Published in final edited form as:

Skin Res Technol. 2010 August ; 16(3): 378–384. doi:10.1111/j.1600-0846.2010.00445.x.

Watershed segmentation of dermoscopy images using a watershed technique

Hanzheng Wang¹, Xiaohe Chen², Randy H. Moss¹, R. Joe Stanley¹, William V. Stoecker³, M. Emre Celebi⁴, Thomas M. Szalapski³, Joseph M. Malters⁵, James M. Grichnik⁶, Ashfaq A. Marghoob⁷, Harold S. Rabinovitz⁸, and Scott W. Menzies⁹

¹Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO, USA

²Apple Inc., Cupertino, CA, USA

³Stoecker and Associates, Rolla, MO, USA

⁴Department of Computer Science, Louisiana State University, Shreveport, LA, USA

⁵The Dermatology Center, Rolla, MO, USA

⁶Department of Dermatology, University of Miami Health System, Miami, FL, USA

⁷Memorial Sloan-Kettering Skin Cancer Center, Hauppauge, NY, USA

⁸Skin and Cancer Associates, Plantation, FL, USA

⁹Sydney Melanoma Diagnostic Centre, Sydney, Australia

Abstract

Background/purpose—Automatic lesion segmentation is an important part of computer-based image analysis of pigmented skin lesions. In this research, a watershed algorithm is developed and investigated for adequacy of skin lesion segmentation in dermoscopy images.

Methods—Hair, black border and vignette removal methods are introduced as preprocessing steps. The flooding variant of the watershed segmentation algorithm was implemented with novel features adapted to this domain. An outer bounding box, determined by a difference function derived from horizontal and vertical projection functions, is added to estimate the lesion area, and the lesion area error is reduced by a linear estimation function. As a post-processing step, a second-order B-Spline smoothing method is introduced to smooth the watershed border.

Results—Using the average of three sets of dermatologist-drawn borders as the ground truth, an overall error of 15.98% was obtained using the watershed technique.

Conclusion—The implementation of the flooding variant of the watershed algorithm presented here allows satisfactory automatic segmentation of pigmented skin lesions.

Keywords

malignant melanoma; watershed; image processing; segmentation

Skin cancer is one of the most common cancers in humans, and malignant melanoma is the most deadly form of skin cancer. The incidence of melanoma cases, especially melanoma *in situ*, is increasing rapidly (1), adding to the importance of further developing early detection methods such as those based on digital image analysis. Lesion segmentation in the early stage of skin cancer is an important part of digital image analysis in detecting malignant melanoma (2, 3). Numerous methods have been proposed for lesion segmentation in dermoscopy images. Schmid (4) used color clustering. Supervised intensity radial profiles from the HSI color space were used by Donadey et al. (5). Deng and Manjunath (6) presented a region-based approach (JSEG), which was implemented for skin lesion segmentation by Celebi et al. (3). Erkol et al. (7) proposed a method involving gradient vector flow (GVF) snakes. Iyatomi et al. (8) used a region-growing approach to obtain a border for the lesion. A comprehensive survey of the border detection methods applied to dermoscopy images can be found in (9).

In this research, we explore an approach based on the watershed algorithm developed by Meyer and Beucher (10). The simple steps of this method are: (1) growing regional minima, (2) labeling regional minima, (3) creating watershed, (4) removing watershed lines, and (5) creating the watershed image from the labeled image. An adaptive watershed object size is applied to accommodate for possible differences in image resolutions. An outer bounding box method is introduced to estimate the rough lesion ratio, which is an estimate of the ratio between the area of the lesion and that of the whole image. An improved lesion ratio estimate (LRE) formula is proposed to establish a relationship between the outer bounding box ratio and the estimate of the lesion ratio. The rest of this paper is organized into the following sections: (1) general description of the watershed algorithm for lesion segmentation, (2) preprocessing, watershed implementation with adaptive area modification, (3) LRE and object merging, (4) border smoothing and overlay of the border and (5) conclusions and future work.

