International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:3, No:12, 2009

Skin Lesion Segmentation Using Color Channel Optimization and Clustering-based Histogram Thresholding

Rahil Garnavi, Mohammad Aldeen, M. Emre Celebi, Alauddin Bhuiyan, Constantinos Dolianitis, and George Varigos

Abstract—Automatic segmentation of skin lesions is the first step towards the automated analysis of malignant melanoma. Although numerous segmentation methods have been developed, few studies have focused on determining the most effective color space for melanoma application. This paper proposes an automatic segmentation algorithm based on color space analysis and clustering-based histogram thresholding, a process which is able to determine the optimal color channel for detecting the borders in dermoscopy images. The algorithm is tested on a set of 30 high resolution dermoscopy images. A comprehensive evaluation of the results is provided, where borders manually drawn by four dermatologists, are compared to automated borders detected by the proposed algorithm, applying three previously used metrics of accuracy, sensitivity, and specificity and a new metric of similarity. By performing ROC analysis and ranking the metrics, it is demonstrated that the best results are obtained with the X and XoYoR color channels, resulting in an accuracy of approximately 97%. The proposed method is also compared with two state-of-the-art skin lesion segmentation methods.

Keywords—Border detection, Color space analysis, Dermoscopy, Histogram thresholding, Melanoma, Segmentation.

I. Introduction

A NUMBER of research schemes have been proposed in the last two decades to enhance the clinical diagnosis of melanoma, by which dermoscopy technique (a non-invasive in vivo clinical examination which allows for a magnified and clear visualization of the morphological structures of the skin that are not visible to the naked eye) and several clinical diagnostic algorithms such as pattern analysis, ABCD rule of dermoscopy, Menzies method, 7-points checklist and the CASH algorithm have been introduced [1], [2]. Numerous studies have shown that using these algorithms along with dermoscopy and digital monitoring improves the diagnosis of melanoma compared to simple naked-eye examination. However, even with the use of dermoscopy and dermoscopic algorithms, clinical diagnosis is still challenging and its accuracy is considered to be limited, especially with difficult cases [3]. Computer aided diagnosis of melanoma provides quantitative and objective evaluation of the skin lesion, as opposed to visual assessment, which is subjective in nature. It allows for reproducible diagnosis by diminishing the inter-observer and intra-observer variabilities that could be found in dermatologists' examinations. It also automates the analysis, and thereby reduces the amount of repetitive and tedious tasks to be done by physicians.

A system for the computer-aided diagnosis of melanoma is generally comprised of four major components: skin image acquisition, lesion segmentation, feature extraction, and lesion classification. Automatic segmentation of lesions in color skin images, which is the main focus of this paper, is one of the most important steps towards the automated analysis and evaluation of dermoscopy images in the computer aided diagnosis of melanoma. The accuracy of the segmentation process is of high importance due to the bias it can impose on the subsequent steps of the diagnosis system.

In this paper a novel automatic segmentation algorithm based on color space analysis and clustering-based histogram thresholding is proposed. It determines the most effective and discriminative color channel for melanoma application. Different color channels from various color spaces have been incorporated to maximize the discrimination between two clusters of pixels within the image, lesion and normal skin pixels, to obtain a more accurate histogram analysis. Each color channel undergoes preprocessing steps such as intensity adjustment and noise removal by the use of pill-box low-pass filter. Clustering-based thresholding is applied, and a set of pixel-based computations and morphological operators are utilized to eventually obtain the segmented lesion. Producing 25 different channel images from each initial RGB skin image, we have determined the color space and the corresponding channel that outperforms others in detecting the lesion borders.

The rest of the paper is organized as follows: Section 2 reviews the related work. Section 3 describes the proposed segmentation algorithm. Section 4 discusses the experimental results and section 5 presents the conclusion.

II. Related Work

Different image features such as shape, color, texture and brightness have been employed to perform skin lesion segmentation. For this purpose, numerous methods have been

R. Garnavi and M. Aldeen are with the Department of Electrical and Electronic Engineering, NICTA Victoria Research Laboratory, University of Melbourne, Australia. e-mail: r.garnavi@ee.unimelb.edu.au, aldeen@unimelb.edu.au

M. E. Celebi is with the Department of Computer Science, Louisiana State University, Shreveport, USA. e-mail: ecelebi@lsus.edu

A. Bhuiyan is with the Centre for Eye Research Australia, University of Melbourne, Australia. e-mail: abhuiyan@unimelb.edu.au

C. Dolianitis and G. Varigos are with the Department of Dermatology, Royal Melbourne Hospital. e-mail: constantinos1@optusnet.com.au, george.varigos@mh.org.au

proposed [4] including histogram thresholding [5], [6], color clustering [7]–[9], JSEG algorithm based on color quantization and spatial segmentation [10], statistical region merging [11], etc. Melli et al. [7] criticized the adaptive thresholding methods for not providing accurate segmentation result due to the problems associated with color calibration and lack of sharp bimodal luminance distribution between the surrounding skin and the lesion. Unsupervised approaches, despite their lower performance compared to supervised trained systems, have been claimed to exhibit more robustness. Moreover, although numerous methods have been developed for lesion segmentation, few studies have focused on determining the effective color space and color channel for melanoma application. In this study, however, we decide to use a threshold-based approach with the main goal of determining the most effective and discriminative color channel for melanoma application. The choice is motivated by simplicity and low computational cost. The improvement in the accuracy of the segmentation algorithm is obtained by using a clustering-based histogram thresholding, where color channels from various color spaces are used to detect more accurate borders.

