

A precise automatic system for the hair assessment in hair-care diagnosis applications

H. Shih^{1,2}

¹Human-Computer Interaction Multimedia Laboratory, Department of Electrical Engineering, Yuan Ze University, Taoyuan, Taiwan and ²The Innovation Center for Big Data and Digital Convergence, Yuan Ze University, Taoyuan, Taiwan

Background/Purpose: One emerging subject in medical image processing is to quantitatively assess the health and the properties of cranial hairs, including density, diameter, length, level of oiliness, and others. This information helps hair specialists with making a more accurate diagnosis and the therapy required. We develop a practical hair counting algorithm. This analytic system calculates the number of hairs on a scalp using a digital microscope camera, providing accurate information for both the hair specialist and the patient. Our proposed hair counting algorithm is substantially more accurate than the Hough-based one, and is robust to curls, oily scalp, noise-corruption, and overlapping hairs, under various levels of illumination. Rather than manually counting the hairs on a person's scalp, the proposed system determines the density, diameter, length, and level of oiliness of the hairs.

Methods: We propose an automated system for counting the amount of hairs in the microscopy images. To reduce the effect of bright spots, we develop a robust morphological algorithm for color to smooth out the color and preserve the fidelity of the hair. Then, we utilize a modified Hough transform algorithm to detect the different hair lengths and to reduce any false detection due to noise. Our proposed system enables us to look at curved hairs as multiple pieces of straight lines. To avoid missing hairs when the thinning process is applied, we use edge information to discover any hidden or overlapping hairs. Finally, we employ a mutually associative regression

method to label a group of line segments into a meaningful 'hair'.

Results: We demonstrated a novel approach for accurately computing the number of hairs, and successfully solved the three main obstacles in automated hair counting, including (i) oily and moist hairs, (ii) wavy and curly hairs, and (iii) underestimation of the number of hairs occurs when hairs cross and occlude each other. The framework of this paper can be seen as the first step toward intelligent computer-aided medical image processing for cosmetic treatment applications.

Conclusions: The goal of this study was to develop an automated hair counting system for clinical application using the microscope image from the hairs. The proposed method reduces the frequent errors and variances encountered when hairs are manually counted by human assessors. This clinical intelligent system can diagnose the health condition of a person's hair and can be applied in therapy recommendations by doctors for their patients.

Key words: hair counting – scalp diagnosis – testing of hair follicles – pathology analysis of a hair follicle – follicle diagnosis

© 2015 John Wiley & Sons A/S. Published by John Wiley & Sons Ltd
Accepted for publication 14 February 2015

CLASSICALLY, THE hair image analysis has been focused mainly on pigmented skin lesion images for hair removal (1–3) and border detection (4–6). Research literature has been focused mainly on hair segmentation and removal because the hairs are the obstacles for the clinical diagnosis of the scalp. Regarding to the skin texture analysis, many researches address on the age-dependent changes in skin surface using digital image analysis approaches (7–11). Huang et al. (12) evaluated a hair removal software, DullRazor (13), and proposed conventional matched filters to enhance the curvilinear structures and improve the detection of

thin hairs and hairs in shadow. The authors also demonstrated that missing hair intersections can be recovered by applying regional growing algorithms with a color-similarity criteria. However, these aforementioned techniques do not take into consideration the condition of the hairs such as their health condition, density, diameter, oiliness, nor the number of hairs on each single follicle.

Nowadays, we live in an environment that suffers from UV rays overexposure, air pollution, and acid rain. As a result, scalp care has become a public health issue. Some people want to quantitatively assess the health condition

of their cranial hair, including density, diameter, and length, because these hair properties enable a hair specialist to make a better informed decision regarding the therapy to use. Some recent literature focuses on the parameters and condition of the hairs and the follicles (14–16). Based on the follicular conditions, many pathological changes need to be considered, including pore obstruction, over-active secretion of sebum, sensitive scalp, and if an unsuitable hair cleanser has been used. Many people are interested in their bodily hair, especially the hair on their scalp, resulting in the rapid growth of the scalp care industry. At the same time, new drugs such as Finasteride are available for hair regrowth (17). In order to support the development of these treatments and drugs, and to inform the patients on their efficacy, dependable mechanisms are required to count the number of hairs as well as determine their average length and diameter.

