

Communications

Counting Moles Automatically From Back Images

Tim K. Lee*, M. Stella Atkins, Michael A. King, Savio Lau, and David I. McLean

Abstract—Density of moles is a strong predictor of malignant melanoma, therefore, enumeration of moles is often an integral part of many studies that look at malignant melanoma. An automatic method of segmenting and counting moles would help standardize studies, compared with manual counting. We have developed an unsupervised algorithm for segmenting and counting moles from two-dimensional color images of the back torso region, as part of a study to evaluate the effectiveness of sunscreen. The method consists of a new variant of mean shift filtering that forms clusters in the image and removes noise, a region growing procedure to select candidates, and a rule-based classifier to identify moles. When this algorithm was compared to an assessment by an expert dermatologist, the algorithm showed a sensitivity rate of 91% and diagnostic accuracy of 90% on the test set, for moles larger than 1.5 mm in diameter.

Index Terms—Adaptive mean shift filters, biomedical image processing, image segmentation, moles, nevi, noise removal.

I. INTRODUCTION

Cutaneous malignant melanoma is a potentially lethal form of skin cancer. The mechanism of melanoma development is not yet determined, but mole density has been reported as the strongest risk factor [1] with about 50% of melanoma originating from pre-existing moles [2], and moles have been considered as the precursor for the disease [3]. The important relationship between moles and melanomas makes mole counting an integral part of many melanoma studies. However, manual counting is costly and subjective and can be inconsistent, depending upon the training of the counter. An automatic method of segmenting and counting moles would help standardize studies, and would aid in the registration and tracking of moles of patients who are at a high risk of developing the disease and who need regular mole examinations.

In related work, the most common mole segmentation methods are based on thresholding skin images [4]–[6]. Other methods require two steps. Step one uses radial lines or profiles [7]–[10] crossing the mole interior into the surrounding skin to determine the threshold, then in step two the true boundary position on the radial lines or profiles is determined. Texture features have also been exploited. Markov random field model parameters were estimated from the first principal component image in [11], while [12] used co-occurrence matrices in conjunction with a pyramid-based region growing method on the intensity image. Another region growing algorithm reported in [11] utilized a nonlinear diffusion technique termed “stabilized inverse diffusion equation”. Neural networks have been used to classify extracted mole features in [13].

Manuscript received June 7, 2004; revised February 13, 2005. This work was supported in part by the Natural Science and Engineering Research Council of Canada (NSERC). *Asterisk indicates corresponding author.*

*T. K. Lee is with the School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada. He is also with the Cancer Control Research, British Columbia Cancer Agency, 600 W 10th Ave., Vancouver, BC V5Z 4E6, Canada, (e-mail: tlee@bccancer.bc.ca).

M. S. Atkins and M. A. King are with the School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada.

S. Lau is with the School of Engineering Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada.

D. I. McLean is with BC Cancer Agency and the University of British Columbia, Vancouver, BC V5Z 4E6, Canada.

Digital Object Identifier 10.1109/TBME.2005.856301

The above methods have achieved various degrees of success. However, these methods were designed to work with images with exactly one mole in the middle of the image surrounded by skin background, and these methods cannot be applied without modifications to our study, which attempts to identify all moles in a back torso. Currently, only a few methods have been described to extract multiple moles from images of large anatomic sites [14], [15]. In [14], the moles in the front and the back torso are detected by thresholding the output of a Sobel operation, which highlights the border of the moles. However, a global threshold value is difficult, if not impossible, to obtain. In [15], a stereoscopic mole mapping system has been proposed, which utilizes a pair of stereoscopic cameras and a high resolution texture camera. This system allows full three-dimensional reconstruction of the trunk using eight views when a patient is rotated on a motorized turntable, but the technical details of locating the moles are not given. Our research team has been investigating the mole counting problem and earlier results have been presented in [16], [17].

