Automated Segmentation and Analysis of the Epidermis Area in Skin Histopathological Images

Cheng Lu¹ and Mrinal Mandal¹, Senior Member, IEEE

Abstract—In the diagnosis of skin melanoma by analyzing histopathological images, the segmentation of the epidermis area is an important step. This paper proposes a computer-aided technique for segmentation and analysis of the epidermis area in the whole slide skin histopathological images. Before the segmentation technique is employed, a monochromatic color channel that provides a good discriminant information between the epidermis and dermis areas is determined. In order to reduce the processing time and perform the analysis efficiently, we employ multi-resolution image analysis in the proposed segmentation technique. At first, a low resolution whole slide image is generated. We then segment the low resolution image using a global threshold method and shape analysis. Based on the segmented epidermis area, the layout of epidermis is determined and the high resolution image tiles of epidermis are generated for further manual or automated analysis. Experimental results on 16 different whole slide skin images show that the proposed technique provides a superior performance, about 92% sensitivity rate, 93% precision and 97% specificity rate.

I. INTRODUCTION

Skin cancer is the most frequent and malignant types of cancer [1] 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 9,000 die from it in the United States every year [2]. The early detection of malignant melanoma will help to lower the mortality from this cancer.

It is known that the histopathological examination is the gold standard for many kinds of cancer diagnosis as the histopathology slides provide a cellular level view of the disease [3]. Traditionally, the histopathology slides are examined under a microscope by pathologists. With the help of the whole slide histology digital scanners, glass slides of tissue specimen can now be digitized at high magnification to create the whole slide image (WSI). Such high resolution images are similar to what a pathologist observes under a microscope to diagnose the biopsy. The pathologists can now do the examination via the WSI instead of using the microscope. Note that the WSI takes a large storage space and huge computing power for processing. For example, one 30mm² glass slide tissue scanned at 40X magnification (0.11625 µm/pixel) will consist of about 2.2 $\times 10^9$ pixels, and it will approximately take about 6.6 Gigabytes in uncompressed color format. Therefore, it is time consuming and difficult to analyze a WSI manually. In addition, the manual diagnosis is subjective and often leads to intra-observer and inter-observer variability [4]. To address this problem, automated computational tool which can provide reliable and reproducible objective results for quantitative analysis is needed.

Several techniques have been proposed for automated segmentation and analysis of the WSI. Mete et al. [5] proposed a block-based supervised technique that can delineate the malignancy regions of histopathological head and neck slides. Roullier et al. [6] proposed a multi-resolution graph-based analysis framework for the WSI of

¹C. Lu and M. Mandal are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 2V4. e-mail: lcheng4@ualberta.ca, mmandal@ualberta.ca. breast cancer. Wang et al. [7] developed an automated computeraided system for the diagnosis of cervical intraepithelial neoplasia. The authors reported that the processing time is about 3 hours for a typical digital slide of $152,000 \times 41,500$ pixels, which is very time consuming and cannot be directly applied in practical clinic setting.



Fig. 1. The anatomy of a skin tissue.

Unlike other types of specimen, a typical skin tissue slide consists of three main parts: epidermis, dermis and sebaceous tissues (see Fig. 1, the lower image shows the manually labeled contour of the epidermis). The epidermis area is the most important observation area for the diagnosis of a skin tissue. In most cases of melanoma, the grading of the cancer can be made through the architectural and morphological features of atypical cells in the epidermis or epidermis-dermis junctional area. Therefore, segmentation of the epidermis area is an important step before further analysis is performed. Once the epidermis area is segmented, the dermis and sebaceous tissues can then be easily located.

Due to the specificity of the skin tissue structure, we do not need to process a WSI pixel by pixel. In this paper, we propose an efficient technique for automatic segmentation and analysis the epidermis area in the skin WSI. First, we determine a monochromatic color channel which can provide a good discriminant information between the epidermis and dermis areas for the subsequence segmentation procedure. We then apply a global threshold and shape analysis technique to segment the epidermis area at low resolution (LR). Finally, the high resolution (HR) image tiles of interest are then generated effectively for further manual or automated analysis.

