Automated Quantification of Clinically Significant Colors in Dermoscopy Images and Its Application to Skin Lesion Classification

M. Emre Celebi, *Senior Member, IEEE*, and Azaria Zornberg

***Abstract*—Dermoscopy is a noninvasive skin imaging technique, which permits visualization of features of pigmented melanocytic neoplasms that are not discernable by examination with the naked eye. Color information is indispensable for the clinical diagnosis malignant melanoma, the most deadly form of skin cancer. For this reason, most of the currently accepted dermoscopic scoring systems either directly or indirectly incorporate color as a diagnostic criterion. For example, both the asymmetry, border, colors, and dermoscopic (ABCD) rule of dermoscopy and the more recent color, architecture, symmetry, and homogeneity (CASH) algorithm include the number of clinically significant colors in their calculation of malignancy scores. In this paper, we present a machine learning approach to the automated quantification of clinically significant colors in dermoscopy images. Given a true-color dermoscopy image with N colors, we first reduce the number of colors in this image to a small number K, i.e., $K \ll N$, using the K-means clustering algorithm incorporating a spatial term. The optimal K value for the image is estimated separately using five commonly used cluster validity criteria. We then train a symbolic regression algorithm using the estimates given by these criteria, which are calculated on a set of 617 images. Finally, the mathematical equation given by the regression algorithm is used for two-class (benign versus malignant) classification. The proposed approach yields a sensitivity of 62% and a specificity of 76% on an independent test set of 297 images.**

***Index Terms*—Clustering, dermoscopy, symbolic regression.**

I. Introduction

INVASIVE and *in situ* malignant melanoma together comprise one of the most rapidly increasing cancers in the world. Invasive melanoma alone has an estimated incidence of 76 100 and an estimated total of 9710 deaths in the USA in 2014 [1]. Early diagnosis is particularly important since melanoma can be cured with a simple excision if detected early.

Dermoscopy, also known as epiluminescence microscopy, is a noninvasive skin imaging technique that uses optical magnification and either liquid immersion and low angle-of-incidence lighting or cross-polarized lighting; this makes subsurface structures more easily visible when compared with conventional clinical images [2]. Dermoscopy allows the identification of dozens of morphological features such as atypical pigment networks, dots/globules, streaks, blue-white areas, and blotches [3]. This reduces screening errors and provides greater differentiation between difficult lesions such as pigmented Spitz nevi and small clinically equivocal lesions [4]. However, it has been demonstrated that dermoscopy may actually lower the diagnostic accuracy in the hands of inexperienced dermatologists [5]. Therefore, in order to minimize the diagnostic errors that result from the difficulty and subjectivity of visual interpretation, the development of computerized image analysis techniques is of paramount importance [6].

Color information is indispensable for the clinical diagnosis of melanoma [7]. The most important chromophore in melanocytic neoplasms is melanin. The color of melanin essentially depends on its localization in the skin: black due to melanin located in the stratum corneum and the upper epidermis, light to dark brown in the epidermis, gray to gray-blue in the papillary dermis, and steel blue in the reticular dermis. Blue occurs when there is melanin localized within the deeper parts of the skin as the portions of visible light with shorter wavelengths are more dispersed than portions with longer wavelengths. Red is associated with an increased number or dilation of blood vessels, trauma, or neovascularization. White is often caused by regression and/or scarring.

Most of the currently accepted dermoscopic scoring systems either directly or indirectly incorporate color as a diagnostic criterion [8]. For example, both the asymmetry, border, colors, and dermoscopic (ABCD) structure rule of dermoscopy [9] and the more recent color, architecture, symmetry, and homogeneity (CASH) algorithm [8] consider six different colors in determining the color score: white, red, light brown, dark brown, blue-gray, and black. In both schemes, the color score contributes more than 33% to the overall score. Menzies' scoring method [10] considers the presence of a single color as a negative feature enabling the exclusion of a melanoma diagnosis, whereas the presence of multiple colors is considered as a positive feature. More specifically, the presence of five to six colors has a specificity of 92% and a sensitivity of 53% for invasive melanoma [11]. Finally, according to pattern analysis [12] and its simplified variant, i.e., the seven-point checklist [13], the presence of colored dermoscopic structures such as brown globules, blue-white areas, and black dots are diagnostically significant.

Manuscript received August 13, 2012; revised February 21, 2013; accepted March 23, 2013. Date of publication April 22, 2014; date of current version August 21, 2014. This work was supported by the U.S. National Science Foundation under Grant 0959583 and Grant 1117457.

