Multimed Tools Appl (2018) 77:9897–9908
https://doi.org/10.1007/s11042-018-5672-7

Flow-guided hair removal for automated skin lesion identification

Dongwann Kang[1] · **Sanggeun Kim**[2] · **Sangoh Park**[3]

Received: 1 July 2017 / Revised: 14 December 2017 / Accepted: 15 January 2018 /
Published online: 25 January 2018

Abstract In this paper, we propose a method for removing hairs automatically from skin lesion images. To achieve this, we employ an edge-detection technique based on edge-tangent flow. To detect only hair-like structures, rather than contour boundaries, we propose a novel, specialized method for detecting hairs. Regardless of the personal characteristics of the hairs, hairy regions are detected because our method detects coherent thin lines of consistent width. We then restore the hairy regions detected by the proposed method by using the texture synthesis method. Our method restores the regions occluded by hairs with very few remarkable artifacts, because we utilize pixels that actually exist in the source image to restore the occluded areas by searching for the best matching pixels.

Keywords Dermatology · Skin lesion · Hair removal · Line detection

1 Introduction

Skin cancer is divided into melanoma, which develops from the pigment-containing cells known as melanocytes, and non-melanoma, including basal cell carcinomas and squamous cell carcinomas, which are generally grouped together to distinguish them from melanomas. Melanoma, also known as malignant melanoma, is the most dangerous type of skin cancer. In 2012, more than 232,000 people were diagnosed with melanoma globally, and among them, 55,000 died [13]. Moreover, the rates of melanoma have been rising consistently for the last three decades. Although the risk of melanoma increases with age, melanoma

✉ Sangoh Park
sopark@cau.ac.kr

[1] Faculty of Science and Technology (SciTech), Bournemouth University, Poole, UK

[2] Department of Computer Science and Engineering, Sungkyul University, Anyang, South Korea

[3] School of Computer Science and Engineering, Chung-Ang University, Seoul, South Korea

is not uncommon even among people under 30. In fact, this type of skin cancer is one of the most common cancers in young adults. On the other hand, non-melanoma skin cancer, which refers to a group of cancers that slowly develop in the upper layers of the skin, have higher incidence rates, but can typically be treated with a relatively simple surgery. The survival rate for people with skin cancer depends upon when they start treatment. Even for melanoma, the cure rate is very high when it is detected in the early stages. Therefore, diagnosing skin cancer in its early stage is the most important factor for increasing the survival rate.

To diagnose possible skin cancer in an early stage, it is recommended to have the skin checked regularly by health professionals. Unfortunately, appropriate diagnosis and treatment of skin are not easy tasks, especially for persons in remote areas. Rural patients with skin diseases have few healthcare options, and may not obtain appropriate treatment without time-consuming and costly journeys to the nearest big city. Consequently, the deficiency of available dermatologic expertise in remote rural areas contributes to skin-related disabilities and high mortality rates. For rural patients, remote diagnosis via automatic identification of skin diseases is one of the most efficient and practical ways of improving skin care services. Recently, the rapid growth of high-performance computing with digital imaging has enabled support for therapeutic or diagnostic decisions, and most dermatology experts also agree that it is viable to automatically identify skin cancers through state-of-the-art image processing and pattern-recognition techniques [6].

Identification of skin lesions from dermatological images can be classified into three major tasks: eliminating noise, detecting the skin lesion area with its boundary, and extracting features for identifying or classifying skin lesion types. Among these tasks, eliminating noise, especially for hairs, is significantly important, because dermatological images usually contain hairs that occlude the lesion area, which can cause inaccurate identification. Therefore, in skin lesion identification, hairs must be removed appropriately to enhance identification performance, but this issue has not been sufficiently treated in this field. Generally, the color, shape, length, and amount of hair depends on the person, so detecting hairs is not an easy task. Moreover, after removing the detected hairs, the region occluded by hairs must be restored. On the one hand, if hairs are over-detected, some of the information of the lesion area may be lost. On the other hand, if the hairs are not fully removed or if the removed hairs are not appropriately restored, this may decrease identification performance.

