

Combination of features from skin pattern and ABCD analysis for lesion classification

Zhishun She^{1,2}, Y. Liu¹ and A. Damatoa¹

¹*Faculty of Technology and Computer Science, NEWI, University of Wales, Wrexham, UK,*

²*School of Electronic & Information Engineering, Southwest University, China*

Background/Purpose: It is known that the standard features for lesion classification are ABCD features, that is, asymmetry, border irregularity, colour variegation and diameter of lesion. However, the observation that skin patterning tends to be disrupted by malignant but not by benign skin lesions suggests that measurements of skin pattern disruption on simply captured white light optical skin images could be a useful contribution to a diagnostic feature set. Previous work using both skin line direction and intensity for lesion classification was encouraging. But these features have not been combined with the ABCD features. This paper explores the possibility of combining features from skin pattern and ABCD analysis to enhance classification performance.

Methods: The skin line direction and intensity were extracted from a local tensor matrix of skin pattern. Meanwhile, ABCD analysis was conducted to generate six features. They were asymmetry, border irregularity, colour (red, green and blue) variegations and diameter of lesion. The eight features of each case were combined using a

principal component analysis (PCA) to produce two dominant features for lesion classification.

Results: A larger set of images containing malignant melanoma (MM) and benign naevi were processed as above and the scatter plot in a two-dimensional dominant feature space showed excellent separation of benign and malignant lesions. An ROC (receiver operating characteristic) plot enclosed an area of 0.94.

Conclusions: The classification results showed that the individual features have a limited discrimination capability and the combined features were promising to distinguish MM from benign lesion.

Key words: skin pattern – ABCD analysis – lesion classification – feature combination – melanoma

© Blackwell Munksgaard, 2007

Accepted for publication 3 January 2006

AS THE survival rate of malignant melanoma (MM) depends on its thickness, diagnosis of MM at an early stage could reduce the risk of mortality and increase the chance of prognosis considerably. In order to achieve this, a computer automatic diagnosis (CAD) system used as a diagnostic aid is required which is of great assistance to primary clinical care.

In the CAD system a feature set is utilised to distinguish between MM and benign lesions. Conventional diagnostic features are related to asymmetry, boundary irregularity, colour and diameter of skin lesion, which is called ABCD analysis (1). However, most areas of the human skin surface are covered with a network of skin lines (glyphic pattern) and this skin pattern is clearly disrupted when an MM disturbs the structure of the dermis (2). This suggests that a measure of skin pattern disruption can be used as part of a feature set to distinguish malignant from

benign skin lesions. In a previously published procedure the skin pattern was extracted from normal white light clinical (WLC) images by high-pass filtering (3) and the skin line direction and skin line intensity were used for lesion classification by processing a small image set (4, 5). Nevertheless these new features have not been combined with the ABCD features yet and the number of clinical cases to be investigated needs to be increased. All the features are required to be combined to enhance classification performance.

The work described in this paper is to combine the features from skin pattern and ABCD analysis for lesion classification using a larger image set. This should eventually result in increased lesion discrimination compared with the use of each feature alone. The paper is organised as follows: in 'skin pattern analysis' two features extracted from skin pattern are discussed briefly with examples, computational algorithms to calculate

the ABCD features are described in 'ABCD analysis', 'lesion classification' conducts the classifications using individual and combined features and finally the conclusions are drawn.

Skin Pattern Analysis

A skin pattern image is extracted from an original WLC image by high-pass filtering (3). Then the gradient vector $\nabla G(m, n) = [G_x(m, n), G_y(m, n)]^T$ is calculated, where $G(m, n)$ is the skin pattern image, $G_x(m, n)$ and $G_y(m, n)$ are the gradients in the horizontal and vertical directions. Finally, a tensor $J(m, n)$ was introduced to find the skin line structure, namely,

$$J(m, n) = \nabla G(m, n) \nabla G^T(m, n) \quad (1)$$

This tensor is estimated by averaging over a local sub-image centred at pixel (i, j) with a size of $M \times N$ pixels, which is mathematically expressed as

$$\bar{J}(i, j) = \frac{1}{MN} \sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} J(i+m, j+n) \quad (2)$$

$\bar{J}(i, j)$ is a 2×2 matrix, that is,

$$\bar{J}(i, j) = \begin{bmatrix} f_{xx}(i, j) & f_{xy}(i, j) \\ f_{xy}(i, j) & f_{yy}(i, j) \end{bmatrix} \quad (3)$$

where

$$f_{xx}(i, j) = \frac{1}{MN} \sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} G_x^2(i+m, j+n),$$

$$f_{yy}(i, j) = \frac{1}{MN} \sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} G_y^2(i+m, j+n)$$

and

$$f_{xy}(i, j) = \frac{1}{MN} \sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} G_x(i+m, j+n) G_y(i+m, j+n).$$

Skin line direction

Skin line direction is determined by the eigenvector corresponding to the minimum eigenvalue of the tensor matrix $\bar{J}(i, j)$ in Eq. (3). The angle of this eigenvector is

$$\phi(i, j) = \frac{1}{2} \tan^{-1} \left[\frac{2f_{xy}(i, j)}{f_{xx}(i, j) - f_{yy}(i, j)} \right] \quad (4)$$

In order to reduce the interference of noise, the skin line direction is averaged over a local area.

Then a snake-based edge detection technique is used to determine the lesion boundary (6). The detected boundary segments the image into skin area A_s and lesion area A_l . Finally, the circular means rather than the arithmetic averages of skin line directions in skin and lesion areas are calculated as the skin line directions are the directional data (7). The absolute difference of skin line direction between skin and lesion areas is used for lesion classification.