In this paper, we describe an automatic method of segmenting and counting moles, using a new unsupervised algorithm for segmenting and counting moles from two-dimensional color images of the back torso region. This will help standardize studies, and can be used to track moles and mole development. The remainder of the paper is structured as follows. The materials and methods are given in Section II, and the results in Section III. Section IV contains a discussion of the results, and the conclusions presented in Section V.

II. MATERIALS AND METHODS

A. Materials

Data was obtained from color slides taken from an epidemiologic study on broad-spectrum sunscreen use and mole development [18]. The study photographed the back of participants during their mole examinations according to the protocol specified in [19], and moles were manually counted according to the protocol specifications, which define countable moles as brown to black spots that are well defined and darker than surrounding skin. The images were digitized at 2000 dpi with 24-bit color so that the resolution was about 4 pixels/mm. A test set of eight digitized images was chosen at random and labeled by a dermatologist trained for mole recognition. One of the test set of images was also digitized at 48-bit color, using the same digitizer.

B. Methods: Finding the Moles

The first stage is to identify the back, the region of interest, from the digitized images.¹ The second stage, to find moles from skin images, consists of three major steps, detailed below. The first step uses a version of the mean shift algorithm [20]–[22] that removes noise from the image while preserving the moles. The second step is a simple merging algorithm that creates large clusters of pixels and identifies the possible moles in the image. The final step classifies each candidate as moles or not.

1) *Noise Removal Using the Mean Shift Algorithm:* The general mean shift algorithm [20] is a good clustering algorithm that preserves edge boundaries as well keeping a representative color or intensity of each cluster. It has been used with extensions [21], [22] for image segmentation. We built another extension with an adaptive kernel for use in segmenting moles, exploiting the fact that moles are clusters of pixels that are darker than the surrounding skin. Our version of the

¹An automatic segmentation program of the back torso was developed, but it is not used in this paper.

mean shift algorithm works on the intensity image of the back, whose range is $[0, 1]$, and is defined as follows.

Let S be a window of pixels centered at pixel x whose intensity is $I(x)$. The sample mean of x is determined by the pixels in S whose intensity values are similar to $I(x)$. The other pixels whose intensity values are far away from $I(x)$ are not used in the sample mean computation. Mathematically, the sample mean $m(x)$ is defined as

$$m(x) = \frac{\sum_{s \in S} K(|I(s) - I(x)|)w(s, x)I(s)}{\sum_{s \in S} K(|I(s) - I(x)|)w(s, x)} \quad (1)$$

where K is an adaptive Gaussian kernel defined in intensity space and w is a weighting function in spatial space. The adaptive Gaussian kernel K has a standard deviation inversely proportional to a function $B(Y)$, which specifies the percentage of pixels having similar intensity to x in a set $Y \subseteq S$ as

$$K(z) = \begin{cases} e^{-B(Y)z^2}, & \text{if } z \leq a \\ 0, & \text{otherwise} \end{cases}$$

$$\text{and } B(Y) = \frac{c \sum_{y \in Y} f(|I(y) - I(x)|)}{T_y},$$

$$\text{and } f(z) = \begin{cases} 1, & \text{if } z \leq a \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where T_y is the number of elements in Y . For our application, we used a window size of 17×17 pixels, and made the simplification that all the pixels in Y are in S . Thus, T_y is $= 289$. The constant a determines the smoothness of the algorithm; as a is increased the smoother it gets (i.e., as a tends toward infinity, it turns into just a mean filter). So the value of a was chosen to balance smoothness while preserving edges; the chosen value was 0.02. The value of c was chosen to allow typical values to go between a flat to a Gaussian kernel; c was chosen to be 5000. Now as $B(Y)$ decreases, the kernel K resembles a flat kernel which tends to smooth small objects and artifacts. On the other hand, moles should have a large $B(Y)$ since there are many mole pixels with similar intensity clustered together. A high $B(Y)$ will result in a sharp or spiky-shaped Gaussian kernel, which does not allow the skin pixels with high intensity to influence the intensity of the mole pixels.