To solve this problem, this paper proposes a method for removing hairs automatically from skin lesion images. Although the appearance of the hair depends on the person, hairs have a common characteristic which is distinguishable from the lesion. In dermatology images, they are usually represented as individual directional thin lines. To detect this, we employ an edge-detection technique which is based on edge-tangent flow. This technique effectively recognizes boundaries in images. However, we need to detect only hair-like structures, rather than boundaries. To achieve this, based on this technique, we propose a novel method specialized in detecting hairs. We then restore the hairy regions detected by the proposed method by using the texture synthesis method.

The main contributions of this paper are as follows. First, the method proposed in this paper is able to detect hair-like regions that consist of coherent thin lines with consistent width. Consequently, regardless of the personal characteristics of hairs, hairy regions are detected automatically. Second, our method, which employs the texture synthesis technique, restores the regions occluded by hairs with few remarkable artifacts, because we utilize the pixels that actually exist in the source image to restore the occluded areas by searching for the best matching pixels.

The remainder of this study is organized as follows. In Section 2, we provide an overview of studies related to hair removal in dermatological imaging. We then present the details of our method for detecting and restoring the hairy regions in Section 3. The results of our method and a comparison of the performance are given in Section 4. Finally, we conclude with a summary of our ideas and discuss future work in Section 5.

2 Related work

In dermatology, especially for skin cancer diagnosis, computer-assisted identification based on image processing and pattern recognition is regarded as an effective and useful tool, so that many researches have been conducted to classify skin lesion images. Schmid [11] proposed a L*u*v* color-based segmentation approach for dermatoscopic images. He calculated a smoothed 2D histogram with the two principal components, then obtained the maxima location and a set of features from the histogram contour lines, which is useful to determine the number of classes. Lesion region was finally separated by using segmentation based on the fuzzy c-means clustering. Grana et al. [7] segmented melanoma images through Otsu's thresholding, and detected smooth lesion boundary by considering luminance along a normal direction of the contour. Their method enabled the evaluation of lesions with irregular shape. Zhang et al. [18] proposed a method for detecting skin tumor borders in clinical dermatology images using a radial search. Before two rounds of radial search in their methods, low-pass filtering was performed to remove the noise including flash reflections and hairs. To detect melanoma skin lesion images automatically, Taouil et al. [14] used segmentation methods based on thresholding, morphology functions and active contours. As a preprocessing step, they used the combination of different morphology functions, such as the top-hat and bottom-hat transforms, to remove noise in the digitized images before the segmentation step. Consequently, this step improved the quality of detected boundary.

As mentioned above, the noise problem especially for hair was one of main concerns of skin lesion detection and identification, but it has not been sufficiently addressed in the computer assisted dermatologic diagnosis field. Lee et al. [10] proposed a method to remove hairs from a dermatologic image using DullRazor®, which is a pre-processing program detected the hair by using morphological closing operator and then replaced the pixels underneath the hair through bilinear interpolation. However, this method was not only effective just for the thick and dark hair, but also led blurry artifact. Nonetheless, due to its simplicity, similar approaches were employed in many researches: median filtering [11], morphological operation [5, 12].

To remove hair from dermatologic image, image processing techniques based on partial differential equation (PDE) were used in a few studies [3, 17]. In these studies, PDE-based erosion operation with line segment eliminates straight structures perpendicular to it. This approach has an advantage that linear geometry in the image is appropriately preserved. However, it also leads noticeable blurry pixels where hair is removed.

A few studies [1, 16, 19] employed texture synthesis-based approach which fills missing pixels to restore the region where hair was removed. These studies are based on the exemplar-based image inpainting technique [4]. In general, this approach effectively removes the hair without leading blurry artifact. In this paper, we also employ a texture synthesis technique to restore hair-occluded region.

3 Hair removal for skin lesion identification

Our method, which removes hair from dermatological images, consists of two parts: detecting hair regions and restoring them. In the following sections, we present the details of each part.

