Published in final edited form as:
Skin Res Technol. 2008 August ; 14(3): 347–353. doi:10.1111/j.1600-0846.2008.00301.x.

Border detection in dermoscopy images using statistical region merging

M. Emre Celebi[1], **Hassan A. Kingravi**[2], **Hitoshi Iyatomi**[3], **Y. Alp Aslandogan**[4], **William V. Stoecker**[5], **Randy H. Moss**[6], **Joseph M. Malters**[7], **James M. Grichnik**[8], **Ashfaq A. Marghoob**[9], **Harold S. Rabinovitz**[10], and **Scott W. Menzies**[11]

[1]Department of Computer Science, Louisiana State University in Shreveport, Shreveport, LA, USA

[2]Department of Computer Science, Texas A&M University, College Station, TX, USA

[3]Department of Electrical Informatics, Hosei University, Tokyo, Japan

[4]Department of Computer Science, Prairie View A&M University, Prairie View, TX, USA

[5]Stoecker & Associates, Rolla, MO, USA

[6]Department of Electrical and Computer Engineering, University of Missouri-Rolla, Rolla, MO, USA

[7]The Dermatology Center, Rolla, MO, USA

[8]Department of Medicine, Duke University Medical Center, Durham, NC, USA

[9]Memorial Sloan-Kettering Skin Cancer Center, Hauppauge, NY, USA

[10]Skin and Cancer Associates, Plantation, FL, USA

[11]Sydney Melanoma Diagnostic Centre, Sydney, Australia

Abstract

Background—As a result of advances in skin imaging technology and the development of suitable image processing techniques, during the last decade, there has been a significant increase of interest in the computer-aided diagnosis of melanoma. Automated border detection is one of the most important steps in this procedure, because the accuracy of the subsequent steps crucially depends on it.

Methods—In this article, we present a fast and unsupervised approach to border detection in dermoscopy images of pigmented skin lesions based on the statistical region merging algorithm.

Results—The method is tested on a set of 90 dermoscopy images. The border detection error is quantified by a metric in which three sets of dermatologist-determined borders are used as the ground-truth. The proposed method is compared with four state-of-the-art automated methods (orientation-sensitive fuzzy c-means, dermatologist-like tumor extraction algorithm, meanshift clustering, and the modified JSEG method).

Conclusion—The results demonstrate that the method presented here achieves both fast and accurate border detection in dermoscopy images.

Address: M. Emre Celebi, Department of Computer Science, Louisiana State University in Shreveport, Technology Center 206, One University Place, Shreveport, LA 71115, USA, Tel: +1 318 797 5324, ecelebi@lsus.edu.

Keywords

computer-aided diagnosis; skin cancer; melanoma; dermoscopy; segmentation; border detection; statistical region merging

Malignant melanoma has consistently had one of the most rapidly increasing incidence of all cancers, with 59,940 new cases and 8110 deaths estimated in the United States in 2007 (1). Early diagnosis is particularly important because melanoma can be cured with a simple excision if detected early.

Dermoscopy is a non-invasive skin imaging technique that uses optical magnification and either liquid immersion and low angle-of-incidence lighting or cross-polarized lighting to make the contact area translucent, making subsurface structures more easily visible when compared with conventional macroscopic (clinical) images. Dermoscopy allows the identification of dozens of morphological features such as pigment networks, dots/globules, streaks, blue-white areas, and blotches (2). This reduces screening errors, and provides greater differentiation between difficult lesions such as pigmented Spitz nevi and small, clinically equivocal lesions (3). However, it has been demonstrated that dermoscopy may actually lower the diagnostic accuracy in the hands of inexperienced dermatologists (4). Therefore, due to the lack of reproducibility and subjectivity of human interpretation, the development of computerized image analysis techniques is of paramount importance (5).

