Total Dermoscopy Score calculation using quantitative measurements in digital dermoscopy

Christos Nikolaos E. Anagnostopoulos, Dimitrios D. Vergados, Ioannis Anagnostopoulos and Panagiotis Mintzias

***Abstract*—The main contributions of this paper are an automated approach for applying the ABCDE rules in a digital dermoscopy platform with fixed settings and a new registration method specially designed for aligning and comparing follow-up digital dermoscopy images in order to evaluate the evolution over time parameter E. Experimental evaluations of the registration method are reported for image pairs acquired during follow-up examinations.**

Keywords: Total Dermoscopy Score, digital dermoscopy, image processing.

I. Introduction

In this paper, we propose a fully automatic method for supporting the ABCDE rules accompanied with a registration and comparison step of melanocytic nevi images acquired from follow-up examinations. For the latter issue, our problem has many common aspects with [1], since the evolution assessment over time for melanocytic nevi, implies that reliable matching (and thus registration) is only possible for a part of the nevi, due to difference in their appearance between examinations. Contrary to [1], for the evolution over time factor, we choose not to exclude SIFT keypoints that lie outside the nivus area, since the digital camera of the dermatoscope offers stable imaging conditions.

II. Materials and Methods

More than 200 images from a total of 48 subjects were selected from the Dematological Unit of the General Hospital of Kalamata. For this study, a Mediscope digital videomicroscope with lens that provide manual focus and polarized LED light and the Optipix Dermatoscopy software were used, which are built for management of dermatoscopy images. The system includes a sophisticated imaging platform for body mapping of nevi. The user may link dermatoscopy images to a predefined body map. One of the advantages of using this software is to store and manually compare follow up images with the previous image. By flagging a location as follow up the physician can easily access the nivus the next examination. However, the automatic comparison of follow-up images is not possible if those images are not registered accordingly [2]. By only this way, a reliable quantitative comparison method could be possible, where even the slightest variation can be identified and assessed. All of the magnifications of the instrumentation are pre-calibrated and can be used to measure important dimensions such as area, circumference, color variation etc.

C.N.E. Anagnostopoulos is with the Cultural Technology and Communication Dpt., University of the Aegean, Mytilene, Lesvos, GR 81100, Greece, (Tel: +30 2251036624; fax: +30 2251036637; e-mail: canag@ct.aegean.gr).

D.D. Vergados is with the Informatics Dpt., University of Piraeus, 18534, Greece (e-mail: vergados@unipi.gr).

I. Anagnostopoulos is with the Computer Science and Biomedical Informatics Dpt., University of Thessaly, Lamia, GR 35100, Greece (e-mail: janag@ucg.gr).

P. Mintzias is Director of the Dermatological Department of the General Hospital of Kalamata, 24100, Greece, (Email: petermitzias@gmail.com).

III. Image processing for the ABCDE rule

To calculate the ABCDE score, the "Asymmetry (A), Border (B), Colors (C), Diameter (D) and Evolution (E)" criteria are assessed quantitatively using image processing methods [3]. Each one of the above criteria is multiplied by a given weight factor to yield the Total Dermoscopy Score (TDS) as shown in Equation 1. TDS values less than 4.75 indicate a benign melanocytic nivus, values between 4.8 and 5.45 indicate a suspicious for melanoma nivus, while values of 5.45 or greater are highly suggestive of melanoma [4],[9].

$$TDS = A*1.3 + B*0.1 + C*0.5 + D*2.5 + E \quad (1)$$

A. Nivus detection

The segmentation method used is a robust active contour algorithm based on Expectation-Maximization Snakes model (EM-Snakes) as proposed in [5]. The EM-Snakes tries to estimate the object contour using elastic models in the presence of cluttered background i.e., some of the features extracted from the image (e.g., edge points) should not attract the model.

In our research, the edge segments are computed as follows. Initially, the original image is thresholded with the popular Otsu method [6]. The white points belong to the nivus and the black ones are considered as skin. The average intensity in both regions (nivus and skin) is computed and a set of radial lines is defined starting from the Center of Mass (CM) of the nivus.

