Published in final edited form as:

Skin Res Technol. 2010 February ; 16(1): 60–65. doi:10.1111/j.1600-0846.2009.00402.x.

Detection of atypical texture features in early malignant melanoma

Bijaya Shrestha¹, Joseph Bishop¹, Keong Kam¹, Xiaohe Chen¹, Randy H. Moss¹, William V. Stoecker^{2,3}, Scott Umbaugh⁴, R. Joe Stanley¹, M. Emre Celebi⁵, Ashfaq A. Marghoob⁶, Giuseppe Argenziano⁷, and H. Peter Soyer⁸

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

²The Dermatology Center, Stoecker & Associates, Rolla, MO, USA

³Department of Internal Medicine, Division of Dermatology, University of Missouri Health Sciences Center, Columbia, MO 65212, USA

⁴Department of Electrical & Computer Engineering, Southern Illinois University Edwardsville, Edwardsville, IL, USA

⁵Department of Computer Science, Louisiana State University in Shreveport, Shreveport, LA, USA

⁶Memorial Sloan-Kettering Skin Cancer Center, Hauppauge, NY, USA

⁷Department of Dermatology, Second University of Naples, Naples, Italy

⁸Dermatology Group, School of Medicine, University of Queensland, Brisbane, Australia

Abstract

Background—The presence of an atypical (irregular) pigment network (APN) can indicate a diagnosis of melanoma. This study sought to analyze the APN with texture measures.

Methods—For 106 dermoscopy images including 28 melanomas and 78 benign dysplastic nevi, the areas of APN were selected manually. Ten texture measures in the CVIptools image analysis system were applied.

Results—Of the 10 texture measures used, correlation average provided the highest discrimination accuracy, an average of 95.4%. Discrimination of melanomas was optimal at a pixel distance of 20 for the 768×512 images, consistent with a melanocytic lesion texel size estimate of 4–5 texels per mm.

Conclusion—Texture analysis, in particular correlation average at an optimized pixel spacing, may afford automatic detection of an irregular pigment network in early malignant melanoma.

Keywords

melanoma; image analysis; texture; texel; pigment network; dermoscopy

Dermoscopy is a non-invasive imaging technique that uses optical magnification and fluid immersion or cross-polarized lighting to allow better clinical assessment of skin lesions (1).

Dermoscopy has been shown to improve the diagnostic accuracy of pigmented lesions in those with formal training (2).

The pigment network was defined at a dermoscopy consensus meeting in 1999 (1). A typical pigment network is a light-to-dark-brown network with small, uniformly spaced network holes and thin network lines distributed more or less regularly throughout the lesion and usually thinning out at the periphery. An atypical pigment network (APN) is a black, brown, or gray network with irregular holes and thick lines. The significance of a typical pigment network, when present as a global pattern, is that a melanocytic nevus is likely present. When an APN is present, the likelihood of melanoma is increased, with a calculated odds ratio of melanoma of 9.0 for an APN (1). However, APNs may be present in a significant percentage of benign melanocytic skin lesions, particularly dysplastic nevi. The discrimination of these benign nevi with irregular pigment networks from early malignant melanoma is difficult.

A pigment network possesses a visible form of texture. Clinicians usually use the term texture to refer to a palpable texture. In this study, we refer to visible texture, with repeated elements that form a visibly discernable pattern. Texture may be analyzed in an attempt to quantify textural qualities such as coarseness and regularity. Statistical texture analysis, the usual analytic method used for texture analysis, allows detection of various features on a gray-level image to allow discrimination of different textures. These features are not scale independent, and reducing the size of an image makes a textured surface appear smoother. Differences in texture can be detected by comparing local gray-level statistics in various regions. The purpose of this study is to identify a method using texture that can discriminate malignant melanoma with an irregular texture, most commonly an APN, from benign dysplastic nevi, which generally do not have an APN, using texture measures alone.

Methods

In this study, we use five of the classical statistical texture measures of Haralick et al. (3) incorporated into the CVPTools software: (4) energy, inertia, correlation, inverse difference, and entropy. Both the average and the range of each of these variables are computed, yielding 10 standard texture measures, numbered from F1 through F10, as shown in Table 1.