As an example the optical image of a benign naevus is shown in Fig. 1a and its skin pattern image is displayed in Fig. 1b. Skin line direction of skin pattern in Fig. 1b was calculated in a local area with 16×16 pixels. The final image of smoothed skin line direction for the skin pattern in Fig. 1b is shown in Fig. 1c.

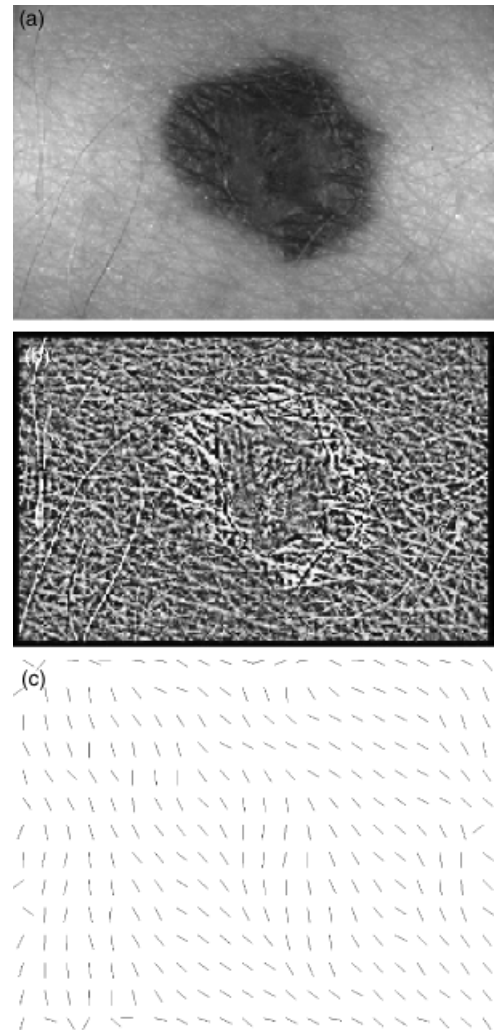


Fig.1. A benign naevus: (a) optical image, (b) skin pattern and (c) skin line directions.

Skin line intensity

Although the skin line direction is determined by the eigenvector corresponding to the minimum eigenvalue of $\bar{J}(i, j)$, the intensity of skin line in the skin line direction is determined by this minimum eigenvalue, that is,

$$I(i, j) = \frac{f_{xx}(i, j) + f_{yy}(i, j) - \sqrt{[f_{xx}(i, j) - f_{yy}(i, j)]^2 + 4f_{xy}^2(i, j)}}{2} \quad (5)$$

Then the arithmetic means of skin line intensity in skin and lesion areas are computed. To calibrate the brightness of individual clinical images, the absolute difference of skin line intensity between skin and lesion areas normalised by the mean of skin line intensity in skin area is used for lesion classification.

For example, the optical image of a MM is shown in Fig. 2a and its corresponding skin pattern is given in Fig. 2b. Skin line intensity of

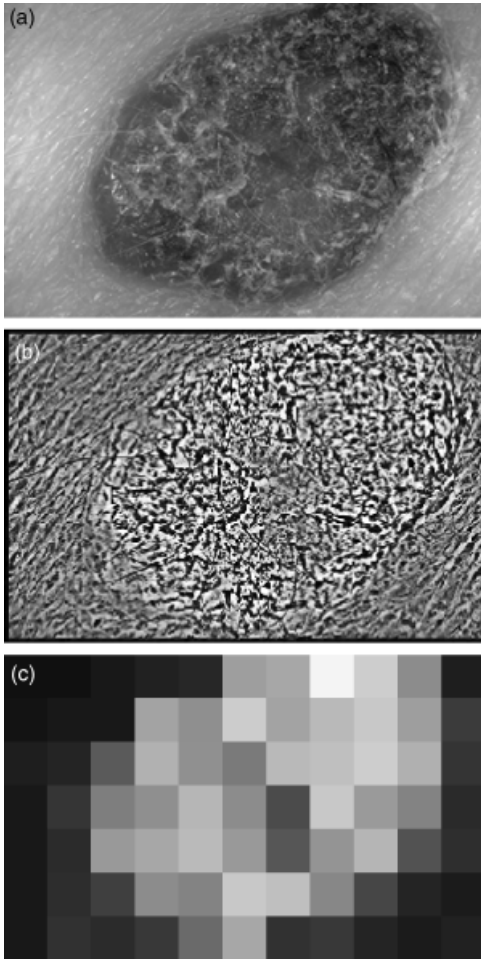


Fig. 2. A malignant melanoma: (a) optical image, (b) skin pattern and (c) skin line intensities.

skin pattern in Fig. 2b was calculated in a local area with 16×16 pixels. The skin line intensity image for the skin pattern in Fig. 2b is shown in Fig. 2c where the grey blocks represent the intensities of each local patch.

ABCD Analysis

Before ABCD features are calculated, shape analysis of skin lesion is conducted. The lesion area A_l is detected by a snake-based algorithm (6). Then a binary image, $A(k, l)$ is generated which satisfies $A(k, l) = 1$ if $(k, l) \in A_l$ and $A(k, l) = 0$ elsewhere. Shape analysis includes three steps to process $A(k, l)$. First, the central position of lesion is located with co-ordinates

$$\bar{k} = \frac{1}{T} \sum_{k=1}^K \sum_{l=1}^L k A(k, l) \quad (6)$$

$$\bar{l} = \frac{1}{T} \sum_{k=1}^K \sum_{l=1}^L l A(k, l) \quad (7)$$

where K and L are the number of image pixels in the horizontal and vertical directions, respectively, and T is the total number of pixels in lesion area, namely,

$$T = \sum_{k=1}^K \sum_{l=1}^L A(k, l) \quad (8)$$

Then the orientation of lesion shape is determined by

$$\theta = \frac{1}{2} \tan^{-1} \left[\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}} \right] \quad (9)$$

where the $\mu_{p,q}$ is the (p, q) order central moment which is defined as

$$\mu_{p,q} = \sum_{k=1}^K \sum_{l=1}^L (k - \bar{k})^p (l - \bar{l})^q A(k, l) \quad (10)$$

