes and measure the distances from the centroid to the boundary points in each melanocyte. Let the maximum distance be denoted by $dist_{max}$, we then chose the $R_{max} = 2 * dist_{max}$. Based on the biological knowledge, the size of the nuclei in skin epidermis should fall within a desired range. The choice of the R_{max} is to ensure that the radial line can reach all the melanocytes and the surrounding halo regions. For determining the parameter $T_{R_{HT}}$, two images are selected and the nuclei and their halo regions within these images are manually labeled. In addition, the value of R_{HN} is calculated for each nucleus. We then perform the trail-and-error test, i.e., we use a set of possible values for $T_{R_{HT}}$ (from the minimum value of all the R_{HN} to the maximum value of all the R_{HN} in the test images) to calculate the PRE and chose the value that achieves the best

PRE as the parameter $T_{R_{HT}}$. The determined value of R_{max} ($R_{max} = 35$) and $T_{R_{HT}}$ ($T_{R_{HT}} = 0.8$) are then used for the evaluation.

3.3. Evaluation on halo region estimation

Since we use the ratio of the estimated halo region area and the nuclei area, R_{HN} defined in Eq. (16), as a criterion for the melanocyte detection, the estimation of the halo region of the nuclei is important for the proposed technique. In this section, we evaluate the performance on the estimation of halo region.

The halo region estimation performance is shown in Table 1. The average value of all the three evaluation metrics for melanocytes and other keratinocytes are shown in the second row and the third row of Table 1, respectively. It is observed that the proposed RLS technique is able to estimate the halo region of the nuclei correctly. This is expected to provide a good performance in the melanocytes detection.

A few examples of the estimated halo regions for the melanocytes (the first row) and other keratinocytes (the second row) are shown in Fig. 8. In each image, the interested nuclei are put in the middle of the images and their halo regions are shown as the contours. The dotted contours represent the estimated halo regions obtained by the proposed RLS technique whereas the solid contours are the manually labeled halo regions. In the case of melanocytes (Fig. 8(a)–(d)), the halo regions are much larger than the nuclei regions. In the case of other keratinocytes (Fig. 8(e)–(h)), the estimated halo regions are very close to the nuclei boundary, which is desired since there is little or no halo region for other keratinocytes. The proposed RLS technique is able to estimate the halo region correctly, with minor distortions compared to the manually labeled halo region (see Fig. 8(a)–(c), (e)-(h)). These minor distortions will not degrade the final melanocytes detection performance since we will use the region ratio criterion R_{HN} (described in Eq. (16)) to differentiate the melanocytes and other keratinocytes. In Fig. 8(d), due to the staining variation, a considerable amount of the halo region cannot be recovered by the proposed RLS technique. This might

Table 1
Performance evaluation of the halo region estimation.