Twenty-eight primary malignant melanoma and malignant melanoma *in situ* images from the EDRA Interactive Atlas of Dermoscopy (5) were used in this study. These images were selected from the 252 malignant melanoma lesions on this disk as having one or more of the following atypical texture features as primary features in the identification of melanoma: APN, branch streaks, radial streaming, and pseudopods. These features were chosen as the forms of pigmentation irregularities with an identifiable texture that are associated with a diagnosis of malignant melanoma. Melanomas were not included if they also displayed any of the following melanoma features without a distinguishing texture: eccentric and irregular blotches (structureless areas), globules, blue areas, white areas, and vascular features. Therefore, for this group of melanomas, the pigment network and other repeating texture features were the only visible means of identification of these lesions as malignant melanomas by dermoscopy. Dermatopathologists confirmed these 28 lesions as early melanomas, with 16 melanomas *in situ*, 11 invasive melanomas with a thickness <0.76 mm, and one invasive melanoma with a thickness of >1.5 mm.

A set of benign lesions was selected at random from the set of dysplastic nevi in the same Interactive Atlas of Dermoscopy (5). Seventy-eight dysplastic lesions were selected so that 10 dysplastic lesions had an APN, or 12.8% of the entire 78-lesion set. This percentage is

similar to the percentage of all dysplastic nevi in the EDRA atlas that had an APN: 13.4% (53 of 396 lesions).

For all 106 images of benign and malignant lesions, a dermatologist (W. V. S.) marked the portion of the lesion with the most irregular texture, the APN area, as defined above. This area was termed the APN area. For the melanomas, these irregular areas contained one or more of the following features: thickened and dark pigment networks, branch streaks, radial streaming, or pseudopods. For the benign lesions, the areas chosen as the APN area were the most irregular areas. Some benign lesions contained one or more of the above features, or some degree of irregularity, and these were selected as the APN areas. For other benign lesions for which no textural irregularities were detected, the APN area was randomly selected. The marked APN areas varied from 2% to 20% of the lesion area, with a median of 9% of the lesion area. Areas of APN for benign and malignant lesions are shown in Figs 1 and 2, respectively. Although the majority of benign lesions show no APN, these figures show that an APN may be present in both malignant and benign lesions. Both examples have thickened lines in the APN area; however, the hole to net unit size ratio is somewhat smaller in the malignant lesion.

For all images, CVIPtools (4) was used to compute two sets of data for each image. The first set of data consisted of all 10 calculated textural features from the APN area. The second set of data consisted of all 10 calculated textural features from the rest of the lesion that was not in the APN area. The dataset for each feature was calculated from a gray-level co-occurrence matrix (GLCM) constructed from the luminance plane of each RGB color image. GLCMs were constructed for each image using pixel distances (d -values) of 6, 12, 20, 30, and 40. Six classifiers (BayesNet, ADTree, DecisionStump, J48, NBTree, and Random Forest) were arbitrarily selected and 10-fold cross-validation was performed for each d -value using these classifiers. A 10-fold cross-validation means that the overall dataset is split into 10 equal-sized bins. Nine bins are used for training and the tenth bin is used as a test set. The trial is repeated nine more times with the test set rotated through the 10 bins. The average results of all 10 trials are reported.

Weka (6, 7), a general data mining tool developed by the University of Waikato in New Zealand, was used as the implementation of the classifiers. BayesNet produces probability estimates based on the Bayes network learning. ADTree builds an alternative decision tree and is optimized for two-class problems. Decision-Stump builds one-level binary decision trees for the datasets. J48 is a standard C4.5 rev 8 implementation of Quinlan's method (6). NBTree builds a decision tree with Naïve Bayes classifiers at the leaves. Random Forest is a statistical classifier that grows many random classification trees (6).

The six classifiers were trained on the data extracted from all 106 images and applied to a dermoscopy image of an early lentigo maligna (melanoma *in situ*, 4.5 mm in diameter, on the cheek of a 69-year-old male). This lesion was difficult to diagnose and was considered to be benign clinically (Fig. 3).