Finally, the best-fit ellipse for the skin lesion is found. Semimajor axis a and semiminor axis b are determined by

$$a = \left(\frac{4}{\pi} \right)^{1/4} \left[\frac{(I'_{\max})^3}{I'_{\min}} \right]^{1/8} \quad (11)$$

$$b = \left(\frac{4}{\pi} \right)^{1/4} \left[\frac{(I'_{\max})^3}{I'_{\min}} \right]^{1/8} \quad (12)$$

where I'_{\max} and I'_{\min} are the maximum and the minimum moments in the orientation of θ . They

are calculated by

$$I'_{\min} = \sum_{k=1}^K \sum_{l=1}^L [(k - \bar{k})\cos\theta - (l - \bar{l})\sin\theta]^2 A(k, l) \quad (13)$$

$$I'_{\max} = \sum_{k=1}^K \sum_{l=1}^L [(k - \bar{k})\cos\theta + (l - \bar{l})\sin\theta]^2 A(k, l) \quad (14)$$

As an example the optical image of a benign naevus is shown in Fig. 3a and its inverted segmented binary image is displayed in Fig. 3b with its lesion centre (the crossing point of the major and minor axes), orientation (direction of the major axis) and the best-fit ellipse.

Asymmetry of lesion shape

Asymmetry is the behaviour of lesion shape about the major axis. One way to measure this is to fold the lesion outline about the major axis of the best-fit ellipse, find the difference (the non-overlapping region) and calculate the percentage

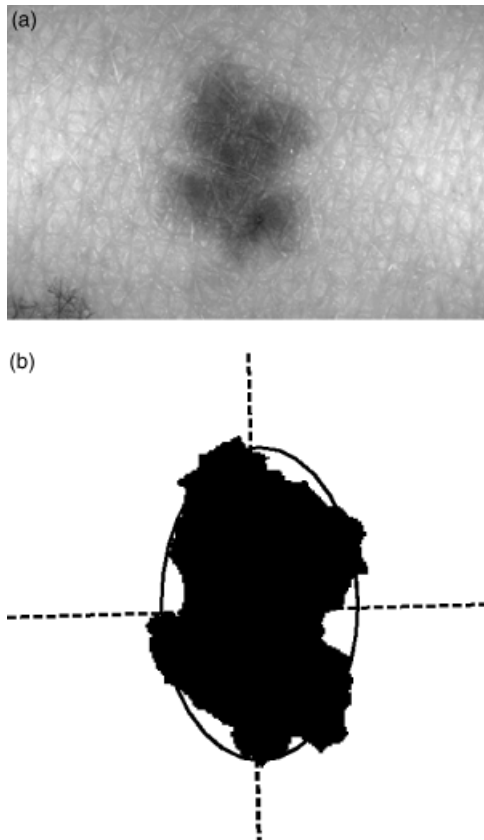


Fig. 3. Lesion shape analysis: (a) optical image and (b) inverted lesion shape with its centre, orientation and the best-fit ellipse.

of this difference over the area of lesion, that is,

$$AS = \frac{\Delta T}{T} \times 100\% \quad (15)$$

where ΔT is the number of pixels in the difference area.

For example, the shape of lesion in Fig. 4a is rotated as shown in Fig. 4b so that the major axis is in the horizontal direction. Then top part is folded in the vertical direction. The folded part and the underneath part are displayed in Fig. 4c and d, respectively. The difference between the folded part and the underneath part is shown in Fig. 4e.

Border irregularity

Border irregularity is measured by the ratio of the square of the perimeter of lesion to the area of lesion. It is computed by

$$B = \frac{P^2}{4\pi T} \quad (16)$$

where P is the perimeter of lesion boundary and T is the lesion area. Border irregularity has the minimum value of one for a circle, the most regular shape.

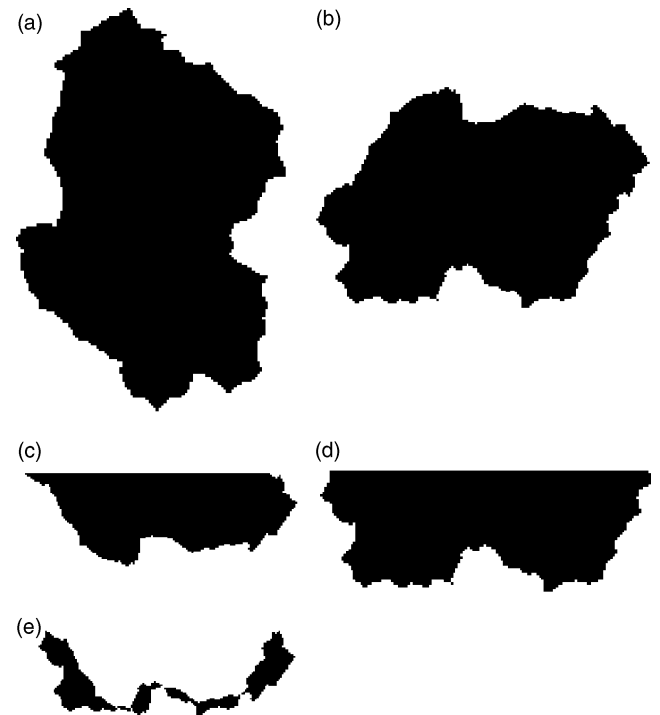


Fig. 4. Asymmetrical measurement: (a) inverted segmented image, (b) rotated inverted segmented image, (c) folded top part, (d) underneath part and (e) difference between (c) and (d).