The edge points are detected along each direction using template matching and they are linked using a standard procedure, as two edge points are linked if they belong to consecutive lines and their distance to the center is similar. Classic active contour algorithms approximate the object boundary by an elastic curve attracted by image features.

But, since many of them are outliers and attract the model towards misleading configurations the EM algorithm is adopted which assigns confidence degree w^j to each segment y^j and updates them accordingly [5]. The elastic curve $u = [u_1,..., u_M]$ is estimated by minimizing an energy function with two terms:

$$E(u) = E_{\text{int}}(u) + \sum_{i=1}^{M} P_a(u_i) \quad (2)$$

where $E_{int}(u)$ is an internal energy which depends on the curve configuration and $P_a(u_i)$ is an adaptive potential given by equation (3):

$$P_a(u_i) = -\sum_j w^j (\sum_j G_\sigma(y_n^j - u_i) \quad (3)$$

where y^j_n is the n^{th} sample of the j^{th} segment and G_σ denotes the Gaussian function [3]. The adaptive potential depends on the segment confidence degrees w^j, which change during the estimation process. The update of the weights w^j and the minimization of the contour energy are performed by the EM algorithm [5] (see Figure 1).

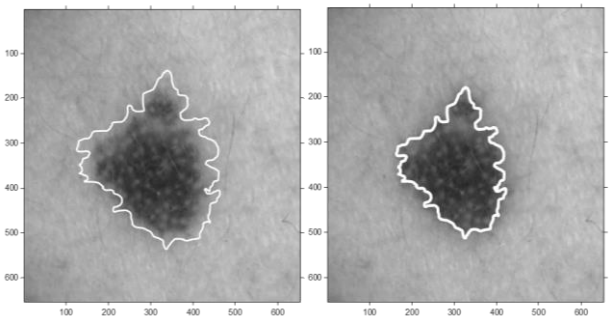

Fig. 1. The nivus detection process: (left) contour during iteration process, (right) final contour/border.

B. Asymmetry

To assess asymmetry, the melanocytic lesion is bisected by two 90º axes that were positioned to produce the lowest possible asymmetry score. If both axes dermocopically show asymmetric contours with regard to shape, colors and/or dermoscopic structures, the asymmetry score is 2. If there is asymmetry on one axis only, the score is 1. If asymmetry is absent with regard to both axes the score is 0. In this work, we evaluate the asymmetry of the nivus identifying possible symmetry in its shape. To measure the nivus symmetry, the Gradient Orientation Histogram (GOH) that was proposed by Sun and Si is used [7].

The Gradient Orientation Histogram (GOH) plots the frequency F(θ) (i.e. number of appearances in angle θ against θ values in the horizontal axis). By looking at the histogram for a specific image, a viewer is able to judge if there is a clear symmetry axis in this image. The angle is measured 0 degrees with reference to the vertical direction, increasing by rotating clockwise. If F(θ) is greater than a specific threshold (e.g. 200), then angle θ is a direction of a symmetry axis.

C. Border

For this factor, the lesion is divided into eighths, and the pigment pattern is assessed. Within each one-eighth segment, a sharp, abrupt cut-off of pigment pattern at the periphery receives a score equal to 1. In contrast, a gradual, indistinct cut-off within the segment receives a score equal to 0. Thus, border can take an integer number from 0 to 8.

In this paper we follow an approach similar to the one described in [8] where a basic parameter characterizing the pigment transition (gradient) can be evaluated effectively. In our method, a strip area of a fixed width (e.g. 100 pixels) is defined using the edges of the nivus boundary and then edge profiles perpendicular to this boundary are calculated.

To this end, the contrast (difference of max and min value), the gradient and the edge zero-crossing are calculated for those profiles and each of these values is assigned to one of the respective 8 regions show in Figure 2. If the gradient of an edge profile is above a certain threshold, this is a strong indication that in this profile there is an abrupt irregular change in border. This yields in characterizing the respective 8-region area as irregular. It should be noted that the pixels in the strip are smoothed to further remove small irregularities.