Materials and Methods

Watershed Algorithm for Lesion Segmentation

Figure 1 illustrates the flowchart of the watershed algorithm. Each flowchart component will be discussed in turn.

Preprocessing: hair removal, vignette removal and black border removal—

Hairs or hair stubbles within the lesion area of an image often affect the watershed segmentation because most of the hairs have a luminance value similar to the pigment in the lesion area. In this research, a hair removal method with the morphological closing operator is introduced to remove hairs if they exist (11). For the purposes of border identification, a more powerful and complete hair removal algorithm is needed. Pixels darker than average in areas of high deviation in luminance are detected. Areas with a high concentration of these pixels are considered to be hair. The algorithm finds much more hair than Dullrazor (11) but does misclassify some features as hair; however, these features are not useful for border identification and can be removed along with the hair without incurring any accuracy loss. Figure 2a shows example images before and after hair removal is applied.

The vignetting effect refers to a position-dependent loss of brightness in the output of an optical system. This effect is seen as the darkening of the outer portions of the images (especially the corners) in some dermoscopy images. The vignetting is detected in the red plane (because it is least affected by the lesion) and normalized for each of the other two planes. Three steps are used to address vignetting: first, define a certain set of concentric circular regions with the width of the circle radius based on a tunable parameter; second, start the procedure with the image center; and third, adjust the brightness of the next circular

region so that the average intensity is the same as the center and continue this for every region. Figure 2b illustrates the images before and after the vignette clearing method is applied. The most noticeable improvement is observed in the image corners.

Similar to the vignetting effect, black borders at image boundaries also adversely affect the image projection histogram. The method chosen to solve this problem is to crop the black border from the innermost point of each black rim. Also, a compensation step that involves resizing the image to the original size by padding is applied before the error calculations. Figure 2c shows the original image and black border-free image.

Luminance and blue plane images—We had initially planned to use the luminance image for segmentation. In the luminance image, the red, green and blue planes are combined into a single plane using the following formula:

$$L=0.30R+0.59G+0.11B$$

where L is the luminance plane, R is the red value, G is the green value and B is the blue value.

After observing that on images with a red lesion rim, larger errors occur using the luminance plane than using the blue plane, the blue plane image was chosen for use. Usually, the blue plane image performs better when a red color occurs at the lesion edge, as with blood vessels and inflammation, which may falsely enlarge the boundary.

Watershed algorithm—The watershed algorithm uses an intensity-based topographical representation, in which the brighter pixels represent higher altitudes or the ‘hills’ and the darker pixels correspond to the ‘valleys,’ which allows for the determination of the path that a falling raindrop would follow (12–14). Watershed lines are the divide lines of ‘domains of attraction’ of water drops. The flooding modification of the watershed algorithm is analogous to immersion of the relief in a lake flooded from holes at minima. The flooding variant is more efficient than the original falling raindrop approach for many applications (14). The basic watershed rainfall topology simulation is used to reach the regional minima by following the raindrop paths, using these basic steps:

1. These regional minima become the flooding start points. We use the same label for the pixels in the rainfall path and the regional minima.
2. Repeat the rainfall simulation on all neighboring pixels to see whether there is another point that could reach the same regional minimum. Mark all the pixels with the same label when the regional minimum is reached; otherwise, a new flooding process will start with a regional minimum proposed for the next flooding.
3. When the different labels of lakes are about to merge, a ‘dam’ is built to prevent the merging. The dam boundaries are finally built when the flooding procedure reaches the global maximum. The dam boundaries here are referred to as watershed lines.

Figure 3 shows a three-dimensional illustration of the flow stages of the flooding process. Figure 4 shows a lesion image and its watershed object’s image.

Adaptive Area Size

For different images, the size of the blobs may affect the total lesion shape. Hence, area size control is necessary in watershed segmentation. In our implementation, an optimized threshold ‘area,’ which stands for the minimum area size of the objects, is introduced into the watershed algorithm:

$$\text{Area}=5I/(1024 \times 768)$$

where area is the minimum size of watershed objects and I is the image size in pixels. Figure 5 shows a comparison of the modified watershed segmentations. Figure 5a shows the original image; the minimum area size is 5 for Fig. 5b and the adaptive area size is used for Fig. 5c.