Colour variegation

Colour variegation is quantified by the normalised standard deviation of red, green and blue components of lesion, respectively. They are expressed by

$$C_r = \frac{\sigma_r}{M_r} \quad (17)$$

$$C_g = \frac{\sigma_g}{M_g} \quad (18)$$

$$C_b = \frac{\sigma_b}{M_b} \quad (19)$$

where σ_r , σ_g and σ_b are the standard deviations of red, green and blue components of lesion area and M_r , M_g and M_b are the maximum values of red, green and blue components in lesion region. The normalisation process is important because objects with higher normal skin pigmentation would also tend to show higher pigmentation

in the lesion itself and this needs to be accounted for in order to make any comparison between cases.

Diameter of lesion

Diameter of lesion is calculated by

$$D = 2a \quad (20)$$

where a is the semi-major axis of the best-fit ellipse. The scale is converted from pixels into millimetre (mm) by knowledge of image pixel parameters and spatial relation at a particular magnification.

Lesion classification

The image set used in the experiment of this technique has been extended considerably and

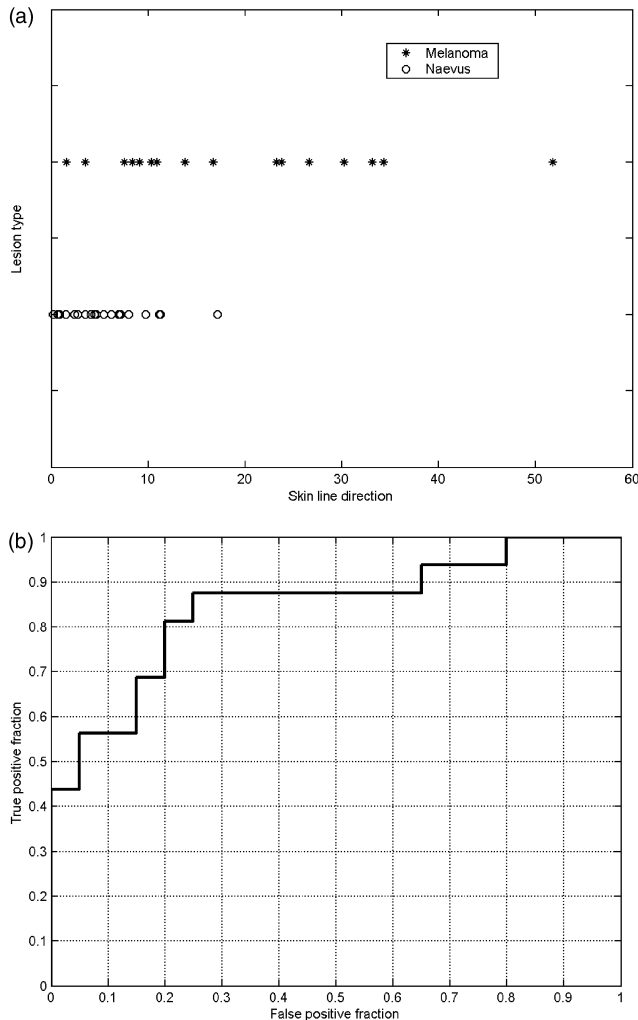


Fig. 5. Lesion classification using skin line direction: (a) scatter plot and (b) ROC curve.

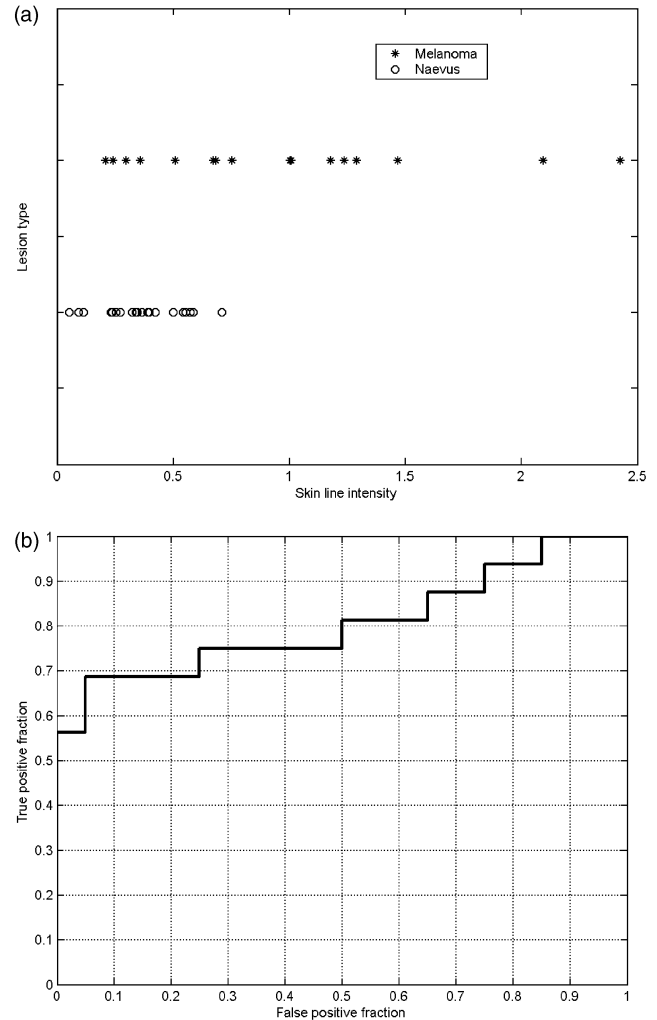


Fig. 6. Lesion classification using skin line intensity: (a) scatter plot and (b) ROC curve.