III. Proposed Segmentation Algorithm

In order to separate the lesion from the surrounding normal skin, a segmentation algorithm, based on clustering-based histogram thresholding, color space analysis, pixel-level computations and morphological operations, has been developed. It consists of several steps, as shown in Figure 1, which are explained next.

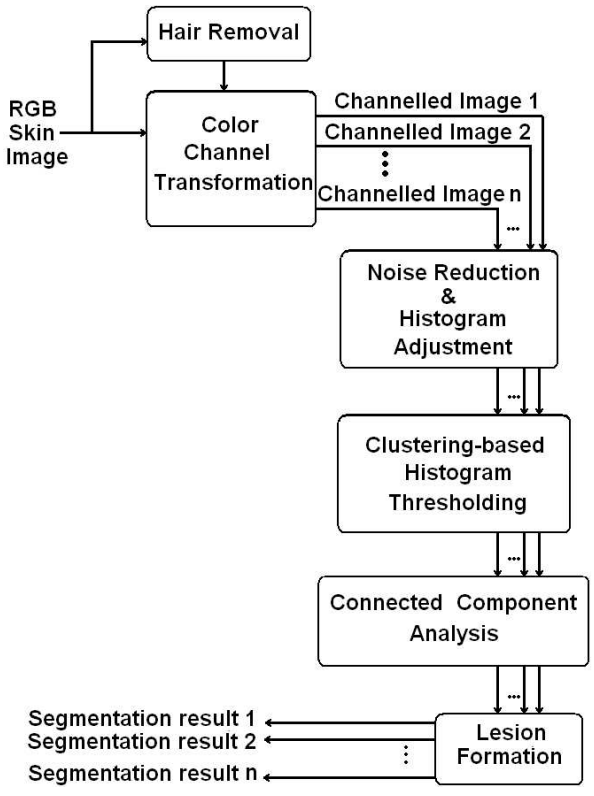

Fig. 1. Segmentation process diagram

A. Hair Removal

Lesions occluded with dark thick hairs can cause problems in the segmentation process. In such cases, the proposed algorithm starts with a hair removal preprocessing, which includes a sequence of steps [12]. These are: (1) localizing dark hairs, using morphological closing operation in vertical, horizontal and diagonal directions, (2) interpolating the removed hair pixels by close non-hair pixels, and (3) smoothing the final result using a median filter to eliminate the remaining thin lines. Figure 2 shows a dermoscopy image before and after hair removal.

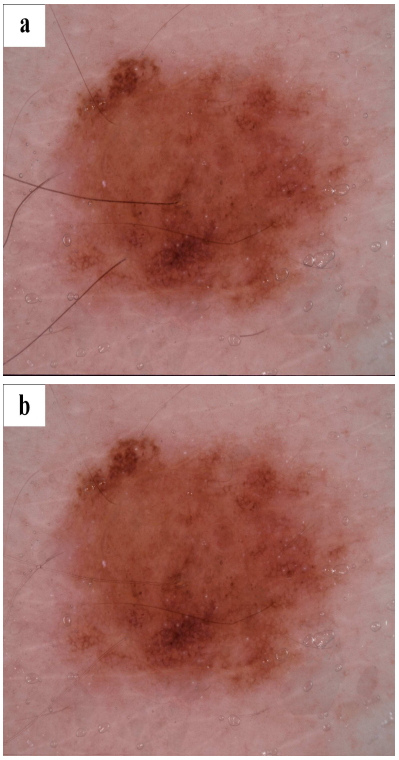

Fig. 2. Dermoscopy image (a) before and (b) after hair removal.

B. Color Space Transformation

Since color information plays a significant role in skin image processing, this step incorporates color information of the skin image into the segmentation process, where the original RGB image is transformed to different color spaces, and the corresponding color channels are extracted. Figure 3 shows a dermoscopy image before and after color space transformation. Although there exists several color spaces, each might outperform others in a particular application [13]. In this study the color spaces we have utilized include RGB, HSV, HSI, CIE-XYZ, CIE-LAB, and YCbCr. In this step of the algorithm, the initial RGB skin image is transformed to a set of 25 various color channel images from above-mentioned color spaces, as shown in Table I. Accordingly, a set of 25 channel images for each original RGB image is produced. These images include both single color channel images such as R from the RGB color space and X from the XYZ color

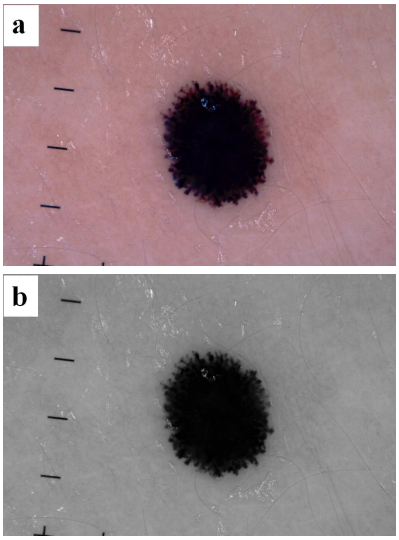

Fig. 3. Dermoscopy image: (a) original RGB image (b) channeled image.

pace, as well as combinations of them such as XoYoR which combines X and Y color channels from the XYZ color space with R color channel from the RGB color space.

TABLE I
COLOR CHANNELED USED IN COLOR SPACE TRANSFORMATION.