II. SELECTION OF THE MONOCHROMATIC COLOR CHANNEL

In the skin tissue, the epidermis area and the dermis area are closely connected with each other (see Fig. 1). In order to segment the epidermis area accurately, we need to first determine a monochrome channel which can provide good distinguishable information between the epidermis and dermis areas. Note that a WSI contains $\{R, G, B\}$ three channels. To obtain a monochrome channel *M*, a linear transform $A = [a_1, a_2, a_3]$ is required, i.e.,

$$M = a_1 R + a_2 G + a_3 B = [R, G, B][a_1, a_2, a_3]^T,$$
(1)
s.t., $a_1 + a_2 + a_3 = 1.$

Denote an image pixel by p and its gray value by g(p). Denote the pixels belonging to epidermis area as class 1 (denoted as Ω_1) and the pixels belonging to the dermis area as class 2 (denoted as Ω_2). Let the mean intensity of the epidermis area, dermis area and dermis + epidermis areas be denoted as μ_1 , μ_2 , and μ_o respectively. In this work, to distinguish the different classes, we use the following metric [8]:

$$R_{II} = \frac{\sigma_{inter}^2}{\sigma_{intra}^2} \tag{2}$$

where

$$\sigma_{inter}^{2} = w_{1}(\mu_{1} - \mu_{o})^{2} + w_{2}(\mu_{2} - \mu_{o})^{2}$$
(3)

$$\sigma_{intra}^{2} = \frac{w_{1}}{N_{1}} \sum_{p \in \Omega_{1}} (g(p) - \mu_{1})^{2} + \frac{w_{2}}{N_{2}} \sum_{p \in \Omega_{2}} (g(p) - \mu_{2})^{2}$$
(4)

where w_1 and w_2 are the size proportion of two classes, i.e., epidermis and dermis, respectively. N_1 and N_2 are the total number of pixels in two classes, respectively. The σ_{inter}^2 and σ_{intra}^2 are the inter classes and intra class variances of the pixel intensities. In this work, we assume that the monochrome channel with a maximum quality metric R_{II} provides the best discriminant information between the epidermis and dermis area. In particular, we need to find an optimal linear transform A^* which can maximize the quality metric R_{II} , i.e., $A^* = \max_{A \in \Omega_A}(R_{II})$, where Ω_A is the space of linear transform A. In this paper, we use a grid search method to find an approximate optimal linear transform \tilde{A} which is described as follows.

- First, we manually label the epidermis and dermis areas for 8 different skin WSIs as the test data set.
- A set of 1-by-3 vectors which can approximately cover the space of the linear transform A, denoted as $\tilde{\Omega}_A$, is initialized. The grid step for each variable, i.e., a_1 , a_2 , and a_3 , in the linear transform is 0.05.
- For each possible linear transform $A \in \overline{\Omega}_A$, the quality metric R_{II} is calculated.
- The linear transform A that corresponds to the maximum quality metric R_{II} is recorded as the approximate optimal linear transform \tilde{A} .

All the approximate optimal linear transforms \tilde{A} for the 8 test WSIs with corresponding quality metric R_{II} are presented in Table I. The 8 WSIs are labeled as WSI1, WSI2, \cdots , WSI8, respectively.

It is observed that the coefficient associated with the red channel, i.e., a_1 , is 1 or close to 1 for all the test WSIs. In other words, the red channel of the RGB color image can provide good distinguishable information between the epidermis and dermis area. Based on such observation, in this work, we choose the red channel as the monochrome channel for latter processing since it is ready to be used without further calculation, even though it may not be always the optimal one. This is further validated with results shown in Table II and Fig. 2. Table II shows the quality metric R_{II} of the R, G, B channels from RGB color model, the gray channel computed

using transform $A_{gray} = [0.299, 0.587, 0.114]$, and H, S, V channels from the HSV color model [9] for the 8 test WSIs. It is noted that the R channel achieve the largest quality metric R_{II} compared to the other channels. The V channel from the HSV color model also provides good discriminating information.