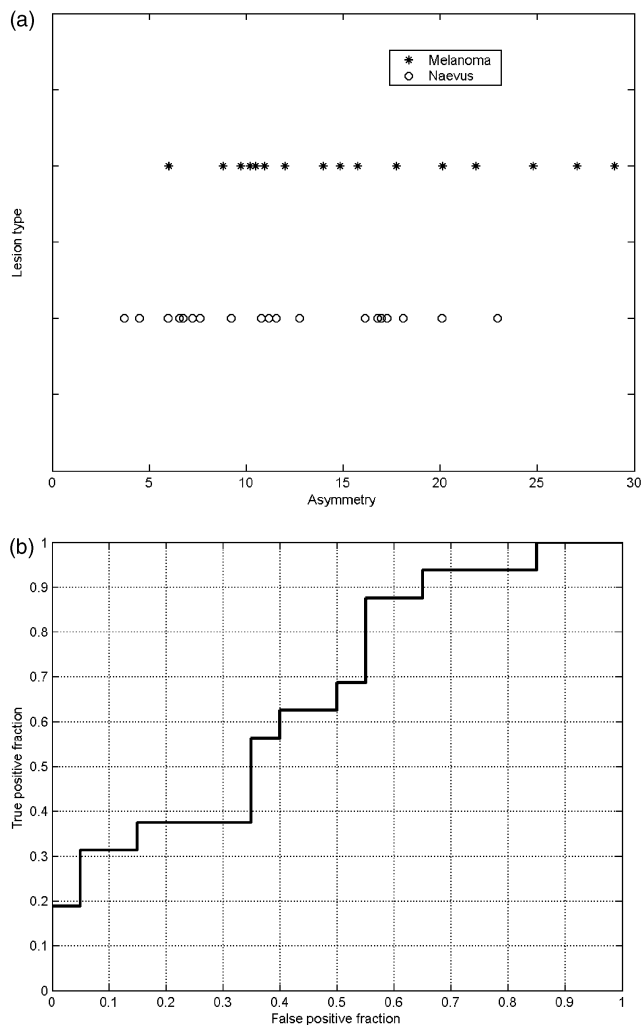


Fig. 7. Lesion classification using asymmetry: (a) scatter plot and (b) ROC curve.

contains examples of several types of lesion including 16 melanomas and 20 compound or junctional naevi. The total number of cases to be acquired is limited by recruitment of patient volunteers in clinical practice. The original images are 24-bit full colour digital format. The size of image is $230 \times 350 \times 3$ pixels. It is chosen to have a reasonable skin area surrounding the lesions. The colour images are converted into grey-level intensity for skin pattern analysis.

For lesion classification, it is important to determine the sample size. It needs to satisfy $Q > 3cp$ (8), where Q is the sample size, c the number of classes and P the number of features. In this study we have two classes (benign naevi and MM) and thus one or two features will be used for lesion classification in order to meet this requirement.

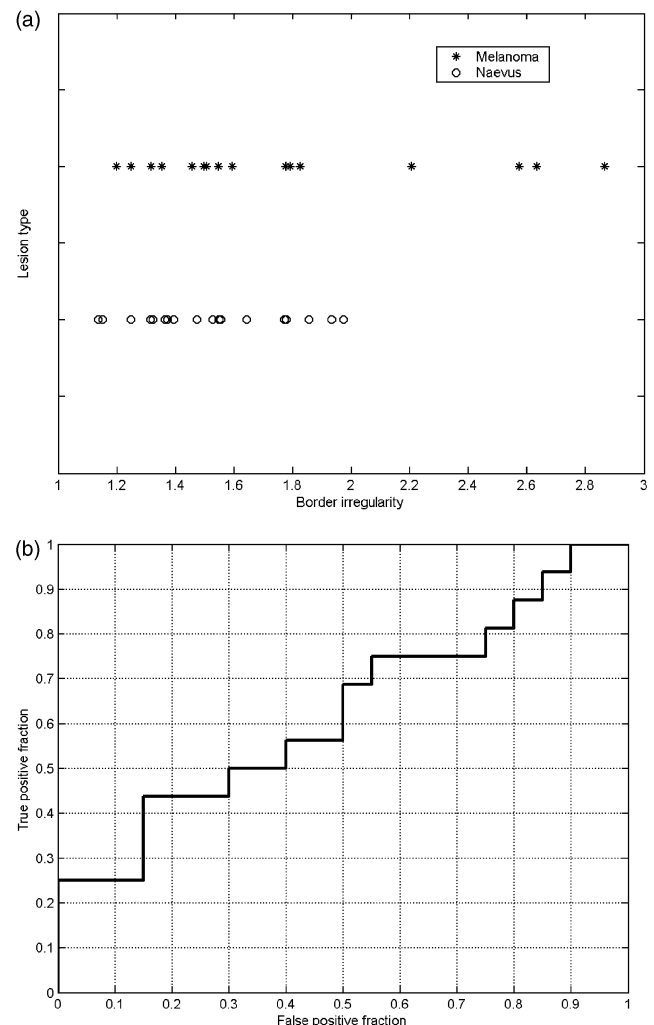


Fig. 8. Lesion classification using border irregularity: (a) scatter plot and (b) ROC curve.

Classification using individual features

Classification using skin line direction

The means of skin line direction for skin and lesion areas and their difference were calculated and the scatter plot of the line direction difference is shown in Fig. 5a, showing a degree of separation between the benign and malignant lesions. Figure 5b shows the receiver operating characteristic (ROC) curve and the area under the curve is approximately 0.84.

