34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Hair detection in dermoscopic images using Percolation

Ana Afonso and Margarida Silveira

Abstract— **The automatic analysis of dermoscopy images is often impaired by artifacts such as air bubbles, specular reflections or dark hair covering the skin lesions. Consequently, an important pre-processing step includes their detection and elimination. The most common and probably the most compromising of these artifacts is the presence of hair and therefore specific algorithms are required for its detection. This paper proposes a method for the detection of hair in dermoscopy images based on an efficient percolation algorithm for image processing recently proposed in [1]. The percolation algorithm locally processes image points by taking into account the intensity and connectivity of neighboring pixels. A cluster of connected points is thus obtained and the shape of this cluster is subsequently analyzed. If the cluster has a shape that is approximately linear then the image point is classified as hair. The performance of the proposed method was investigated on real dermoscopy images and compared with the DullRazor software [2]. Our results indicate that the method provides effective hair detection outperforming the DullRazor method by more than 10%, both in terms of false positive and false negative rates.**

I. INTRODUCTION

Dermoscopy is a non-invasive diagnosis technique used in dermatology for the in vivo observation of pigmented skin lesions [3]. Dermoscopic images have great potential in the early diagnosis of malignant melanoma, but their interpretation is subjective and time consuming and therefore there is currently a great interest in the development of computer-aided diagnosis (CAD) systems, based on image processing and pattern recognition algorithms, that can assist dermatologists in their evaluation.

However, the automatic analysis of dermoscopy images is often impaired by artifacts such as air bubbles, specular reflections or dark hair covering the skin lesions. Therefore, an important pre-processing step includes the detection and elimination of these artifacts. For this task, several researchers have used general purpose noise removal techniques such as Gaussian filtering [4], median filtering [5] or grey-scale morphological closing [6]. Other authors, on the other hand, recognizing that the presence of hair is the most common and probably the most compromising of these artifacts, as Fig. 1 illustrates, have developed image processing algorithms specifically for hair detection and removal.

Many of these techniques are based on mathematical morphological filtering. They assume that hairs are, locally, thin linear structures and use erosion/dilation [7], closing [2] or the top-hat operator [8], with a linear structuring element to detect it. Usually the operation is repeated several times with the structuring element at different orientations, for example [7] uses horizontal and vertical straight lines, [2] uses three orientations $(0^o, 45^o, 90^o)$ and [8] uses four $(0^o, 45^o, 90, 135^o)$. Finally, if the response of the morphological operators exceeds a threshold Th at a pixel (x, y), then that pixel is classified as hair.

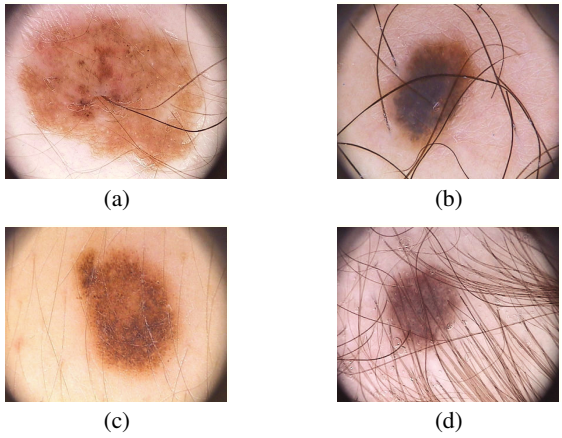

Fig. 1. Examples of the presence of hair in dermoscopic images.

Another approach is to use specific filters. For example [9] used matched filtering with an universal Gaussian kernel and [10] used matched filtering with derivative of gaussian filters, followed by hair refinement based of morphological edge-based techniques. A similar approach is described in [11] where difference of Gaussian filters are proposed, such that one of the filters has parameter σ_x much greater than σ_y in order to create an elongated shape. In [11], since the hair direction is not known, a bank of directional filters is used in order to detect an elongated structure at different directions.

Recently, an approach based on supervised learning, namely maximum a posteriori (MAP) estimation, has also been proposed in [12] but the results do not improve those of the previous approaches and the algorithm has greater complexity.

The main problem with the above mentioned techniques is that, besides detecting the hairs that are present in the images, these techniques usually also detect dermoscopic features in the lesion area that have a textured structure and in particular those that have linear shapes such as streaks, pigment network or vascular network.

In this paper, we propose a method that aims to reduce the falsely detected hairs by using percolation, a technique

A. Afonso and M. Silveira are with Institute for Systems and Robotics - Instituto Superior Tecnico, Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal ana.afonso@ist.utl.pt msilveira@ist.utl.pt

This work was supported by FCT projects PEst-OE/EEI/LA0009/2011 and PTDC/SAU-BEB/103471/2008.