Fig. 2 provides a visual evaluation example, where the RGB color image with R, G, B, gray, H, S, V channels of the WSI1 skin histopathological image are shown. It is observed that in the red channel and V channel image (shown in Fig. 2(b) and (j), respectively), the epidermis area is easily distinguishable from its neighboring dermis area. The corresponding quantitative results are shown in Table II.



Fig. 2. An example of the WSI. (a) Original H&E stained color image. (b), (c), and (d) shows the R, G, and B channel of the image, respectively. (e) shows the gray channel image. (f), (i), (j) shows the H, S, V channels.

III. SEGMENTATION OF THE EPIDERMIS AREA

In order to reduce the processing time and efficiently perform the analysis, we apply multi-resolution image analysis. The schematic of the proposed technique is shown in Fig. 3. It is observed that there are four modules in the proposed technique. Module one generates a LR WSI using down sampling method. Module two segment the LR image using a global threshold and shape analysis. In the last two modules, the layout of epidermis is determined, and the HR image tiles of epidermis are generated for further analysis. The details of these modules are discussed below.



Fig. 3. The schematic for the WSI segmentation.

TABLE I The approximate optimal linear transform \tilde{A} with its R_{II} value for 8 test WSIs.

	WSI1	WSI2	WSI3	WSI4	WSI5	WSI6	WSI7	WSI8
Ã	[1,0,0]	[1,0,0]	[0.9, 0.1, 0]	[0.95, 0.05, 0]	[0.85, 0.15, 0]	[1,0,0]	[0.9, 0.1, 0]	[0.95, 0.05, 0]
R_{II}	0.36	0.53	0.37	0.53	0.53	0.45	0.25	0.24

TABLE II THE QUALITY METRIC R_{II} for the R, G, B, Gray, H, S, V channels of 8 WSIs.

	WSI1	WSI2	WSI3	WSI4	WSI5	WSI6	WSI7	WSI8
R	0.36	0.53	0.34	0.51	0.49	0.45	0.22	0.23
G	0.06	0.11	0.25	0.14	0.21	0.22	0.04	0.03
В	0.10	0.21	0.10	0.27	0.20	0.21	0.03	0.01
Gray	0.11	0.20	0.33	0.32	0.33	0.34	0.09	0.07
Η	0.03	0.03	0.07	0.22	0.05	0.05	0.05	0.04
S	0.07	0.11	0.26	0.19	0.24	0.23	0.07	0.05
V	0.35	0.50	0.29	0.49	0.49	0.41	0.19	0.21

A. Image Down-sampling

Given a high resolution WSI I_h as the input image with size $r \times c$, we employ down-sampling [10] to create a low resolution image I_l with size $(r/k) \times (c/k)$, where k is a down sampling factor. (In order to reduce the aliasing, a low pass filter is usually applied to the image before the down sampling.) In this work we have used k=32, which can reduce a 2.2×10^9 pixel image to a 2.15×10^6 pixels LR image.

B. Segmentation of the Epidermis using Global Threshold and Shape Analysis

As explained in Section II, the red channel is chosen for the segmentation. It is observed in Fig. 2(b) that the epidermis area has darker intensity value compared to the remaining part of a WSI (see Fig 3). Therefore, we can use a simple thresholding method combined with the shape analysis to segment the epidermis area. The thresholding method groups the pixels into two classes: foreground and background, based on the pixel values. Specifically, let τ be a selected threshold for the segmentation. The classification of pixel p with gray value g(p) is then done as follows:

$$p = \begin{cases} foreground, & if \ g(p) \le \tau \\ background, & if \ g(p) > \tau \end{cases}$$
(5)

In our problem, the foreground corresponds to the epidermis area. We apply Otsu's thresholding [11] method to determine the optimal threshold for segmentation.

Before the thresholding, the white background (e.g., the pixels with gray value greater than 240) are removed in order to estimate more accurate result. Three examples of the threshold results on different skin tissues are shown in Fig. 4(b), (e), and (h). Note that the epidermis areas are segmented as well as other low intensity components, such as hair follicles, blood vessels and other unrelated tissues. We label all the regions in the binary image via 8-connected criterion [9] and denote all the 8-connected regions as $\{C_i\}_{i=1...n}$, where *n* is the total number of the regions. In a typical image, we may result in several hundreds candidate regions. A further test is required to eliminate the unrelated false regions.