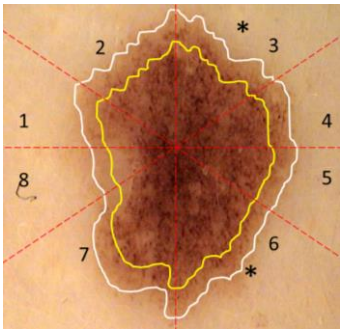

Fig. 2. Examples of Border score calculation. Asterisks indicate abrupt change in border.

TABLE I
ORIGINAL AND CALIBRATED RGB TRIPLETS FOR EVALUATING COLOR SCORES

Color	RGB values as proposed in [5]	RGB calibrated colors
Black	0,0,0	0,0,0
Red	204,51,51	176, 54, 27
Light Brown	153,102,0	144,76,0
Dark Brown	51,0,0	40,0,0
Gray-Blue	51, 153, 153	83, 167, 174
White	255,255,255	255,255,255
Other	Remaining hues	Remaining hues

A. Color

Six distinct colors are considered for the evaluation of the color score, namely white, red, light brown, dark brown, blue-gray, and black [9]. For each color present, a one is added to the score. White is counted only if the area is lighter than the adjacent skin. The maximum color score is 6, and the minimum score is 1.

Thus, for the color factor, the occurrence of these colors in the digital dermoscopy image is measured. Alcon et al [10] provides RGB triplets indicative for skin lesions that are well

accepted by the literature (see Table I, central column). However, since the proposed method is applied in polarized LED environment, those values are calibrated accordingly [11].

In addition, color counters B, R, LB, DB, GB, W and O are normalized (divided) by the nivus area A, and generate the 6-feature vector with normalized counters indicated with subscript n, as shown in Equation 4.

$$Color = \{B_n, R_n, LB_n, DB_n, GB_n, W_n, O_n\} \quad (4)$$

Finally, a 1 is added to the overall color score of the TDS index, if a normalized counter (except O_n) occupies more than 1% of the nivus area.

B. Diameter and differential structures

According to this factor, a lesion is suspicious if it has a diameter greater than 6 mm. In our approach, the diameter is specified by the equivalent diameter of a circle with the same area as the segmented nivus area A. As a result, the equivalent diameter d_e is computed as shown in Equation 5:

$$d_e = \sqrt{(4 \cdot A)/\pi} \quad (5)$$

Therefore, possible values are zero or one (i.e. in case that diameter is lower of higher than 6 mm respectively).

C. Evolution

This is the last and most complex parameter to be calculated, which is related to evolution over time and thus it is the only criterion that needs at least one follow-up examination. Any change in size, shape, color, elevation, or another trait is a significant indication for malignancy and E=1.2 in the TDS score (Equation 1). In contrast, if there is no change, then E=-0.8. Hence, prior to classification, a robust registration module is necessary. In this paper, the main contribution is a fully automatic method for successful registration (matching) of melanocytic nevi images to ensure reliable quantitative comparison of them.

A modified SIFT algorithm (called ROI-SIFT) is used to localize and match correspondence interest points, which will be used to compute affine transformations to map a dermatological image to the respective follow-up examination image [12]. But, since the mapping procedure can be misled by erroneous correspondences, RANSAC is used to identify outliers and maximize the registration accuracy. Then follows rigid image registration, based on the best homography found.

Rigid transformation method is used since it is global in nature and it does not model local geometric differences between images. Thus, local geometric differences in nevi between follow-up images will be exposed successfully.

1) Salient feature detection using ROI-SIFT

In this paper, the parameter h is not set with a constant value as in original SIFT. Specifically, the algorithm begins with a "hard" value for h (h=0.01) and the respective feature points are identified. Then, according to the position, the scale and the orientation of the keypoints, we define the expanded Region of Interests (RoIs). The next step implements the execution of the SIFT algorithm with a "soft" value for h (h=0.03) producing a larger number of keypoints. In our method, only the keypoints belonging to the ROIs are kept, while those that fall outside the ROIs areas are eliminated.