The first step in the computerized analysis of skin lesion images is the detection of the lesion borders. The importance of border detection for the analysis is two-fold. First, the border structure provides important information for accurate diagnosis. Many clinical features such as asymmetry, border irregularity, and abrupt border cutoff are calculated from the border. Second, the extraction of other important clinical features such as atypical pigment networks, globules, and blue-white areas critically depends on the accuracy of border detection. Automated border detection in dermoscopy images is a challenging task due to several reasons: (i) low contrast between the lesion and the surrounding skin; (ii) irregular and fuzzy lesion borders; (iii) artifacts such as skin lines, air bubbles and hairs; and (iv) variegated coloring inside the lesion.

Numerous methods have been developed for border detection in pigmented skin lesion images earlier; most of these dealt with clinical images (6). However, recent research has focused more on dermoscopy images. Gao et al. (7) proposed two methods: one based on stabilized inverse diffusion equations, a form of non-linear diffusion and another one based on Markov random fields in which the model parameters are estimated using the mean field theory. Pagadala (8) described a method based on optimized histogram thresholding. Schmid (6) developed a technique based on color clustering. First, a 2D histogram is calculated from the first two principal components of the *CIE L*u*v** color space. The histogram is then smoothed and initial cluster centers are determined from the peaks using a perceptron classifier. Finally, the lesion image is segmented using a modified version of the fuzzy c-means (FCM) clustering algorithm. Donadey et al. (9) presented a supervised method based on intensity radial profiles calculated from the *I* (intensity) component of the *HSI* space. Cucchiara et al. (10) presented a recursive FCM clustering technique that augments Schmid's method using topological information. Erkol et al. (11) proposed a method based on the gradient vector flow (GVF) snakes with an automatic initialization. Iyatomi et al. (12) described a method called the dermatologist-like tumor extraction algorithm (DTEA) that is based on thresholding followed by iterative region growing. Melli et al. (13) compared four different color clustering algorithms: median cut, k-means, FCM, and meanshift. They concluded that the meanshift algorithm gave the best results. Celebi et al. (14) developed a

method based on the JSEG segmentation algorithm. Their method involves an algorithm for approximate lesion localization that reduces the computational time and improves the accuracy by focusing the border detection process on the immediate neighborhood of the lesion rather than the whole image.

In this paper, we present an unsupervised approach to border detection in dermoscopy images based on the statistical region merging (SRM) algorithm (15). The SRM algorithm is adapted to this problem due to its simplicity, computational efficiency, and excellent performance in a variety of image domains.

Materials and Methods

Overview of the SRM algorithm

SRM is a recent color image segmentation technique based on region growing and merging. The method models segmentation as an inference problem, in which the image is treated as an observed instance of an unknown theoretical image, whose statistical (true) regions are to be reconstructed. The advantages of this method include its simplicity, computational efficiency, and excellent performance without the use of quantization or color space transformations.

Let I be an observed image that contains $|I|$ pixels, each of which is comprised of R, G, B color channel values belonging to the set $\{0, 1, \ldots, g - 1\}$ (where $g = 256$ for 24-bit RGB images with 8 bits per color channel). I is an observation of a true image I^* in which pixels are perfectly represented by a family of distributions from which each of the observed color channel is sampled. The optimal statistical regions in I^* share a *homogeneity property* such that inside any statistical region and given any color channel, the statistical pixels have the same expectation, whereas the expectations of adjacent statistical regions differ in at least one color channel.

I is obtained from I^* by sampling each statistical pixel for observed RGB values. The color channel values for every pixel in I^* is replaced by a set of Q independent random variables, which take on values from $(0, g/Q)$. It is to be noted that the Q parameter can be used to quantify the statistical complexity of I^*, the generality of the model, and the statistical difficulty of the problem. Higher values of Q result in finer segmentation and thus the generation of more regions.

Like other region growing algorithms, SRM is based on two major components; a merging predicate and the order followed in testing this predicate. The predicate is defined as (15):

$$P(R,R')=\begin{cases}\text{true} & \text{if } \forall_a \in \{\mathbf{R},\mathbf{G},\mathbf{B}\} |\bar{R}'_a - \bar{R}_a| \le |\sqrt{b^2(R)+b^2(R')}| \\ \text{false} & \text{otherwise}\end{cases}$$
$$b(R)=g\sqrt{\frac{1}{2Q|R|}\ln(6|I|^2 R_{|R|})} \quad (1)$$

where R and R' represent the two regions being tested, $\bar{R}_a$ denotes the observed average for color channel a in region R and $R_{|p|}$ is the set of regions with p pixels. The order of region merging follows an invariant A, which implies that when any test between two parts of true regions occurs, all tests inside each of the two regions have previously occurred.