Classification using skin line intensity

The means of skin line intensity for skin and lesion areas and their normalised difference were computed and the distribution of the normalised difference is shown in Fig. 6a. It indicates a tendency to a greater skin line intensity deviation in the MM images compared with that in the benign lesion images. Figure 6b shows the ROC

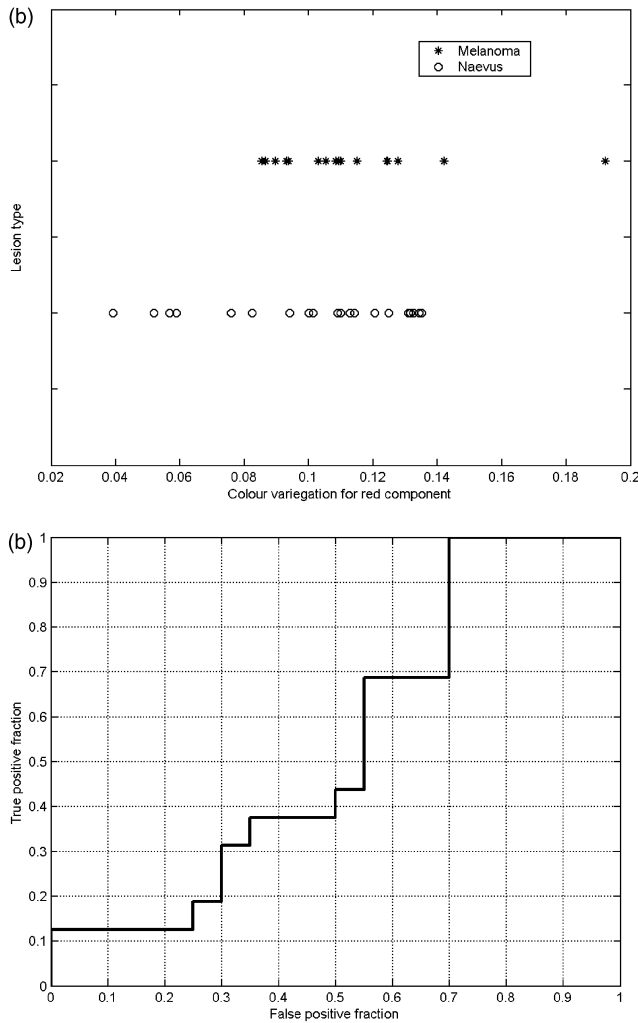


Fig. 9. Lesion classification using red component: (a) scatter plot and (b) ROC curve.

curve and the area underneath the ROC curve is about 0.80.

Classification using asymmetry

Figure 7a is the scatter plot of asymmetry of skin lesion. Its corresponding ROC curve is shown in Fig. 7b. The area under the curve is approximately 0.66, showing some separation between MM and benign lesion.

Classification using border irregularity

The distribution of border irregularity is given in Fig. 8a. Figure 8b is the ROC curve where the area under the curve is approximately 0.62 which indicates a modest separation between MM and benign lesion.

Classification using colour variegation

Classification using red component. The scatter plot for colour variegation of red component and

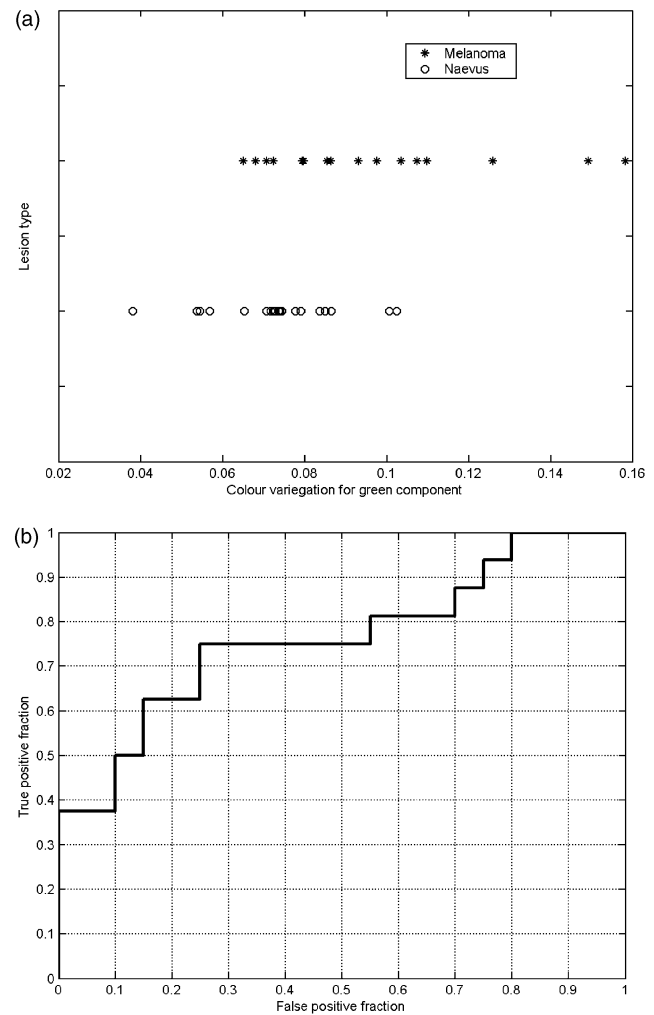


Fig. 10. Lesion classification using green component: (a) scatter plot and (b) ROC curve.

its corresponding ROC plot are shown in Fig. 9a and b, respectively. The area under the ROC plot is approximately 0.54 which presents an average separation of the two classes.

Classification using green component. Figure 10a is the distribution of colour variegation for green component. Its corresponding ROC curve is shown in Fig. 10b. The area under the curve is approximately 0.76, showing a potential separation between MM and benign lesion.

Classification using blue component. The scatter plot for colour variegation of blue component is shown in Fig. 11a. Figure 11b is the ROC curve where the area under the curve is approximately 0.78 which indicates a positive separation between the two classes.