M. E. Celebi is with the Department of Computer Science, Louisiana State University in Shreveport, Shreveport, LA 71115 USA (e-mail: ecelebi@lsus.edu).

A. Zornberg is with Half Hollow Hills High School West, Dix Hills, NY 11746 USA.

Digital Object Identifier 10.1109/JSYST.2014.2313671

Numerous studies utilized color information for the classification of skin lesions in dermoscopy images [14]–[19]. Most of these studies, however, involved the use of simple low-level cues such as mean and variance of various color channels to characterize the color content of lesions. A notable exception is the study of Seidenari *et al.* [20], where a reference palette with 23 colors was determined over a training set of 369 images (326 nevi and 43 melanomas). On an independent test set of 243 images (200 nevi and 43 melanomas), the authors obtained between 63% and 79% sensitivity and between 62% and 85% specificity depending on the discriminant analysis threshold. It should be noted that the approach of Seidenari *et al.* requires both a calibrated image acquisition system and a border detection step prior to color analysis.

In this paper, we present a machine learning approach to the automated quantification of clinically significant colors in dermoscopy images. The rest of this paper is organized as follows. Section II describes the conventional K-means (KM) clustering algorithm and its application to dermoscopy images. Section III describes the image set used in the study and presents the experimental results, whereas Section IV gives the conclusions and future work.

II. Color Quantization Using KM Clustering Algorithm

A. KM Clustering Algorithm

KM is undoubtedly the most widely used partitional clustering algorithm [21]. Given a data set $\mathcal{X} \in \mathbb{R}^D$, the objective of KM is to partition $\mathcal{X}$ into K exhaustive and mutually exclusive clusters by minimizing the sum of squared error. This problem is known to be NP-hard even for $K = 2$ [22] or $D = 2$ [23], but a heuristic method developed by Lloyd [24] offers a simple solution. Lloyd's algorithm starts with K arbitrary centers. Each point is assigned to the nearest center, and then each center is recalculated as the mean of all points assigned to it. These two steps are repeated until a predefined termination criterion is met.

The complexity of KM is $\mathcal{O}(NK)$ per iteration for a fixed D value. In image processing applications, D is typically equal to three since the clustering procedure is often performed in 3-D color spaces such as RGB or CIELAB [25].

Due to its gradient descent nature, KM often converges to a local minimum of the criterion function [26]. For the same reason, it is highly sensitive to the selection of the initial centers [21]. Adverse effects of improper initialization include empty clusters, slower convergence, and a higher chance of getting stuck in bad local minima [27]. KM is often initialized by K cluster centers chosen uniformly at random from the data points. Such random initialization not only exhibits the aforementioned problems but also leads to clusterings with highly variable quality. In this paper, the deterministic *maximin* algorithm of Gonzalez [28] is used to initialize the cluster centers. This algorithm chooses the first center arbitrarily, and the ith ($i \in \{2, 3, \ldots, K\}$) center is chosen to be the point that has the greatest minimum distance to the previously selected centers. Despite the fact that it was originally developed as a 2-approximation to the K-center clustering problem,[1] this method is commonly used as a KM initializer. In this paper, the first center is chosen as the centroid of the data set. Note that this choice is optimal when $K = 1$ [29].

B. Application of K-Means Clustering Algorithm to Dermoscopy Images

In order to reduce the number of colors in a dermoscopy image using KM, two issues need to be addressed. First, a suitable representation for the image pixels needs be chosen. Second, an appropriate value for the K parameter needs be determined. In this paper, each pixel is represented by a $D = 5$ dimensional vector (x, y, r, g, b), where subvectors (x, y) and (r, g, b) denote the pixel's spatial and chromaticity (in the RGB color space) coordinates, respectively. The distance between two pixels $\mathbf{p}_1 = (x_1, y_1, r_1, g_1, b_1)$ and $\mathbf{p}_2 = (x_2, y_2, r_2, g_2, b_2)$ is then calculated using the following weighted squared Euclidean formula:

$$d(\mathbf{p}_1, \mathbf{p}_2) = w \cdot \left\| \left(\frac{x_1 - x_2}{W}, \frac{y_1 - y_2}{H} \right) \right\|^2 + (1 - w) \cdot \left\| \left(\frac{r_1 - r_2}{L}, \frac{g_1 - g_2}{L}, \frac{b_1 - b_2}{L} \right) \right\|^2 \quad (1)$$

where $w \in [0, 1]$, W, H, and L denote a spatial weight term, image width, image height, and maximum channel value (for 8-bit-per-channel RGB images $L = 2^8 - 1 = 255$), respectively. Note that the normalization of spatial and chromaticity coordinates is necessary unless $W = H = L$.