The function $w(s, x)$ of (1) controls the weight of contributing pixels in the spatial domain and is defined as

$$w(s, x) = \frac{1}{1 + D(s, x)} \quad (3)$$

where $D(s, x)$ denotes the Euclidean distance between s and x . Since moles are always solid in shape and without holes, giving nearby pixels higher weight will help the edges stay within the mole cluster.

The mean shift filtering operation is an iterative algorithm, where every pixel x is replaced by its sample mean $m(x)$ simultaneously for all pixels in the image in an iteration. The algorithm repeats until the mean change between an iteration is less 0.002, which is 10% of an intensity quantum as defined in the next step.

2) *Selection of Candidate Moles:* After the image has been processed using the mean shift algorithm, an image with many clusters of pixels with similar value is produced. We further quantize the output to 50 levels, so the distance between quanta is 0.02. All minimum intensity regions are located. A cluster of such minimum intensity regions is considered as a mole candidate if the cluster has a lower intensity than any of the other clusters surrounding it, and it is not against the border of the image. To find the entire area of a mole candidate, it is important not to underestimate large moles where the center may be significantly darker than the outer parts, so an iterative procedure is performed to merge outer mole clusters to the central cluster. This growing stage is now applied, by expanding each candidate region outward to include the neighboring pixels if these pixel values are within a quantum distance of 0.02 from the minimum region. If there are no such pixels,

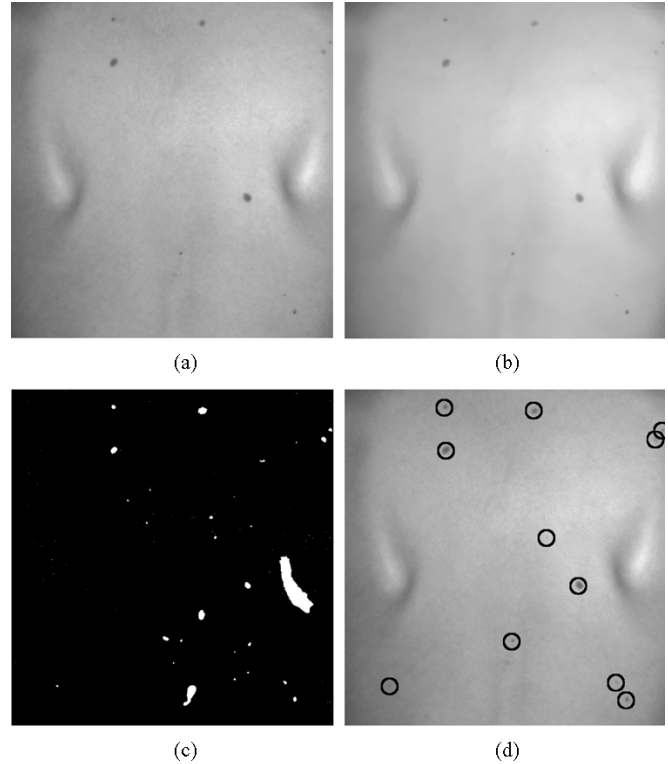


Fig. 1. Illustration of the three-step mole finding algorithm. (a) The selected region of interest for one of the test images. (b) The smoothed image after three iterations of mean shift filtering. (c) The selected candidates after the growing step 2). (d) The final detected moles, highlighted in black.

growing is stopped for that region. Otherwise, growing is repeated for a predefined number of times. The best number of iterations has been determined experimentally to be six.

3) *Use of Feature Space to Identify True Moles:* Our program assumes that a mole is 1.5–23 mm wide, dark, and has an elliptical shape. We computed a set of properties for each candidate mole cluster and stored them in an associated feature vector. The important features are size (defined as the area), major axis length (which defines the diameter of the circle containing the mole) and contrast (defined as the average intensity of the pixels in the candidate divided by the average intensity of the pixels neighboring the border of the candidate). We discard small candidates whose size is less than 1 mm^2 as noise. The major axis length must be between 1.5 mm and 23 mm, to enforce the assumption about the diameter of the mole. Another supporting feature that is used is the candidate's shape, where the elliptical measure is defined as the number of pixels within the mole's convex hull divided by its size, and elongation is the ratio between the major axis and the minor axis of the mole. Setting the limits of both ellipticity and elongation to 10 prevents errors from body folds and shadows from the shoulder blades, which are typically long, thin, and dark. The minimum contrast between the mole and its neighbor is set experimentally to 0.955, as discussed in Section III-C. The program examines every candidate and removes those that do not fall into the range. The rest are considered true moles.