	Color Channel	Color Space
1	R	RGB
2	G	RGB
3	B	RGB
4	RGB	RGB
5	RoB	RGB
6	GoB	RGB
7	RoG	RGB
8	RoGoB	RGB
9	RGBoR	RGB
10	RGBoG	RGB
11	RGBoB	RGB
12	RGBoRoGoB	RGB
13	I	HSI
14	V	HSV
15	L	LAB
16	Y	YCbCr
17	X	XYZ
18	Y	XYZ
19	Z	XYZ
20	XoY	XYZ
21	XoZ	XYZ
22	YoZ	XYZ
23	XoYoZ	XYZ
24	XoYoR	XYZ and RGB
25	XoYoZoR	XYZ and RGB

1) Color Spaces and Color Channels: The most frequently used color presentation in image processing is the RGB color space where colors are represented by their red, green, and blue spectral wavelength responses in an orthogonal Cartesian space. Consequently, each pixel has a 3D vector with component values ranging from 0 to 255. To overcome the limitation of the RGB color space in high level processing, other color spaces have been developed based on mathematical transformation of the original RGB color channels. HSV and HSI (equation (1)) [14] color spaces was proposed to mimic the human visual perception of color in terms of *hue*, *saturation* and *intensity* (*value*). The hue component is proportional to the average wavelength of the color, saturation indicates the amount of white in the color and intensity represents the brightness or the amount of energy in the color.
HSI color space is expressed as follows:

$$I = \frac{R+G+B}{3} \tag{1}$$

$$S = 1 - \frac{3}{R+G+B}[\min(R,G,B)]$$

$$W = cos^{-1}\left(\frac{R - \frac{1}{2}(G+B)}{\sqrt{[(R-G)^2 + (R-B)(G-B)]}}\right)$$

where

$$H = \begin{cases} W & \text{if} \quad G > B \\ 2\pi - W & \text{if} \quad G < B \end{cases}$$

However, the color spaces outlined above do not provide perceptual uniformity; i.e. color vectors with a particular distance in color space are placed in the same distance when perceived by human visual system. To meet the uniformity requirement CIE-LAB color space were proposed. CIE-XYZ (equation (2) and (3)) [14] color space, which is one of the first mathematically defined color spaces created by the International Commission on Illumination (CIE) in 1931, is another color space investigated in this study. The transformation from RGB color space to XYZ color space can be achieved through the following relationship:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = T \times \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{2}$$

where T is a 3×3 constant matrix, the parameters of which may differ from one application to another according to the reference white value. In this paper we have choosen the following parameters, for the illuminant D65:

$$T = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2127 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9502 \end{bmatrix} \tag{3}$$

C. Noise Filtering

To elevate the accuracy of the segmentation result and save computation time, it is useful to eliminate the artifacts that might be present in the image. In dermoscopy images, external artifacts include skin lines, air bubbles or other random noise caused by the imaging process. To this end, the skin image is smoothed through a circular averaging low-pass filter with radius of 5, using the pill-box point spread function shown in Figure 4. The result of applying this filter on the dermoscopy image of Figure 3 is shown in Figure 5.

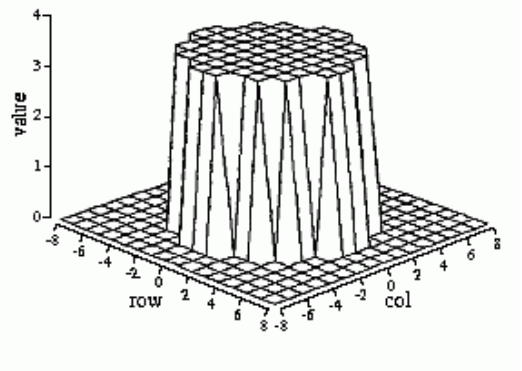

Fig. 4. Pill-box point spread function

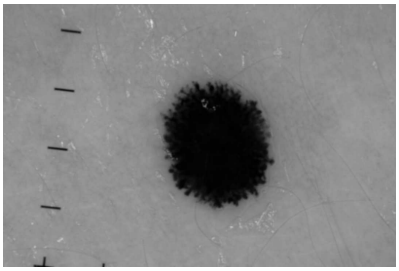

Fig. 5. Result of applying noise filtering step on the dermoscopy image of Figure 3.

D. Intensity Adjustment

This step is essentially an enhancement process in which the dynamic range of pixel values of the image is mapped into a new range. The purpose is to smooth and stretch the image histogram and increase the contrast of the image to be able to find a more precise threshold value in the succeeding step of clustering-based histogram thresholding. Intensity adjustment works by scaling the intensity values in the original image to cover the entire dynamic range of [0 1]. This linear scaling is done such that 1% of data is saturated at lowest (0) and highest (1) intensities. Figure 6 illustrates the result of applying this intensity adjustment step on the dermoscopy image of Figure 5.

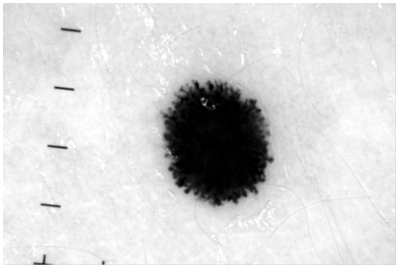

Fig. 6. Result of applying intensity adjustment step on the dermoscopy image of Figure 5.