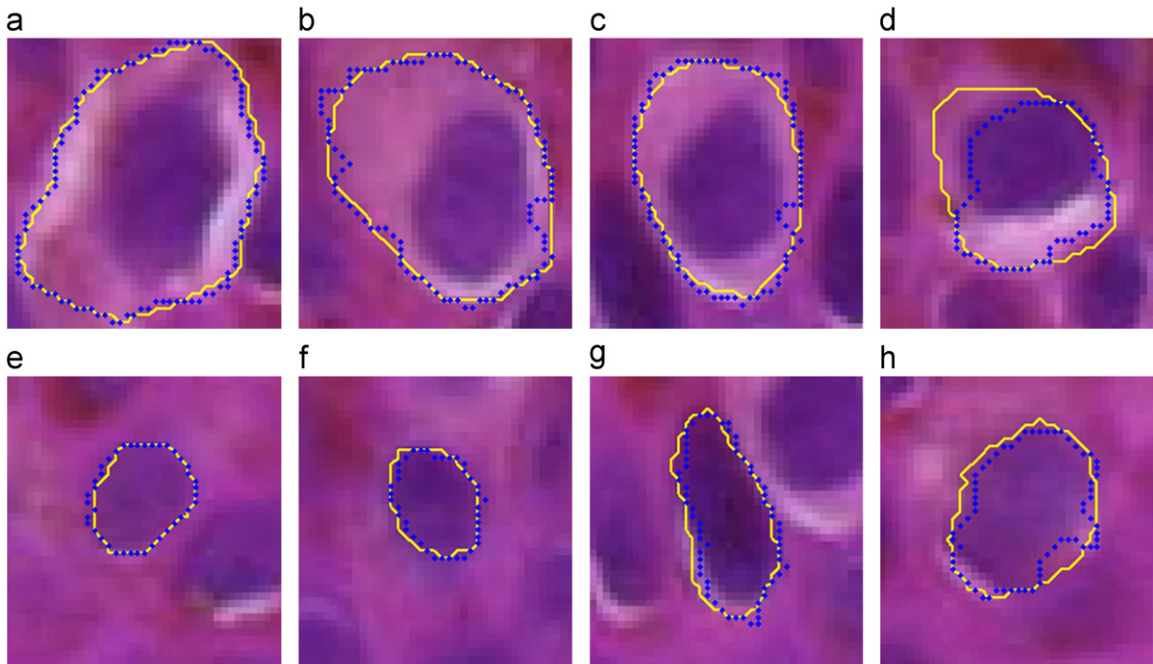| Cell Type	$\mathcal{A}_{SEN}$ (%)	$\mathcal{A}_{PRE}$ (%)	MAD (pixels)
Melanocytes	84.69	93.80	1.51
Other keratinocytes	86.29	96.16	0.72

Fig. 8. Examples of the estimated halo regions compared to manually labeled halo regions for melanocytes and other keratinocytes. The first and second rows show the melanocyte examples and other keratinocytes examples, respectively. The dotted contours represent the estimated halo regions obtained by the proposed RLS technique whereas the solid contours are the manually labeled halo regions.

lead to a miss detection of this melanocyte since the region ratio R_{HN} will have a small value.

3.4. Quantitative evaluation on melanocyte detection

Since this is the first technique to solve the melanocyte segmentation problem, there is no other equivalent technique in the literature that can be compared with the proposed technique. Therefore, we compare the proposed technique with a possible alternative technique, known as template matching (TM). Note that the TM technique is a widely used technique for pattern detection in medical imaging [15,21]. A set of 30 templates are designed to capture the "MPattern" and selected 10 templates are shown in Fig. 9. The templates are designed according to 30 manually selected melanocytes. Note that we design 30 templates with different sizes in terms of the template size (15×15, 20×20, and 25×25 pixels), center heart radii (8, 9, 10, 11, and 12 pixels), and the ring thickness (2 and 3 pixels), in order to create the variations. The intensity for the heart of the template and the ring of the template are assigned according to the mean value on the corresponding regions in the 30 samples of manually selected melanocytes. In the TM implementation, each template is first applied on the image by using the normalized cross correlation (NCC) [11]. Denote the NCC output of the ith template as O_i, and the 2D coordinate in the image as (u, v). The potential pixel (u, v) belong to the melanocyte region $J_i(u,v)$ obtained from the ith template is as follows:

$$J_i(u,v) = \begin{cases} J_i(u,v) & \text{if } O_i(u,v) \geq 0.80, \\ 0 & \text{otherwise}. \end{cases} \qquad (23)$$

The accumulated map O_{acc} for all 30 templates are calculated as follows:

$$O_{acc} = \sum_{i=1}^{30} \sum_{(u,v) \in \Omega} J_i(u,v), \qquad (24)$$

where Ω represents the image domain. The final result is a binary image which is determined using a threshold T_{TM} on O_{acc}.

