Digital dermoscopy improves the accuracy of melanoma diagnosis. The aim of this study was to develop and validate software for assessment of asymmetry in melanocytic lesion images, based on evaluation of colour symmetry, and to compare it with assessment by human observers. An image analysis program enabling numerical assessment of asymmetry in melanocytic lesions, based on the evaluation and comparison of CIE L*a*b* colour components (CIE L*a*b* is the name of a colour space defined by the Commission Internationale de l’Eclairage) inside image colour blocks, was employed on the recorded lesion images. Clinical evaluation of asymmetry in dermoscopic images was performed on the same image set employing a 0–1 scoring system. Asymmetry judgement was expressed by the clinicians for 12.8% of benign naevi, 44.7% of atypical naevi and 64.2% of malignant melanomas, whereas the computer identified as asymmetric 6.3%, 33.3% and 82.2%, respectively. Numerical parameters referring to malignant melanomas were significantly higher, both with respect to benign naevi and atypical naevi. The numerical parameters produced could be effectively employed for computer-aided melanoma diagnosis.

Key words: Bhattacharyya distance; pattern analysis; multivariate Gaussian distribution.

(Accepted October 4, 2005.)


Costantino Grana, Department of Computer Engineering, University of Modena and Reggio Emilia, Via Vignolese 905/b, IT-41100 Modena, Italy. E-mail: grana.costantino@unimore.it

The incidence of malignant melanoma (MM) is increasing, and early diagnosis with surgical removal of thin lesions is the only curative treatment. Surface microscopy, which employs incident light magnification systems associated with the epiluminescence technique or polarized light, improves diagnostic accuracy for MM, especially for difficult-to-diagnose lesions (1–3). Moreover, semi-quantitative methods on dermoscopic images have been introduced to guide the clinician to diagnosis in a systematic way (4–10).

Asymmetry of melanocytic lesions is an important indicator of MM and contributes substantially to the diagnosis based on these semi-quantitative algorithms.

According to Stolz et al. (4) asymmetry is defined as the asymmetric distribution of dermoscopic structures, colours and shape with regard to two orthogonal mirror axes crossing at the gravity centre of the lesion (6, 7). These two axes have to be rotated so that the lowest possible degree of asymmetry is obtained. Asymmetry is then evaluated on two axes and scored 0–2, and the score is multiplied by 1.3. A modified ABC-point list of dermoscopy was recently introduced by Blum et al. (10). The outer shape of a melanocytic lesion was analysed if there was an asymmetry in one or two axes. In addition, the asymmetry of the structure inside the lesion in at least one axis was scored.

Reproducibility of clinical judgement on dermoscopic images is often unsatisfactory. Inter-rater agreement on asymmetry showed a kappa value of 0.41 in a consensus net meeting on dermoscopy via the Internet (11). Although a simultaneous consensus of several observers may establish a probably correct diagnosis, the individual judgement may depend on observers’ experience and subjectivity. Therefore, no totally objective clinical method and no golden standard for diagnosis of asymmetry exist.

To overcome unavoidable subjectivity and variability in the interpretation of dermoscopic images, programs for image analysis, enabling the numerical description of the morphology of pigmented skin lesion images, have been developed recently (12–23). Some of these have provided a reproducible quantification of lesion features and an aid to clinical diagnosis. Automated image assessment is based on a new mathematical semioLOGY, defining the geometry, texture, pigmentation and colours of the lesion.

The aim of this study was automatically to assess colour asymmetry in images referring to MM, atypical naevi (AN) and clearly benign naevi (BN), to evaluate the diagnostic capability of numerical parameters of asymmetry, and to compare the computer data with those provided by the evaluation of asymmetry performed on the same images by two clinicians.

MATERIALS AND METHODS

Image database

A total of 459 melanocytic lesion images, comprising 76 AN, 288 clearly BN and 95 MM, were studied. These images were...
subdivided into these three subgroups according to clinical and dermoscopic evaluation. Histopathological examination was performed on AN and MM, whereas only 30% of BN were excised and examined.