As described in the previous section, KM requires the number of clusters K to be specified *a priori*, which clearly conflicts with our objective of automatically determining the number of clinically significant colors. A straightforward approach to address this drawback consists of obtaining a set of clusterings with different K values, i.e., numbers of clusters, and then selecting the clustering and thus the corresponding K value that provides the "optimal" result according to a specific quality (validity) criterion [30], [31]. In this paper, we use the following five relative cluster validity criteria [32]: Calinski–Harabasz (CH) index [33], WB index [34], Davies–Bouldin (DB) index [35], Ray–Turi (RT) index [36], and Pakhira–Bandyopadhyay–Maulik (PBM) index [37].

For each image, KM was successively executed to obtain locally optimal clusterings of the pixel data for K values between 2 and 16. These 15 clusterings were then evaluated using the aforementioned five validity criteria, resulting in a set of five "optimal" K value estimates per image.

III. Experimental Results and Discussion

A. Image Set Description

The image set used in this paper consists of 914 digital dermoscopy images obtained from EDRA Interactive Atlas

[1] Given a set of N points in a metric space, the goal of K-center clustering is to find K representative points (centers) such that the maximum distance of a point to a center is minimized.

of Dermoscopy [2], a collection of images acquired in three institutions: University of Naples Federico II, Naples, Italy; University of Graz, Graz, Austria; and University of Florence, Florence, Italy. These were true-color images with a typical resolution of 768 pixels $\times$ 512 pixels. The diagnostic distribution of the cases was as follows: 272 melanomas, 28 blue nevi, 405 dysplastic nevi, 13 combined nevi, 17 congenital nevi, 33 dermal nevi, 20 dermatofibroma, 79 Reed/Spitz nevi, and 47 seborrheic keratoses. The lesions were biopsied and diagnosed histopathologically in cases where significant risk for melanoma was present; otherwise, they were diagnosed by follow-up examination. The image set was randomly divided into two mutually exclusive sets: a training set of size 617 and a test set of size 297. The split was performed in a manner that ensures approximately the same proportion of benign and malignant cases in the two subsets. We adopted such a stratified hold-out validation scheme for two main reasons. First, given its relatively large size, it was reasonable to randomly split the data set into two disjoint subsets at a ratio of 2 : 1. Second, the computationally intensive training phase (see the next section) prohibited the use of more sophisticated validation schemes such as leave one out or k-fold cross validation [38].

B. Symbolic Regression

Symbolic regression [39] is an established method based on evolutionary computation for searching the space of mathematical expressions while minimizing various error metrics [40]. Unlike traditional linear and nonlinear regression methods that fit parameters to an equation of a given form, symbolic regression searches for both the parameters and the form of equations simultaneously. Initial expressions are formed by randomly combining mathematical building blocks such as algebraic operators, analytical functions, constants, and state variables. New equations are formed by recombining previous equations and probabilistically varying their subexpressions. The algorithm retains equations that better model the experimental data and discards unpromising solutions. After equations reach a desired level of accuracy, the algorithm terminates, returning a set of equations that are most likely to correspond to the intrinsic mechanisms underlying the observed system.

In this paper, we used the `Eureqa` software [41] to perform symbolic regression on our training data set, which consists of five estimated K values for each training set image given by the cluster validity criteria mentioned in Section II-B. In order to discover simple mathematical equations with high generalization capability, we used only the most basic building blocks, including the four algebraic operators, i.e., $\{+, -, \times, \div\}$, and constants. Four different spatial weight w values were tested in the experiments, i.e., 0.125, 0.250, 0.375, and 0.500. In each run, the program was terminated after one million generations. The equations generated as a result of each run were evaluated on the test data set using the geometric mean of *sensitivity* and *specificity*. Equations that gave the best results are shown in Tables I and II. Here, the `logistic` (sigmoid) function is given by $\text{logistic}(x) = 1/(1 + e^{-x})$.