2) Feature matching, validation and registration

Once the interest points have been localized, their SIFT description has to be computed and they have to be matched in a robust way in order to proceed to the registration. However, with SIFT descriptors we have correct and incorrect matches between the pair of images. A good way to tackle with this problem is to implement the RANdom SAmpling Consensus method [13] (RANSAC). The method achieves its goal by iteratively selecting a random subset of the original data. These data are hypothetical inliers and this hypothesis is then tested as follows: (a) a model is fitted to the sample of hypothetical inliers, (b) all other data are then tested against the fitted model and, those points that fit the estimated model well are considered as part of the consensus set, (c) the estimated model is reasonably good if sufficiently many points have been classified as part of the consensus set and (d) the model may be improved by re-estimating it using all members of the consensus set.

3) Registration performance

In order to assess the performance of the proposed registration procedure, we tested our method on two image sets. The first set consists of 64 pairs of images only with the nevi under examination, while the second one (33 pairs of images) contains visible artifacts (mostly hairs). The evaluation in two distinctive sets was decided due to the fact that, unfortunately, the proposed registration method is distracted by the presence of dense hair (straight or intersecting to each other). However, the presence of thin and sparse hair segments do not affect the performance of the method, since the ROI-SIFT method and the RANSAC algorithm usually are tolerant to this kind of artifacts.

TABLE II
REGISTRATION PERFORMANCE IN 2 SETS
(ALL VALUES ARE AVERAGE VALUES)

Set	Initial key-points (ROI-SIFT)	Final key-points after RANSAC	Average distance error of final keypoints	Average distance error of 10 manually selected pair of points	PNSR/ MSE
Set 1	67.4	9.8	4.2	6.7	36.6032/ 14.2156
Set 2	97.2	4.3	5.8	14.1	12.2569/ 45.4361

The registration error is measured by defining manually 10 pairs of matching points for each follow-up pairs of images and measuring the average pixel distance achieved with the estimated homography. The error is computed as: (i) the average Euclidean distance between the matched points by RANSAC (Table II, column 4), (ii) average Euclidean

distance between the 10 manually selected pair of points (Table II, column 5) and (iii) global image metrics like Peak signal-to-noise ratio (PNSR) and Mean Squared Error (MSE) of the initial and the registered image (Table II, column 6).

IV. Results

Tests of correct implementation of the dermoscopy ABCDE factors were carried out on the 64 pairs of images. For the evaluation we have chosen images that do not contain hairs or artifacts in the examined area. The nevi segmentation method was found to be highly accurate when compared to ground truth images provided by an expert. The evaluation was performed calculating the Hammoude distance [14], which makes a pixel by pixel comparison of the pixels annotated as nevi or skin:

$$HM\ (SR, GT) = \frac{\#(SR \cup GT) - \#(SR \cap GT)}{\#(SR \cup GT)} \quad (6)$$

where SR and GT denote the result of the proposed segmentation method and the ground truth segmentation obtained by the medical expert. Both SR and GT are binary images such that all the pixels inside a nivus have label 1 and all the others in the skin have label 0.

The border error for 64 medical images was slightly above 10 %, which is a value that did not affect the next factors in the ABCDE rules. The percentage of agreement between expert opinion and symmetry detection algorithm was also high reaching 90% (58/64 cases). The percentage of perfect matching in border irregularity score was sufficient enough reaching 54.7%, which means that for every nivus the algorithm correctly classified the eights of the 8-regions. However, this is acceptable since this score ranges from 0 to 8. Moreover, only in five cases there was a difference more than 3 in the border irregularity score. Similarly, there was a significant agreement in the color score in 46/64 cases (72%), bearing in mind that the scoring scale is 0 to 6. Finally, there was a complete agreement in measuring the diameter.