At present, the diagnosis of a person's hair and scalp relies mainly on the skill and knowledge of professional scalp assessors. The density and diameter of a person's hairs usually reflects the health condition of that person's scalp. Hair counting is usually performed manually by a human assessor, and often with unreliable results. In addition, manual counting cannot tell us the diameter, length, and oiliness of the hairs. An automated system capable of counting body hairs was proposed in (18). In the present study, we focused on hair counting on the scalp. Performing a detailed scalp assessment is very time consuming. In an attempt to solve this problem, this study proposed an automated scalp diagnosis system to reduce the time required for making a detailed scalp assessment. Testing images were obtained by means of digital microscopy and then used to, together with the knowledge of professional scalp assessors and doctors, accurately diagnose the hair condition and determine the optimal scalp treatment.

Material and Methods

Hair segments extraction scheme

The only assumption made in this study is that the hair color is darker than the color of the skin. To simplify the system, we used black and brown hairs for the test images. It should be noted, however, that the outcome for hairs with

other colors will be the same as long as the eye can make the distinction between the color of the hair and the scalp. Most noise in a test image could be effectively depressed. However, a few discrepancies could not be fully eliminated using the fundamental image processing algorithms. For instance, hairs with a gap in their length suffer from oily spot reflection. When using the conventional region growing method, it will produce many false mergers between individual hairs. To solve this 'gap' problem, we present a morphology-based bright spot removal procedure. Table 1 shows the detailed tasks performed in each step.

Step 1: pre-processing

Contrast is the main factor for determining the quality of the image. However, images often lack sufficient contrast due to the variations in the environment. We aimed to increase the color distinction between hair and scalp pixels. We therefore stretched the middle-intensity level, and kept the levels of the low-intensity and high-intensity so as to prevent creating false colors. The two transit points $\{(r_1, s_1), (r_2, s_2)\}$ are empirically assigned. The contrast stretching is performed without changing the original hair pixels and without exaggerating the oily bright pixels. As a result, the intensity of the scalp pixels was reduced and the color distinction between scalp and hair was increased. In addition, the pixels of the bright spots remained unchanged, but they were toned down by the bright spot removal process proposed in this study.

In this study, we used the Karhunen-Loève transform (19) to convert the color image to a gray-scale image, with each color component keeping the highest energy component. In the

TABLE 1. Four major steps of hair segments extraction

Step 1: pre-processing	Image enhancement Color-to-gray conversion Image binarization and denoizing
Step 2: bright spot removal	Color morphology Hybrid ordering approach
Step 3: multi-scale line detection (MSLD)	Thinning process Hough transform Image rescaling
Step 4: parallel line bundling (PLB)	Edge detection Distance measurement for neighboring line segments

image binarization step, we used Otsu thresholding (20) (referring to the energy of the brightness) to obtain a reliable binary image. This process not only increases the process efficiency, but also purifies the resulting image.

Step 2: bright spot removal

In the image acquisition step, the oily and moist hair produces bright spots in the middle portions of the hair. The reason for the appearance of these oily spots is due to the light reflecting from the built-in LEDs of the digital microscopy device. The surfaces of the moist hair and the oily scalp reflect the supplementary light into the Complementary Metal-Oxide-Semiconductor/Charge-coupled Device sensors of the digital microscopy device. This not only results in early sensing saturation, but also replaces the original color with a white spot. Thus, we need to eliminate any bright spots prior to counting the hairs, and avoid the situation of 'breaks' in the hairs. Unfortunately, it is impossible to precisely know the locations of all hair pixels before the hair is detected. To preserve the scalp information such as the intrinsic color pattern and skin texture for further applications, we utilized color morphological processing approach rather than the binary or gray-scale one for handling this kind of defect.