The procedure to classify a new image is as follows. First, the image is clustered using KM successively with K values between 2 and 16. The resulting 15 clusterings are then evaluated using the five cluster validity criteria, and the optimal K values given by these criteria are plugged into the appropriate equation from Table I. Finally, the input image is classified as benign if the output of the `logistic` function is less than 0.5 and malignant (melanoma) otherwise.

TABLE I
SYMBOLIC REGRESSION EQUATIONS

Weight (w)	Best Equation
0.125	$\text{logistic}(1.413\text{e}4\, K_{\text{CH}} - 1.981\text{e}5)$
0.250	$\text{logistic}(778.7\, K_{\text{CH}}^2 - 31220\, K_{\text{PBM}})$
0.375	$\text{logistic}(9.232\text{e}7\, K_{\text{CH}} + 7.098\text{e}7\, K_{\text{DB}} - 7.809\text{e}7\, K_{\text{WB}})$
0.500	$\text{logistic}\left(\frac{5.609\text{e}5\, K_{\text{WB}} - 5.929\text{e}6}{K_{\text{DB}} - 8.401}\right)$

TABLE II
PERFORMANCE OF SYMBOLIC REGRESSION EQUATIONS

Weight (w)	Accuracy (%)	Sensitivity (%)	Specificity (%)
0.125	56.23	60.47	54.50
0.250	71.72	61.63	75.83
0.375	59.26	61.63	58.29
0.500	52.53	69.77	45.50

It can be seen that $w = 0.250$ achieves the best tradeoff between *sensitivity* (61.63%) and *specificity* (75.83%), leading to the highest overall *accuracy* (71.72%). It should be noted that automated diagnosis based on color information alone is not likely to reach the accuracy of experienced clinicians. This is because clinicians evaluate lesions based on various factors, including size (diameter), shape, color, texture, patient history, etc., and therefore obtain higher confidence in their diagnoses. In fact, in a recent large-scale clinical study, with a simple threshold of 6 mm on the lesion diameter, the authors obtained 75% sensitivity and 54% specificity [42]. Therefore, single-feature approaches such as the one presented here should be combined with complementary automated approaches to provide diagnostic accuracy comparable to those of clinicians [14], [16], [18]–[20].

IV. CONCLUSION AND FUTURE WORK

In this paper, we have presented a machine learning method for the automated quantification of clinically significant colors in dermoscopy images of skin lesions. Given a true-color dermoscopy image, our method reduces the number of colors in this image using a deterministically initialized spatially constrained variant of the KM clustering algorithm guided by several relative cluster validity criteria whose outputs (i.e., estimates of the optimal number of clusters) are input to an explicit mathematical function derived using an evolutionary computation-based symbolic regression algorithm trained on a set of 617 randomly selected dermoscopy images with known diagnoses. The output of the derived function is then thresholded to estimate the diagnosis (benign or malignant) of the input image. The proposed method yielded a sensitivity of 62%, a specificity of 76%, and an overall accuracy value of 72% on an independent test set of 297 images. Our method is segmentation free, easy to implement, and efficient, and unlike some of the existing dermoscopy image color classification methods, it does not require a calibrated image acquisition system.

Future work includes the validation of the proposed method using dermatologist-determined ground truth and the estimation of its generalization error on a larger image set using k-fold cross validation.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their insightful suggestions and constructive comments that improved the quality and presentation of this paper.

REFERENCES

[1] R. Siegel, J. Ma, Z. Zou, and A. Jemal, "Cancer Statistics, 2014," *CA: Cancer J. Clinicians*, vol. 64, no. 1, pp. 9–29, 2014.

[2] G. Argenziano, H. Soyer, V. De Giorgi, D. Piccolo, P. Carli, and M. Delfino, *Dermoscopy: A Tutorial*. Milan, Italy: EDRA Medical Publishing & New Media, 2002.

[3] S. Menzies, K. Crotty, C. Ingvar, and W. McCarthy, *Dermoscopy: An Atlas*, 3rd ed. New York, NY, USA: McGraw-Hill, 2009.

[4] K. Steiner, M. Binder, M. Schemper, K. Wolff, and H. Pehamberger, "Statistical evaluation of epiluminescence microscopy criteria for melanocytic pigmented skin lesions," *J. Amer. Academy Dermatology*, vol. 29, no. 4, pp. 581–588, Oct. 1993.

[5] M. Binder, M. Schwarz, A. Winkler, A. Steiner, A. Kaider, K. Wolff, and H. Pehamberger, "Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists," *Archives Dermatology*, vol. 131, no. 3, pp. 286–291, Mar. 1995.