**Image acquisition system**

Images were acquired by means of a digital videomicroscope (VMS-110A, Scalar Mitsubishi, Tama-shi, Tokyo, Japan), with a 20-fold magnification enabling the whole lesion to be included in the monitor area. The instrument has been described elsewhere (16–18). The images were digitized by means of a Matrox Orion frame board and stored by an image acquisition program (VideoCap 8.09, DS-Medic a, Milan, Italy), which runs in Microsoft Windows. The digitized images offer a spatial resolution of 768×576 pixels and a resolution of 16 million colours. The camera system is calibrated monthly on a set of colour patches with known colour properties (Gretag Machbet® ColorChecker Chart) and the resulting colour profile is adjusted for the white test patch, between each patient examination, according to a well-defined procedure (24).

**Dermatologists’ evaluation of lesion asymmetry**

The images were evaluated by two clinicians employing the videomicroscopic technique on a regular basis. The examiners performed their evaluations simultaneously and a consensus was reached on each estimate after a thorough discussion. For each image the presence of structural and colour asymmetry along at least one axis was assessed and the agreed judgement was expressed by a 0–1 score. Reproducibility of the method of assessment was tested on a set of 50 images not considered in this study, and was found to be very high. Investigators’ evaluations were directly entered into the computer by means of an electronic form and were immediately ready for statistical elaboration.

**Image analysis program for asymmetry assessment**

After detection of the lesion border (22) and extraction of reference geometrical measures, such as centroid and main inertia axes, according to standard algorithms, lesion asymmetry was assessed by a three-step procedure. The first step consists of an image rotation along the direction specified by the major inertia axis, so that the major axis is aligned with the horizontal direction and the minor axis with the vertical one. The contour is also rotated accordingly. The second step consists of the elimination of colour details in the image and in its simplification into colour blocks. This is achieved by subdividing the image with a grid, aligned with the axis, after having selected the block sizes we are interested in (Fig. 1). Each colour block is described assuming that the colours are distributed as a multivariate Gaussian distribution, by collecting the mean CIE L* a* b* vector and its corresponding covariance matrix. Colour blocks are considered valid only if more than 25% of their area is constituted by lesion points, whereas they are excluded if less than 25% is comprised inside the lesion border (excluded blocks are shown in white in Fig. 1(d–f)).

As a third step, colour symmetry is computed, assessing the distance between each block and its symmetric one with respect to major and minor axis. The distance is computed by means of the Bhattacharyya distance (30), that is:

\[ b(a,b) = \frac{1}{8} (m_a - m_b)^T \Sigma_{\text{maj}}^{-1} (m_a - m_b) + \frac{1}{2} \log \left( \frac{\Sigma_{\text{maj}}}{\Sigma_{\text{maj}} + \Sigma_{\text{min}}} \right) \]

where \( m_a \) and \( m_b \) are the mean vectors and \( \Sigma_{\text{maj}} \) and \( \Sigma_{\text{min}} \) are the covariance matrices of the distributions of the two blocks. Blocks without a corresponding symmetric one are excluded from the analysis, to avoid including geometrical properties in the evaluation. The total distance obtained is averaged over the number of blocks that contributed to the resulting value.

This procedure was employed on images composed by 92×96, 50×48, 19×19 and 12×12 pixels per block, equivalent to a resizing of the image to 1%, 2%, 5% and 8%. This process does not produce a very high number of comparisons between blocks, since it only needs to make a comparison for each valid block. Following this procedure we obtained a set of two parameters, one for the minor axis (MIN) and the other for the major one (MAJ), for each block size, corresponding to eight measurements per lesion (MIN_1%, MAJ_1%, MIN_2%, MAJ_2%, MIN_5%, MAJ_5%, MIN_8% and MAJ_8%).

**Comparison between clinical evaluation and computer assessment**

Sensitivity and specificity, and asymmetry frequency for computer and clinicians were compared.

**Statistics**

For statistical analysis the SPSS statistical package (release 12.0, 2003; SPSS Inc., Chicago, IL, USA) was used. As basic statistics, mean and standard deviation of computer values, i.e. average distance along the major and minor axis, were calculated for MM, AN and BN. Differences between values referring to the three lesion groups were evaluated using the Mann-Whitney U test for independent samples.