3.1 Hair detection using flow-guided line detection

In many studies that remove hairs from dermatological images, traditional edge detectors in the image processing field are utilized to detect the hair. Generally, edge detectors use first- and second-order derivatives of the image gradient to detect edges in images. Methods that use the first-order derivative usually compute the strength of edges by obtaining the gradient magnitude, and search for local maxima of gradients. The other methods search for zero-crossings in the second-order derivates. These methods quite effectively find edges, and have been consequently employed in various fields. However, generally, edge detectors do not distinguish actual lines and boundaries of contours. Therefore, for hair removal, which must focus on detecting actual lines rather than boundaries, these methods are not very suitable.

In this paper, we propose an edge-detection method suitable for actual lines to remove hairs. It is not easy to distinguish the boundaries of contours and lines, because distinguishing lines with width from contours is ambiguous. Especially in dermatological images, in the case of a low-resolution image, hairs can be represented as thin lines; however, in cases of high resolution images, hairs are likely to be represented as directional long contours. To cope with this, we define hairs as bi-directional regions with regular width that are distinguishable from the surrounding regions, and propose a specialized algorithm for detecting them.

The edge tangent flow (ETF) algorithm [8], which is widely used to draw coherent lines following directional features in the stylization field [9], generates a coherent directional vector field by weighted-averaging edge-tangent vectors that are perpendicular to image gradients, and effectively detects and draws long edges by applying the edge detection filter, deformed non-linearly along the field. In this point of view, this method is very suitable for our study. Therefore, in this paper, we detect hairs coherently by utilizing an ETF field. First, we obtain a vector which is tangent to the edge direction on each pixel by using (1).

$$v_t(x) = R_{90} v_g(x) \tag{1}$$

where, $v_g(x)$ and $v_t(x)$ refer to a gradient vector and a tangent vector of pixel x, respectively, and R_{90} indicates a matrix that rotates a vector by 90 degree. By using (2), we then weight-average the direction of the vectors using their magnitudes and directions, and consequently find the smoothed vectors v_s that are robust for noise and follow the neighboring strong-edge directions.

where, $v_g(x)$ and $v_t(x)$ refer to a gradient vector and a tangent vector of pixel x, respectively, and R_{90} indicates a matrix that rotates a vector by 90 degree. By using (2), we then weight-average the gradient directions using their magnitudes, and consequently find the smoothed directions v_s that are robust for noise and follow the neighboring strong-edge directions.

$$v_s(x) = \frac{1}{k} \sum_{y \in \Omega(x)} \phi(x, y) v_t(x) \omega_s(x, y) \omega_m(x, y) \omega_d(x, y) \tag{2}$$

where k is the normalizing term, and $\Omega(x)$ denotes the set of neighbors of x. ω_s, ω_m and ω_d mean the spatial-, magnitude-, and direction weight function, respectively. These functions are defined as follows.

$$\omega_s(x, y) = \begin{cases} 1 & \text{if } ||x - y|| < r, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

$$\omega_m(x, y) = \frac{1}{2}(1 + tanh[\eta(\hat{g}(y) - \hat{g}(x))]) \tag{4}$$

$$\omega_d(x, y) = |v_t(x) \cdot v_t(y)| \tag{5}$$

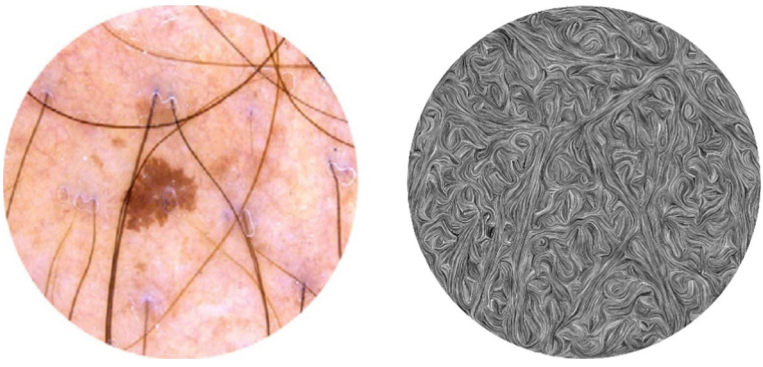
$$\phi(x, y) = \begin{cases} 1 & \text{if } v_t(x) \cdot v_t(y) > 0, \\ -1 & \text{otherwise.} \end{cases} \tag{6}$$

where $\hat{g}(x)$ denotes the normalized gradient magnitude at x. Please refer to [8] for more details. Figure 1 shows the flow field calculated by ETF.

