CrossMark

ORIGINAL ARTICLE

Combination of 3D skin surface texture features and 2D ABCD features for improved melanoma diagnosis

Yi Ding[1] · Nigel W. John[1] · Lyndon Smith[2] · Jiuai Sun[2] · Melvyn Smith[2]

Abstract Two-dimensional asymmetry, border irregularity, colour variegation and diameter (ABCD) features are important indicators currently used for computer-assisted diagnosis of malignant melanoma (MM); however, they often prove to be insufficient to make a convincing diagnosis. Previous work has demonstrated that 3D skin surface normal features in the form of tilt and slant pattern disruptions are promising new features independent from the existing 2D ABCD features. This work investigates that whether improved lesion classification can be achieved by combining the 3D features with the 2D ABCD features. Experiments using a nonlinear support vector machine classifier show that many combinations of the 2D ABCD features and the 3D features can give substantially better classification accuracy than using (1) single features and (2) many combinations of the 2D ABCD features. The best 2D and 3D feature combination includes the overall 3D skin surface disruption, the asymmetry and all the three colour channel features. It gives an overall 87.8 % successful classification, which is better than the best single feature with 78.0 % and the best 2D feature combination with 83.1 %. These demonstrate that (1) the 3D features have additive values to improve the existing lesion classification and (2) combining the 3D feature with all the 2D features does not lead to the best lesion classification. The two ABCD features not selected by the best 2D and 3D combination, namely (1) the border feature and (2) the diameter feature, were also studied in separate experiments. It found that inclusion of either feature in the 2D and 3D combination can successfully classify 3 out of 4 lesion groups. The only one group not accurately classified by either feature can be classified satisfactorily by the other. In both cases, they have shown better classification performances than those without the 3D feature in the combinations. This further demonstrates that (1) the 3D feature can be used to improve the existing 2D-based diagnosis and (2) including the 3D feature with subsets of the 2D features can be used in distinguishing different benign lesion classes from MM. It is envisaged that classification performance may be further improved if different 2D and 3D feature subsets demonstrated in this study are used in different stages to target different benign lesion classes in future studies.

Keywords Malignant melanoma · 3D skin surface texture features · 2D ABCD features · Feature combination · Skin lesion classification

1 Introduction

Malignant melanoma (MM) is one of the most life-threatening skin cancers. Although it is the least common of all skin cancers, its incidence rates have risen faster than any other common cancers during the last 30 years. Recent statistics show that more than 1800 people in the UK are killed by this disease every year; its incidence rates have quadrupled since the 1970s [5]. Fortunately, MM can be treated successfully if it is detected and excised at an early stage.

There has been an increasing interest in early diagnosis of malignant melanoma using computer-assisted techniques in recent years [3, 7, 20, 24, 28]. Most computer-assisted

✉ Yi Ding
  robin_ding@hotmail.com

1 School of Computer Science, Bangor University,
  Bangor (Gwynedd) LL57 1UT, UK

2 Centre for Machine Vision, Bristol Robotics Laboratory,
  University of the West of England, Coldharbour lane,
  Bristol BS16 1QY, UK

diagnosis systems are based on the ABCD features of malignant melanoma, i.e. asymmetry of lesion shape [10–12, 25, 26, 32], border irregularity [1, 9, 22, 26, 27], colour variegation [2, 6, 8] and large diameter (typically over 6 mm). Although the discriminating capabilities of the ABCD features are indicative, they are far from convincing [31]. This may due to the fact that they are primarily 2D features, which are prone to environmental effects and cannot fully describe a lesion's distinctive 3D characteristic. Therefore, new features that can provide additional information are needed to be fused for an improved diagnosis.