Let S_I be a set that contains all pairs of adjacent pixels in the image based on 4-connectivity, p and p' be pixels in image I, and $R(p)$ stand for the current region to which a pixel p belongs. The SRM algorithm first sorts these pairs in increasing order according to a function $f(p, p')$. After the sorting is completed, the order is traversed only once, performing

the merging test $P(R(p), R(p'))$ for any pair of pixels (p, p') for which $R(p) \neq R(p')$, and merging $R(p)$ and $R(p')$ if it returns true. The complexity of the ordering phase is $O(|I| \log(g))$ when radix sort is used with color differences as the keys. Similarly, the merging phase can be performed in linear time using an efficient union-find algorithm.

Because the model of image generation presented above assumes that the observed color variations within true regions should be significantly smaller than those between the regions, one way to approximate A is to calculate the between-pixel local gradients and then compute their maximum per-channel variation in $f(.)$, i.e.

$$f(p, p') = \max_{a \in \{R,G,B\}} f_a(p, p').$$

The simplest choice for $f(.)$ is to use the pixel channel values (p_a and p_a') directly:

$$f(p, p') = |p_a - p_a'| \quad (2)$$

Pre-processing

Black frame removal—Dermoscopy images often contain black frames that are introduced during the digitization process. These need to be removed because they might interfere with the subsequent border detection steps. In order to determine the darkness of a pixel with (R, G, B) coordinates, the lightness component of the HSL color space (16) is utilized:

$$L = \frac{\max(R, G, B) + \min(R, G, B)}{2} \quad (3)$$

In particular, a pixel is considered to be black if its lightness value is <20. Using this criterion, the image is scanned row-by-row starting from the top. A particular row is labeled as part of the black frame if it contains 60% black pixels. The top-to-bottom scan terminates when a row that contains less than the threshold percentage of pixels is encountered. The same scanning procedure is repeated for the other three main directions.

Image smoothing—Dermoscopy images often contain extraneous artifacts such as skin lines, air bubbles and hairs around the lesion. These reduce the accuracy of the border detection and increase the computational time. In order to mitigate the detrimental effects of these artifacts, the images should be pre-processed with a smoothing filter.

The median filter is one of the most common smoothing filters in the literature. Median filtering with a mask of appropriate size can eliminate most of the artifacts in a dermoscopy image (6). Note that the mask size should be proportional to the image size for optimal results. In this study, given an M by N image, the mask size n is determined by:

$$n = \text{floor}\left(5 \cdot \sqrt{(M/768) \cdot (N/512)}\right) \quad (4)$$

Equation (4) is based on the observation that for a typical 768 × 512 image $n = 5$ is a good choice and it ensures that when the image size changes, this is reflected on the mask size proportionally.

Post-processing—The segmented image often contains regions that are part of the background skin. In order to eliminate these regions, the background skin color needs to be determined. In this study, the approach described in (17) is adopted. Four patches of size 20 × 20 pixels from the corners of the image are taken and the mean *R*, *G*, and *B* of the pixels is calculated. This mean color is taken as an estimate of the background skin color. The light-colored regions, i.e. the regions whose mean color has a distance <60 to the background skin color, are then eliminated. In addition, the regions that touch the image frame and those with rectangular borders are discarded. The initial border detection result is obtained by removing the isolated regions and then merging the remaining regions. Figure 1a–c illustrate this procedure.