There are many spatial filtering operators that can suppress the white spots. By treating a white spot as a salt and pepper noise (i.e., impulse noise), we can apply a nonlinear median filter to eliminate it. However, the size of the filtering mask is critical. Another approach is the spatial smooth filter, which can successfully reduce the intensity of the white spot. The drawback, however, is that the non-hair regions of the test image will also be blurred. To solve these difficulties we used the color-based mathematical morphology method, and used it as an ordering process. The mathematical morphology defines the correlation between pixels in the spatial structure, and compares the order. This, however, requires the ordering and analysis of multi-dimensions. In addition, a Red-Green-Blue color image does not have a definite ordering pattern when compared to a gray scale image. Common sequential ordering can be classified into four categories (21) including Marginal-ordering (M-ordering), Conditional-

ordering (C-ordering), Partial-ordering (P-ordering), and Reduced-ordering (R-ordering). Each ordering approach has its pros and cons. For instance, M-ordering is quick and simple but it is error prone when it comes to color accuracy, R-ordering decreases computation time, but it lacks integrity, and while C-ordering completely avoids the ambiguous situations caused by the sequential comparison error, it has a high computational complexity. In this article we applied our previous study (22) on a hybrid ordering approach based on the M-ordering and C-ordering methods that will solve aforementioned problems.

Step 3: multi-scale line detection

The Hough transform (HT) (23, 24) is one of the most used frameworks for line detection. The principle of the HT is based on the scatter position points of an image for identifying the parameters of a certain shape (e.g. straight line or circle). Every point is mapped onto a few other candidate points so as to generate parameters of possibility. Then all the values are collected and used to decide the most probable shape. The Hough line indicates a vectorized line segment reversed from the feature space. Generally speaking, its two major forms are slope-intercept and angle-radius.

However, a large amount of hair segments can be lost when the conventional single-scale HT is applied. At the same time, having to choose local maximal values in parametric space is an annoying problem. It is difficult to achieve a convincing result for hair detection because some hairs curl, cross-over each other, and there may be dandruff and dust present. First, we propose a multi-scale framework to improve the accuracy of hair detection. Because the parameters of the HT are sensitive to the length of a line segment, we employed three scales of the images [i.e., 1024×768 (original scale), 512×384 , and 256×192] to the HT, while preserving the different lengths of the hairs. Two processing methods are applied to the scaled images, the edge detection and the thinning process. In addition, we applied a parallel line bundling (PLB) algorithm (see Step 4) for edge image to restore any hidden or overlapped hairs. Finally, the vectorized line segments from thinning process are used as input data for the hair labeling and counting module.

Step 4: parallel line bundling

In the image binarization step, the scalp pixels located between two hairs are labeled as part of a hair, due to the incorrectly assumed connection of two individual hairs. Moreover, when two hairs are too close together or overlap each other, one or both hairs will be missed if the thinning algorithm is applied directly. Therefore, we used edge information to discover any hidden or overlapping hairs. From the edge viewpoint, two parallel edges can be extracted from a hidden hair or multiple hairs that overlap. Based on the distance between two parallel lines, we can determine if there are any hairs that are hidden or are overlapping. First, we apply the Canny edge detector (25) to obtain the edge map. Suppose that two parallel lines A and B are detected expressed by $ax + by + c_a = 0$ and $ax + by + c_b = 0$, where d denotes the distance between edge lines A and B, and where $d = |c_a - c_b|/\sqrt{a^2 + b^2}$. Second, we compute the average distance d_{avg} between the parallel lines that sandwich a thinning line. If the distance value is large than d_{avg} , we will continue discovering hidden hairs, when $(d - d_{avg}) > w_{th}$, where w_{th} denotes the bounding factor. This factor will be empirically modified by the resolution of the scalp image.

In summary, the proposed PLB algorithm is applied to the edge image to discover the missing line segments, and the HT is applied to the thinning images to extract the line segments. By taking advantage of the PLB algorithm, the hidden and overlapping hairs can be restored. Finally, the vectorized line segments are rescaled into their original size of 1024×768 , and integrated by a logical OR operator.