Previous research [14–18, 34] has found useful MM features through analysis of 3D surface textures, in the form of surface normals in the tilt and slant direction, the so-called skin tilt pattern and skin slant pattern. It has demonstrated better classification results than traditional 2D-textured-based features [17]. However, whether a combination of the 3D features with the classic 2D ABCD features can improve the existing diagnosis based purely on ABCD features has not been studied before. This motivates this study to carry a multivariate study on the combinations of the 3D surface texture features with the classic 2D ABCD features. The multivariate study firstly assesses the discriminating capability of each individual feature; secondly, it uses a forward selection scheme to select the best subset features. It is envisaged that the 3D skin surface texture features (3D surface normal features), which are related to a lesion inherent topographic information, will be complementary to the 2D ABCD features and that the fusion will be useful for improving the existing computer-assisted diagnosis of MM based on the ABCD features. In addition to the previously proposed 3D features, namely the overall skin tilt/slant pattern disruptions, this work will propose two new 3D features, so-called the most tilt/slant pattern disruptions. Both the previous and the proposed 3D features will be used in this feature combination study. Also, a comprehensive feature enhancement scheme consisting of both a preprocessing Gaussian filter and a postprocessing feature-preserving anisotropic filter is proposed here.

2 Methods

2.1 Photometric stereo and 3D surface texture

The 3D skin texture was acquired from a six-light photometric stereo device, whose theory is explained briefly here. For an ideal Lambertian surface, the image irradiance equation can be expressed as [19]:

$$i(m, n) = \rho(m, n) \cdot \frac{-p(m, n) \cos\alpha \sin\beta - q(m, n) \sin\alpha \sin\beta + \cos\beta}{\sqrt{p^2(m, n) + q^2(m, n) + 1}} \quad (1)$$

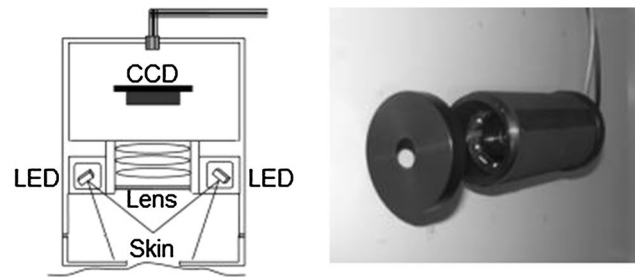**Fig. 1** *Left* Schematic to scale; *Right* developed hand-held colour photometric stereo device known as the 'Skin Analyser'. Only two of the six LED light sources are shown in either picture

where α and β are slant and tilt directions of the illuminants, the partial derivatives, $p = dz/dx$ and $q = dz/dy$ are the x-axis (indexed by m) component and y-axis (indexed by n) component of the surface gradients at image position (m, n), respectively, and ρ is the surface reflection rate (albedo). At least three images with each acquired under a different illuminant are required to solve the three variables, i.e. p, q and ρ in Eq. 1. Since there are three extra images under another three different illuminants, those abundant information can be used to detect problematic pixels under specular and shadows and remove them from the computation. As a result, the recovered surface normals and reflectance images are free from those environmental effects.

Figure 1 depicts a six-light photometric stereo device known as the Skin Analyser [33, 34], which is used as the data acquisition system. Its schematic configuration is illustrated in Fig. 1(left), and the developed device is shown in Fig. 1(right). When used in clinical trials, it is placed with its axis perpendicular to the skin surface and a camera takes six images with each under a different LED illumination. The entire operation takes less than 1 s, so it meets the demand of the static set-up required by photometric stereo. All of the following experiments were carried out using this device. Figure 2 illustrates the actual six-light photometric stereo system on the left with a sample lesion image in the middle and its recovered surface normals on the right.

2.2 3D skin surface texture features

2.2.1 Skin tilt pattern

It has been observed that MM tends to disrupt skin's naturally formed and regularly shaped surface patterns by forming new irregularities or disruptions [23, 30]. As an example, Fig. 3 illustrates a MM's 2D image and its 3D reconstructed image using the surface normals acquired by the six-light photometric stereo device. 3D skin surface disruptions can be clearly seen in the 3D image. In order