Note that in Fig. 1c the automatic border is mostly contained inside the manual border. This is observed in many cases because during the border determination procedure, the computer algorithms tend to find the sharpest pigment change, whereas the dermatologists choose the outmost detectable pigment. In order to bring the automatic border closer to the manual one, three different methods are applied to the initial border. The first one is majority filtering (18) in which the center pixel in an $n \times n$ neighborhood is assigned the majority label (object or background). Because the initial borders are jagged, this operation smoothes the borders, at the same time expanding them. The second method is morphological dilation (18) with a circular structuring element of size n. The third one is based on the Euclidean distance transform (19) in which the initial border is expanded until the lesion diameter reaches a certain percentage, i.e. $(100 + k)\%$, of its initial value. For the first two methods, the operator size n is chosen as $n = \text{floor}(k \cdot (d/500))$, where d is the diameter of the lesion and k is a scaling factor. Figure 1d–f shows a comparison of these three expansion methods on a sample border. It can be seen that after these operations, the automatic borders are much closer to the manual border. The quantification of the border detection error will be explained in the next section.

Results and Discussion

The proposed method is tested on a set of 90 dermoscopy images (23 invasive malignant melanoma and 67 benign) obtained from the EDRA Interactive Atlas of Dermoscopy (20) and the dermatology practices of Dr Ashfaq Marghoob (New York, NY), Dr Harold Rabinovitz (Plantation, FL), and Dr Scott Menzies (Sydney, Australia). These are 24-bit RGB color images with dimensions ranging from 577 × 397 pixels to 1921 × 1285 pixels. The benign lesions include nevocellular nevi and dysplastic nevi.

As a ground truth for the evaluation of the border detection error, manual borders were obtained by selecting a number of points on the lesion border, connecting these points by a second-order B-spline and finally filling the resulting closed curve. Three sets of manual borders were obtained by dermatologists Dr William Stoecker, Dr Joseph Malters, and Dr James Grichnik using this method.

Using the dermatologist-determined borders, the automatic borders obtained from the five automated methods [orientation-sensitive fuzzy c-means (OSFCM) (6), DTEA (12), meanshift clustering (13), JSEG (14), and SRM] are compared using the grading system developed by Hance et al. (21). Here, the percentage border error is given by:

$$\text{Border Error} = \frac{\text{Area}(\text{AutomaticBorder} \oplus \text{ManualBorder})}{\text{Area}(\text{ManualBorder})} \cdot 100\% \quad (5)$$

where AutomaticBorder is the binary image obtained by filling the computer detected border, ManualBorder is the binary image described above, $\oplus$ is the exclusive-OR operation

that gives the pixels for which the AutomaticBorder and ManualBorder disagree, and Area(*I*) denotes the number of pixels in the binary image *I*.

For the SRM, in order to determine the most effective post-processing method, the *k* parameter is varied between 1 and 10 and the mean error (average of the mean percentage border error over each border set) values are calculated as plotted in Fig. 2. As the figure shows, majority filtering is not very effective because it is not capable of expanding the borders to cause a significant reduction in the mean error. On the other hand, morphological dilation reduces the error rates particularly when *k* is between 3 and 6. For comparison purposes, morphological dilation with $k = 6$ will be used as the post-processing method.

Table 1 gives the mean and standard deviation border error for the five automated methods. The best entry (lowest error value) in each row is shown in bold. It can be seen that the results vary significantly across the border sets, highlighting the subjectivity of human experts in the border determination procedure. Overall, the proposed SRM method achieves the best results followed by the DTEA and the JSEG methods. It is to be noted that, with the exception of the SRM, the error rates increase in the melanoma group possibly due to the presence of higher border irregularity and color variegation in these lesions.

Figure 3 shows sample border detection results for the SRM method. It can be seen that the method performs well even in the presence of complicating factors such as diffuse edges, blood vessels and skin lines.

Conclusion

In this paper, a fast and unsupervised approach to border detection in dermoscopy images based on the SRM algorithm is presented. The proposed approach is comprised of three main phases: preprocessing, segmentation, and post-processing. The pre-processing phase includes black frame removal and image smoothing. The segmentation phase includes pixel couple ordering and region merging. Finally, the post-processing phase consists of eliminating the regions that belong to the background skin, removing the isolated regions, merging the remaining regions, and expanding the initial border by morphological dilation to obtain the final result. The execution time of the proposed method is about 0.4 s for a typical image of size 768 × 512 pixels on an Intel Centrino 1.6 GHz computer (Santa Clara, CA, USA). This can be further reduced by using faster algorithms for median filtering (22) and morphological dilation (23).