Experimental Results

Figure 4 shows the percentage of correctly classified instances for each classifier and d -value using 10-fold cross-validation using the 106-image set and all 20 inputs: 10 APN area texture features and 10 non-APN area texture features.

Table 2 shows the percentage of correctly classified instances for each classifier and the d -value with 10-fold cross-validation using 10 texture features from the APN area only for the 106-image set. The results show little change from the results using both the APN and the non-APN area, as shown in Fig. 4. Table 3 shows two classification accuracies in each table

location. The table entries on the left show the percentage of correctly classified instances for all 20 variables for each of the 106 images: 10 texture features from the APN area and 10 texture features from the non-APN area. The table entries on the right show the percentage of correctly classified instances using only a single calculated texture feature for each of the 106 images: the correlation average from the APN area. Table 3 shows no overall loss of accuracy using only the correlation average over the APN area, instead of using all 20-texture features to classify lesions.

Table 4 compares the accuracy of the classification of the melanoma *in situ* image (Fig. 3) for each classifier and d -value using all 20 features (10 APN and 10 non-APN features) vs. using the APN area correlation average only. In the table, Y represents the correct classification and N represents the incorrect classification. The same d -values were used here as those used previously for the 106-image set. Four classifier results show a loss in diagnostic accuracy for this lesion using only the correlation average compared with 20 features; however, three classifier results show a gain in diagnostic accuracy, for a net loss of one classifier decision in 30 decisions. For this difficult case, there appears to be no significant overall loss in classifier accuracy using the correlation average alone, in comparison with using all 20 calculated texture features. Overall, when comparing the classification results for different d -values and different classifiers, the choice of the d -value appears to be more important than the choice of the classifier.

Discussion

Visible texture has been analyzed for automatic discrimination of melanoma in digital images using the circular simultaneous autoregressive model (8), fractal dimension (9, 10), neighborhood gray-level dependence matrix (NGLDM) (9, 11), Laws energy filters (9, 11), Cohen orthogonal masks (12), and wavelet functions (13). Anantha et al. (11) compared Laws and NGLDM texture analysis techniques for dermoscopy images and found better results with the Laws technique. Harris compared melanoma discrimination results for several methods, and found that for clinical images, NGLDM and Laws techniques allowed better texture discrimination than fractal and simultaneous autoregressive models, although no method yielded high melanoma discrimination for clinical images (9). No general comparison of methods for clinical or dermoscopy image discrimination has been made for the large number of texture methods available, although visible texture has been incorporated among other features in digital image analysis systems (14). For the study presented here, the Haralick features in Table 1 were used because of their readily available implementation in CVIP-tools.

A previous study by Deshabhoina et al. (15) used the Haralick features to discriminate melanoma from seborrheic keratoses in clinical images, and found a lower success rate in discrimination than the current study. Of the texture features in the Deshabhoina study, texture correlation and energy features proved to be the most useful features; however, the overall discrimination accuracy for melanomas and seborrheic keratoses did not exceed 80%. Energy is not strictly a texture measure in the sense of repeating subunits, as it measures mainly the brightness of the image. Results are not strictly comparable to the current study, as the Deshabhoina study of clinical images averaged texture measures over entire lesions, thus averaging areas of widely varying texture.

In this work, we have found that the correlation average was the single most useful CVIP-tools texture feature in discriminating APN areas in malignant melanomas from the network found in benign dysplastic lesions. The correlation average, applied to the most atypical network area, was able to discriminate malignant lesions with an accuracy equivalent to that obtained with 10 features from the APN area and 10 from the rest of the lesion, a total of 20

features. This result is unexpected, because visible texture seen in pigmented lesions varies from lesion to lesion, and the visible texture deviation from normal in the atypical area would be expected to be a relative texture change. Therefore, it would be expected that the texture information supplied by texture analysis in the normal or the non-APN area would contribute to the accuracy obtained by texture analysis.