The performance of the proposed technique is evaluated using the receiver operating characteristic (ROC) and the precision-recall curve (PRC). The ROC and PRC curves of the proposed technique and the TM technique are shown in Fig. 10. The ROC and PRC curves of the proposed technique are generated by varying the parameter $T_{R_{HN}}$ from to 0 to 2 with a step of 0.1. The ROC and PRC curves of the TM technique are generated by varying the parameter T_{TM} from to 0 to 1 with a step of 0.1. It is noted that in the ROC curve, the curve that is close to the upper left corner is better than that is far away from the upper left corner. It is observed from Fig. 10(a) that the proposed technique can provide the TPR around 90% with the FPR at 95%. As for the TM technique, the TM technique provide a poor performance because it is difficult to distinguish the melanocytes from other keratinocytes by using the templates. In the case of the PRC curve, the curve that is close to the upper right corner is better than that is far away from the upper right corner. In Fig. 10(b), it shows that the proposed technique can provide 90% recall rate with around 80% precision rate. As for the TM technique, it usually generate a large number of false detections (results in low precision rate) because the pattern of the other cytological components is very similar to the templates. It is observed from both ROC and PRC curves that the proposed technique provides a superior
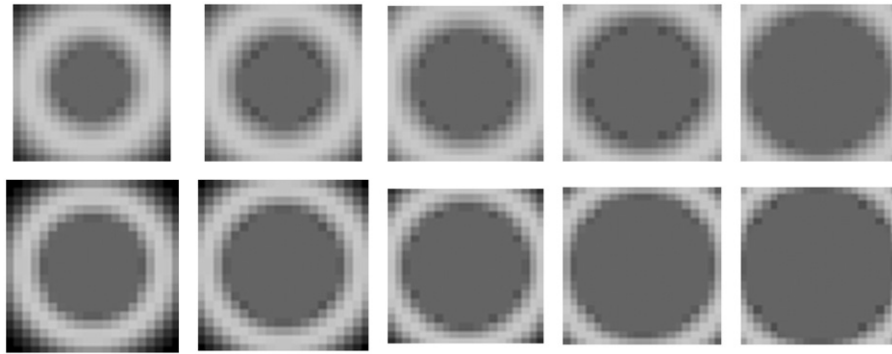

Fig. 9. Ten templates with size 20×20 used in the template matching techniques. Note the variation of the center heart radii and the ring thickness. The templates are magnified for demonstration purpose.
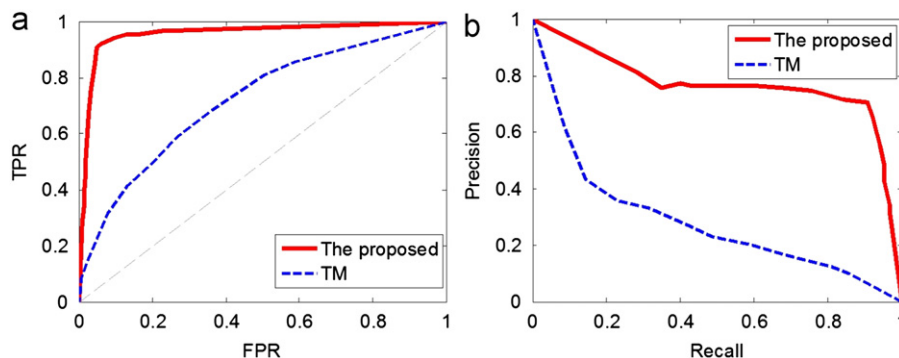