LRE and Object Merging

LRE based on the outer bounding box algorithm—After performing the preliminary watershed segmentation, the LRE is determined. The LRE is based on the outer bounding box ratio, which is given by the ratio of the area of the outer bounding box to the whole image area. Three formulas are used to determine the outer bounding box: the vertical projection of the image is a function of the horizontal index j :

$$p_j = \sum_{i=1}^n I_{ij} \quad (1)$$

Let p_f be the best-fit quadratic function created from the projection curve p_j :

$$p_f = a_2 x^2 + a_1 x + a_0 \quad (2)$$

To obtain the final subtraction equation, the means of the curves $\overline{p_f}$ and $\overline{p_j}$ are used to normalize the curve, yielding the final equation of the bounding curve B_j :

$$B_j = [p_j - (p_f - \overline{p_f})] \frac{\overline{p_j}}{\overline{p_j}} \quad (3)$$

The maxima of B_j on each side of the global minimum provide the horizontal limits of the bounding box. A similar procedure provides the vertical limits of the bounding box. Figure 6a shows the outer bounding box determination curves, and Fig. 6b shows the bounding box superimposed on a lesion image.

Instead of processing the whole image as was done in Chen's outer bounding box finding method (12), 30 lines in from the image edge in each direction are ignored. Also, we look for the global minimum only in the center 25% of the image area. That is, we start 25% of the image width from one edge of the image and search to the point that is 75% of the image width from the starting edge. A similar procedure is used vertically. This technique can potentially save up to 75% of running time (Fig. 7).

Based on the outer bounding box, we observed the relationship between the outer bounding box ratio and the LRE. Figure 8 shows the relation of the outer bounding box to the actual lesion ratio and the best-fit line (triangles), used to find the LRE.

Object merging—Using the LRE, the watershed objects are merged based on the watershed object histogram. After locating the global maximum, the merging application is started at the peak point of the maximum. Merging stops when the merged area equals the estimated skin area, $(1 - LRE)I$. Figure 9 shows the object merging procedure.

Border Smoothing

B-Spline smoothing—Second-order B-Spline closed curve fitting (15) is applied to the border of the merged watershed segmentation. Average control points for the B-Spline are calculated at a distance of 32 pixels (in either the x -direction or y -direction, whichever occurs first) along the border (based on experimentation using 8-, 16- and 32-pixel distances) for the B-Spline smoothing interval. Figure 10 shows the border mask after B-Spline smoothing.

Final border overlay—In Fig. 11, the solid line represents the watershed border and the dashed line represents the average of three dermatologist-drawn borders.

Results and Discussion

When compared with the average dermatologist border, our watershed method had a mean percentage border error of 15.98% based on the 100-image set. We compared the mean errors on the same set of lesions using previously implemented border detection methods: Pagadala's method (16), the GVF snake method (7) and the JSEG method (3). The watershed error (benign error of 15.72%, melanoma error of 16.66% and overall error of 15.98%) is lower than the error from Pagadala's method (benign 19.87%, melanoma 21.96% and overall 21.49%), similar to the error obtained using the GVF snake method (benign 13.77%, melanoma 19.76% and overall 15.59%). The borders from these other two methods were compared with the border of only one dermatologist. The watershed error was higher than that obtained using the JSEG method on the same set of manual borders (borders obtained from the averages of all three dermatologists) (benign 10.78%, melanoma 14.91% and overall 12.02%). The error metric used was the method developed by Hance et al. (17). Let M_{Lesion} represent the area of a manually segmented skin lesion. Let A denote the automatically segmented lesion and $\oplus$ represent the exclusive-OR operation. Then, the

percentage border error E is given by
$$E = \frac{A \oplus M_{\text{Lesion}}}{M_{\text{Lesion}}} \times 100\%.$$

Conclusion

The watershed algorithm provides a useful method for lesion segmentation. The LRE and bounding box methods allow improvements in the basic watershed method by controlling the lesion size. In the preprocessing stage, hair, black border and vignette removal clears some of the noise that affects the lesion segmentation, thus helping the watershed algorithm to segment the image similar to the way dermatologists do. The B-Spline smoothing yields a visually pleasing and satisfactory final segmentation result. Future improvement could be obtained by another segmentation iteration using the LRE as an index to allow further correction of errors made at the first iteration.