In order to eliminate the false regions, we use two criteria: area and the shape. We first assume that the epidermis area is greater than a predefined threshold T_{area} . We set $T_{area} = N \times 0.5\%$, where N is the total pixel number of the skin tissue, and the fraction 0.5%, determined based on the domain prior and experiment results, represents the portion of the epidermis area in the WSI. Secondly,



Fig. 4. Three examples of epidermis segmentation. (a), (d), and (g) are the original red channel image; (b), (e), and (h) show the results after thresholding; (c), (f), and (i) show the final binary mask for epidermis.

from the observation, the epidermis area is generally a long and narrow region. This shape feature can be captured by the major axis length (l_1) to minor axis length (l_2) ratio of a best fit ellipse [12].

Given a 8-connected region C_i , we use the above mentioned two criteria based on the domain knowledge to determine if it is epidermis or other tissue components.

$$C_{i} = \begin{cases} epidermis, & if \quad A(C_{i}) \ge T_{area} \quad and \quad l_{1}/l_{2} \ge T_{l} \\ others, & otherwise \end{cases}$$
(6)

where $A(\cdot)$ is the area operator, $A(C_i)$ represents the area of region C_i . Threshold T_i is set to 4 empirically to make sure that the epidermis is long and narrow.

Finally, the holes inside the region are filled to generate the mask for the epidermis. Three examples of the binary masks for the epidermis are shown in Fig. 4(c), (f), and (i). Note that the undesirable regions are removed.

C. Determine the Layout of Epidermis and Generate Image Tiles

The epidermis mask is now available for further processing. In this section, we divide the epidermis region into several image tiles. These image tiles will be used for further analysis at high resolution. The reason behind dividing the epidermis is two-fold:

- Accuracy analysis. The binary mask of the epidermis is obtained from the low resolution image I_l based on the global threshold and shape analysis as the local intensity variants may incur inaccuracy segmentation results (usually under segmentation). By analyzing the epidermis mask locally, it is expected that we can achieve better segmentation and analysis results.
- Efficient computation. Normally, the epidermis area which taking about 1% to 10% of the whole image. That means about 22 220 Mega pixels need to be manipulated for a $30mm^2$ tissue, which will take a considerable amount of computational resources. By dividing the rough epidermis mask into image tiles, we can use the parallel computation technique [13] to distribute the processing load into each node which will greatly help to do the computation in an efficient way.

First, we need to divide the whole image into non-overlapping image blocks. A grid G_l which indicates the non-overlap blocks are defined. The grid has the same size as that of the LR image I_l , and the width and height of the block are denoted as b_w and b_h , respectively. We set $b_w = 2000/k$ and $b_h = 1000/k$, remind that k is the down sampling factor for the high resolution image I_h . This grid is overlapped onto the binary image as shown in Fig. 5(a). Note that we are only interest in the image blocks which contain the epidermis mask.



Fig. 5. An example for determining the layout of the epidermis. (a) shows the grid on to the binary image; (b) shows the blocks which contains the rough epidermis area (white blocks) as well as the fitted line (red solid line) based on the coordinates of the value 1 in \mathcal{L} ; (c) illustrates the acute angle θ between the fitted line and the referenced horizontal line.

In the manual examination of the epidermis, a pathologist usually takes into consider the tissue structure and cytological features vertically across the epidermis. In the computerized image analysis, it is also reasonable to analyze the cytological components vertically across the epidermis. The idea is shown in Fig. 6. A horizontal image tile that consists of several blocks is shown in Fig. 6(b). In contrast, Fig. 6(c) shows an image tile consisting of several blocks generated vertically. It is clear that the image tile obtained vertically in this example preserves all the sub-layer information of the epidermis very well. On the other hand, it is difficult to analyze the image tile shown in Fig. 6(b) since it provides only partial information from different location of the epidermis. Since



Fig. 6. An illustration of the horizontal image tile and vertical image tile. (a) shows the lower portion of Fig. 5(a). Note the blue rectangle that points to (c) represents an image tile obtained horizontally whereas the yellow rectangle that points to (c) represents an image tile obtained vertically. The image tile is consisted of several image blocks.

the epidermis in this example is lying horizontally, it is easy to note that generation of the image tile vertically provides intact information rather than generation of the image tile horizontally.