III. RESULTS

A. Segmentation of the Moles

The above three-step mole finding algorithm was implemented in Matlab. To test the program, we manually selected the back from our test images, processed the selected region of interest by the program and compared the detected moles with the dermatologist's labels, which were considered as gold standard. Fig. 1(a) shows one of the test images and Fig. 1(b) shows the corresponding smoothed intensity

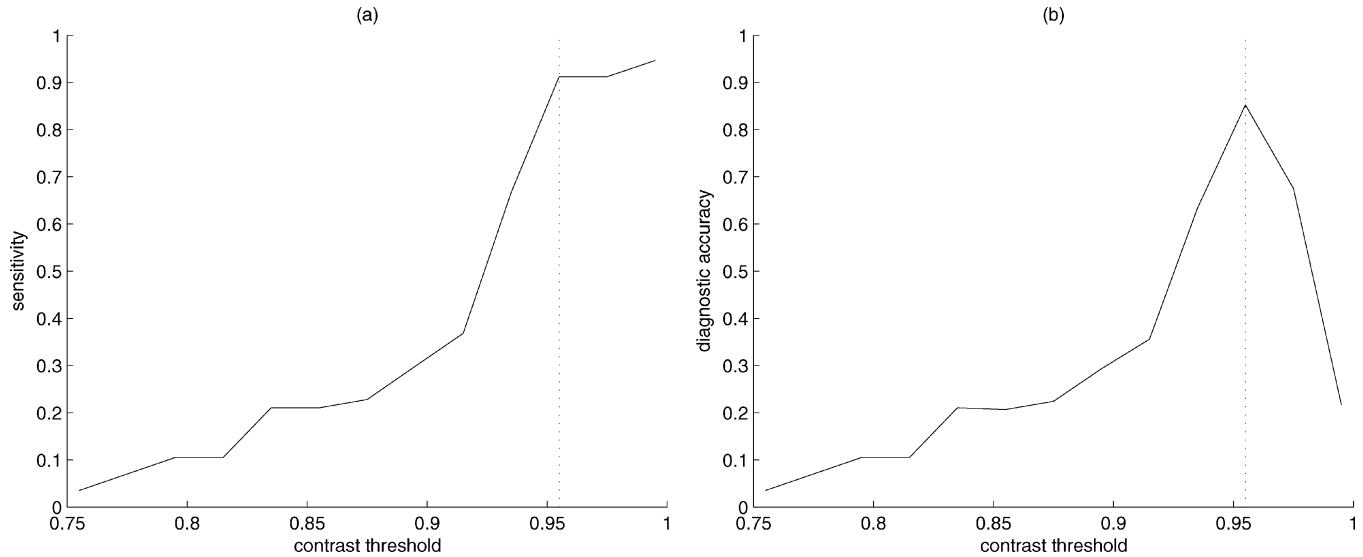


Fig. 2. (a) Sensitivity and (b) diagnostic accuracy of the program for different contrast threshold values. Sensitivity is plotted in the range of 0 and 1.

image after three iterations of mean shift filtering. The mean changes for the three iterations were 0.003, 0.0021, and 0.0011, respectively.² Fig. 1(c) depicts all mole candidates after the growing step. After examining each candidate's feature vector, 11 moles were detected by the program. These moles are outlined in black in Fig. 1(d). Comparing these detected moles with the gold standard, we found the detected moles were all true positives (TPs), with no false positives (FPs) or false negatives (FNs).