E. Clustering-based Histogram Thresholding

Thresholding is the process of classifying the pixels of a grayscale image into two classes, so that the image can be converted to a binary image by assigning each pixel either a 0 or 1, depending on the gray level. The thresholding procedure used in this paper is based on the well-known Otsu's thresholding method [15]. The basic premise of this method is the assumption that an image, in this application a skin lesion image, contains two clusters of pixels e.g. foreground and background which correspond to lesion and surrounding normal skin, respectively. To identify these two clusters adequately, an algorithm is used to search for an optimal threshold level using discriminant analysis, where zero-th and first-order cumulative moments of the color histogram are calculated and used to define a measure of separability between the two clusters. An optimal threshold level separating the two clusters is achieved when the within-cluster variance (σ_ω^2) is minimal. The within-cluster variance defined as a weighted sum of variances of the two clusters:

$$\sigma_\omega^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t) \quad (4)$$

where weights ω_i are the probabilities of the two clusters separated by a threshold t and σ_i^2 are variances of these clusters. It can be shown [15] that minimizing the within-cluster variance (σ_ω^2) is equivalent to maximizing between-cluster variance (σ_b^2):

$$\sigma_b^2(t) = \sigma^2 - \sigma_\omega^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2 \quad (5)$$

where μ_i are the mean values of the two clusters. Starting from an initial threshold value of $t = 1$, ω_i and σ_i^2 are updated iteratively and in each iteration $\sigma_b^2(t)$ is calculated. The optimal threshold corresponds to the maximum value of $\sigma_b^2(t)$. The output binary image has values of 1 (white pixel) for all pixels in the input image with luminance greater than the threshold level and 0 (black pixel) for the remaining pixels. The result of applying the thresholding step on the dermoscopy image of Figure 6 is given in Figure 7.

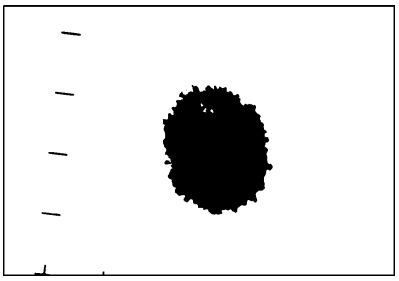

Fig. 7. Result of applying thresholding step on the dermoscopy image of Figure 6.

F. Connected Component Analysis

In some of the skin images extra objects appear in the surrounding skin area, such as blue marks made by dermatologists when examining the patient's skin. These objects, which have not been eliminated in the noise removal step, appear with intensity values similar to that of the lesion and may be misclassified as lesion. The purpose of this step is to exclude these objects from the segmentation output. To this end, the number of connected objects within the image

is counted using the run-length encoding technique [16] at the same time labeling the connected objects. Finally, the two largest areas (i.e. lesion and normal skin) are kept and all other components are discarded. Figure 8 shows the result of applying this step on the dermoscopy image of Figure 7. Finally, in order to obtain the lesion object the holes inside the boundary are filled using the morphological filling operation on the binary image. Figure 9 shows the final segmentation result and Figure 10 shows the whole segmentation sequence on another sample dermoscopy image. Figure 11 shows more samples of the detected border by the proposed algorithm. As it is shown in Figure 11, the image set contains a variety range of dermoscopy images in terms of color, texture and shape, which makes the algorithm generic and applicable to real world applications.

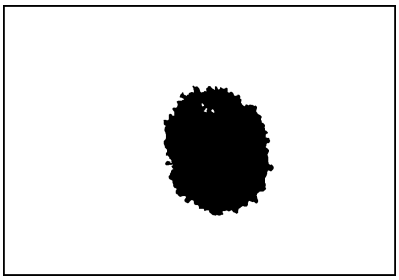

Fig. 8. Result of connected component analysis on the dermoscopy image of Figure 7.

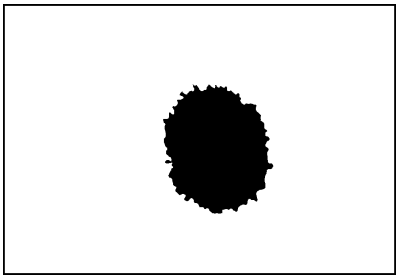

Fig. 9. Applying morphological operator on the dermoscopy image of Figure 8 to form the final segmentation result.

IV. Experimental Results

The proposed algorithm in this paper is tested on a set of 30 high resolution dermoscopy images. The images are 24-bit RGB images with 8 bits per color channel. As the ground truth for the evaluation of the detected border, four manual borders for each lesion were independently drawn by four experts; two experienced dermatologists and two dermatologist registrars. The manual borders are drawn using an A4 size graphical Tablet PC to provide an accurate result under controlled conditions. Figure 12 illustrates four different manual borders drawn by the dermatologists for a sample dermoscopy image.

A. Evaluation Metrics

To quantitatively compare the borders drawn by dermatologists with the computer derived borders, different metrics have been utilized [17]. In this paper, three statistical measurements of sensitivity (6), specificity (7) and accuracy (8) is applied. Furthermore, we have applied a metric called similarity which has not been used before in dermoscopy applications.

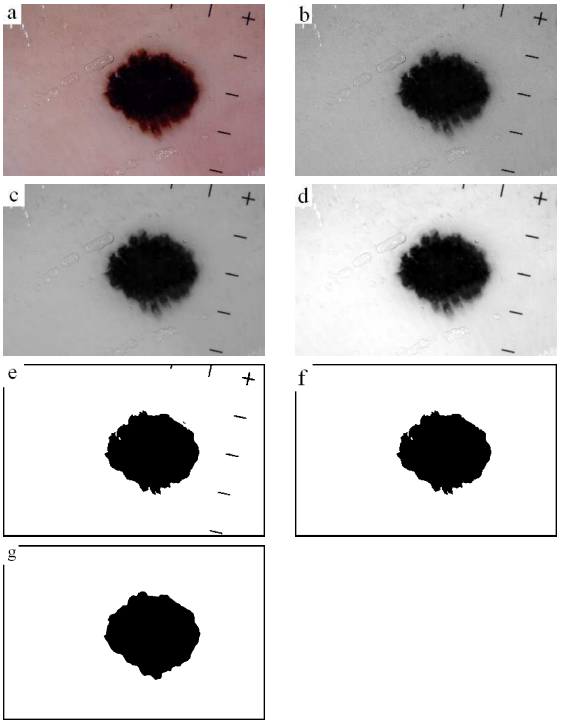

Fig. 10. Segmentation algorithm: (a) Original image, (b) Color space transformation, (c) Noise removal, (d) Intensity adjustment, (e) Thresholding, (f) Connected component analysis, (g) Morphological filling.