Once the ETF field is computed, we apply the filter for obtaining hair-like regions along the field. In the original method [8], a flow-based difference of Gaussian (FDoG) filter, which is an anisotropic DoG filter based on flow, is employed to detect edges. This filter effectively detects directional edges, even if source image contains plenty of noise, because it calculates the first-order derivatives along the direction perpendicular to flow direction, and accumulates them along the flow. However, FDoG detects not only lines but also the boundaries of contours, whereas we need to use a specific filter suitable for recognizing only lines to find the hair regions.

In general, hair region is narrow and relatively darker than skin region. Consequently, in a cutaway view of hair (Fig. 2), the brightness is drastically decreased at both sides of boundaries of hair region toward their center. Based on this knowledge, we propose the flow-guided hair detector. Figure 3 illustrates our detection scheme. In the figure, $c_x(s)$ denotes the flow line at x, where s is an arc-length parameter. If x is at the center of curve, then it is equivalent to $c_x(0)$. In our filtering scheme, 1D filter F is applied traveling along c_x. At this time, $F(s)$ is located on the line l_s perpendicular to $v_t(c_x(s))$. When the radius of hair is given as T, a 1-dimensional filter $F(s, t)$ for detecting hair is defined as follows:

$$F(s, T) = \int_{-T}^{T} f(t, l_s(t))dt \tag{7}$$



(a) skin lesion image (b) ETF field

Fig. 1 ETF field visualized by using line integral convolution [2]

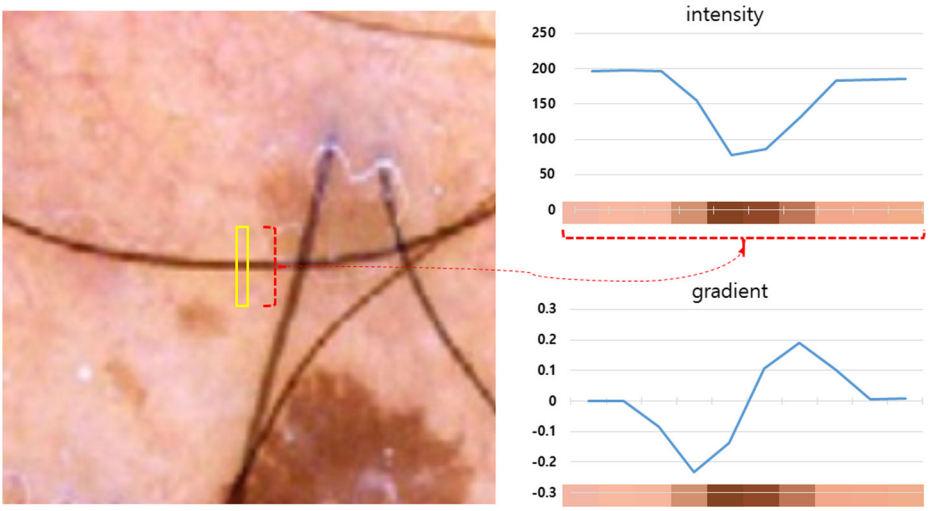


Fig. 2 A cutaway view of hair region

where t and $l_s(t)$ denote an arc-length parameter and the point on the line l_s at parameter t, respectively. Note that l_s is parallel to the gradient vector $v_g(c_x(s))$. f calculates the difference between gradients of $l_s(t)$ and the ideal hair region:

$$f(t, l_s(t)) = \begin{cases} |\hat{g}(l_s(t))| & \text{if } t = 0, \\ -|t|/t \times \hat{g}(l_s(t)) \times 0.5 + 0.5 & \text{otherwise.} \end{cases} \tag{8}$$