Fig. 10. The ROC and PRC curves of the proposed technique and the TM technique. (a) The ROC curve. (b) PRC curve.
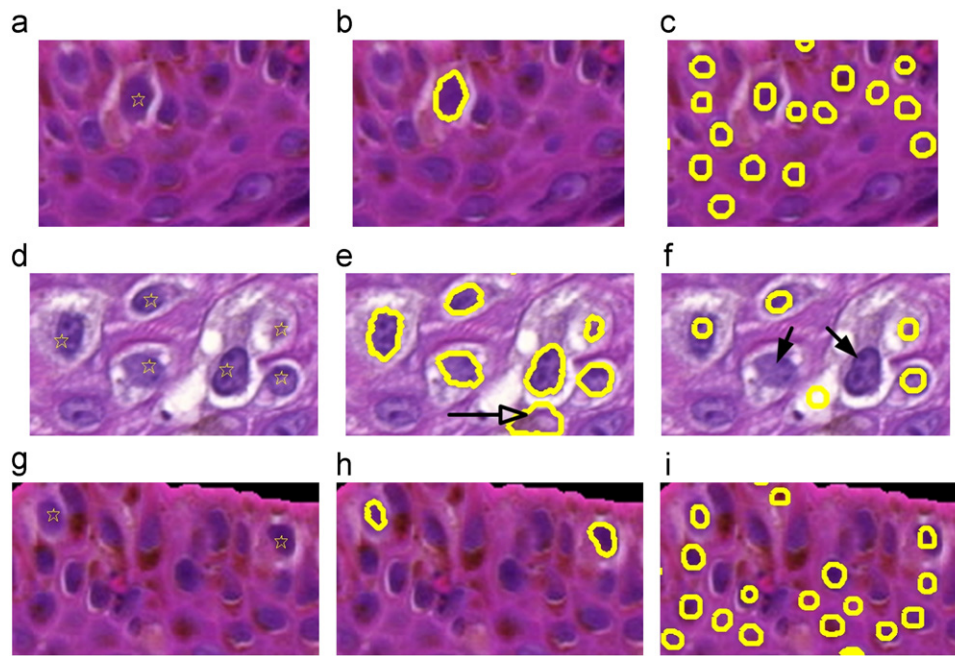
Fig. 11. Three examples for qualitative evaluation. (a), (b), and (g) show three original color histopathological images. The ground truths melanocytes locations are indicated by stars. (b), (e), and (h) show the result provided by the proposed technique. (c), (f), and (i) show the result provided by ML technique. Note that detected regions are presented as the thick contours. The solid arrows in (f) indicate the miss detection by the ML technique whereas the hollow arrow in (e) indicates the false detection by the proposed technique. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

performance compare to the TM technique. Note that the template matching technique is a general technique for the object detection problem, and it could not be expected to provide a better performance compared to a specifically designed technique. Since there is no other available technique for the task at hand, i.e., melanocytes detection in the skin epidermis, we used the TM technique with specifically design templates as the baseline technique for the comparison.

3.5. Qualitative evaluation on melanocyte detection

For the qualitative evaluation, three close-up examples are shown in Fig. 11. These images are captured from different skin tissues, and the parameter $T_{R_{HN}} = 0.8$ and $T_{TM} = 0.2$ were used for the proposed technique and the TM technique, respectively. Note that Fig. 11(a), (b), and (g) shows three original color histopathological images, where the ground truth melanocytes locations are indicated by stars. The last two columns show the results obtained using the proposed technique and the TM technique, respectively. The detected regions are presented as the thick contours. It is observed that the proposed technique (shown in Fig. 11(b), (e), and (h)) can detect all the melanocytes and has one false positive. In Fig. 11(e), there is a false detection which indicates by a hollow arrow. This false detection is mainly because the candidate nuclei region is very close to a typical melanocyte where its surrounding cytoplasm are retracted.

It is observed that in the results provided by the TM technique have many false detections (shown in Fig. 11(c), (f), and (i)) since there are many cytological components similar to that of the templates. In addition, because it is very difficult to model a good template set that can capture the size and shape variations of the natural melanocytes, it misses some of the ground truths melanocytes (see shown in Fig. 11(f), the solid arrows indicate the miss detection). Overall, the proposed technique is able to provide a good performance in the test images set.

4. Conclusion

This paper presents a simple but effective computer-aided technique for the detection of the melanocyte in the skin histopathological image. The candidate nuclei regions are first segmented through an adaptive threshold method. A radial lines scanning algorithm is then proposed to estimate the halo regions of melanocytes. Finally, the melanocytes are detected based on the ratio of the estimated halo region area and the nuclei area. The evaluation using 40 histopathological images (with more than 300 melanocytes) shows the efficiency of the proposed technique. In future, we plan to analyze the cytological and architectural features of the detected melanocytes in order to grade and diagnose the skin tissues.