The left-most columns of Table I compare the program results for all test images with the gold standard. Note that for test image 5, there were no moles larger than 1.5 mm in diameter and the program successfully found no moles. However, because TP, FP, and FN were all zero we could not compute the sensitivity, which is defined as $TP / (TP + FN) * 100\%$, and diagnostic accuracy, which is defined as $TP / (TP + FP + FN) * 100\%$, for this image. Furthermore, the specificity of the algorithm cannot be estimated, since we cannot numerate the number of the true negatives (TNs). All normal skin area without incorrectly identified mole is considered as TN; however, it is impossible to quantify such a measure. Instead, we summed up all TPs, FPs, and FNs for the entire set and computed the overall sensitivity and diagnostic accuracy. With 52 TPs, 1 FP, and 5 FNs, the overall sensitivity and the diagnostic accuracy for the test set were determined as 91% and 90%, respectively.

B. Comparison of Our Algorithm With a Simple Mole Finding Algorithm

It is of interest to compare our algorithm with a simple mole finder consisting of a 3×3 median filtering step followed by a local thresholding step. We repeated the median filtering until the mean change between iterations was less than 0.002 as in our algorithm.³ We implemented local thresholding by dividing the image into a fixed set of windows each of 65×65 pixels. For each window, the pixels whose intensities were three standard deviations lower than the window's mean intensity were selected as candidate moles. Finally the candidates were classified by their feature vectors into moles, similar to our algorithm. The right-most columns of Table I report the results of the simple mole finder, which performed worse than our algorithm for all test images, except for test image 3 where the simple mole finder had 1 TP and 1 FN, same our algorithm. The overall sensitivity for the simple mole

²All test images required two to three iterations of mean shift filtering.

³All test images required two iterations of median filtering.

TABLE I
COMPARISON WITH THE GOLD STANDARD OF THE NUMBER OF MOLES DETECTED USING TWO ALGORITHMS, FOR EACH TEST IMAGE

Test image	Detecting moles by our 3-step mole finding algorithm			Detecting moles by median filter, local threshold and classification		
	TP	FP	FN	TP	FP	FN
1	11	0	0	10	1	1
2	5	0	0	3	2	2
3	1	0	1	1	0	1
4	7	0	2	4	1	5
5	0	0	0	0	2	0
6	3	1	0	3	3	0
7	10	0	1	8	2	3
8	15	0	1	10	1	6
Overall	52	1	5	39	12	18

finder was 68% and the diagnostic accuracy was 57%. We realized that the local thresholding might increase FPs and might not capture the entire mole since the center of a mole is often darker than the other parts. Therefore, we replaced the local thresholding with the candidate selection step, step 2) of our algorithm. Such modification detected one more TP and reduced the FPs and FNs by 2 and 1, respectively. The overall sensitivity and diagnostic accuracy were improved slightly to 70% and 60%, respectively, but they were still inferior to our algorithm using the adaptive mean shift filter.

C. Sensitivity Analyses of the Program Parameters

Similar to many image analysis programs, our algorithm is controlled by a set of parameters. The most important parameter is the contrast threshold, which controls whether a faint candidate, compared to its neighbor, is accepted or rejected. Fig. 2(a) and (b) shows the overall sensitivity and diagnostic accuracy for our algorithm when the contrast threshold was varied from 0.75 to 0.995. The overall sensitivity increased as the threshold increased, because more faint moles were detected. On the other hand, the diagnostic accuracy dropped after it peaked at the threshold used by our program (marked as dotted line) because of the increasing FPs. We found experimentally that choosing a contrast threshold of 0.955 was appropriate for all our test images. Furthermore, changing the contrast threshold changes the results in a

predictable way, hence, if it was considered important to eliminate FPs, it would be easy to decrease the contrast threshold a little.

The program behaved stably (i.e., the program results changed slowly and predictably) when the other parameters (minimum lesion size, major axis length, ellipticity, and elongation) changed.

D. The 48-bit Image

One of the test images that had many small moles around 1.5 mm in diameter was scanned using 24-bit and 48-bit color resolution with the same scanner in one session. Processing the images by the program found that there was no difference in results. Potentially, the re-quantization of the image to 50 levels in the candidate selection step removed the advantages of the 48-bit color image. Hence, we continued with experiments only on the more compact 24-bit data.