Hair labeling and counting

Using the multi-scale line detection (MSLD) module, we derived a set of line segments. The hairs were realized as a cluster of piecewise line vectors with different lengths depending on the curvature and the orientation of the hair flow. This aim of this study was to accurately count the number of hairs on a scalp. This can be viewed as a clustering and labeling problem. The goal is to assemble a group of line segments into a semantic 'hair' and assign a unique label. Because each hair consists of a cluster of mutually correlated line segments, we assigned a unique label to each cluster. We

adopted the Relaxation labeling algorithm (RLA) (26) for the identification of each individual line segment to determine which line segments it was associated with.

Relaxation labeling algorithm is a mutually associative regression method, which uses symbols to describe the shape of model. It aims to match the target object (i.e., line segment in this article) with the symbol or so-called marker (i.e., hair label). The RLA first assigns a set of random markers. Then, through iterative calculations a more accurate and precise set of markers is obtained. In this study, the RLA is treated as a clustering method for labeling every line segment. A compatibility factor in the RLA provides the clue for converging the grouping operation. Here, the compatibility is in terms of orientation and distance from the origin. This labeling algorithm is an iterative parallel procedure analogous to the label-discarding rule used in probabilistic relaxation. The operator iteratively adjusts the label weights based on the other weights and compatibilities. For each line segment and each label, a new weight will be updated iteratively until the compatibilities stop varying or converge to 1. Each line segment is iteratively verified until it has been assigned to the correct semantic label 'hair'.

Results

Experimental setup

To evaluate our system, we captured 40 scale images with a resolution of 1024×768 from the UPMOST (UPG622) digital microscope as the dataset for testing. We used a zoom magnification of $85\times$ to capture images. Equivalently, the captured area on the scalp is around 0.25×0.19 inch².

Evaluation of the hair labeling algorithm

The result of hair labeling is shown in Fig. 1. The speed of the convergence depends on how many line segments there were with correlated orientations and small distances. For instance, label #L7 converged first and labels #L4 and #L17 converged last. The hair label #L7 converged fastest to 1, because it was an isolated single line segment with less hair competing for its label.

The horizontal axis in Fig. 2 denotes the amount of hairs that were found, while the vertical axis denotes the number of line segments

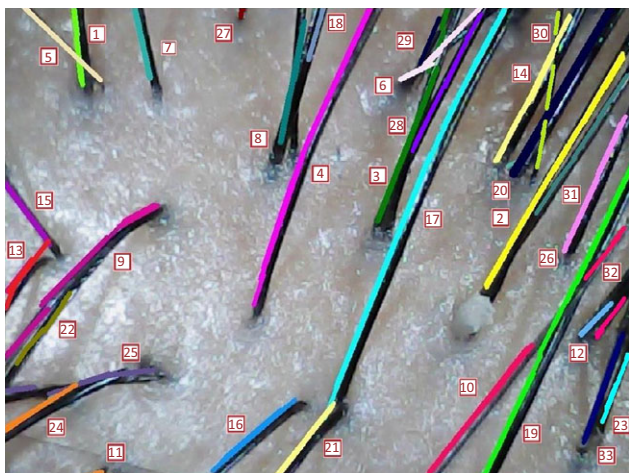


Fig. 1. Result of hair labeling.

that corresponded to one hair label. The histogram shows that there were 33 hairs detected. Hair labels #L4 and #L17 were composed of 13

line segments due to the long length and the curves of the hairs. Nevertheless, our proposed algorithm successfully grouped all line segments into the corresponding hair label.

Evaluation of the line segment detection

To evaluate the performance of the MSLD algorithm, the testing images we applied three different scales for the HT for extracting line segments. The hair labeling mechanism was then used to determine the number of the hairs. As shown in Fig. 3, the colored lines represent the label of the hair. In other words, hairs can be labeled accurately even when they are crossed or overlapped.