To examine the discriminant power of our numerical parameters for differentiating between naevi and melamomas, discriminant analysis was performed on a training set comprising 50% of BN, AN and MM (229 lesions). Discriminant analysis enables the identification of variables, which are important for distinction among the groups and develops a procedure

Fig. 1. Image analysis process for the assessment of structural asymmetry in melanocytic lesions images. The digital image (a, c) is subdivided by a grid (b, f). The multivariate Gaussian distribution of each colour block is estimated (c, g), considering only pixels belonging to the lesion. Excluded blocks (when less than 25% is comprised inside the lesion border) are shown in white in d–h, while the colour shown is the mean colour of the distribution.
for group classification based on a score attribution. A linear combination of independent variables is formed and serves as a basis for assigning cases to groups. A D score, obtained for each lesion by the linear discriminant equation, is employed for the attribution of cases to groups. A receiver operating characteristic (ROC) analysis was performed to investigate sensitivity and specificity of the discriminant equation on classification of melanocytic lesions belonging to the test set, comprising the remaining 230 lesions. Diagnostic accuracy was estimated by the ratio between the percentage of the sum of true positives and true negatives, and the total number of lesions, and was calculated for each threshold (D) value. The area under the curve (AUC) and its 95% confidence interval (95% CI) were employed to estimate the likelihood of correctly classifying the lesions into benign and malignant.

To establish the threshold for classifying lesions into asymmetric or symmetric according to computer values, we selected the D score associated to the best diagnostic accuracy (melanoma/naevus) on the training set.

The frequencies of asymmetry in MM, AN and BN, both established by clinicians and the computer, were compared using the $\chi^2$ test. A $p$-value $< 0.01$ was considered significant. Concordance between clinical and computer evaluation was calculated employing the Spearman’s correlation coefficient, both considering single numerical parameters and overall assessment, whereas reproducibility of judgement was assessed by Cohen’s Kappa index. For a melanoma risk estimate, the odds ratio (OR) calculation was performed on the overall assessment of asymmetry.

RESULTS

Clinical evaluation

Asymmetry judgement was expressed by clinicians for 12.8% of BN, 44.7% of AN and 64.2% of MM (Table I). Differences between the three populations were significant. Based on asymmetry assessment alone, sensitivity for clinical judgement was 64.2%, whereas specificity was 80.5%.

Computer evaluation

Mean and standard deviation of the mathematical parameters calculated for BN, AN and MM are listed in Table II. Mean distance, variance and maximum distance were significantly higher in MM both with respect to BN and AN.

The parameters selected among numerical ones calculated by the computer by means of the discriminant equation for distinguishing between naevi and MM on the training set, were MAJ_2% and MIN_8%. On

| Table I | Percentage of asymmetric lesions according to clinicians and computer. For the computer case the threshold was selected on the training set |
|---------|-------------------------------------------------------------------------------------------------|-----------------|-----------------|-----------------|
| Clearly benign lesions (288)                  | Atypical nevi (76)                         | Melanomas (95)  |
| Clinicians | Computer | Clinicians | Computer | Clinicians | Computer |
| % asymmetry | 12.8   | 7.6       | 44.7$a$  | 28.9$a$   | 64.2$a,b$ | 89.5$a,b$ |

$a$Significant with respect to clearly benign naevi.

$b$Significant with respect to atypical naevi.

The frequencies of asymmetry in MM, AN and BN, both established by clinicians and the computer, were significantly higher in MM both with respect to clearly benign naevi ($p < 0.001$) and atypical naevi ($p < 0.001$). Both the Kappa value and Spearman’s correlation coefficient remained significant.

Computer asymmetry judgement corresponded to an OR of 34.133 for MM.

Correlation between clinical and computer evaluation

Eighty-one percent of symmetric/asymmetric evaluations were concordant between clinician and computer. Both the Kappa value and Spearman’s correlation coefficient were 0.530.