Then each filter f is accumulated along c_x:

$$H(x, T) = \int_{-S}^{S} G_\sigma(s) F'(s, T) ds \tag{9}$$

$$F'(s, T) = \max(F(s, i) : i = T - \alpha, \cdots, T + \alpha) \tag{10}$$

where G_σ denotes a a1-dimensional Gaussian function of variance σ. α controls a tolerance of hair consistency. In this paper, we set $G_\sigma = 7$ and $\alpha = 1$. H effectively detects hair regions when the width of hair is given. However, the width depends on the resolution of image so that we apply H with various T, and then find minima:

$$H(x) = \max(H(x, u) : u = 1, \cdots, T_{max}) \tag{11}$$

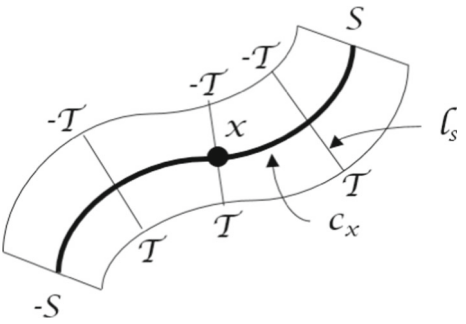


Fig. 3 Our filtering scheme for detecting hair regions

(a) the result of H (b) the hair map

Fig. 4 The hair we detected

where T_{max} denotes user-defined minimum value of T. We also can find the radius which minimizes H:

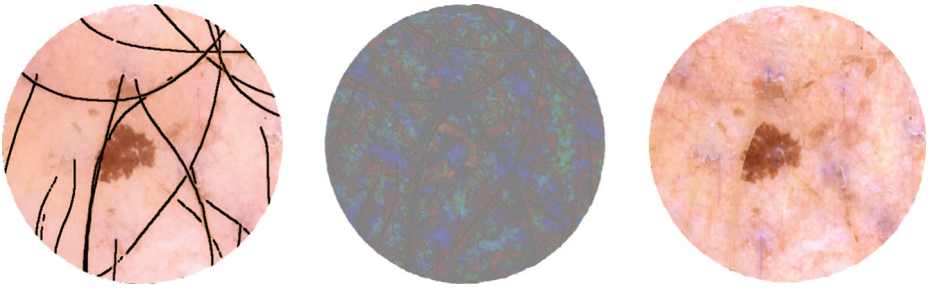
$$R(x) = \arg\min_{T} H(x, T) \tag{12}$$

Figure 4a shows the result of H. However, the result does not match the radius we found in (12). To overcome this, we skeletonize the result of H by finding local minima of H along $l_s(t)$, where $R \leq t \leq R$, and then expand its radius by using R. We finally obtain a hair map consists of pixel of which H value is lower than a user-defined threshold. Figure 4b shows the hair map.

3.2 Hair restoration using texture synthesis

In previous studies, various methods for restoring the removed hair region, including averaging, blurring, and median filtering, are used. However, these methods occasionally generate blurred textures. As a result, the restored regions can have artifacts, which significantly decrease the performance of skin lesion identification. To avoid this, we propose a method based on texture synthesis.

First, we decompose input image into two layers; base layer and texture layer. To extract the base layer (Fig. 5a) from input image, we employ bilateral filtering [15]. We then



(a) base layer (b) texture layer (c) hair restored image

Fig. 5 Our hair restoration result

obtain the texture layer (Fig. 5b) by subtracting the base layer from input image in *lab* color space. Second, on the base layer, we remove the pixels indicated as hair in the hair map obtained in Section 3.1, and then fill holes by using morphological dilation operation.