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin.* 2009; 59:225–249. [PubMed: 19474385]
2. Zhang Z, Stoecker WV, Moss RH. Border detection on digitized skin tumor images. *IEEE Trans Med Imaging.* 2000; 19:1128–1143. [PubMed: 11204850]
3. Celebi ME, Aslandogan YA, Stoecker WV, Iyatomi H, Hiroshi Oka H, Chen X. Unsupervised border detection in dermoscopy images. *Skin Res Technol.* 2007; 13:454–462. [PubMed: 17908199]
4. Schmid P. Segmentation of digitized dermatoscopic images by two-dimensional color clustering. *IEEE Trans Med Imaging.* 1999; 18:164–171. [PubMed: 10232673]

5. 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.
6. Deng Y, Manjunath BS. Unsupervised segmentation of color-texture regions in images and video. *IEEE Trans Pattern Anal Mach Intell*. 2001; 23:800–810.
7. 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]
8. Iyatomi H, Oka H, Celebi ME, Hashimoto M, Hagiwara M, Tanaka M, Ogawa K. An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. *Comput Med Imaging Graph*. 2008; 32:566–579. [PubMed: 18703311]
9. Celebi ME, Iyatomi H, Schaefer G, Stoecker WV. Lesion border detection in dermoscopy images. *Comput Med Imaging Graph*. 2009; 33:148–153. [PubMed: 19121917]
10. Meyer F, Beucher S. Morphological segmentation. *J Vis Commun Image R*. 1990; 1:21–46.
11. Lee TK, Ng V, Gallagher R, Coldman A, McLean D. Dullrazor: a software approach to hair removal from images. *Comput Biol Med*. 1997; 27:533–543. [PubMed: 9437554]
12. Chen, X. PhD dissertation. Department of Electrical and Computer Engineering, University of Missouri-Rolla; 2007. Skin lesion segmentation by an adaptive watershed flooding approach.
13. Bleau AL, Leon LJ. Watershed-based segmentation and region merging. *Comput Vis Image Underst*. 2000; 77:317–370.
14. Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans Pattern Anal Mach Intell*. 1991; 13:583–598.
15. de Boor, C. A practical guide to splines. New York: Springer-Verlag; 2001.
16. Pagadala, P. MS thesis. Department of Electrical and Computer Engineering, University of Missouri-Rolla; 1998. Tumor border detection in epiluminescence microscopy images.
17. 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.

Fig. 1.
Lesion border segmentation using the watershed algorithm.

Fig. 2.
Preprocessing steps: (a) hair removal images; (b) vignette removal images; and (c) black border cropped images.

Fig. 3. Flooding procedure. Original image with a square indicated in white on the left top inset, the magnification of that square area in the right bottom inset and the flooding procedure for that square area shown in the center.

Fig. 4. Original image and watershed segmentation. (a) Original image and (b) watershed image.

Fig. 5. Modified watershed segmentation and the original image. (a) Original image; (b) object minimum area size of 5; and (c) adaptive minimum area size.

Fig. 6. Outer bounding box. (a) Determination curves; (b) superimposed outer bounding box.

Fig. 7.
Improved global minimum searching.

Fig. 8.
Linear relationship between the actual lesion ratio and the LRE.

Fig. 9. Object merging. (a) Merging method from the global peak (12); (b) watershed objects merged.

Fig. 10.
B-Spline smoothing.

Fig. 11. Final mask overlay: dermatologist border (dashed) watershed (solid). Note the false peninsulas.