[6] M. G. Fleming, C. Steger, J. Zhang, J. Gao, A. B. Cognetta, I. Pollak, and C. R. Dyer, "Techniques for structural analysis of dermatoscopic imagery," *Comput. Med. Imag. Graph.*, vol. 22, no. 5, pp. 375–389, Sep./Oct. 1998.

[7] R. P. Braun, H. S. Rabinovitz, M. Oliviero, A. W. Kopf, and J. H. Saurat, "Dermoscopy of pigmented skin lesions," *J. Amer. Academy Dermatology*, vol. 52, no. 1, pp. 109–121, 2005.

[8] J. S. Henning, S. W. Dusza, S. Q. Wang, A. A. Marghoob, H. S. Rabinovitz, D. Polsky, and A. W. Kopf, "The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy," *J. Amer. Academy Dermatology*, vol. 56, no. 1, pp. 45–52, Jan. 2007.

[9] F. Nachbar, W. Stolz, T. Merkle, A. B. Cognetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun-Falco, and G. Plewig, "The ABCD rule of dermatoscopy. High prospective value in the diagnosis of doubtful melanocytic skin lesions," *J. Amer. Academy Dermatology*, vol. 30, no. 4, pp. 551–559, Apr. 1994.

[10] S. W. Menzies, C. Ingvar, K. A. Crotty, and W. H. McCarthy, "Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features," *Archives Dermatology*, vol. 132, no. 10, pp. 1178–1182, Oct. 1996.

[11] S. W. Menzies, C. Ingvar, and W. H. McCarthy, "A sensitivity and specificity analysis of the surface microscopy features of invasive melanoma," *Melanoma Res.*, vol. 6, no. 1, pp. 55–62, Feb. 1996.

[12] H. Pehamberger, A. Steiner, and K. Wolff, "In vivo epiluminescence microscopy of pigmented skin lesions. I. Pattern analysis of pigmented skin lesions," *J. Amer. Academy Dermatology*, vol. 17, no. 4, pp. 571–583, Oct. 1987.

[13] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino, "Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis," *Archives Dermatology*, vol. 134, no. 12, pp. 1563–1570, Dec. 1998.

[14] H. Ganster, A. Pinz, R. Rohrer, E. Wildling, M. Binder, and H. Kittler, "Automated melanoma recognition," *IEEE Trans. Med. Imag.*, vol. 20, no. 3, pp. 233–239, Mar. 2001.

[15] H. Iyatomi, H. Oka, M. Saito, A. Miyake, M. Kimoto, J. Yamagami, S. Kobayashi, A. Tanikawa, M. Hagiwara, K. Ogawa, G. Argenziano, H. P. Soyer, and M. Tanaka, "Quantitative assessment of tumor extraction from dermoscopy images and evaluation of computer-based extraction methods for automatic melanoma diagnostic system," *Melanoma Res.*, vol. 16, no. 2, pp. 183–190, Apr. 2006.

[16] M. E. Celebi, H. A. Kingravi, B. Uddin, H. Iyatomi, A. Aslandogan, W. V. Stoecker, and R. H. Moss, "A methodological approach to the classification of dermoscopy images," *Comput. Med. Imag. Graph.*, vol. 31, no. 6, pp. 362–373, Sep. 2007.

[17] 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," *Comput. Med. Imag. Graph.*, vol. 32, no. 7, pp. 566–579, Oct. 2008.

[18] D. Ruiz, V. Berenguer, A. Soriano, and B. Sanchez, "A decision support system for the diagnosis of melanoma: A comparative approach," *Expert Syst. Appl.*, vol. 38, no. 12, pp. 15 217–15 223, Nov./Dec. 2011.

[19] A. G. Isasi, B. G. Zapirain, and A. M. Zorrilla, "Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms," *Comput. Biol. Med.*, vol. 41, no. 9, pp. 742–755, Sep. 2011.

[20] S. Seidenari, C. Grana, and G. Pellacani, "Colour clusters for computer diagnosis of melanocytic lesions," *Dermatology*, vol. 214, no. 2, pp. 137–143, 2007.

[21] M. E. Celebi, H. Kingravi, and P. A. Vela, "A comparative study of efficient initialization methods for the K-means clustering algorithm," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 200–210, Jan. 2013.

[22] D. Aloise, A. Deshpande, P. Hansen, and P. Popat, "NP-hardness of Euclidean sum-of-squares clustering," *Mach. Learn.*, vol. 75, no. 2, pp. 245–248, May 2009.