The method was tested on a set of 90 dermoscopy images. Three sets of dermatologist-determined borders were used as the ground-truth. The border detection error was quantified by a metric developed by Hance et al. (21) computed as the number of pixels for which the automatic and manual borders disagree divided by the number of pixels in the manual border. The results were compared with four other automated methods. The implementation of the SRM algorithm will be made publicly available as part of the Fourier image processing and analysis library, which can be downloaded from http://sourceforge.net/projects/fourier-ipal

Acknowledgments

This work was supported by grants from NSF (#0216500-EIA), Texas Workforce Commission (#3204600182), James A. Schlipmann Melanoma Cancer Foundation, and NIH (SBIR #2R44 CA-101639-02A2).

References

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer Statistics 2007. Cancer J Clin. 2007; 57:43–66.
2. Menzies, SW.; Crotty, KA.; Ingwar, C.; McCarthy, WH. An atlas of surface microscopy of pigmented skin lesions: dermoscopy. Sydney, Australia: McGraw-Hill; 2003.
3. Steiner K, Binder M, Schemper M, Wolff K, Pehamberger H. Statistical evaluation of epiluminescence dermoscopy criteria for melanocytic pigmented lesions. J Am Acad Dermatol. 1993; 29:581–588. [PubMed: 8408794]
4. Binder M, Schwarz M, Winkler A, Steiner A, Kaider A, Wolff K, Pehamberger H. Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. Arch Dermatol. 1995; 131:286–291. [PubMed: 7887657]
5. Fleming MG, Steger C, Zhang J, Gao J, Cognetta AB, Pollak I, Dyer CR. Techniques for a structural analysis of dermatoscopic imagery. Comput Med Imaging Graphics. 1998; 22:375–389.
6. Schmid P. Segmentation of digitized dermatoscopic images by two-dimensional color clustering. IEEE Trans Med Imaging. 1999; 18:164–171. [PubMed: 10232673]
7. Gao, J.; Zhang, J.; Fleming, MG.; Pollak, I.; Cognetta, AB. Segmentation of dermatoscopic images by stabilized inverse diffusion equations; Proc IEEE Int Conf. Image Process; 1998. p. 823-827.
8. Pagadala, P. MS Thesis. Rolla: Department of Electrical and Computer Engineering, University of Missouri; 1998. Tumor border detection in epiluminescence microscopy images.
9. Donadey T, Serruys C, Giron A, Aitken G, Vignali J-P, Triller R, Fertil B. Boundary detection of black skin tumors using an adaptive radial-based approach. SPIE Med Imaging. 2000; 3379:810–816.
10. Cucchiara R, Grana C, Seidenari S, Pellacani G. Exploiting color and topological features for region segmentation with recursive fuzzy C-means. Machine Graphics Vision. 2002; 11:169–182.
11. Erkol B, Moss RH, Stanley RJ, Stoecker WV, Hvatum E. Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes. Skin Res Technol. 2005; 11:17–26. [PubMed: 15691255]
12. Iyatomi H, Oka H, Saito M, et al. Quantitative assessment of tumor extraction from dermoscopy images and evaluation of computer-based extraction methods for automatic melanoma diagnostic system. Melanoma Res. 2006; 16:183–190. [PubMed: 16567974]
13. Melli R, Grana C, Cucchiara R. Comparison of color clustering algorithms for segmentation of dermatological images. SPIE Med Imaging. 2006; 6144:3S1–3S9.
14. Celebi ME, Aslandogan YA, Stoecker WV, Iyatomi H, Oka H, Chen X. Unsupervised border detection in dermoscopy images. Skin Res Technol. 2007; 13:454–462. [PubMed: 17908199]
15. Nock R, Nielsen F. Statistical region merging. IEEE Trans Pattern Anal Machine Intelligence. 2004; 26:1452–1458.
16. Levkowitz H, Herman GT. GLHS: a generalized lightness, hue, and saturation color model. Graphic Models Image Process. 1993; 55:271–285.
17. Xu L, Jackowski M, Goshtasby A, Roseman D, Bines S, Yu C, Dhawan A, Huntley A. Segmentation of skin cancer images. Image Vision Comput. 1999; 17:65–74.
18. Pratt, WK. Digital image processing: PIKS inside. Hoboken, NJ: John Wiley & Sons; 2007.
19. Felzenszwalb, PF.; Huttenlocher, DP. Distance transforms of sampled functions. Cornell Computing and Information Science TR2004-1963. Available at: http://people.-cs.uchicago.edu/~pff/papers/dt.pdf
20. Argenziano, G.; Soyer, HP.; De Giorgi, V., et al. Dermoscopy: a tutorial. Milan, Italy: EDRA Medical Publishing & New Media; 2002.
21. Hance GA, Umbaugh SE, Moss RH, Stoecker WV. Unsupervised color image segmentation with application to skin tumor borders. IEEE Eng Med Biol. 1996; 15:104–111.
22. Weiss B. Fast median and bilateral filtering. ACM Trans Graphics. 2006; 25:519–526.
23. Van Droogenbroeck M, Buckley M. Morphological erosions and openings: fast algorithms based on anchors. J Math Imaging Vision. 2005; 22:121–142.