References

[1] W. Allsbrook, K. Mangold, M. Johnson, R. Lane, C. Lane, J. Epstein, Inter-observer reproducibility of Gleason grading of prostatic carcinoma: general pathologist, Human Pathology 32 (1) (2001) 81–88.

[2] A.N. Basavanhally, S. Ganesan, S. Agner, J.P. Monaco, M.D. Feldman, J.E. Tomaszewski, G. Bhanot, A. Madabhushi, Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology, IEEE Transactions on Biomedical Engineering 57 (3) (2010) 642–653.

[3] L.E. Boucheron, Object- and Spatial-Level Quantitative Analysis of Multi-spectral Histopathology Images for Detection and Characterization of Cancer, Ph.D. Thesis, University of California, Santa Barbara, March 2008.

[4] T. Elsheikh, S. Asa, J. Chan, R. DeLellis, C. Heffess, V. LiVolsi, B. Wenig, Interobserver and intraobserver variation among experts in the diagnosis of thyroid follicular lesions with borderline nuclear features of papillary carcinoma, American Journal of Clinical Pathology 130 (5) (2008) 736–744.

[5] R. Gonzalez, R. Woods, Digital Image Processing, 2002.

[6] M. Gurcan, T. Pan, H. Shimada, J. Saltz, Image analysis for neuroblastoma classification: segmentation of cell nuclei, in: Proceedings of the 28th IEEE Annual International Conference of EMBS, 2006, pp. 4844–4847.

[7] S. Ismail, A. Colclough, J. Dinnen, D. Eakins, D. Evans, E. Gradwell, J. O'Sullivan, J. Summerell, R. Newcombe, Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia, British Medical Journal 298 (6675) (1989) 707.

C. Lu et al. / Pattern Recognition 46 (2013) 509–518

[8] C. King, J. Long, Prostate biopsy grading errors: a sampling problem, International Journal of Cancer 90 (2000) 326–330.

[9] V. Korde, H. Bartels, J. Barton, J. Ranger-Moore, Automatic segmentation of cell nuclei in bladder and skin tissue for karyometric analysis, Analytical and Quantitative Cytology and Histology 31 (2) (2009) 83.

[10] V. Kumar, A. Abbas, N. Fausto, et al., Robbins and Cotran Pathologic Basis of Disease, Elsevier, Saunders, Philadelphia, 2005.

[11] J. Lewis, Fast normalized cross-correlation, in: Vision Interface, vol. 10, Citeseer, 1995, pp. 120–123.

[12] C. Lu, M. Mandal, Automated segmentation and analysis of the epidermis area in skin histopathological images, in: Proceeding of the 31th IEEE Annual International Conference of EMBS, 2012.

[13] C. Lu, M. Mahmood, N. Jha, M. Mandal, A robust automatic nuclei segmentation technique for quantitative histopathological image analysis, Analytical and Quantitative Cytology and Histology. In Press.

[14] I. Maglogiannis, C. Doukas, Overview of advanced computer vision systems for skin lesions characterization, IEEE Transactions on Information Technology in Biomedicine 13 (5) (2009) 721–733.

[15] S. Naik, S. Doyle, S. Agner, A. Madabhushi, M. Feldman, J. Tomaszewski, Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology, in: Proceedings of the Fifth IEEE International Symposium on Biomedical Imaging: From Nano to Macro ISBI 2008, pp. 284–287.

[16] J. Pawley, Handbook of Biological Confocal Microscopy, Springer-Verlag, 2006.

[17] S. Petushi, F.U. Garcia, M.M. Haber, C. Katsinis, A. Tozeren, Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer, BMC Medical Imaging 6 (2006) 14.

[18] D. Rigel, J. Russak, R. Friedman, The evolution of melanoma diagnosis: 25 years beyond the ABCDs, CA: A Cancer Journal for Clinicians 60 (5) (2010) 301–316.

[19] E. Ruijter, G. van Leenders, G. Miller, F. Debruyne, C. van de Kaa, Errors in histological grading by prostatic needle biopsy specimens: frequency and predisposing factors, The Journal of Pathology 192 (2) (2000) 229–233.