Classification using diameter

The distribution for diameter of skin lesion and its corresponding ROC plot are shown in Fig. 12a

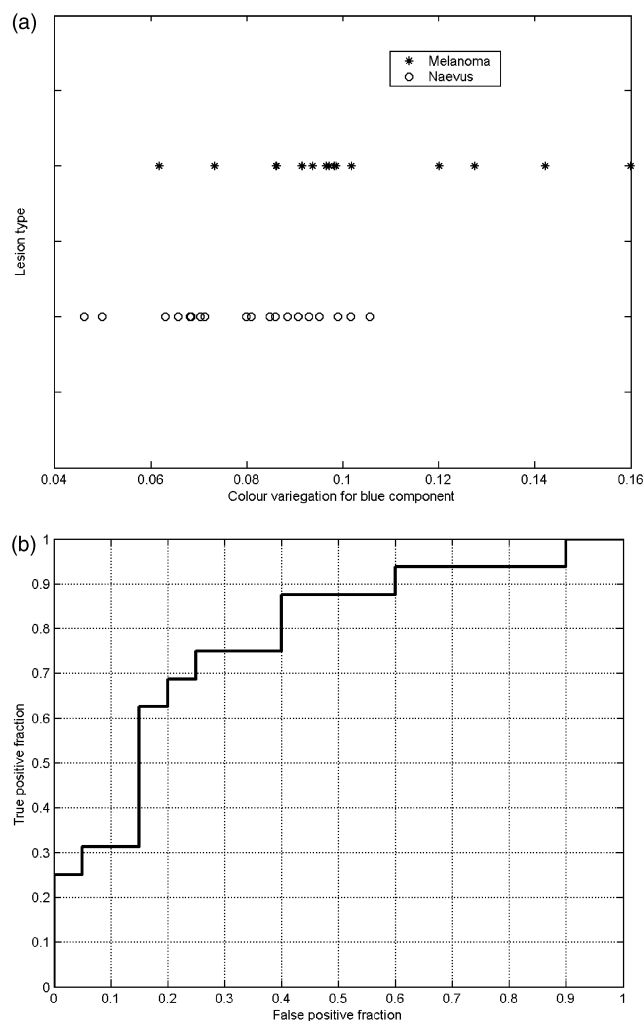


Fig. 11. Lesion classification using blue component: (a) scatter plot and (b) ROC curve.

and b, respectively. The area under the ROC plot is approximately 0.62, indicating an average separation between MM and benign lesion.

Classification using combined features

Features extracted from skin pattern are combined with those from ABCD analysis to enhance the classification accuracy. The principal component analysis (PCA) is used for this combination (9). It includes four steps: (1) for each skin lesion a feature vector is formed by the eight features, that is skin line direction, skin line intensity, asymmetry, border irregularity, red component variegation, green component variegation, blue component variegation and diameter of lesion; (2) a covariance matrix of feature vector is computed by sample averaging; (3) eigenvalues and eigenvectors of the covariance matrix are calculated by singular value analysis (10) and a trans-

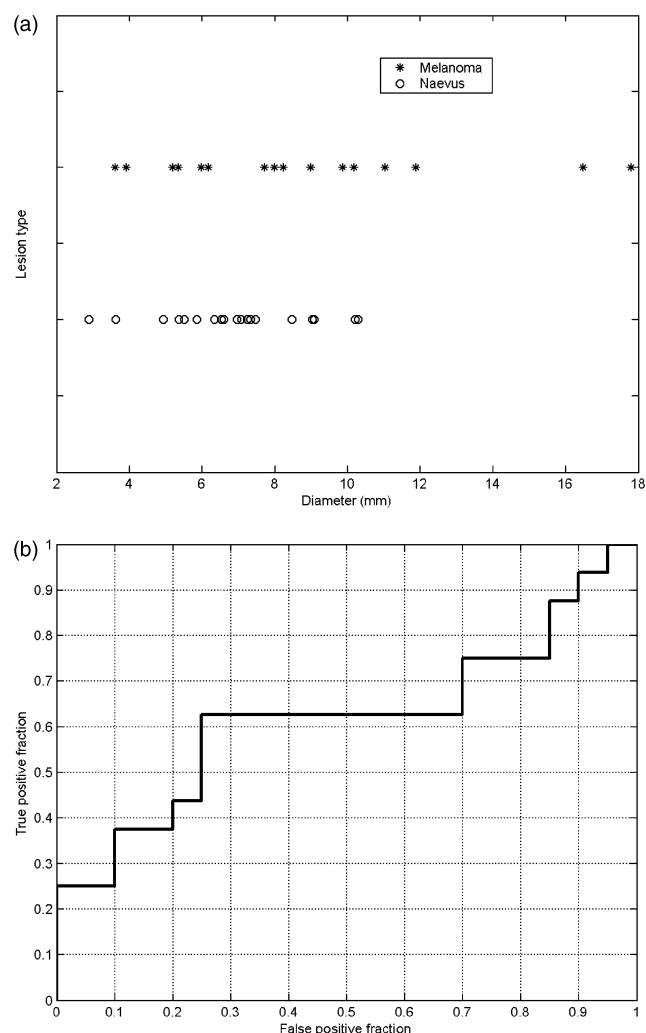


Fig. 12. Lesion classification using diameter of lesion: (a) scatter plot and (b) ROC curve.

form matrix is constructed by the eigenvectors corresponding to the first and second dominant eigenvalues of the covariance matrix; and (4) the dimension of feature space is reduced from eight to two by multiplying the transform matrix with each feature vector. This transformation results in two dominant features for each case. These two dominant features are uncorrelated and are the linear combinations of the eight features extracted from skin pattern and ABCD analysis. These two dominant features are chosen for lesion classification.