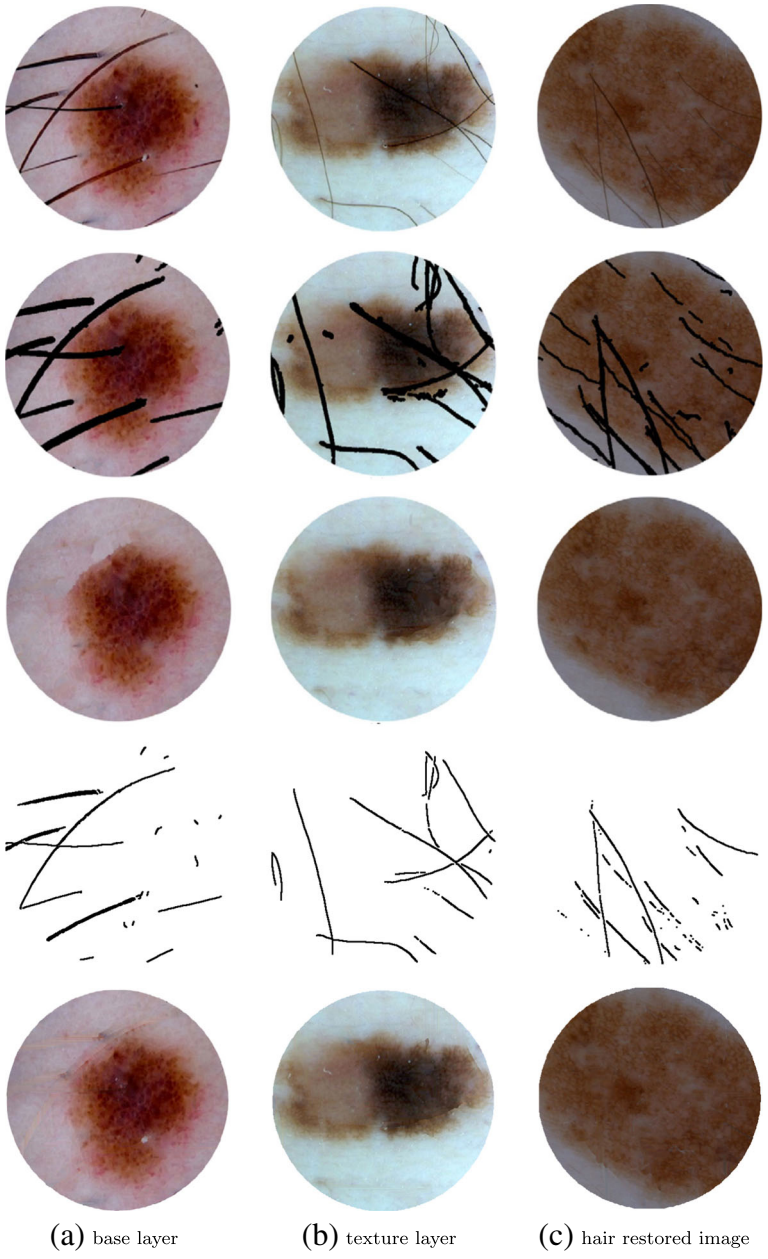


Fig. 6 Result comparison. From top to bottom: input images, [1]'s hair detection results, [1]'s hair region restoration results, our hair detection results, our hair region restoration results

Finally, we transfer the texture of non-hair pixels in the texture layer to the base layer by using (13).

$$D(p,q) = \frac{h}{H} \sum_{x \in N(p), y \in N(q)} k(x,y)|c_b(x) - c_b(y)| \cdot ||c_t(x) - c_t(y)|| \tag{13}$$

where $N()$ denotes neighboring function, $c_b(x)$ and $c_t(x)$ denotes the color vector at x in base layer and texture layer, $k(p,q)$ denotes the indicator whether both p and q were not indicated by the hair map, h denotes the number of pixels of which $k(p,q)$ value is 1, and H denotes the number of pixels in $N()$. For each pixel indicated by the hair map, we search for the best restoring pixel among all the pixels in the same texture layer. The best restoring pixel is defined as the pixel which minimizes $D(p,q)$ because this pixel has a local texture that is the most similar to the pixel indicated by the hair map. We replace the pixels in the hair region in the texture layer with the pixel found to be the best restoring pixel. We then synthesize two layers by adding their color values. Figure 5c shows our hair restoration result.

4 Results

In our experiments, a PC with an Intel i5 6700 CPU (8MB cache, 3.4 GHz), 16 GB RAM, and OpenCV image processing library were used to implement our method. Our proposed framework was written in C++. It took up to 2–3 s for calculating H of (11) from an input image with VGA resolution. Once H was obtained, users adjusted a user-defined threshold, then the hair map was generated and displayed in real-time. Because the region near hair is slightly darker than other skin region, we dilated hair map before applying restoration process to prevent that hair region is restored with dark pixels. The process which restored hair region took 1–2 min approximately.