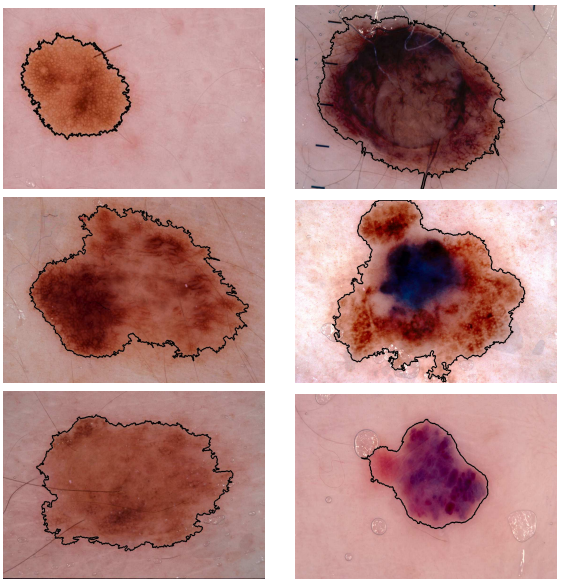

Fig. 11. Sample segmentation results.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{6}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{7}$$

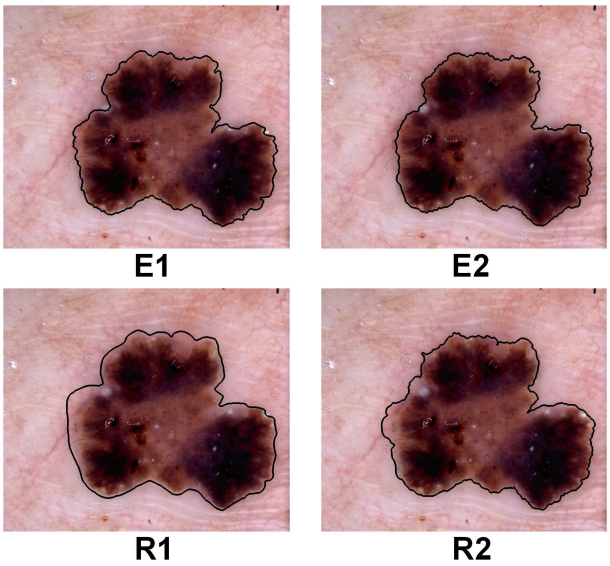

Fig. 12. Manual borders of the same lesion drawn by four dermatologists; E1 and E2: experienced dermatologists, R1 and R2: dermatologist registrars.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{8}$$

where TP, TN, FP, and FN refer to true positive, true negative, false positive and false negative, respectively. Manual segmentation by the dermatologists is used as the *ground truth* when calculating the above measures and for the remainder of this paper we will refer to any result obtained by the use of our algorithm as *automatic*. TP shows the number of pixels which are classified as lesion in both manual and automatic segmented images. TN represents the number of pixels which are classified as surrounding normal skin in both manual and automatic borders. FP indicates the number of pixels which are classified as lesion in automatic segmentation but are labelled as normal skin in manual segmentation. Finally, FN shows the number of pixels which are classified as normal skin in the automatic border but are labelled as lesion in the ground truth image, i.e. manual segmentation.

As previously mentioned, in the above metrics, the manual borders drawn by dermatologists are taken as the ground truth and the calculated measures are used to quantify the discrepancy between the automatic borders and the manual ones. Another perspective in the analysis is to evaluate the inter-observer variabilities among the four dermatologists, as demonstrated in Figure 12, and also to investigate which automatic border is closer to each manual one. To this aim, another metric called the *similarity* is applied to quantify the degree of similarity or agreement between any two borders, without taking any of them as the ground truth. The Sorensen similarity index [18], is a measure used for comparing the similarity of two objects or samples. It was developed by the botanist Thorvald Sorensen and published in 1948. This metric (9) has been widely used in different domains to measure the similarity, yet it has not been used before in dermoscopy applications.

$$\text{Similarity} = \frac{2 \times A_{1 \cap 2}}{A_1 + A_2} \tag{9}$$

where A_i is the areas of any two segmented image and $A_{1 \cap 2}$ is the number of pixels the two images share (intersection of the two images). This measure can also be expressed in terms of TP, TN, FP, and FN, given by:

$$\text{Similarity} = \frac{2 \times TP}{2 \times TP + FN + FP} \tag{10}$$

Table II shows the mean and standard deviation of similarity among the manual borders drawn by the four different dermatologists for the image set of 30 dermoscopy skin lesions. The similarity values indicate that there are higher similarities between the two experienced dermatologists and two dermatologist registrars. The overall similarity is high which indicates the reliability of the ground truth.

TABLE II
MEAN ± STANDARD DEVIATION FOR SIMILARITY (%) BETWEEN DERMATOLOGISTS; E1 AND E2: EXPERIENCED DERMATOLOGISTS, R1 AND R2: DERMATOLOGY REGISTRARS.