Fig. 2. Preprocessing of the image in Fig. 1 (a): (left) channel with highest entropy (right) mask obtained for corner removal.

that locally processes image points and finds clusters of connected neighboring pixels with similar intensity. The cluster obtained by percolation at each candidate pixel is then analyzed and the candidate point is classified as hair if the cluster shape is approximately linear.

For the removal of the detected hairs, different methodologies have also been used. Some of the most popular are inpainting methods such as the one proposed by [13] or the one based on Partial Differential Equations (PDE) proposed by [14]. These techniques are well studied and provide good results and therefore the topic of hair removal will not be addressed in this paper. See [15] for a recent review of hair removal algorithms.

The remainder of this paper is organized as follows: section II describes the preprocessing for channel selection and corner removal and for the generation of candidate points for percolation, section III describes the proposed method for hair detection based on percolation, section IV describes the experimental results and section V concludes the paper.

II. Preprocessing

Several preprocessing steps are performed before the hair detection algorithm is applied. First, the selection of the image channel and the removal of the corners of the image. Second, the detection of candidate pixels for the percolation.

A. Corner removal and Channel selection

The dermoscopy images are color images (RBG) but only one of the three color channels is used for hair detection. We select the channel with highest entropy, assuming that it is the channel where the hair is most visible. Then, the dark regions in the 4 corners of the images are removed. To remove the corners, the images are first filtered with a median filter (5x5 neighborhood) to reduce noise and gray-level thresholding is performed using Otsu's method [16] which calculates the optimal threshold that minimizes intra class variance (or equivalently maximizes the inter class variance). After that, the binary components that are connected to each of the four corners of the image are eliminated. Fig. 2 illustrates these preprocessing steps.

B. Generation of candidate pixels

In order to generate a set of candidate hair pixels that will be used as seeds in the percolation algorithm, we use the method proposed by [8] which is based on top-hat filtering. Assuming that hairs are darker that the surrounding skin, the morphological black top-hat filtering is applied which calculates the difference between the image and its morphological closing:

$$T_i = I - C(I, e_i) \tag{1}$$

In this expression I represents the input image and $C(I,e_i)$ represents the morphological closing of image I using e_i as the structuring element. In case of light hairs, the black top-hat filtering should be replaced by the white top-hat where the difference between the image and its opening is calculated.

Considering that hairs are thin linear structures, the structuring element will be a straight line. Since the hair direction is not known, the top-hat filtering operation is performed four times using a straight line with four different orientations $\theta_i = (0^o, 45^o, 90^o, 135^o)$ and the maximum of the four top-hat responses is calculated.

Finally, a binary image is obtained by thresholding the maximum top-hat response using Otsu's threshold [16] calculated on the region image without the corners. Fig. 3 illustrates the candidate generation step.

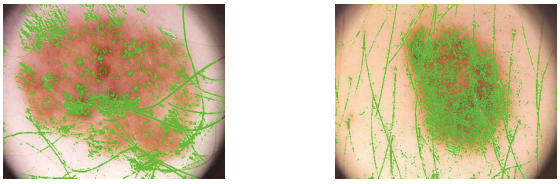

Fig. 3. Candidates pixels obtained for images (a) and (c) of Fig. 1.

It can be seen that a large number of the candidate pixels obtained by this method are false positives that correspond to dermoscopic features in the lesion area.

III. Proposed Method

The proposed technique for hair detection is based on an efficient percolation algorithm applied to a set of candidate hair pixels. Percolation is performed at each candidate pixel followed by the analysis of the shape of the obtained cluster. The candidate point is classified as hair if the obtained cluster is considered to be approximately linearly shaped. These steps are described in the following subsections.

A. Percolation

Percolation is a mathematical model used in physics to describe the movement and filtering of fluids through porous materials [17]. It can be applied in image processing to find clusters of connected neighboring pixels with similar intensity. In this work we use the percolation algorithm proposed by [1]. The algorithm requires the definition of the following parameters: the size N of an initial window, the maximum size M of that window and a threshold acceleration parameter, w. The algorithm is composed of two similar main cycles. The steps of the two cycles are presented next:

1) Percolation starts by placing an initial window of size NxN around a candidate pixel p_s, which is added to the set of percolated pixels, denoted by D_p. Additionally, a percolation

threshold T is set to the intensity of that center pixel, $T = I(p)$.

2) The first cycle begins in this step where the threshold T is updated as follows:

$$T = \max(\max_{p \in D_p} I(p), T) + w \quad (2)$$

where w is the predefined acceleration parameter.

3) The eight neighboring pixels of the elements in D_p are added to the set of candidate pixels denoted by D_c. The pixels in D_c with intensity smaller than the threshold are percolated and included in D_p. If there are no pixels in this condition then only the pixel with smallest intensity is added.

4) if D_p did not reach the boundary of the NxN window then percolation goes back to step 2). Otherwise N is incremented to $N+2$ and continues to step 5).