Fig. 2 *Left* Six-light photometric stereo imaging system in operation; *Middle* one of six skin lesion images taken by the Skin Analyser, *Right* recovered surface normal map of the lesion
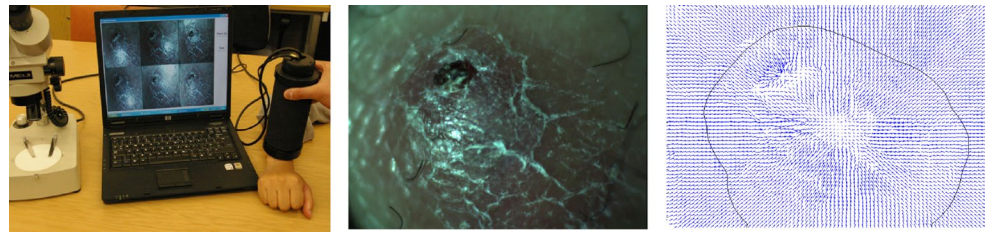
Fig. 3 *Left* A malignant melanoma's 2D colour image, *Right* its reconstructed 3D profile using the surface normals acquired by our six-light photometric stereo device
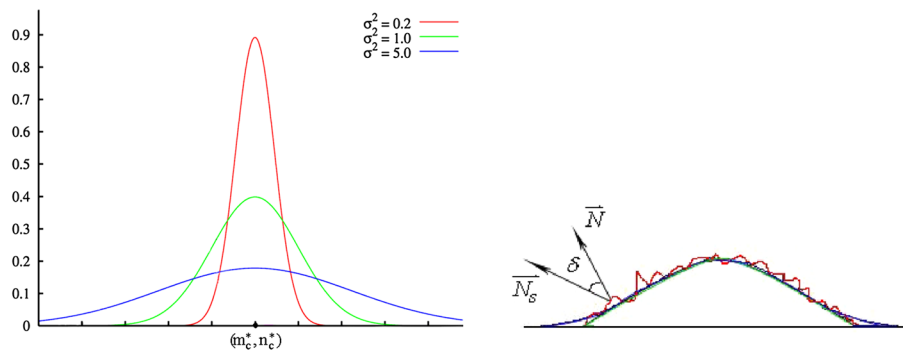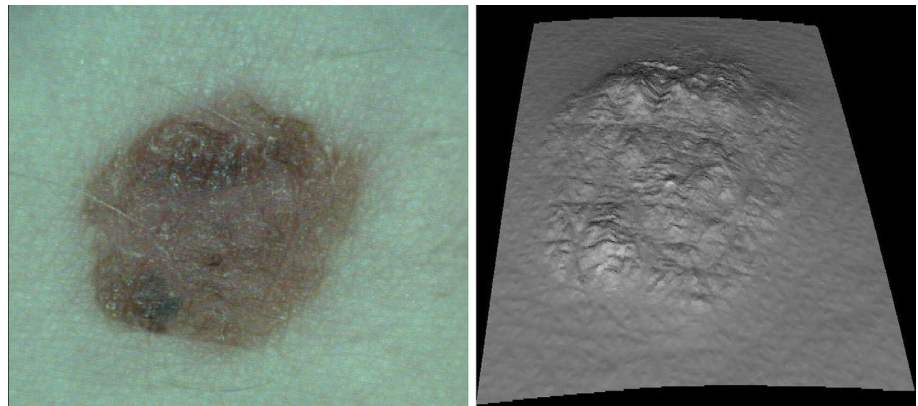
Fig. 4 *Left* Cross section of 2D Gaussian envelopes can be very curved providing small variance(s) to near flat providing large variance(s). *Right* The disruptions in 3D surface normals are estimated as the difference in direction δ between a lesion's acquired surface normal $\vec{N}$ and the corresponding simulated surface normal $\vec{N_s}$ by a possible best-fit 3D skin model generated by a Gaussian envelope. Here, the *green curve* indicates the lesion topography, while the overlapped *blue curve* indicates a possible best-fit Gaussian envelope and the irregular *red curve* indicates the surface disruptions (colour figure online)

to estimate these skin surface disruptions, a reference skin model is needed whose surface normals can be used as a reference to be compared with the actual surface normals of a lesion.

A natural choice would be a 2D Gaussian function, which allows us to adaptively select the best (or closest) model according to the surface characteristics for each lesion. Reasons for this can be explained as follows. Firstly, it is the most frequent distribution in real life and is widely used in various parametric statistical hypothesis and analyses; it is envisaged that anything abnormal such as MMs is likely to exhibit large deviations from those normal statistics. Secondly, as shown in Fig. 4(left), its flexibility and variability allow it to approximate a wide range of 3D topographies, including topographies with sharp protrusions where the variances are set small and the amplitude large, near-flat topographies where the variances are large and the amplitude low, and hemispherical topographies where only the central part of the Gaussian envelope is used. Thirdly, a Gaussian distribution has a symmetrical contour, so it allows an asymmetry analysis of the 3D data. Due to the abnormal reproduction of melanocytic cells, it is envisaged that many MMs tend to have asymmetrical and irregular shape, so the symmetrical

contour of the Gaussian distribution is capable of detecting these abnormalities. Fourthly, the transition in surface gradient from pixel to pixel on a Gaussian envelope is smooth; therefore, the similarity between neighbouring surface normal patterns is high, which is useful to simulate the regular skin patterns.