The results presented here suggest that texture discrimination is critically dependent on the pixel distance (d -value) used in texture analysis. Anantha et al. (11) found considerable differences in the results depending on the pixel distance used in texture calculations. Texture discrimination was optimized at a texture unit width of about 10 pixels in the 512-pixel width images used, which corresponded to a texel size of 0.22 mm on the skin for both Laws and NGLDM methods. Our results show the best melanoma discrimination at about $d = 20$, which corresponds to a similar 0.2-mm texel width in our 768-pixel width images. In Fig. 5, both branch streaks at the top of the figure and the APN on the right of the figure show approximately four to five repeating malignant texture structures per millimeter, yielding a texture width similar to that found in the Anantha study. Deshabhoina found an optimum pixel distance (d -value) of 2 for unmagnified images. For 10-power images in the current study, an optimum pixel distance in the range of 20 pixels would be expected, and this is the optimum value of all the d -values tested (Fig. 4).

The finding of an average 95.4% discrimination of melanoma from benign lesions by the six classifiers using the correlation average alone is encouraging, as the study set is of greater than average diagnostic difficulty, with the majority of the malignant lesions at the early *in situ* stage, and all but one of the remaining malignant lesions thinner than 0.76 mm. However, the study reported here used only a relatively small number of lesions that were selected for texture abnormalities. The APN areas were selected manually from the lesion, and automatic detection of the most irregular area is difficult. For future work, automatic selection of abnormal pigment network areas can be undertaken, and the techniques reported here can be applied to a larger set of lesions.

Acknowledgments

This publication was made possible by Grant number SBIR R44 CA-101639-02A2 of the National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NIH. The authors wish to thank Jason Hagerty, who served as system administrator.

References

1. Argenziano G, Soyer HP, Chimenti S, et al. Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet. *J Am Acad Dermatol*. 2003; 48:679–693. [PubMed: 12734496]
2. Binder M, Schwarz M, Winkler A, et al. Epiluminescence microscopy: a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. *Arch Dermatol*. 1995; 13:286–291. [PubMed: 7887657]
3. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst, Man, Cybern*. 1973; SMC-3:610–621.
4. Umbaugh, S. [accessed November 13, 2007] CVIPtools, a software package for the exploration of computer vision and image processing. Available at <http://www.ee.siue.edu/CVIPtools>
5. Argenziano, G.; Soyer, HP.; De Giorgi, V., et al. Interactive atlas of dermoscopy. Milan, Italy: EDRA Medical Publishing; 2000.
6. Witten, IH.; Frank, E. Data mining: practical machine learning tools and techniques. San Francisco: Morgan Kaufmann; 2005.
7. Frank, E.; Hall, M.; Trigg, L. [accessed June 19 2007] Weka. Available at <http://www.cs.waikato.ac.nz/ml/weka/>

8. Stoecker WV, Chiang CS, Moss RH. Texture in skin images: comparison of three methods to determine smoothness. *Comput Med Imag Graph.* 1992; 16:179–190.
9. Harris, DE. PhD dissertation. Department of Electrical Engineering, University of Missouri – Rolla; Rolla, MO: 1994. Texture analysis of skin cancer images.
10. Manousaki AG, Manios AG, Tsompanaki EI, Tosca AD. Use of color texture in determining the nature of melanocytic skin lesions—a qualitative and quantitative approach. *Comput Biol Med.* 2006; 36:419–427. [PubMed: 16488774]
11. Anantha M, Moss RH, Stoecker WV. Detection of pigment network in dermatoscopy images using texture analysis. *Comput Med Imaging Graph.* 2004; 28:225–234. [PubMed: 15249068]
12. Chang, J-C. MS thesis. Department of Electrical Engineering, University of Missouri – Rolla; Rolla, MO: 1992. Computer vision techniques for texture analysis in skin cancer diagnosis.
13. Cheng, Y-T. MS thesis. Department of Electrical Engineering, University of Missouri – Rolla; Rolla, MO: 1994. Texture classification for skin tumor images using the adaptive wavelet packets transform.
14. Wollina U, Burroni M, Torricelli R, Gilardi S, Dell'Eva G, Helm C, Bardey W. Digital dermoscopy in clinical practise: a three-centre analysis. *Skin Res Technol.* 2007; 13:133–142. [PubMed: 17374053]
15. Deshabhoina SV, Umbaugh SE, Stoecker WV, Moss RH, Srinivasan SK. Melanoma and seborrheic keratosis differentiation using texture features. *Skin Res Technol.* 2003; 9:348–356. [PubMed: 14641886]

Fig. 1.
Benign dysplastic nevus, with the atypical network area marked.