DISCUSSION

Asymmetry is considered a crucial clue for MM identification. In the dermoscopic ABCD rule for melanocytic lesions, the minimum final dermoscopic score indicating possible melanoma corresponds to 4.75 and the minimum score for a melanocytic lesion is 1 (i.e. at least one colour and one dermoscopic structure are mandatory) (5). Thus, assessment of asymmetry of microscopic features (including asymmetric distribution of differential structures and colours) contributes by one-third (one axis-asymmetry = 1 × 1.3) or two-thirds (two axes-asymmetry = 2 × 1.3) to MM diagnosis. A modified ABCD rule for melanoma diagnosis was recently introduced by Blum et al. (10) The outer shape of a
melanocytic lesion is analysed if there is an asymmetry in one or two axes. In addition, the asymmetry of the structure inside the lesion in at least one axis is scored. Simple points are given for asymmetry of outer shape and for asymmetry of differential structures inside the lesion in at least one axis, abrupt cut-off of network at the border in at least one-quarter of the circumference, three or more colours, three or more differential structures, or visible change in the last 3 months. With this method structural asymmetry contributes by 25% to the overall score for a lesion to be considered a MM. Asymmetry of the outer shape in 1 or 2 axes was found in 77% of benign lesions and in 98% of MM, whereas asymmetry inside the lesion was found in 63% of benign lesions and in 96% of MM (10).

Lorentzen et al. (25) assessed the sensitivity of axis asymmetry using latent class analysis, in order to minimize observer dependence. They analysed ratings from four experts in dermatoscopy of 232 pigmented lesions. Possible ratings were no asymmetry, asymmetry on one axis and asymmetry on two axes. No asymmetry sensitivity was 40–77%, one-axis asymmetry 40–70%, and two-axes asymmetry 77–92%. Melanomas proved significantly more asymmetric than melanocytic naevi.

Asymmetry of pigment distribution was also employed for the classification of AN by Hoffmann-Wellnhofer et al. (26). Central and peripheral hyper- or hypo-pigmentation, arranged in a localized and multifocal distribution, were considered. The authors drew the attention to the peripherally hyper-pigmented type, mimicking an \textit{in situ} MM.

Asymmetry can also be assessed by mirroring dermatoscopic structures in the centre of gravity of the lesion, as suggested by Menzies et al. (6) who incorporated the presence of point and plane symmetry of an indicator of non-MM in their diagnostic algorithm.

For a parameter to be diagnostic value a sufficiently high reproducibility rate is necessary: a high degree of observer disagreement indicates uncertainty. In the consensus net meeting on dermoscopy inter-observer agreement showed a kappa value of 0.41, which is barely satisfactory (11). Since for biological structures perfect symmetry does not exist, when assessing lesion images by dermoscopy, each observer has to decide if the personal threshold for asymmetry has been reached, in order to classify a lesion symmetric or asymmetric. This represents a highly subjective process, exposed to changeable bias. Moreover, as for other diagnostic parameters, the definition of symmetry cannot be precise enough to be unequivocally interpreted by different observers, and no fully reproducible threshold can be set, even if it represents the result of a consensus. In order to overcome subjectivity and variability in the interpretation of dermoscopic images, several image analysis programs have been recently introduced as a possible support to clinical diagnosis (12–23). The emphasis has been put on assessment of lesion size, shape, colour and texture, which are expressed by mathematical parameters. For asymmetry assessment, both the shape and structure of the lesion were considered. Andreassi et al. (19) employed both circularity (defined as the percentage of lesion not overlapping a circle of equal area), and imbalance of dark areas. We used the polar moments of inertia, calculated as the distance of dark areas within the lesion from the barycentre, and showed that these measures have discriminant power (16, 18). Recently, we compared two different methods for the description of distribution of dark areas in dermoscopic melanocytic lesion images (23). Significant differences in dark area distribution between MMs and naevi were observed employing both methods, permitting a good discrimination of melanocytic lesions, with a diagnostic accuracy ranging from 71% to 75%.