The following section demonstrates the objective measurement of hair counting regarding precision and recall rates. We compared four cases using combinations of the BSR and MSLD

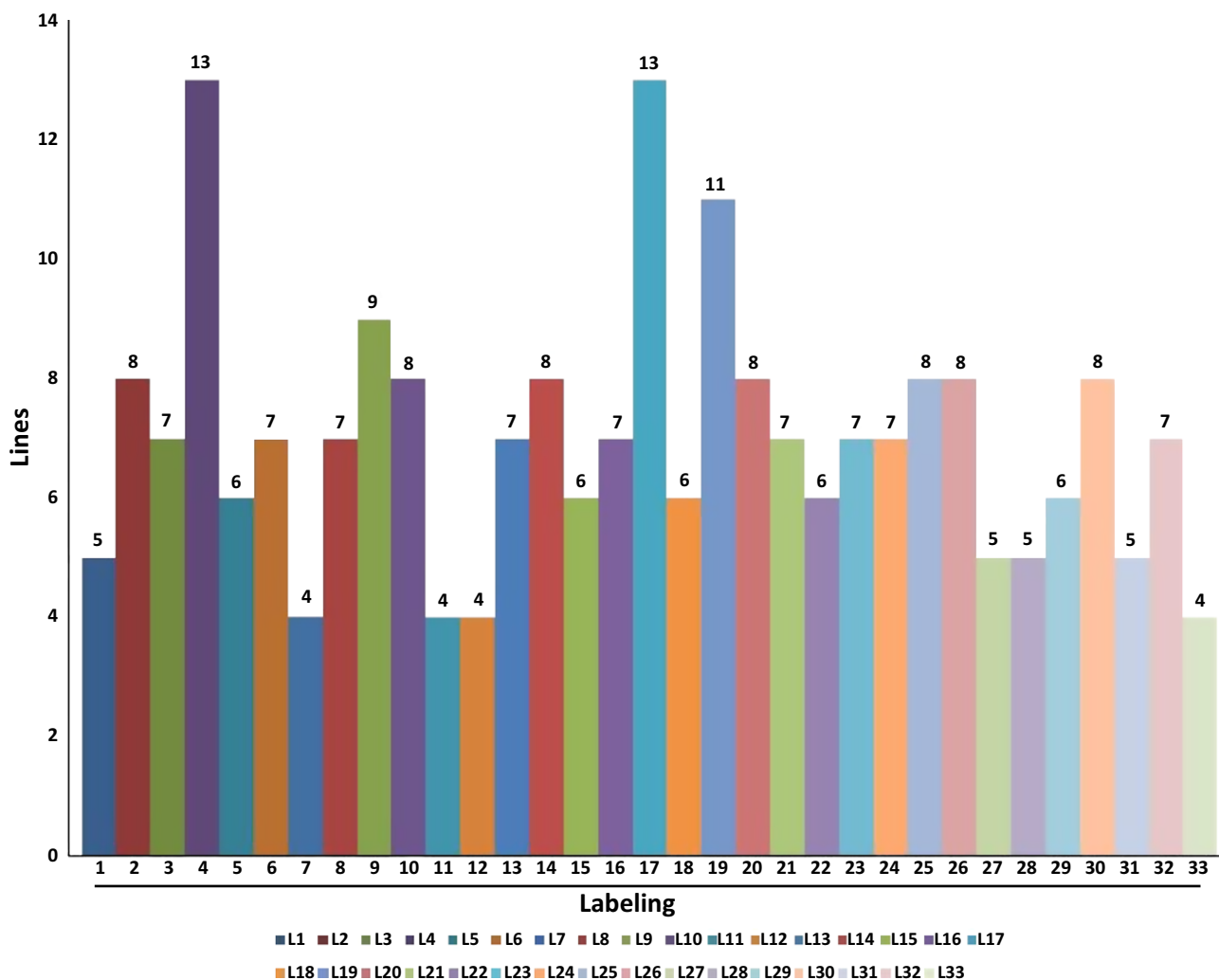


Fig. 2. The number of line segments for the corresponding label.