	E1	E2	R1	R2
E1	-	95.92 ±1.76	95.09 ±1.75	94.78 ±2.25
E2	-	-	94.68 ±1.73	94.78 ±2.22
R1	-	-	-	95.86 ±1.25

For the image illustrated in Figure 12, the similarity is shown in Table III, which again shows very high percentage of similarity or agreement.

TABLE III
SIMILARITY (%) BETWEEN DERMATOLOGISTS FOR THE DERMOSCOPY IMAGE OF FIGURE 12; E1 AND E2: EXPERIENCED DERMATOLOGISTS, R1 AND R2: DERMATOLOGIST REGISTRARS.

	E1	E2	R1	R2
E1	-	94.86	95.96	96.89
E2	-	-	92.63	93.03
R1	-	-	-	96.83

B. Analysis of the Evaluation Metrics

The main purpose of this analysis is to find those discriminative color channel(s) that lead to the most accurate border in the skin lesion segmentation algorithm. To achieve this aim, 25 different color channels listed in Table I are extracted from the 30 images, resulting in 25 different segmentation results per image. Each resultant border (out of 25×30), is then separately compared to each of the four manually drawn borders (ground truth) in terms of the four metrics of sensitivity, specificity, accuracy and similarity. The 30 values of each metric are averaged, resulting in 25 average measures of sensitivity, specificity, accuracy and similarity. The maximum values of each of these metrics and their corresponding color channels are then identified, resulting in four color channels for each metric, as shown in Table IV. As shown in Table IV, the initial 25 color channels are narrowed

down to the four color channels of X, XoYoR, XoYoZoR and R.

TABLE IV
COLOR CHANNELS WITH THE LARGEST MEAN FOR FOUR GROUND TRUTHS; E1 AND E2: EXPERIENCED DERMATOLOGISTS, R1 AND R2: DERMATOLOGY REGISTRARS.

	Accuracy	Similarity	Sensitivity	Specificity
E1	XoYoR (96.03)	X (91.45)	XoYoZoR (93.52)	R (99.96)
E2	XoYoR (96.01)	X (91.55)	XoYoZoR (94.33)	R (99.92)
R1	XoYoR (95.14)	XoYoZoR (90.61)	XoYoZoR (90.60)	R (99.99)
R2	XoYoR (94.89)	XoYoZoR (90.56)	XoYoZoR (90.37)	R (99.99)

The results presented in Table IV have been calculated using each dermatologist's manually drawn border as the ground truth, hence four sets of results. Another useful analysis is performed when a single ground truth for each image is used. This single ground truth is obtained by finding the intersection of the four manually drawn borders (in the ensuing, we will refer to these as *common manual borders*). Then the common manual border is used as the ground truth and applied the four color channels shown in Table IV on the each of the 30 images. The average values of the four metrics is shown in Table V. Table V suggests that in terms of accuracy metric, XoYoR color channel provides the highest score and X color channel follows it closely. With respect to similarity, X color channel gains the best and XoYoR achieves the second best result.

A more informative analysis of the results of Table V can be done by calculating the AUC (Area Under ROC Curve) value, obtained from the two measures of sensitivity and specificity. This has been achieved by drawing the ROC graph and calculating the corresponding AUC value for R, X, XoYoZ and XoYoRoZ color channels. The AUC values are calculated [19] according to the sensitivity and specificity values for all images in the image set, taking the common manual border as the ground truth; i.e. $AUC_R = 0.991$, $AUC_X = 0.998$, $AUC_{XoYoR} = 0.998$, $AUC_{XoYoZoR} = 0.995$. Figure 13 shows the ROC curves for the four color channel X, R, XoYoY and XoYoZoR. Table VI shows the performance ranking of the color channels according to different evaluation metrics, which shows that color channels X and XoYoR provide the best overall results. Figures 14 and 15 show the accuracy and similarity percentage of these two color channels over the image set of 30 dermoscopy images, taking the common manual border as the ground truth.

It is well known that XYZ is a standard color space defined by CIE. However, in the domain of skin lesion segmentation, we have shown the appropriateness of XYZ color space and more specifically the X and Y color channels when combined with the R channel.

C. Comparison with other automated methods

The proposed method is compared with two state-of-the-art skin lesion segmentation methods, namely; statistical region merging [11] and JSEG method [10]. Table VII shows the

TABLE V
SEGMENTATION RESULTS (MEAN ±MARGIN OF ERROR) FOR OPTIMAL COLOR CHANNELS, THE COMMON MANUAL BORDERS AS GROUND TRUTH.

	Accuracy	Sensitivity	Specificity	Similarity
X	96.80 ±0.01	90.49 ±0.03	99.02 ±0.00	93.18 ±1.82
XoYoR	96.94 ±0.01	92.86 ±0.02	98.51 ±0.00	92.89 ±3.15
XoYoZoR	95.15 ±0.03	96.80 ±0.01	94.47 ±0.05	91.63 ±5.15
R	92.42 ±0.03	73.40 ±0.07	99.90 ±0.00	82.91 ±5.45

TABLE VI
PERFORMANCE RANKING OF THE COLOR CHANNELS.