5) The second cycle begins in this step where the threshold is updated using equation (2).

6) The eight neighboring pixels of pixels in D_p are added to the set of candidate pixels denoted by D_c. The pixels in D_c with intensity smaller than the threshold are percolated and included in D_p. If there are no pixels in this condition the percolation process is terminated.

7) if D_p did not reach the boundary of the NxN window then percolation goes back to step 5). Otherwise N is incremented to $N+2$ and continues to step 8).

8) if N is larger than the maximum window of size M the percolation process is terminated. Otherwise it continues to step 5).

The result of the percolation process is obtained in set D_p.

B. Analysis of cluster shape

The shape of the cluster D_p obtained by the percolation algorithm at each candidate pixel is then analyzed by estimating its circularity F_c as follows:

$$F_c = \frac{4\#(Dp)}{\pi N^2} \quad (3)$$

where $\#(Dp)$ denotes the number of pixels in D_p and N^2 is the maximum number of pixels that D_p can have. If F_c is close to 1 then the cluster is approximately circular whereas small values of F_c mean that the cluster is close to linear. Therefore, if the value of F_c is smaller than a small threshold T_c then the center pixel p_s is classified as hair.

To make the percolation algorithm computationally more efficient, the analysis of the cluster shape described in III-B is performed not only at the end of the entire process but also at the end of step 4). At that point, if the circularity F_c of the clusters at those steps is already greater than the threshold T_c, then the percolation is immediately stopped without proceeding to the second cycle.

IV. Experimental Validation

This section presents results of the application of the proposed method to real dermoscopy data and compares them with the results obtained using the DullRazor software [2] available at (`http://www.dermweb.com/dull_razor/`) which has been described in section I.

The dataset comprised 70 dermoscopy images from the clinical database of Hospital Pedro Hispano (HPH), in Portugal. Images were colored (RGB) with 24 bits and 760x560 pixels in size, 15 of these images contained moderate or high amount of hair covering the lesion. Our results were obtained in the following conditions: the length of the straight line used as structuring element for candidate generation was l=15, for the percolation, initial window size was $N = 12$ while the maximum size was $M = 24$, the acceleration parameter was $w = 1$ and the threshold for the circularity was $T_c = 0.09$.

A. Hair Detection Results

Figure 4 shows examples of the hair pixels detected by the proposed method and by DullRazor. In these images the detected hair is shown in green, superimposed on the original images.

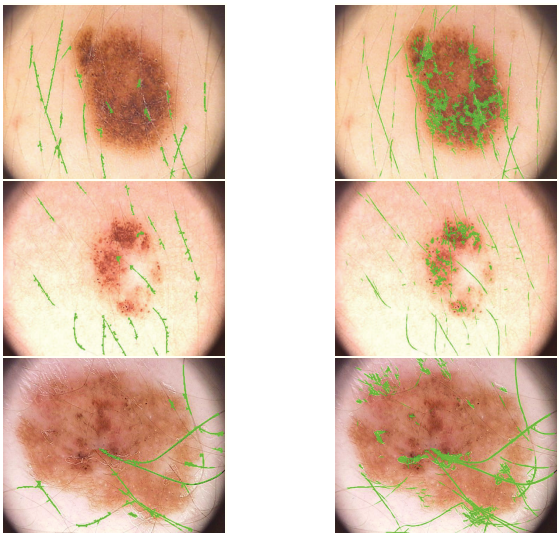

Fig. 4. Hairs detected in dermoscopic images (left) using Percolation (right) using DullRazor.

It can be seen that, although the candidate selection method described in section II-B originated a great number of false positives (see Fig. 3), the percolation algorithm is able to find the correct hair pixels and discard the majority of the falsely detected candidates. The figure also illustrates that the most important difference between the two methods

(percolation and DullRazor) is that the percolation method originates less false positives detected in the lesion area.

B. Hair Detection Quantitative Evaluation

The differences between these results were compared quantitatively using as ground truth (GT), images of the hair present in each dermoscopic image drawn independently by one of the authors (AA), from visual inspection. Two types of error were calculated, the false positive rate (FPR) and the false negative rate (FNR). The FPR measures the rate of pixels classified as hair by the algorithm that were not classified as such in the ground truth while the FNR measures the rate of pixels classified as hair in the ground truth that were missed by the algorithm. Usually there is a tradeoff between the two types of error which are calculated as follows:

$$FPR(AD,GT) = \frac{\#\left(AD \cap \overline{GT}\right)}{\#(GT)} \quad (4)$$

$$FNR(AD,GT) = \frac{\#\left(\overline{AD} \cap GT\right)}{\#(GT)} \quad (5)$$

where AD denotes the image of the automatic detection and GT denotes the ground truth hair detection. Both AD and GT are binary images such that all the hair pixels have label 1 and all the others have label 0. The results are shown in table I.