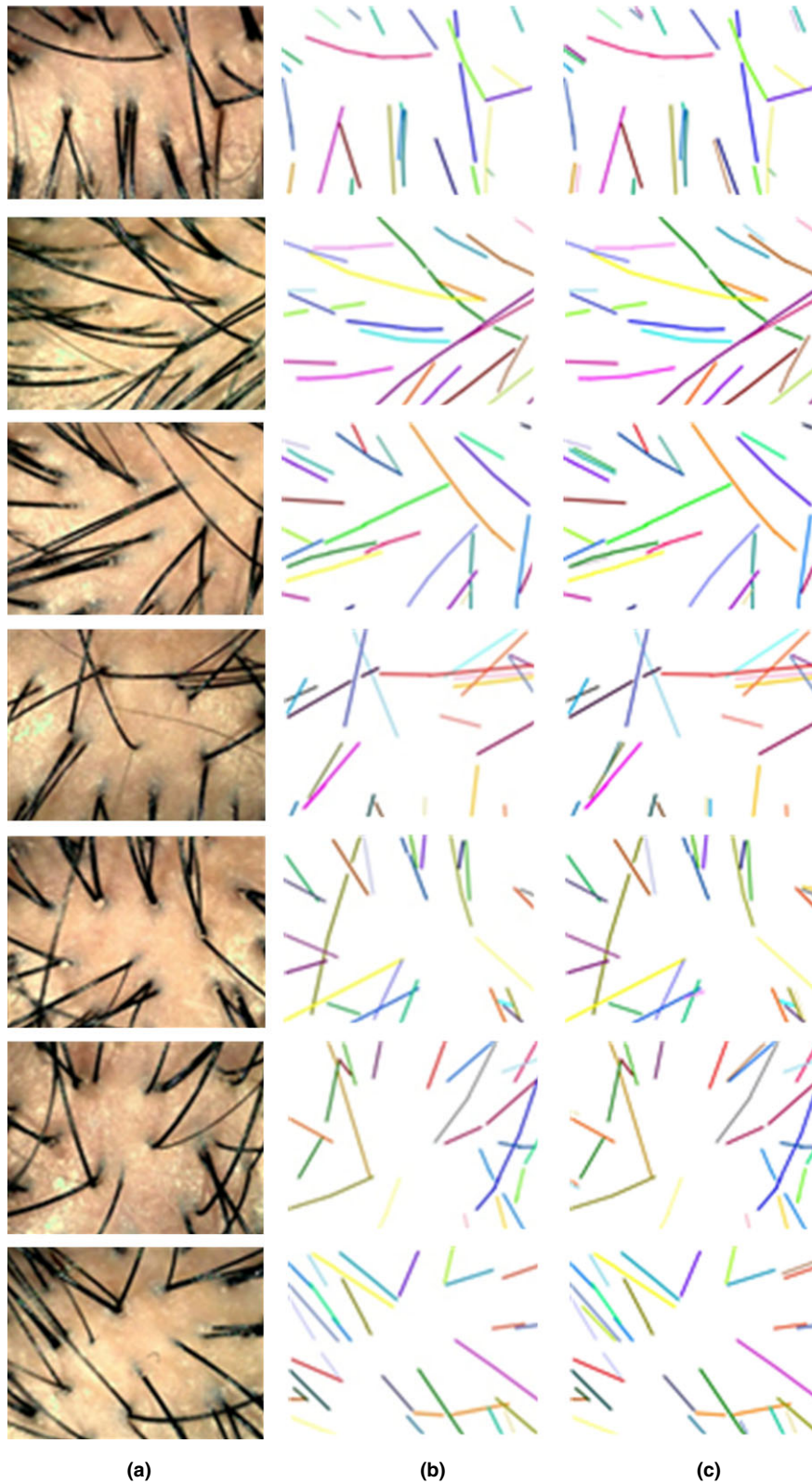


Fig. 3. Result images for line detection. (a) The original images, (b) result images of the BSR and multi-scale, (c) result images of the PLB algorithm.

modules, including BSR+MSLD, without both BSR and MSLD, BSR-only, and MSLD-only.