Based on the discussion above, in this work, we propose to generate reasonable image tiles for further analysis. In order to do that, the layout of the epidermis mask should be first determined. Generally, the skin tissues on the glass slides have two different layouts: horizontal or vertical. In order to determine the layout, we calculate a matrix \mathcal{L} as follows:

$$\mathcal{L}(u,v) = \begin{cases} 1, & if \ \sum_{i=(u-1)b_h+1}^{ub_h} \sum_{j=(v-1)b_w+1}^{vb_w} \mathcal{B}(i,j) \ge T_b \\ 0, & if \ \sum_{i=(u-1)b_h+1}^{ub_h} \sum_{j=(v-1)b_w+1}^{vb_w} \mathcal{B}(i,j) < T_b \end{cases}$$
(7)

where \mathcal{B} is the binary image. u, v and i, j are indices for the 2-D matrix \mathcal{L} and binary image \mathcal{B} , respectively. Threshold $T_b = 0.1 \times b_w \times b_h$ considers the noise effect. After calculating the matrix \mathcal{L} , we use first order polynomial fitting to determine the fitted line. This is shown in Fig. 5(b), where the white blocks correspond to the value 1 in \mathcal{L} while black blocks correspond to the value 0. The solid line across the white blocks is the fitted line based on the coordinates of the value 1 in \mathcal{L} . The layout of the epidermis can now be determined by the acute angle θ between the fitted line and



Fig. 7. An example of generating the reasonable image tiles. Note the layout of the epidermis is horizontal in this example, so the image tiles are generated vertically. The rectangles indicate the generated image tiles for further processing. Some of the snapshot of the image tiles are present.

the referenced horizontal line which defined as follows:

$$Layout = \begin{cases} horizontal, & if \ 0^{\circ} \le \theta \le 45^{\circ} \\ vertical, & if \ 45^{\circ} < \theta \le 90^{\circ} \end{cases}$$
(8)

The idea for determining the layout based on the angle θ is illustrated in Fig. 5(c), where the dash line represents the referenced horizontal line.

Based on the layout of the epidermis, we can now generate the image tiles for further analysis. If the layout is vertical, we generate the image tiles by stitching the image blocks with epidermis mask horizontally. On the other hand, if the layout is horizontal, we generate the image tiles by stitching the image blocks with epidermis mask vertically. An example is illustrated in Fig. 7. Note that the grid G_l for the LR image should be mapped to the HR image domain in order to generate the HR image tiles.

IV. EVALUATION ON THE EPIDERMIS SEGMENTATION

We have evaluated the proposed technique on 16 different skin WSIs. These images are captured from different skin tissue samples which contain normal, nevus, and melanoma. These images are captured under 40X magnification on Carl Zeiss MIRAX MIDI Scanning system.

In the evaluation, the manually labeled epidermis areas (pixellevel contour of the epidermis) are treated as the ground truths. We define GT as the ground truth region of the epidermis, SEG as the segmented region obtained by the proposed technique. Three evaluation metrics: sensitivity (SEN), specificity (SPE), precision (PRE) are defined as follows:

$$SEN = \frac{|GT \cap SEG|}{|GT|} \times 100\%$$
(9)

$$PRE = \frac{|GT \cap SEG|}{|SEG|} \times 100\% \tag{10}$$

$$SPE = \frac{|\overline{GT} \cap \overline{SEG}|}{|\overline{GT}|} \times 100\%$$
(11)

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

$$MAD(C^{GT}, C^{SEG}) = \frac{1}{M} \sum_{i=1}^{M} [min||C_i^{GT} - C_j^{SEG}||]$$
(12)

where C^{GT} and C^{SEG} are the contour of the ground truth region and segmented region, respectively. C_i^{GT} and C_j^{SEG} represent the *i*th pixel and the *j*th pixel on the contour of the ground truth region and segmented region, respectively. *M* is the total number of pixels on the contour of the ground truth region.