Let (m_c^*, n_c^*) be the centre of a 2D Gaussian envelope; by projecting the Gaussian envelope onto the image plane, we obtain an isotropic distribution of tilt directions centred at (m_c^*, n_c^*). In relation to the skin reference model, our objective is to find a surface description that is closest to a lesion, so the centre has to be the one whose Gaussian envelope best fits a lesion's tilt directions acquired from the Skin Analyser. This can be described as

$$\left(m_c^*, n_c^*\right) = \arg\min \sum_{\{(m,n)\in S_l\}} \left\|\varphi(m,n) - \varphi^*(m,n)\right\| \quad (2)$$

where (m_c^*, n_c^*) is the estimated centre of the Gaussian distribution using least-square estimation, S_l denotes the lesion region, ‖.‖ denotes the Euclidean distance, φ and φ^* are the acquired and the estimated tilt direction (pattern). The star sign "*" denotes an estimated variable. Upon the estimation of the distribution centre, the associated differences in the tilt direction are used to estimate the skin tilt pattern disruptions. The overall disruptions in skin tilt pattern (OT) is defined as (the average of) the sum of differences between the skin tilt patterns $\varphi_{\min}$ of the best-fit Gaussian function and the acquired skin tilt patterns φ.

$$\text{OT} = \frac{\sum\limits_{(m,n)\in S_l} \left\|\varphi_{\min}(m,n) - \varphi(m,n)\right\|}{S_l} \quad (3)$$

where S_l is the number of pixels within the lesion. Another feature called the most disrupted tilt (MT) is estimated as

$$\text{MT} = \left\|\varphi_{\min}(m,n) - \varphi(m,n)\right\| \quad (4)$$

2.2.2 Skin slant pattern

So far, for finding the centre location of the Gaussian distribution, all the computations have been limited to the x–y plane and the tilt direction. To determine the exact topography of the Gaussian distribution, the slant directions should also be used, since the topography of a Gaussian distribution is dependent on its variance and amplitude. Accordingly, the best-fit Gaussian topography can be estimated as

$$(A, \sigma) = \arg\min \sum_{(u,v)\in S_l} \|\theta(m,n) - \theta*(m,n)\| \quad (5)$$

where θ denotes the acquired lesion's skin slant pattern and θ^* denotes the estimated skin slant pattern. Both patterns can be represented in terms of surface gradients as

$$\theta = \cos^{-1}\left(\frac{1}{p^2 + q^2 + 1}\right) \quad \text{and}$$

$$\theta^* = \cos^{-1}\left(\frac{1}{\sqrt{\left(Z_x^*\right)^2 + \left(Z_y^*\right)^2 + 1}}\right) \quad (6)$$

where (p, q) are the acquired surface gradients of a lesion from the Skin Analyser, (Z_x^*, Z_y^*) are the estimated surface gradients in the x-axis and y-axis of the Gaussian functions, S_l denotes the lesion region, ‖.‖ denotes the Euclidean distance, A^* and σ^* are estimated amplitude and variance of the Gaussian function.

In estimating the parameters of the best-fit Gaussian function, a nonlinear optimisation method, Levenberg–Marquardt (LM) method was used to solve the problem. Levenberg–Marquardt method refers to a standard routine optimisation scheme that is highly efficient in estimating the parameters that solve the least-square estimation problems. In certain computational software that enables the nonlinear numerical analysis such as MATLAB®, LM method has already been implemented and included in the options of the least-squared-based curve-fitting functions, such as *lsqcurvefit* in MATLAB. To speed up the estimation process, sometimes a good initialisation of the parameters is needed. This can be achieved by firstly searching through several values within the possible range of the parameters, and using the parameters with the lowest estimation error to provide an initial guess of the parameters.