The scatter plot of the 36 skin lesions on the two-dimensional dominant feature space is given in Fig. 13a, which demonstrates that malignant lesions usually have greater disturbances in the space of the two principal features and thus they can be discriminated from benign lesions. An ROC curve was plotted using a straight line with a slope of 5.555 (a linear classifier) to segment the

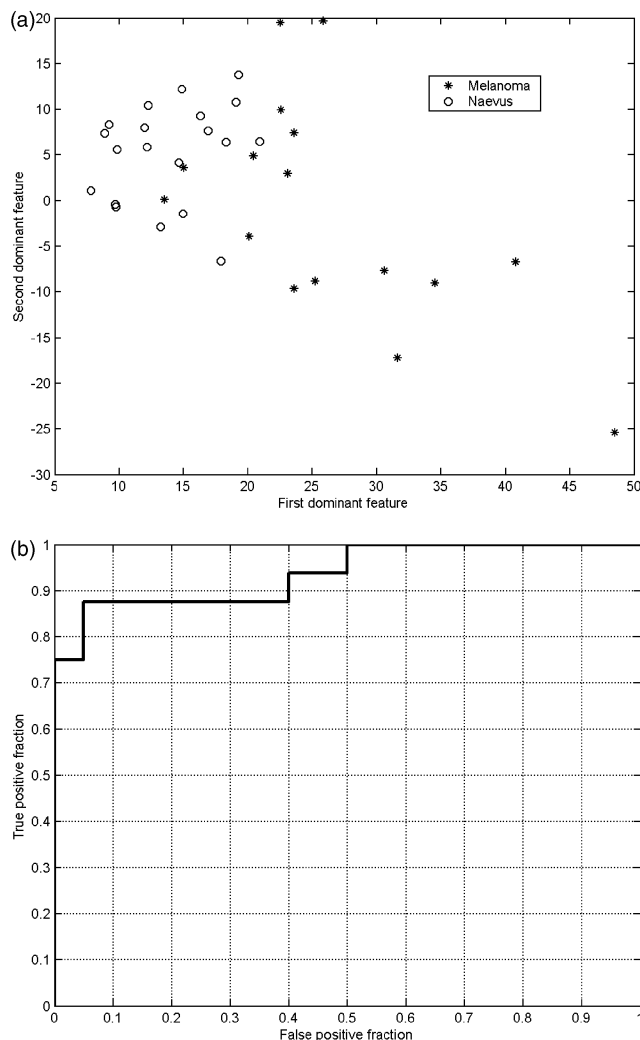


Fig. 13. Lesion classification using combined features: (a) scatter plot and (b) ROC curve.

scatter plot into classified-as-melanoma (below and to the right of the line) and classified-as-benign (above and to the left of the line) and varying the distance of the line from the origin. Figure 13b shows the ROC curve where the area under the curve is approximately 0.94, indicating an excellent classification result.

Conclusions

Two complementary sets of features have been successfully combined to increase diagnostic accuracy. One feature set is extracted from skin pattern, which is characterised by the skin line direction and the skin line intensity. The other feature set is acquired by ABCD analysis of lesion, which produces six features. They are asymmetry, border irregularity, colour (red, green and blue) variegations and diameter of lesion.

Classification results indicate that the individual features have limited discrimination ability, and combination of features from skin pattern and ABCD analysis enhances the classification results significantly, suggesting that combining features from skin pattern and ABCD analysis is promising in the design of CAD system, which could eventually be used as a diagnostic aid at the primary care level.

Acknowledgements

The optical skin lesion images were provided by M. Dickson, V. Wallace and Dr J. Bamber of the Physics Department, Clinical Research Centre, Royal Marsden Hospital, Sutton. Permission to use them is gratefully acknowledged. This work was funded by EPSRC grants GR/M72371 and GR/M72289. Recently it has been supported by natural science foundation project of CQ CSTC.

References

1. Mackie RM. An illustrated guide to the recognition of early malignant melanoma. Glasgow: Blackwood Pillans & Wilson Ltd., Edinburgh and Department of Dermatology, University of Glasgow, 1986.
2. Hall PN. Clinical diagnosis of melanoma. In: Diagnosis and management of melanoma in clinical practice. New York: Springer-Verlag, 1992: 35–52.
3. Round AJ, Duller AWG, Fish PJ. Lesion classification using skin patterning. *Skin Res Technol* 2000; 6: 183–192.
4. She Z, Fish PJ. Analysis of skin line pattern for lesion classification. *Skin Res Technol* 2003; 9: 73–80.
5. Liu Y, She Z. Skin pattern analysis for lesion classification using skin line intensity. *Proceedings of Medical Image Understanding and Analysis*, July 2005: 207–210.
6. She Z, Fish PJ. Boundary detection of skin lesion using a fast snake algorithm. *Proceedings of 16th Biennial International EURASIP Conference*, June 2002: 295–297.
7. Mardia KV. Statistics of directional data. London: Academic Press, 1972.
8. Foley DH. Consideration of sample and feature size. *IEEE Transactions on Information Theory* 1972; 18: 618–626.
9. Theodoridis S, Koutroumbas K. Pattern recognition. Boston: Academic Press, 2003.
10. Golub GH, Loan CFV. Matrix computations. Baltimore, MD: Johns Hopkins University Press, 1989.
11. Jain AK. Fundamentals of digital image processing. Englewood Cliffs, NJ: Prentice-Hall Inc., 1989.

Address:
Zhishun She
Faculty of Technology and Computer Science
University of Wales
NEWI
Wrexham
LL11 2AW, UK

Tel: +44 1978 293414
Fax: +44 1978 293168
e-mail: z.she@newi.ac.uk