Method	FPR (%)	FNR (%)
Candidate	4.73	0.30
Percolation	0.19	0.52
DullRazor	0.30	0.63

TABLE I
PERFORMANCE OF HAIR DETECTION METHODS.

From this table we can confirm that the proposed method is able to discard the false candidate hairs greatly reducing the FPR when compared to the candidates. For both methods, FNR is higher than FPR because several hair borders are not detected correctly producing detected hairs that are thinner than the real hairs. We can also conclude that the proposed method is substantially more accurate than DullRazor since it achieves improvements greater than 10% in both FPR and FNR.

V. CONCLUSIONS

We proposed a method for hair detection in dermoscopy images. The proposed method is based on an efficient percolation algorithm for image processing recently proposed in [1]. The percolation algorithm locally processes image points by taking into account the connectivity among neighboring pixels. A cluster of connected points is thus obtained and the shape of this cluster is subsequently analyzed. If the cluster has a shape that is approximately linear then the image point is classified as hair. We compared our proposed technique with the well-known and freely available DullRazor software [2]. The proposed method outperformed DullRazor, achieving substantially lower false negative and false positive rates of detected hair.

REFERENCES

[1] T. Yamaguchi, S. Nakamura, and S. Hashimoto, "An efficient crack detection method using percolation-based image processing," pp. 1875 –1880, june 2008.

[2] T. Lee, "Dullrazor®: A software approach to hair removal from images," *Computers in Biology and Medicine*, vol. 27, no. 6, pp. 533–543, Nov. 1997.

[3] G. Argenziano, H. P. Soyer, and V. et al De Giorgi, *Dermoscopy, an interactive atlas*, EDRA Medical Publishing. (http://www.dermoscopy.org), 2000.

[4] B. Erkol, R. H. Moss, R. J. Stanley, W. V. Stoecker, and E. Hva-tum, "Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes," *Skin Research & Technology*, vol. 11, pp. 17–26, 2005.

[5] M. Celebi, Y. Aslandogan, and P. Bergstresser, "Unsupervised border detection of skin lesion images," vol. 2, pp. 123–128, April 2005.

[6] H. Ganster, P. Pinz, R. Rohrer, E. Wildling, M. Binder, and H. Kittler, "Automated melanoma recognition," *IEEE Trans. Med. Imag.*, vol. 20, no. 3, pp. 233–239, Mar 2001.

[7] Do Hyun Chung and G. Sapiro, "Segmenting skin lesions with partial-differential-equations-based image processing algorithms," *IEEE Trans. Med. Imag.*, vol. 19, no. 7, pp. 763–767, Jul. 2000.

[8] M. Fiorese, E. Peserico, and A. Silletti, "Virtualshave: Automated hair removal from digital dermatoscopic images," *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, pp. 5145 –5148, 30 2011-sept. 3 2011.

[9] Nhi H. Nguyen, Tim K. Lee, and M. Stella Atkins, "Segmentation of light and dark hair in dermoscopic images: a hybrid approach using a universal kernel," *Medical Imaging 2010: Image Processing*, vol. 7623, no. 1, pp. 76234N, 2010.

[10] Qaisar Abbas, Irene Fondon Garcia, M. Emre Celebi, and Waqar Ahmad, "A feature-preserving hair removal algorithm for dermoscopy images," *Skin Research and Technology*, pp. n/a–n/a, 2011.

[11] C. Barata, J.S. Marques, and J. Rozeira, "Engineering in medicine and biology society,embc, 2011 annual international conference of the ieee," pp. 5120 –5123, 30 2011-sept. 3 2011.

[12] P. Wighton, T.K. Lee, H. Lui, D.I. McLean, and M.S. Atkins, "Generalizing common tasks in automated skin lesion diagnosis," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 15, no. 4, pp. 622 –629, july 2011.

[13] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester, "Image inpainting," *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 417–424, 2000.

[14] Feng-Ying Xie, Shi-Yin Qin, Zhi-Guo Jiang, and Ru-Song Meng, "Pde-based unsupervised repair of hair-occluded information in dermoscopy images of melanoma," *Computerized Medical Imaging and Graphics*, vol. 33, no. 4, pp. 275 – 282, 2009.

[15] Qaisar Abbas, M.E. Celebi, and Irene Fondón García, "Hair removal methods: A comparative study for dermoscopy images," *Biomedical Signal Processing and Control*, vol. 6, no. 4, pp. 395 – 404, 2011.

[16] N. Otsu, "A threshold selection method from gray-level histograms," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 9, no. 1, pp. 62 –66, jan. 1979.

[17] Dietrich Stauffer and Ammon Aharony, *Introduction To Percolation Theory*, CRC Press, 2 edition, July 1994.