IV. DISCUSSION

Sometimes it is difficult to distinguish moles from other skin conditions, such as lentigos. Also, differentiating small moles (those less than 2 mm in diameter) from freckles is not easy even clinically. Lentigos and freckles often occur on sun-exposed areas of the upper back, face, back of hands, and forearms. For our algorithm, these problems are minimized since we focus on the back torso, not a maximally sun-exposed area.

When our dermatologist labeled the test images, he could identify moles less than 1.5 mm in diameter. Our program could also detect these small moles, but with many FPs and FNs. At the current spatial resolution, we set the size limit to a more realistic level of 1.5-mm diameter.

We used a small number of test images to develop the program. However, these images show a large variety of scenarios. One of the images had no moles larger than 1.5 mm in diameter; many images had small moles around 1.5 mm, and many of these were very faint with a low contrast. The largest mole was 5 mm in diameter. The program seems to work well with these images, although a rigid validation with more images is required to fully test the program.

The program has the advantage that no prior knowledge is needed about the number of moles present on the skin. The adaptive mean shift filter, unlike many other de-noising algorithms, preserves the shape of the moles, allowing for accurate statistics to be gathered. The filter also allows a large smoothing kernel for noise removal. The candidate selection steps works locally where a decision is made in a small area without global thresholds. Thus, the resultant program provides a reliable segmentation and enumeration of the moles without any human intervention.

V. CONCLUSION AND FUTURE WORK

This paper presents an algorithm to segment moles from images of backs. The images were obtained using an ordinary 35 mm slide camera with the intention for manual counting of the moles. We have shown the feasibility of counting the moles accurately and reliably using a computer aided approach on the digitized images. With 52 TP moles, one FP, and 5 FNs for a test set, the overall sensitivity and the diagnostic accuracy were determined as 91% and 90%, respectively.

To fully automate the enumeration process, we developed an objective method to segment moles from the torso skin images, using an extended mean shift filter, a region growing procedure, and a rule-based classifier. The program does not require prior knowledge of the number of moles present on the skin. Results so far are very promising, as the identification of moles correlates well with the eight images manually labeled by an expert dermatologist. In addition, analyses have shown that our algorithm outperformed a simple mole finder consisting of a median filter and local thresholding. Furthermore the sensitivity of the results to the program parameters was stable and predictable.

Future work will include further validation of the segmentation of moles in the data, and registration of the more recent images with the earlier images, in order to track mole development.

Ultimately, the program may incorporate diagnosis of color and shape of the lesions to make a complete analysis for the diagnosis of melanoma.

ACKNOWLEDGMENT

The authors would like to thank S. Filipovic, H. Mojtabaei, and A. Lu for testing out the program.