Table 2 compares system sensitivity based on module usage. We proposed the morphology-

based BSR and MSLD mechanisms for improving the performance of hair detection, obtaining rates of precision of 96.52%. Compared with the conventional HT line detection method this was

TABLE 2. Comparison of the system sensitivity for the module usages in the line detection

Module usages	Precision (%)	Recall (%)
With BSR & MSLD	96.52	84.53
Without BSR & MSLD	94.65	74.97
With BSR only	95.07	79.55
With MSLD only	94.58	77.68

TABLE 3. System performance for with/without PLB

With/without PLB	Precision (%)	Recall (%)
Only with BSR & MSLD	96.52	84.53
With PLB	96.89	88.86

an improvement of 2%. From a recall rate point of view it was also higher than other module combinations. On average, we achieved a 9.5% improvement in recall rate. Based on our observations after applying BSR the precision rate was boosted.

Refinement using the PLB algorithm

Table 3 shows the efficiency of the PLB algorithm. Our proposed PLB method enabled the system to extract the missing hairs, with a 0.3% improvement in the precision rate, and a 5% improvement in the recall rate. The PLB algorithm can serve as a system fine-tuning process, increasing the system performance to 96.89% on average. This is reasonable because hidden hairs are not a frequent occurrence. In addition, we can control the resolution and the angle of the hair images from the microscope to avoid this situation.

Discussion

The goal of this study was to develop an automated hair counting system for clinical application using the microscope image from the hairs. The proposed method reduces the frequent errors and variances encountered when hairs are manually counted by human assessors. This

References

1. Debeir O, Decaestecker C, Pasteels J, Salmon I, Kiss R, Van Ham P. Computer-assisted analysis of epifluorescence microscopy images of pigmented skin lesions. *Cytometry* 1999; 37: 255–266.

2. Schmid-Saugeon P, Guillod J, Thiran J. Towards a computeraided diagnosis system for pigmented skin lesions. *Comput Med Imag Graph* 2003; 27: 65–78.
3. Nguyen N, Lee T, Atkins M. Segmentation of light and dark hair in dermoscopic images: a hybrid

clinical intelligent system can diagnose the health condition of a person's hair and can be applied in therapy recommendations by doctors for their patients.

First, we focused on removing the bright spots resulting from oily and moist hairs because it interferes with detecting the thinning process prior to the line detection process. Second, the thinning process suffers from contacting and overlapping hairs being recognized as a single hair, resulting in two or more overlapping hairs being counted as one. In addition, hairs are not always straight, they may be bent or curled making the conventional Hough-based line detection algorithm inappropriate. Third, the threshold of the HT is sensitive to the minimum length of the line segment. In the present study we proposed a MSLD mechanism to deal with the variation in hair length. Finally, we viewed the hair counting problem as a labeling problem using the RLA. It integrated a group of line segments into a semantic 'hair' and assigned a unique label to it.

The framework of this paper can be seen as the first step toward intelligent computer-aided medical image processing for cosmetic treatment applications. In future works we will address the clustering of inflamed follicles and scalp pathological changes using pattern analysis techniques. This analytic system takes into account both texture and color features and applies machine learning algorithms. The development of a medical treatment system for follicle inflammation can provide suggestions to the doctors on the medical treatment for diverse scalp diseases.

Acknowledgments

Authors would like to thank Mr. Eric Chen and Mrs. Mimi Chien from Macrohi LTD., Taiwan, for providing the hair database and technical consultant, and Mr. Bo-Syun Lin to develop the program implementation initially and test the samples.

- approach using a universal kernel. *Proc SPIE Med Imag Conf* 2010; 7623: 76 234N-1–76 234N-8.
4. Celebi ME, Kingravi H, Iyatomi H, Aslandogan Y, Stoecker W, Moss R, Malters J, Grichnik J, Marghoob A, Rabinovitz H. Border detection in dermoscopy images using statistical