[20] O. Sertel, U. Catalyurek, H. Shimada, M. Guican, Computer-aided prognosis of neuroblastoma: detection of mitosis and karyorrhexis cells in digitized histological images, in: Proceeding of the 31th IEEE Annual International Conference of EMBS, IEEE, 2009, pp. 1433–1436.

[21] I. Sintorn, M. Homman-Loudiyi, C. Söderberg-Nauclér, G. Borgefors, A refined circular template matching method for classification of human cytomegalovirus capsids in TEM images, Computer Methods and Programs in Biomedicine 76 (2) (2004) 95–102.

[22] A.C. Society, What Are the Key Statistics About Melanoma? Technical Report, American Cancer Society, 2008.

[23] D. Weedon, G. Strutton, Skin Pathology, vol. 430, Churchill Livingstone, New York, 2002.

Cheng Lu received his BSc and MSc degree in Computer Engineering from Information Engineering College, Northwest A&F University, Shaanxi, China, in 2006 and 2008. He is currently working toward his PhD degree in electrical engineering in University of Alberta, Edmonton, AB, Canada, from 2008 till now. In 2008, he joined the Multimedia Computing and Communication Lab, Department of Electrical Engineering, University of Alberta, as a PhD candidate, research assistant and the teaching assistant. Mr. Lu is the recipient of the Chinese Scholarship Council for his PhD studies. His research interest includes computer vision, medical image analysis, pattern recognition and super resolution imaging. He is an author or coauthor of several papers in leading international journals and conferences.

Muhammad Mahmood received his medical degree from Aga Khan University, Medical College, Karachi, Pakistan. He did his Anatomical Pathology/Clinical Pathology residency from Henry Ford Hospital, Detroit, Michigan, USA. Further, he did two fellowships; an Advanced Selective Pathology fellowship from Barnes-Jewish Hospital, Washington University Medical Center, Saint Louis, MO, USA, and a Dermatopathology fellowship from Ackerman Academy of Dermatopathology, New York, NY, USA. He is an American Board Certified Anatomical and Clinical Pathologist and American Board Certified Dermatopathologist. For last six years, he is a practising Anatomical Pathologist/Dermatopathologist in Canada. He is currently working at Department of Lab Medicine and Pathology, University of Alberta Hospital, Edmonton, Alberta, Canada. His interests include melanocytic cutaneous lesions, cutaneous tumors and inflammatory dermatoses. He is also interested in developing post-graduate medical teaching for Pathologists and Dermatologists in training.

Naresh Jha graduated in 1979 with Bachelor of Medicine and Bachelor of Surgery from Delhi University in India. He participated residency training in Radiation Oncology at the Cross Cancer Institute, Edmonton, Alberta, Canada. He received Fellow of the Royal College of Physicians of Canada (FRCPC) in Radiation Oncology in 1987. He is a Senior Radiation Oncologist at the Cross Cancer Institute since 1988. He specializes in head and neck, thyroid and skin cancers. He is also a pioneer of the submandibular salivary gland transfer—a ground breaking technique to prevent radiation-induced xerostomia.

Mrinal Mandal is a Full Professor and Associate Chair in the Department of Electrical and Computer Engineering and is the Director of the Multimedia Computing and Communications Laboratory at the University of Alberta, Edmonton, Canada. He has authored the book Multimedia Signals and Systems (Kluwer Academic), and coauthored the book Continuous and Discrete Time Signals and Systems (Cambridge University Press). His current research interests include Multimedia, Image and Video Processing, Multimedia Communications, Medical Image Analysis. He has published over 140 papers in refereed journals and conferences, and has a US patent on lifting wavelet transform architecture. He has been the Principal Investigator of projects funded by Canadian Networks of Centers of Excellence such as CITR and MICRONET, and is currently the Principal Investigator of a project funded by the NSERC. He was a recipient of Canadian Commonwealth Fellowship from 1993 to 1998, and Humboldt Research Fellowship from 2005 to 2006 at Technical University of Berlin.