REFERENCES

- [1] E. A. Holly, J. W. Kelly, S. N. Shpall, and S. H. Chiu, "Number of melanocytic nevi as a major risk factor for malignant melanoma," *J. Am. Acad. Dermatol.*, vol. 17, pp. 459–468, 1987.
- [2] T. M. Skender-Kalnenas, D. R. English, and P. J. Heenan, "Benign melanocytic lesions: Risk markers or precursors of cutaneous melanoma?," *J. Am. Acad. Dermatol.*, vol. 33, pp. 1000–1007, 1995.
- [3] R. P. Gallagher and T. K. Lee, "Skin," in *Cancer Precursors: Epidemiology, Detection, and Prevention*, Franco and Rohan, Eds. New York: Springer-Verlag, 2002, pp. 220–231.
- [4] F. Ercal, M. Moganti, W. V. Stoecker, and R. H. Moss, "Detection of skin tumor boundaries in color images," *IEEE Trans. Med. Imag.*, vol. 12, no. 3, pp. 624–627, Sep. 1993.
- [5] T. Lee, V. Ng, D. McLean, A. Coldman, R. Gallagher, and J. Sale, "A multi-stage segmentation method for images of skin lesions," presented at the IEEE Pacific Rim Conf. Communications, Computers and Signal Processing, Victoria, BC, Canada, 1995.
- [6] D. Guthowicz-Krusin, M. Elbaum, P. Szwaykowski, and A. W. Kopf, "Can early malignant melanoma be differentiated from atypical melanocytic nevus by *in vivo* techniques? Part II. Automatic machine vision classification," *Skin Res. Technol.*, vol. 3, pp. 15–22, 1997.
- [7] J. E. Golston, R. H. Moss, and W. V. Stoecker, "Boundary detection in skin tumor images: An overall approach and a radial search algorithm," *Pattern Recogn.*, vol. 23, pp. 1235–1247, 1990.
- [8] S. E. Umbaugh, R. H. Moss, and W. V. Stoecker, "An automatic color segmentation algorithm with application to identification of skin tumor borders," *Comput. Med. Imag. Graphics*, vol. 16, pp. 227–235, 1992.
- [9] T. Donadey, C. Serruys, A. Giron, G. Aitken, J.-P. Vignali, R. Triller, and B. Fertl, "Boundary detection of black skin tumors using an adaptive radial-based approach," presented at the SPIE Medical Imaging, San Diego, CA, 2000.
- [10] E. Claridge and A. Orun, "Modeling of edge profiles in pigmented skin lesions," in *Medical Image Understanding and Analysis*, U.K.: British Machine Vision Association (BMVA), 2002.
- [11] J. Gao, J. Zhang, M. G. Fleming, I. Pollak, and A. B. Cognetta, "Segmentation of dermatoscopic images by stabilized inverse diffusion equations," presented at the Int. Conf. Image Processing, Chicago, IL, 1998.
- [12] A. P. Dhawan and A. Sicsu, "Segmentation of images of skin lesions using color and texture information of surface pigmentation," *Comput. Med. Imag. Graphics*, vol. 16, pp. 163–177, 1992.
- [13] D. P. Filiberti, P. Bellutta, P. Ngan, and D. A. Perednia, "Efficient segmentation of large-area skin images: An overview of image processing," *Skin Res. Technol.*, vol. 1, pp. 200–208, 1995.
- [14] H. Voigt and R. Classen, "Topodermatographic image analysis for melanoma screening and the quantitative assessment of tumor dimension parameters of the skin," *Cancer*, vol. 75, pp. 981–988, 1995.
- [15] J. Smolle, G. Parr, and A. Gerger, "Skin mapping using a low-cost digital imaging," *Skin Res. Technol.*, vol. 9, p. 201, 2003.
- [16] T. K. Lee, M. King, S. Lau, M. S. Atkins, and D. I. McLean, "Automatic nevus counting: A pilot study," *Skin Res. Technol.*, vol. 9, p. 191, 2003.
- [17] M. King, T. K. Lee, M. S. Atkins, and D. I. McLean, "Automatic nevi segmentation using adaptive mean shift filters and feature analysis," in *SPIE Medical Imaging 2004*, San Diego, CA, 2004.
- [18] R. P. Gallagher, J. K. Rivers, T. K. Lee, C. D. Bajdik, D. I. McLean, and A. J. Coldman, "Broad-spectrum sunscreen use and the development of new nevi in white children: A randomized controlled trial," *J. Am. Med. Assoc.*, vol. 283, pp. 2955–2960, 2000.
- [19] D. R. English, R. MacLennan, J. K. Rivers, J. Kelly, and B. K. Armstrong, "Epidemiological Studies of Melanocytic Nevi: Protocol for Identifying and Recording Nevi," Int. Agency Res. Cancer, Lyon, France, IARC Internal Rep. 90/002, 1990.
- [20] Y. Cheng, "Mean shift, mode seeking and clustering," *IEEE Trans. PAMI*, vol. 17, pp. 790–799, 1995.
- [21] D. Comaniciu and P. Meer, "Mean shift analysis and applications," presented at the 7th Int. Conf. Computer Vision, Greece, 1999.
- [22] —, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–618, May 2002.