- region merging. *Skin Res Technol* 2008; 14: 347–353.
5. Celebi ME, Iyatomi H, Schaefer G, Stoecker W. Lesion border detection in dermoscopy images. *Comput Med Imag Graph* 2009; 33: 148–153.
 6. Donadey T, Serruys C, Giron A, Aitken G, Vignali J, Triller R, Fertil B. Boundary detection of black skin tumors using an adaptive radial-based approach. *Proc SPIE Med Imag Conf* 2000; 3979: 810–816.
 7. Tanaka H, Nakagami G, Sanada H, Sari Y, Kobayashi H, Kishi K, Konya C, Tadaka E. Quantitative evaluation of elderly skin based on digital image analysis. *Skin Res Technol* 2008; 14: 192–200.
 8. Zou Y, Song E, Jin R. Age-dependent changes in skin surface assessed by a novel two-dimensional image analysis. *Skin Res Technol* 2009; 15: 399–406.
 9. Lagarde JM, Rouvrais C, Black D. Topography and anisotropy of the skin surface with ageing. *Skin Res Technol* 2005; 11: 110–119.
 10. Boyer G, Laquière L, Le Bot A, Laquière S, Zahouani H. Dynamic indentation on human skin in vivo: ageing effects. *Skin Res Technol* 2009; 15: 55–67.
 11. Choi Y-H, Kim D, Hwang E, Kim BJ. Skin texture aging trend analysis using demoscropy images. *Skin Res Technol* 2014; 1–12.
 12. Huang A, Kwan S-Y, Chang W-Y, Liu M-Y, Chi M-H, Chen G-S. A robust hair segmentation and removal approach for clinical images for skin lesions. In *Proc. of the 35th Annual International Conference of the IEEE EMBS, Osaka, Japan, 3 - 7 July, 2013*, pp. 3315–3318.
 13. Lee T, Ng V, Gallagher R, Coldman A, McLean D. Dullrazor: a software approach to hair removal from images. *Comput Biol Med* 1997; 27: 533–543. Available at: http://www.dermweb.com/dull_razor/
 14. Tsai RY, Lee SH, Chan HL. The distribution of follicular units in the Chinese scalp: implications for reconstruction of natural-appearing hairlines in Orientals. *Dermatol Surg* 2002; 28: 500–503.
 15. Vallotton P, Thomas N. Automated body hair counting and length measurement. *Skin Res Technol* 2008; 14: 493–497.
 16. Amornsiripanitch S. Total Quality of hair parameters measuring method. US Patent, pub. No. 6,389,150 B1. May 14 2002.
 17. Leavitt M, Perez-Meza D, Rao NA, Barusco M, Kaufman KD, Ziering C. Effects of finasteride (1 mg) on hair transplant. *Dermatol Surg* 2005; 31: 1268–1276.
 18. Hoffmann R. TrichoScan. A new instrument for digital hair analysis. *Hautarzt* 2002; 53: 798–804.
 19. Rao KR, Yip PC. *The transform and data compression handbook*. Danvers, MA: CRC Press LLC, 2000.
 20. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Sys Man Cyber* 1979; 9: 62–66.
 21. Angulo J. Morphological colour operators in totally ordered lattices based on distances: application to image filtering, enhancement and analysis. *Comput Vis Image Underst* 2007; 107: 56–73.
 22. Shih HC, Liu E-R. Quartile-based region merging algorithm for morphological color image segmentation. *Proc. of Asian Conference on Computer Vision, Singapore, Nov. 2014*.
 23. Hough PVC. Machine analysis of bubble chamber pictures. *Proc. Int. Conf. High Energy Accelerators and Instrumentation, 1959*.
 24. Duda RO, Hart RE. Use of the Hough transform to detect lines and curves in pictures. *Comm ACM* 1972; 15: 11–15.
 25. Canny A. Computational approach to edge detection. *IEEE Trans Pattern Analysis Machine Intelligence* 1986; 8: 679–698.
 26. Rosenfeld A, Hummel R, Zucker S. Scene labeling by relaxation operations. *IEEE Transactions on Systems, Man and Cybernetics* 1976; 6: 420–433.

Address:
 Huang-Chia Shih
 Rm. 70630, Bldg. 7, 135, Yuandong Rd.
 Taoyuan 32003
 Taiwan
 Tel: +886-3-4638800 ext. 7122
 Fax: +886-3-4639355
 e-mail: hcshih@saturn.yzu.edu.tw