Overall, the overall classification results in the 3 classes (benign, suspicious for melanoma and melanoma) for the implemented ABCDE rules were statistically measured by sensitivity and specificity. From the experimental set, the ABCDE method based on the proposed image processing and registration methods compared to expert opinion, gave 88.2 % of sensitivity and 82.7 % specificity. In different papers the sensitivity ranges between 85-91% and the specificity between 75-90 %.

V. Conclusions

In this paper, we propose a fully automatic method for supporting the ABCDE rules accompanied with a registration and comparison step of melanocytic nevi images acquired from follow-up examinations. At present there is great interest in the prospects of an early screening system for teledermatoscopy, based on the automatic analysis of dermatoscopic images. The benefits of such systems are two-fold: (1) to provide quantitative information about a nivus that can be relevant for the clinician; (2) to be used as a stand-alone early warning tool, with the inherent advantages of time effectiveness and low cost procedures of diagnosis and treatment. To sum up, the results of the above preliminary tests indicate that automatic image processing techniques applied for ABCDE scoring in digital dermoscopy could be valuable assistive diagnostic tool for clinicians in the field.

Acknowledgement

This work is partly is partly supported by University of Piraeus Research Centre (UPRC).

References

[1] S. Zambanini, G. Langs, R. Sablatnig, H. Maier, Automatic Robust Registration of Cutaneous Hemangiomas for Follow-up Examinations, Proc. of 31st AAPR/OAGM Workshop, 2007.

[2] Optipix Video Dermatoscopy, Optilia, Sweden, www.optilia.eu,

[3] H. Kittler, M. Seltenheim, M. Dawid, H Pehamberger, K Wolff, M Binder Morphological changes of pigmented skin lesions: A useful extension of the ABCD rule for dermatology, J Am Acad Dermatol, vol. 40, pp. 558-562, 1999.

[4] Online: http://www.dermoscopy.org/consensus/2b.asp, last access: June 3rd, 2014.

[5] J. Nascimento, J. S. Marques, Adaptive Snakes Using the EM Algorithm, IEEE Trans. Image Processing, pp. 1678-1686, 2005.

[6] N. Otsu, A threshold selection method from gray-level histograms. IEEE Trans. Sys., Man., Cyber. 9 (1), pp. 62–66, 1979.

[7] C. Sun, D. Si, Fast Reflectional Symmetry Detection Using Orientation Histograms, Real-Time Imaging 5 (1), 63-74, 1999.

[8] E. Claridge, A. Orun, Modelling of edge profiles in pigmented skin lesions, Proc. of Medical Image Understanding and Analysis, pp. 53–56, 2002.

[9] F. Nachbar, W.Stolz, T. Merkle, A.B. Cognetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun-Falco, G. Plewig, The ABCD rule of dermatoscopy. High prospective value in the diagnosis of doubtful melanocytic skin lesions, Journal of the American Academy of Dermatology, 30 (4) pp. 551-559, 1994.

[10] J. F. Alcon, C. Ciuhu, W. ten Kate, A. Heinrich, N. Uzunbajakava, G. Krekels, D. Siem, G. de Haan, Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. IEEE Journal of Selected Topics in Signal Processing, 3, pp. 14–25, 2009.

[11] C. Grana, G. Pellacani, S. Seidenari, Practical color calibration for dermoscopy, applied to a digital epiluminescence microscope, Skin Research and Technology, vol. 11, pp. 242–247, 2005.

[12] C.N.E. Anagnostopoulos, D.D. Vergados, P. Mintzias, "Image registration of follow-up examinations in digital dermoscopy", Proc. of 13th IEEE International Conference on BioInformatics and BioEngineering, 2013.

[13] M. Fischler, R. Bolles,"Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Comm. of the ACM, vol. 24, pp. 381–395, 1981.

[14] M. Silveira, J.C. Nascimento, J.S. Marques, A.R.S. Marcal, T. Mendonca, S. Yamauchi, J. Maeda, J. Rozeira, "Comparison of segmentation methods for melanoma diagnosis in dermoscopy images", IEEE Journal of Selected Topics in Signal Processing, 3 (1) pp. 35-45, 2009.