	Accuracy	AUC	Similarity
1	XoYoR	X, XoYoR	X
2	X	XoYoZoR	XoYoR
3	XoYoZoR	R	XoYoZoR
4	R		R

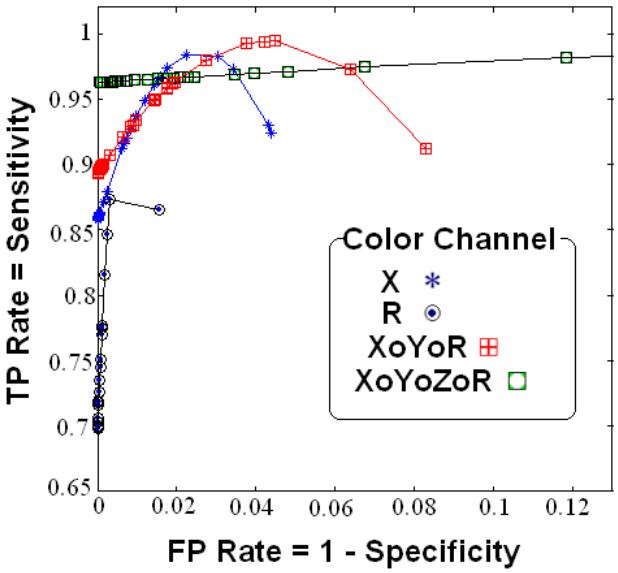

Fig. 13. ROC curves for four color channels X, R, XoYoR and XoYoZoR.

segmentation evaluation result obtained by SRM (statistical region merging) and JSEG, and the proposed thresholding method when applied to the X and XoYoR color channels. The results demonstrate that the proposed thresholding-based method, in spite of its simplicity, with a proper choice of suggested color channels is highly competitive with other well-known skin lesion segmentation methods. X channel outperforms other automated methods with respect to specificity and along with JSEG method, it obtains the best performance in terms of similarity, while XoYoR follows them closely. The proposed method in both color channels gains the highest AUC value, while JSEG follows them closely, and SRM comes last. XoYoR color channel gain the highest sensitivity and accuracy on average, followed closely by others methods. In addition, our method is potentially faster since it mainly involves scalar processing as opposed to the SRM and JSEG methods, which are based on vector processing.

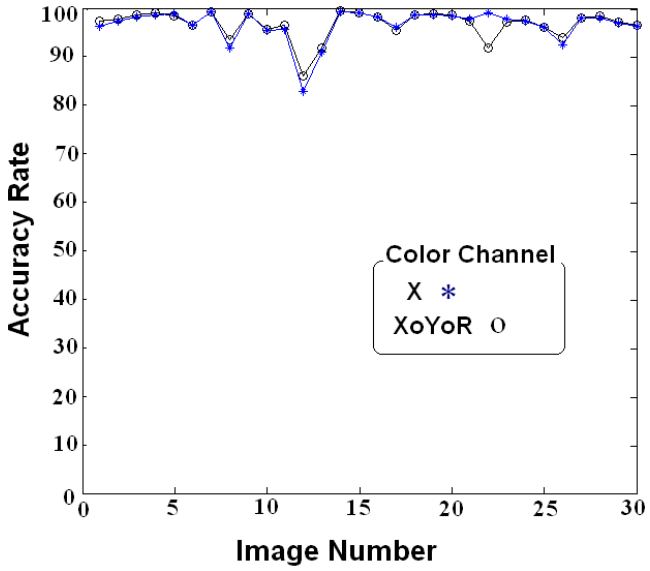

Fig. 14. Accuracy of the optimal color channels, X and XoYoR, over the image set.

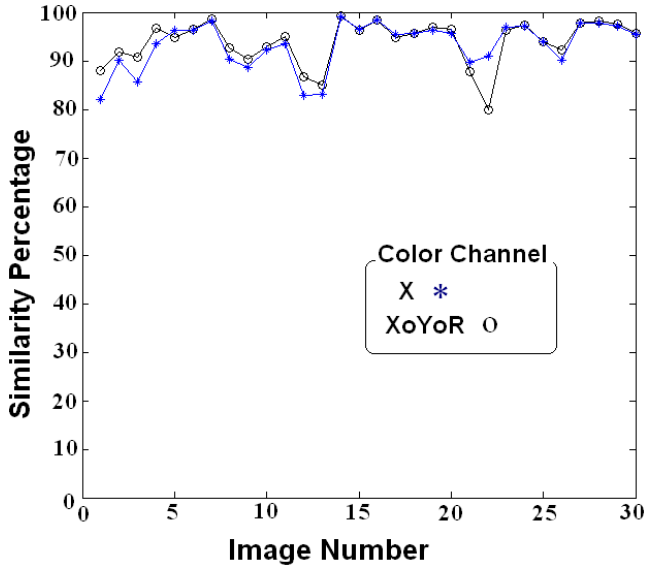

Fig. 15. Similarity of the optimal color channels, X and XoYoR, over the image set.

TABLE VII
SEGMENTATION EVALUATION RESULTS (MEAN ±MARGIN OF ERROR): COMPARATIVE STUDY.

	Accuracy	Similarity	Sensitivity	Specificity	AUC
SRM	96.27 ±0.01	91.55 ±2.32	92.45 ±0.03	97.43 ±0.01	0.990
JSEG	96.70 ±0.01	93.23 ±3.14	92.05 ±0.04	98.35 ±0.01	0.997
X	96.80 ±0.01	93.18 ±1.82	90.49 ±0.03	99.02 ±0.00	0.998
XoYoR	96.94 ±0.01	92.89 ±3.15	92.86 ±0.02	98.51 ±0.00	0.998

V. CONCLUSION

This paper presents a fast accurate automatic segmentation algorithm based on color space analysis and clustering-based histogram thresholding which determines the most effective and discriminative color channel for detecting the borders in dermoscopy images. Various color channels from different color spaces are incorporated to maximize the discrimination between two clusters of pixels within the image, lesion and surrounding skin pixels, to obtain a more accurate histogram analysis, and consequently to detect a more accurate border. Each color channel undergoes preprocessing steps, clustering-based histogram thresholding and a set of pixel-based computations and morphological operators to eventually identify the border of the lesion.