Fig. 2.
Melanoma *in situ*, with the atypical network area marked.

Fig. 3.
Early lentigo maligna, with the atypical network area marked.

Fig. 4.
Classification results for all 10 APN and 10 non-APN features, 106 images

Fig. 5. Melanoma *in situ* image with about four to five repeating texture units per millimeter (5). (White scale lines indicate 1 mm distances.)

TABLE 1

CVIPtools texture measures from (3, 4) used in this study

Type of CVIPtools texture measure		
Number	Texture measure	Description
F1	Texture energy average	Average gray level in a texture region. This texture measure represents brightness
F2	Texture energy range	Variation of the energy along four cardinal orientations in the texture region (superior, inferior, left, and right)
F3	Inertia average	Average contrast in a texture region
F4	Inertia range	Variation of inertia along four orientations
F5	Correlation average	Average measure of similarity between pixels at a given separation (d)
F6	Correlation range	Variation of correlation along the four orientations
F7	Inverse difference average	Average local homogeneity of the texture
F8	Inverse difference range	Variation of homogeneity along the four orientations
F9	Texture entropy average	Average information content of the texture
F10	Texture entropy range	Variation of entropy along the four orientations

\$watermark-text

\$watermark-text

\$watermark-text

TABLE 2
Classification results for all 10 APN (excluding non-APN) features, 106 images

All 10 APN features						
% correctly classified for 10-fold cross-validation						
d	AD-Tree	Decision-Stump	J48	NB-Tree	Random-Forest	Bayes-Net
6	59.8	71.0	67.3	72.0	67.3	73.8
12	59.8	72.0	72.9	73.8	70.1	73.8
20	94.4	96.3	92.5	95.3	94.4	95.3
30	72.0	70.1	73.8	76.6	74.8	73.8
40	70.1	72.0	68.2	74.8	74.8	73.8

APN, atypical pigment network.

\$watermark-text

\$watermark-text

\$watermark-text

TABLE 3
Classification results for all 20 features compared with the APN correlation for 106 images

All 10 APN and 10 non-APN features compared with the APN correlation only % correctly classified for 10-fold cross-validation									
All 20 APN features/APN correlation average only									
D	ADTree	Decision Stump	J48	NBTree	Random Forest	Bayes Net			
6	67.3	67.9	71.0	73.6	73.6	67.3	66.0	73.8	73.6
12	57.9	64.1	73.8	70.8	73.6	73.8	72.9	73.8	73.6
20	93.5	94.3	96.3	96.2	92.5	95.3	96.2	94.3	96.2
30	66.4	67.9	70.1	70.8	72.0	70.8	75.7	70.8	73.6
40	66.4	64.2	72.0	73.6	68.2	73.6	70.1	58.5	73.8

APN, atypical pigment network.

\$watermark-text

\$watermark-text

\$watermark-text

TABLE 4

Classification results for melanoma in situ test image using all 20 APN and non-APN features compared with one feature: APN correlation average

Classification of melanoma <i>in situ</i> test image, Y =correct, N = incorrect						
All 20 Features/APN correlation average only						
<i>d</i>	AD-Tree	Decision-Stump	J48	NB-Tree	Random-Forest	Bayes-Net
6	Y Y	N N	N N	N N	N Y	N N
12	N N	N N	N N	N N	N N	N N
20	Y Y	Y Y	Y Y	Y Y	Y Y	Y Y
30	N Y	Y N	Y N	Y N	N Y	N N
40	Y N	N N	N N	N N	Y Y	N N

APN, atypical pigment network.