Figures 4b and 5b show our hair map and hair restoration result. As shown in the figure, most hair in the image has been correctly detected on the map while ignoring the boundaries of the lesion area. It is also shown that the region detected by the hair map is restored with few remarkable artifacts.

Figure 6 shows the result comparison between [1] and our results. In contrast to [1], it is shown that our method detected only dark hair region. This was led by our hair detection function presented in (8).

5 Conclusions

In this paper, we proposed a method for removing hairs automatically from skin lesion images. To achieve this, we employed an edge-detection technique based on edge-tangent flow. To detect only hair-like structures and not contour boundaries, we proposed a novel method specialized in detecting hairs. Regardless of the personal characteristics of the hairs, the hairy regions are reasonably detected because our method detects coherent thin lines with consistent widths. We then restored the hairy regions detected by the proposed method by using the texture synthesis method. Our method restored the regions occluded by hairs with very few remarkable artifacts, because we utilize pixels from the source image to restore the occluded area by searching for the best matching pixels.

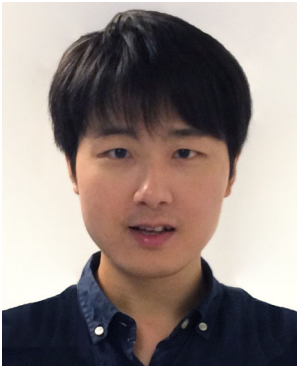
In future work, we will consider a comparison of the performance of lesion image identification by using various hair-removal techniques. Moreover, we will also compare hair removal performance by using ground-truth dermatological images obtained by physical shaving.

Acknowledgements This research was supported by the Chung-Ang University Research Grant in 2017.