The method for asymmetry assessment described in this study subdivides the image into homogeneous blocks, and translates structural differences into colour differences. Details such as globules, dots, regression structures, areas of increased vascularization, or grey-blue areas, etc. have not been considered by image analysis programs so far, whereas for assessment of the network, methods still under evaluation have been proposed (27). According to our procedure, differential structures contribute to the colour values within the image block in which they are included. The distribution of these structures is then assessed by comparing the colour distributions pertaining to different blocks. A large colour distance between blocks signifies that lesion structure is non-homogeneous and architecture is complex. This corresponds to structural asymmetry as assessed by the clinician when employing semi-quantitative methods.

In fact, significant differences between parameter values were observed for different melanocytic lesion populations. In MM, distances between colour distributions referring to different image blocks were significantly higher than in naevi, indicating colour

Table II. Computer parameters in clearly benign naevi (BN), atypical naevi (AN) and malignant melanomas (MM)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>BN (288)</th>
<th>AN (76)</th>
<th>MM (95)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>MAJ_1%</td>
<td>0.490</td>
<td>0.187</td>
<td>0.591a</td>
</tr>
<tr>
<td>MIN_1%</td>
<td>0.581</td>
<td>0.278</td>
<td>0.664a</td>
</tr>
<tr>
<td>MAJ_2%</td>
<td>0.774</td>
<td>0.222</td>
<td>0.941a</td>
</tr>
<tr>
<td>MIN_2%</td>
<td>0.861</td>
<td>0.296</td>
<td>1.024a</td>
</tr>
<tr>
<td>MAJ_5%</td>
<td>1.424</td>
<td>0.335</td>
<td>1.718a</td>
</tr>
<tr>
<td>MIN_5%</td>
<td>1.537</td>
<td>0.429</td>
<td>1.848a</td>
</tr>
<tr>
<td>MAJ_8%</td>
<td>1.904</td>
<td>0.413</td>
<td>2.282a</td>
</tr>
<tr>
<td>MIN_8%</td>
<td>2.031</td>
<td>0.518</td>
<td>2.429a</td>
</tr>
</tbody>
</table>

SD, standard deviation; MIN, parameters for the minor axis; MAJ, parameters for the major axis.

aSignificant with respect to clearly benign naevi.

bSignificant with respect to atypical naevi.

Acta Derm Venereol 86
variegation and complex architecture. Moreover, AN were characterized by intermediate values between those referring to MM and those belonging to naevi. Diagnostic accuracy was higher for overall computer assessment than for clinical judgement of asymmetry. Moreover, according to computer evaluation, more MM and fewer naevi were asymmetric. However, the automated asymmetry assessment alone is not sufficient for the final diagnostic judgement, that has to be based on numerous clinical and dermoscopic descriptors.

Pattern analysis represents the basis for dermoscopic diagnosis of MM (8, 11), however, in early MM differential structures which are diagnostic for MM may be lacking, and diagnostic structures are seldom identified as an early marker of malignant transformation (28, 29). A total of 1862 sequentially recorded dermoscopic images of melanocytic lesions were followed-up for 12 months by Kittler et al. (28). Seventy-five of these showed substantial modifications over time and were excised. Eight changing lesions were histologically diagnosed as early MM, but substantial structural dermoscopic modifications were observed in only 50% of these. Menzies et al. (29) performed digital follow-up of 318 moderately or mildly atypical melanocytic lesions with a history of change, which were monitored during an average period of 3 months. Of the 61 lesions that showed morphological changes, 7 were found to be early MM. None of these MM developed any classic surface microscopic feature of MM, and therefore could only be identified by morphologic change. Whereas a change in size or shape was present in 5 out of 7 lesions and a colour change in 4 out of 7, an architectural change was observed in all lesions.

Based on these observations we conclude that architectural modifications represent the earliest signs of malignant change, and that automated asymmetry assessment could be particularly useful when performing follow-up of AN, as a complement to digital monitoring and dermoscopic observation.

ACKNOWLEDGEMENT

This study was partially supported by Ministero dell’Istruzione dell’Università e della Ricerca (MIUR) grant number 2001068929.

REFERENCES