Segmentation results are quantitatively evaluated by comparing automated results to manual borders independently drawn by four dermatologists. The comparison is done with respect to four different metrics of accuracy, sensitivity, specificity and similarity, incorporating ROC analysis. Experimental results indicate that X and XoYoR color channels obtain the highest overall performance with an accuracy of approximately 97%. The results are also compared with two state-of-the-art automated methods, which demonstrate that the proposed thresholding-based method, in spite of its simplicity, with a proper choice of suggested color channels is highly competitive with the well-known skin lesion segmentation methods, and outperforms them with respect to accuracy, specificity, and AUC metrics. Furthermore, the proposed method is potentially faster since it mainly involves scalar processing as opposed to vector processing performed in those methods.

For the future, we plan to continue this research to improve the accuracy and also robustness of the algorithm against image noise and variabilities in imaging parameters so as to make the method more applicable for real world problem. Moreover, we plan to perform the experiments on a large dermoscopy image set and also to investigate more evaluation metrics.

ACKNOWLEDGMENT

This research is supported by NICTA Victoria Research Laboratory, Australia.

REFERENCES

[1] G. Argenziano, H. P. Soyer, S. Chimenti, and R. T. et al., "Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet," *Journal of the American Academy of Dermatology*, vol. 48, pp. 679–693, 2003.
[2] P. Braun, H. Rabinovitz, M. Oliviero, A. Kopf, and J. Saurat, "Dermoscopy of pigmented lesions," *Journal of the American Academy of Dermatology*, vol. 52, no. 1, pp. 109–121, 2005.
[3] A. Perrinaud, O. Gaide, L. French, J.-H. Saurat, A. Marghoob, and R. Braun, "Can automated dermoscopy image analysis instruments provide added benefit for the dermatologist? A study comparing the results of three systems," *British Journal of Dermatology*, vol. 157, pp. 926–933, 2007.
[4] M. E. Celebi, H. Iyatomi, G. Schaefer, and W. V. Stoecker, "Lesion border detection in dermoscopy images," *Computerized Medical Imaging and Graphics*, vol. 33, no. 2, pp. 148–153, 2009.
[5] H. Iyatomi, H. Oka, M. E. Celebi, M. Hashimoto, M. Hagiwara, M. Tanaka, and K. Ogawa, "An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm," *Computerized Medical Imaging and Graphics*, vol. 32, no. 7, pp. 566–579, 2008.

[6] M. Hintz-Madsen, L. K. Hansen, J. Larsen, and K. T. Drzewiecki, "A probabilistic neural network framework for the detection of malignant melanoma." *Artificial neural networks in cancer diagnosis. Prognosis and Patient Management*, pp. 141–183, 2001.
[7] R. Melli, C. Grana, and R. Cucchiara, "Comparison of color clustering algorithms for segmentation of dermatological images," in *SPIE Medical Imaging*, vol. 6144, 2006, pp. 3S1–3S9.
[8] G. Hance, S. Umbaugh, R. Moss, and W. V. Stoecker, "Unsupervised color image segmentation: With application to skin tumor borders," *IEEE Engineering in Medicine and Biology Magazine*, vol. 15, pp. 104–111, 1996.
[9] P. Schmid, "Segmentation of digitized dermatoscopic images by two-dimensional color clustering," *IEEE Transactions on Medical Imaging*, vol. 18, no. 2, pp. 164–171, 1999.
[10] M. E. Celebi, Y. A. Aslandogan, W. V. Stoecker, H. Iyatomi, H. Oka, and X. Chen, "Unsupervised border detection in dermoscopy images," *Skin Research and Technology*, vol. 13, pp. 454–462, 2007.
[11] M. E. Celebi, H. A. Kingravi, H. Iyatomi, Y. A. Aslandogan, W. V. Stoecker, R. H. Moss, J. M. Malters, J. M. Grichnik, A. A. Marghoob, H. S. Rabinovitz, and S. W. Menzies, "Border detection in dermoscopy images using statistical region merging," *Skin Research and Technology*, vol. 14, pp. 347–353, 2008.
[12] T. Lee, , V. Ng, R. Gallagher, A. Coldman, and D. McLean, "Dullrazor: A software approach to hair removal from images," *Computers in Biology and Medicine*, vol. 27, pp. 533–543, 1997.
[13] L. Lucchese and S. Mitra, "Color image segmentation: A state-of-the-art survey," in *Proceedings of Indian National Science Academy Part A, PINSA2001*, 2001, pp. 207–221.
[14] K. N. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*. Springer, 2000.
[15] 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.
[16] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*. Addison-Wesley, 1992, vol. 1.
[17] M. E. Celebi, G. Schaefer, H. Iyatomi, W. V. Stoecker, J. M. Malters, and J. M. Grichnik, "An improved objective evaluation measure for border detection in dermoscopy images," *to appear in Skin Research and Technology*.
[18] T. Sorensen, "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons." *Royal Danish Academy of Sciences and Letters*, vol. 5, pp. 1–34, 1948.
[19] J. Davis and M. Goadrich, "The relationship between precision-recall and roc curves," in *Proceeding of 23rd International Conference on Machine Learning (ICML)*, vol. 148, 2006, pp. 233–240.