Upon the estimation of the parameters, A^* and σ^* of the resultant Gaussian topography or portions of the Gaussian topography, the associated differences in the slant direction are used to estimate skin slant pattern disruptions. The overall disruptions in skin slant pattern (OS) are defined as (the average of) the sum of the differences between the skin slant patterns $\theta_{\min}$ of the best-fit Gaussian function and the acquired skin slant patterns θ.

$$\text{OS} = \frac{\sum\limits_{(m,n)\in S_l} \left\|\theta_{\min}(m,n) - \theta(m,n)\right\|}{S_l} \quad (7)$$

where S_l is the number of pixels within the lesion. Another feature called the most disrupted slant (MT) region is estimated as

$$\text{MS} = \left\|\theta_{\min}(m,n) - \theta(m,n)\right\| \quad (8)$$

To depict this skin disruption estimation process more vividly, Fig. 4(right) illustrates that the acquired topography of the lesion (shown as a green curve) can be approximated by a 2D Gaussian function (shown as a blue curve), which is the 3D skin model best fits the acquired surface normals. Through this way, the actual disruptions in the 3D surface normals (shown as an irregular red curve) can
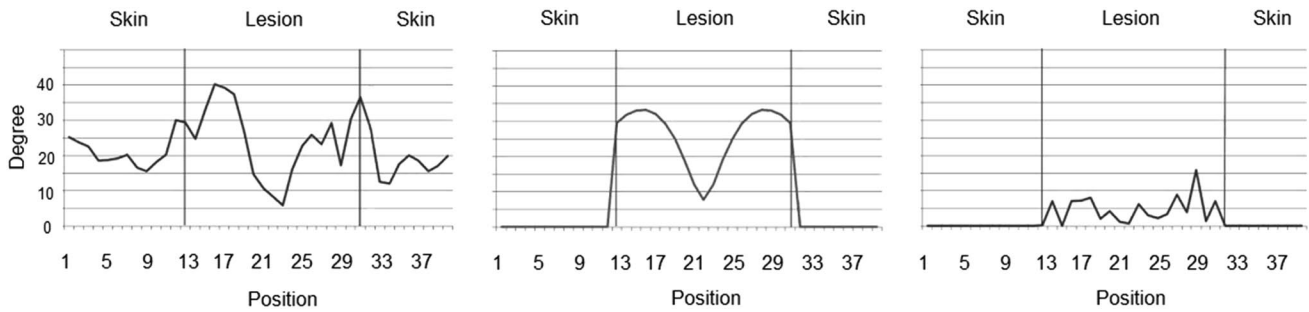
Fig. 5 *Left* Skin slant patterns along a sample row of an elevated lesion of nodular shape; *Middle* its corresponding skin slant patterns generated by a best-fit 2D Gaussian function for the lesion region; *Right* the skin tilt pattern disruptions are estimated as the differences between the *Left* and the *Middle* figures for the lesion region

be estimated by subtracting the acquired surface normals with those simulated by the best-fit skin model, without the influence of the underlying non-flat topography.

As an example from a real lesion, Fig. 5 illustrates the skin slant patterns along a sample row of a non-flat lesion of domed shape with its simulated best-fit skin slant patterns generated by a 2D Gaussian function and the estimated skin slant pattern disruptions on top of the lesion, estimated as the Euclidean differences between the acquired and the simulated skin slant patterns. Surrounding skin is not used as the skin slant pattern disruptions can be found by comparing with the best-fit slant pattern model. Being smooth, symmetrical while fitting closely to a lesion's acquired skin slant patterns, the simulated best-fit skin slant patterns will be able to sense and detect the subtle variations in skin slant patterns without the influence of non-flat surface topography.