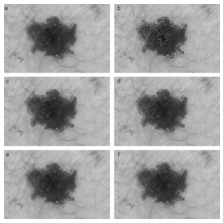

Fig. 1.
(a) Original image, (b) SRM segmentation result, (c) initial border detection result (E = 11.499%), (d) majority filtering (E = 11.477%), (e) morphological dilation (E = 7.081%), and (f) distance transform (E = 7.486%). Green, manual border; blue, automatic border; E, error.

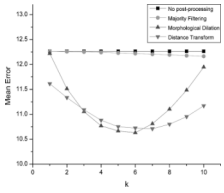

Fig. 2.
Comparison of the post-processing methods.

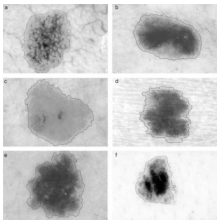

Fig. 3.
Error = (a) 4.186% (melanoma), (b) 5.216% (melanoma), (c) 5.285% (melanoma), (d) 8.290% (benign), (e) 10.245% (benign), and (f) 10.419% (melanoma).

TABLE 1

Percentage border error statistics

Dermatologist	Diagnosis	Statistic	OSFCM	DTEA	Meanshift	JSEG	SRM
W. S.	Benign	Mean	22.995	**10.513**	11.527	10.832	11.384
		SD	12.614	**4.728**	9.737	6.359	6.232
	Melanoma	Mean	28.311	11.853	13.292	13.745	**10.294**
		SD	15.245	5.998	7.418	7.590	**5.838**
	All	Mean	24.354	**10.855**	11.978	11.577	11.106
		SD	13.449	**5.081**	9.193	6.772	6.120
J. M.	Benign	Mean	25.535	10.367	10.802	10.816	**10.186**
		SD	11.734	**3.771**	6.332	5.227	5.683
	Melanoma	Mean	26.743	10.874	12.592	12.981	**10.500**
		SD	14.508	**5.016**	7.202	6.316	8.137
	All	Mean	25.843	10.496	11.259	11.370	**10.266**
		SD	12.426	**4.101**	6.571	5.570	6.351
J. G.	Benign	Mean	27.506	12.091	12.224	12.257	**10.561**
		SD	12.789	5.220	7.393	6.588	**5.152**
	Melanoma	Mean	27.574	12.675	12.168	13.414	**10.411**
		SD	15.836	6.865	7.479	7.379	**5.860**
	All	Mean	27.523	12.240	12.210	12.553	**10.523**
		SD	13.538	5.650	7.373	6.775	**5.308**

DTEA, dermatologist-like tumor extraction algorithm; OSFCM, orientation-sensitive fuzzy c-means; SD, standard deviation; SRM, statistical region merging.