References

1. Abbas Q, Fondón I, Rashid M (2011) Unsupervised skin lesions border detection via two-dimensional image analysis. Comput Methods Prog Biomed 104(3):e1–e15. https://doi.org/10.1016/j.cmpb.2010.06.016. http://www.sciencedirect.com/science/article/pii/S0169260710001690
2. Cabral B, Leedom LC (1993) Imaging vector fields using line integral convolution. In: Proceedings of the 20th annual conference on computer graphics and interactive techniques, SIGGRAPH '93. ACM, New York, pp 263–270. https://doi.org/10.1145/166117.166151. http://doi.acm.org/10.1145/166117.166151
3. Chung DH, Sapiro G (2000) Segmenting skin lesions with partial-differential-equations-based image processing algorithms. IEEE Trans Med Imaging 19(7):763–767. https://doi.org/10.1109/42.875204
4. Criminisi A, Perez P, Toyama K (2004) Region filling and object removal by exemplar-based image inpainting. IEEE Trans Image Process 13(9):1200–1212. https://doi.org/10.1109/TIP.2004.833105
5. Fleming MG, Steger C, Zhang J, Gao J, Cognetta AB, Pollak L, Dyer CR (1998) Techniques for a structural analysis of dermatoscopic imagery. Comput Med Imaging Graph 22(5):375–389. https://doi.org/10.1016/S0895-6111(98)00048-2. http://www.sciencedirect.com/science/article/pii/S0895611198000482
6. Ganster H, Pinz P, Rohrer R, Wildling E, Binder M, Kittler H (2001) Automated melanoma recognition. IEEE Trans Med Imaging 20(3):233–239. https://doi.org/10.1109/42.918473
7. Grana C, Pellacani G, Cucchiara R, Seidenari S (2003) A new algorithm for border description of polarized light surface microscopic images of pigmented skin lesions. IEEE Trans Med Imaging 22(8):959–964. https://doi.org/10.1109/TMI.2003.815901
8. Kang H, Lee S, Chui CK (2007) Coherent line drawing. In: Proceedings of the 5th international symposium on non-photorealistic animation and rendering. ACM, pp 43–50
9. Kyprianidis JE, Collomosse J, Wang T, Isenberg T (2013) State of the "art": a taxonomy of artistic stylization techniques for images and video. IEEE Trans Vis Comput Graph 19(5):866–885. https://doi.org/10.1109/TVCG.2012.160
10. Lee T, Ng V, Gallagher R, Coldman A, McLean D (1997) Dullrazor: a software approach to hair removal from images. Comput Biol Med 27(6):533–543. https://doi.org/10.1016/S0010-4825(97)00020-6. http://www.sciencedirect.com/science/article/pii/S0010482597000206
11. Schmid P (1999) Segmentation of digitized dermatoscopic images by two-dimensional color clustering. IEEE Trans Med Imaging 18(2):164–171. https://doi.org/10.1109/42.759124
12. Schmid-Saugeon P, Guillod J, Thiran JP (2003) Towards a computer-aided diagnosis system for pigmented skin lesions. Comput Med Imaging Graph 27(1):65–78. https://doi.org/10.1016/S0895-6111(02)00048-4. http://www.sciencedirect.com/science/article/pii/S0895611102000484
13. Society AC. Key statistics for melanoma skin cancer. http://www.cancer.org/cancer/skincancer-melanoma/detailedguide/melanoma-skin-cancer-key-statistics
14. Taouil K, Romdhane NB, Bouhlel MS (2006) A new automatic approach for edge detection of skin lesion images. In: 2006 2nd international conference on information communication technologies, vol 1, pp 212–220. https://doi.org/10.1109/ICTTA.2006.1684373
15. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: Sixth international conference on computer vision (IEEE Cat. No.98CH36271), pp 839–846. https://doi.org/10.1109/ICCV.1998.710815
16. Wighton P, Lee TK, Atkins MS (2008) Dermascopic hair disocclusion using inpainting. In: SPIE medical imaging, vol 6914. https://doi.org/10.1117/12.770776
17. Xie FY, Qin SY, Jiang ZG, Meng RS (2009) Pde-based unsupervised repair of hair-occluded information in dermoscopy images of melanoma. Comput Med Imaging Graph 33(4):275–282. https://doi.org/10.1016/j.compmedimag.2009.01.003. http://www.sciencedirect.com/science/article/pii/S0895611109000056

18. Zhang Z, Stoecker WV, Moss RH (2000) Border detection on digitized skin tumor images. IEEE Trans Med Imaging 19(11):1128–1143. https://doi.org/10.1109/42.896789
19. Zhou H, Chen M, Gass R, Rehg JM, Ferris L, Ho J, Drogowski L (2008) Feature-preserving artifact removal from dermoscopy images. In: SPIE medical imaging, vol 6914. https://doi.org/10.1117/12.770824



Dongwann Kang received his Ph.D. degree at Chung-Ang University, Korea in 2013. He also received his B.S. degree in Computer Science and Engineering from Chung-Ang University in 2006. He was a research fellow at Chung-Ang University from Mar. 2013 to Jun. 2015. Now, he is a Marie Skłodowska-Curie fellow at the faculty of Science and Technology (SciTech), Bournemouth University, UK. His research interests include stylization, emotional computing, image manipulation and GPU processing.

Sanggeun Kim received his PhD, MS and BS in Computer Science from the ChungAng University in 1996, 1989 and 1987, respectively. He is currently a faculty of Department of Computer Science and Engineering at Sungkyul University, Anyang, Rep. of Korea. His research interests include: S/E, Cloud Computing, Ubiquitous Computing, image processing.

Sangoh Park received the B.S., M.S., and Ph.D. degrees from the School of Computer Science and Engineering at Chung-Ang University in 2005, 2007, and 2010, respectively. He has been serving as an Assistant Professor of the School of Computer Science and Engineering at Chung-Ang University since 2017. He served as a Senior Researcher of Global Science experimental Data hub Center at Korea Institute of Science and Technology Information from 2012 to 2017 and a Research Professor at the School of Computer Science and Engineering. His research interests include high-performance computing, big data system, tape storage system, embedded system, cyber physical system, home network, smart factory and linux system.