[23] M. Mahajan, P. Nimbhorkar, and K. Varadarajan, "The planar K-means problem is NP-hard," *Theoretical Comput. Sci.*, vol. 442, pp. 13–21, Jul. 2012.

[24] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–136, Mar. 1982.

[25] M. E. Celebi, H. A. Kingravi, and F. Celiker, "Fast color space transformations using minimax approximations," *IET Image Process.*, vol. 4, no. 2, pp. 70–79, Apr. 2010.

[26] S. Z. Selim and M. A. Ismail, "K-means-type algorithms: A generalized convergence theorem and characterization of local optimality," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 1, pp. 81–87, Jan. 1984.

[27] M. E. Celebi, "Improving the performance of K-means for color quantization," *Image Vis. Comput.*, vol. 29, no. 4, pp. 260–271, Mar. 2011.

[28] T. Gonzalez, "Clustering to minimize the maximum intercluster distance," *Theoretical Comput. Sci.*, vol. 38, no. 2/3, pp. 293–306, 1985.

[29] M. E. Celebi and H. Kingravi, "Deterministic initialization of the K-means algorithm using hierarchical clustering," *Int. J. Pattern Recog. Artif. Intell.*, vol. 26, no. 7, pp. 1250018-1–1250018-25, Nov. 2012.

[30] L. Vendramin, R. J. G. B. Campello, and E. R. Hruschka, "Relative clustering validity criteria: A comparative overview," *Statist. Anal. Data Mining*, vol. 3, no. 4, pp. 209–235, Aug. 2010.

[31] O. Arbelaitz, I. F. Gurrutxaga, J. Muguerza, J. M. Perez, and I. Perona, "An extensive comparative study of cluster validity indices," *Pattern Recog.*, vol. 46, no. 1, pp. 243–256, Jan. 2013.

[32] J. Baarsch and M. E. Celebi, "Investigation of internal validity measures for K-means clustering," in *Proc. IAENG ICDMA*, 2012, pp. 471–476.

[33] R. B. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Commun. Statist.–Theory Methods*, vol. 3, no. 1, pp. 1–27, 1974.

[34] Q. Zhao, M. Xu, and P. Franti, "Sum-of-squares based cluster validity index and significance analysis," in *Proc. ICANNGA*, 2009, pp. 313–322.

[35] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 1, no. 2, pp. 224–227, Apr. 1979.

[36] S. Ray and R. H. Turi, "Determination of number of clusters in K-means clustering and application in colour image segmentation," in *Proc. 4th ICAPRDT*, 1999, pp. 137–143.

[37] M. K. Pakhira, S. Bandyopadhyay, and U. Maulik, "Validity index for crisp and fuzzy clusters," *Pattern Recog.*, vol. 37, no. 3, pp. 487–501, Mar. 2004.

[38] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statist. Surveys*, vol. 4, pp. 40–79, 2010.

[39] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA, USA: MIT Press, 1992.

[40] M. Schmidt and H. Lipson, "Distilling free-form natural laws from experimental data," *Science*, vol. 324, no. 5923, pp. 81–85, Apr. 2009.

[41] Nutonian Inc., Eureqa Desktop, 2014. [Online]. Available: www.nutonian.com

[42] N. R. Abbasi, M. Yancovitz, D. Gutkowicz-Krusin, K. S. Panageas, M. C. Mihm, P. Googe, R. King, V. Prieto, I. Osman, R. J. Friedman, D. S. Rigel, A. W. Kopf, and D. Polsky, "Utility of lesion diameter in the clinical diagnosis of cutaneous melanoma," *Archives Dermatology*, vol. 144, no. 4, pp. 469–474, 2008.

M. Emre Celebi (S'05–M'07–SM'11) received the B.Sc. degree in computer engineering from the Middle East Technical University, Ankara, Turkey, in 2002 and the M.Sc. and Ph.D. degrees in computer science and engineering from The University of Texas at Arlington, Arlington, TX, USA, in 2003 and 2006, respectively.

He is currently an Associate Professor with the Department of Computer Science, Louisiana State University in Shreveport, Shreveport, LA, USA. He has pursued research in the field of image processing and analysis. He has published more than 120 articles in journals and conference proceedings. His recent research is funded by grants from the National Science Foundation.

Azaria Zornberg, photograph and biography not available at the time of publication.