TABLE III Performance evaluation of the epidermis segmentation.

	PRE	SPE	SEN	MAD
Red channel	93.51%	96.95%	92.13%	1.78
V channel	94.06%	96.97%	86.82%	3.61
Gray channel	4.51%	89.13%	36.98%	4528

The performance of the epidermis segmentation technique on the red channel, V channel, and Gray channel images for 16 test WSIs are shown in Table III. It is observed that the proposed technique provides a good performance on both red channel and V channel images since these two channels provide good discriminant information as discussed in Section II. The proposed segmentation technique if applied on the gray channel image generates a poor performance because in the gray channel image the epidermis and dermis areas are hard to differentiate by using a threshold. It noted that other monochromatic color channels (discussed in Section II) also have poor performance (the results are not shown here).

The average processing time for segment the epidermis area for an image with 2800×3200 pixels is about 2.38 seconds.

V. CONCLUSIONS

In this work, in order to perform an effective segmentation, a monochromatic color channel is first determined. By using the pre-determined monochromatic channel, a hybrid technique based on global threshold and shape analysis is employed to segment the epidermis from the WSI. We also discuss how to generate the HR image tiles of the epidermis for further analysis. Experimental results show that the proposed technique provides a good performance. The proposed technique is an important pre-requisite procedure towards the automated skin histopathological image analysis since the epidermis area is the key area to be examined. In addition, the morphological features, e.g., depth and symmetry, are useful factors in the skin cancer diagnosis procedure. In the future, further quantitative analysis, e.g., melanocytes detection in the epidermis, will be carried out based on this work.

ACKNOWLEDGMENT

We thank Dr. Naresh Jha, and Dr. Muhammad Mahmood of the University of Alberta Hospital for providing the images and giving helpful advice.

References

- I. Maglogiannis and C. Doukas, "Overview of advanced computer vision systems for skin lesions characterization," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 721–733, 2009.
- [2] A. C. Society, "What are the key statistics about melanoma?" American Cancer Society, Tech. Rep., 2008.
- [3] V. Kumar, A. Abbas, N. Fausto, et al., Robbins and Cotran pathologic basis of disease. Elsevier Saunders Philadelphia:, 2005.
- [4] S. Ismail, A. Colclough, J. Dinnen, D. Eakins, D. Evans, E. Gradwell, J. O'Sullivan, J. Summerell, and R. Newcombe, "Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia." *British Medical Journal*, vol. 298, no. 6675, p. 707, 1989.
- [5] M. Mete, X. Xu, C.-Y. Fan, and G. Shafirstein, "Automatic delineation of malignancy in histopathological head and neck slides." *BMC Bioinformatics*, vol. 8 Suppl 7, p. S17, 2007. [Online]. Available: http://dx.doi.org/10.1186/1471-2105-8-S7-S17
- [6] V. Roullier, L. O., T. V.T., and E. A., "Multi-resolution graphbased analysis of histopathological whole slide images: Application to mitotic cell extraction and visualization," *Computer medical imaging* and graph, 2011.
- [7] Y. Wang, D. Crookes, O. S. Eldin, S. Wang, P. Hamilton, and J. Diamond, "Assisted diagnosis of cervical intraepithelial neoplasia (cin)," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 112–121, 2009.
- [8] R. Duda, P. Hart, and D. Stork, *Pattern classification*. Citeseer, 2001, vol. 2.
- [9] R. Gonzalez and R. Woods, Digital image processing, 2002.
- [10] P. Burt and E. Adelson, "The laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532– 540, 1983.
- [11] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [12] M. Fitzgibbon, A. W.and Pilu and R. B. Fisher, "Direct least-squares fitting of ellipses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 476–480, May 1999.
- [13] B. Cambazoglu, O. Sertel, J. Kong, J. Saltz, M. Gurcan, and U. Catalyurek, "Efficient processing of pathological images using the grid: computer-aided prognosis of neuroblastoma," in *Proceedings* of the 5th IEEE workshop on Challenges of large applications in distributed environments. ACM, 2007, pp. 35–41.