2.2.3 Feature enhancement

In view of the noise effects, an enhancement scheme, which consists of a preprocessing Gaussian filter and a postprocessing anisotropic nonlinear diffusion, is employed to enhance both the tilt and the slant pattern features, respectively. In the preprocessing step, the idea of applying Gaussian smoothing is to reduce the very high-frequency noise only at the expense of slightly reduced 3D skin surface texture information. Although there is a trade-off between reducing the high-frequency noise and preserving high-frequency 3D skin texture, it is envisaged that by properly choosing the smoothness scale (i.e. the variances), the 3D skin texture can be enhanced without losing much useful information. Here, the Gaussian smoothing function is applied directly to the three channels of surface normals, (n_x, n_y, n_z) separately, i.e.

$$n_x^* = n_x * G(u, v)$$
$$n_y^* = n_y * G(u, v)$$
$$n_z^* = n_y * G(u, v)$$

(9)

where $*$ denotes the convolution operator and $G(u, v, (x_c, y_c), \sigma)$ is a 2D Gaussian function, which has the following form,

$$G(u, v) = e^{-\left(\frac{(u-x_c)^2 + (v-y_c)^2}{2\sigma^2} \right)}$$

(10)

where (x_c, y_c) is the centre of the Gaussian window function, and σ is the variance that controls the strength of the Gaussian smoothness function. The size of Gaussian window should be small enough to be sensitive in reducing noise locally and big enough to generate a smooth Gaussian envelope. We use the $(2\sigma + 1)$ rule (which covers 95 % of a Gaussian envelope) to select the window size as 5. To reduce the local skin surface noise while preserving the local 3D surface texture features, it is important to choose a small σ. In our experiments, σ is chosen as 1.

In the postprocessing step, anisotropic nonlinear diffusion is applied here to reduce the noise effects on the skin tilt/slant pattern disruption [16]. Anisotropic diffusion refers to an iterative technique that is able to detect and enhance a local surface's prominent features in both homogeneous and inhomogeneous texture regions [35]. This is a very attractive and an important property considering the fact that many benign lesions, which are covered by a fine network of skin patterns, are likely to have a homogeneous surface texture. Also many MMs, which can cause erosions and disruptions of skin patterns and forming new lines of varying directions, are likely to have inhomogeneous surface texture.

This approach involves adaptively choosing the filtering smoothness strength so that intra-regions become smooth, while edges of inter-regions are preserved. The degree of smoothness is an often decided by a non-negative decreasing function such as a sigmoid function, in which a threshold is used to judge whether the local feature is a signal or noise. A diffusion equation is used here to smooth out the noise within the local skin tilt/slant pattern disruption in successive iterations.

$$(X_\Delta)_t = \text{div}(\mathbf{D} \cdot \nabla X_\Delta) = \text{div}\left(\begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \cdot \begin{bmatrix} \nabla(X_\Delta)_x \\ \nabla(X_\Delta)_y \end{bmatrix}\right) \quad (11)$$

where $\nabla()_x$ and $\nabla()_y$ denote the gradient operator in x-axis and y-axis, X_Δ denotes skin tilt/slant pattern feature, $\mathbf{D}$ is the diffusion tensor, which controls the smoothing strength, and is defined as a function of the structure tensor, i.e.

$$\mathbf{D} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \cdot \begin{bmatrix} \chi & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}^T = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \quad (12)$$

where $\mathbf{v}_1$ is the principal direction vector of the local signal variation, χ is the smoothing strength, which is in the range of [0 1] and is defined using the exponential curve as

$$\chi = e^{-\frac{|\nabla X_\Delta|}{K}} \quad (13)$$

Here, the smoothing strength is chosen adaptively according to magnitude of local signal variation, represented by $|\nabla X_\Delta|$. The parameter K is a threshold to judge whether the local structure is a feature or noise. In our experiment, K is empirically chosen as 0.4. For $|\nabla X_\Delta| << K$, the local structure is deemed to be noise, and a large smoothing strength is applied, while for $|\nabla X_\Delta| >> K$, the local structure is seen as local feature, and very small or no smoothing should be applied to preserve the local feature. Finally, using the Euler forward difference approximation, the diffusion equation of Eq. 11 is expanded as

$$X_\Delta^{(t+1)} = X_\Delta^{(t)} + \tau\left(\frac{\partial}{\partial x}\left(D_{11}(X_\Delta)_x^{(t)}\right) + \frac{\partial}{\partial x}\left(D_{12}(X_\Delta)_y^{(t)}\right)\right.$$
$$\left. + \frac{\partial}{\partial y}\left(D_{21}(X_\Delta)_x^{(t)}\right) + \frac{\partial}{\partial y}\left(D_{22}(X_\Delta)_y^{(t)}\right)\right) \quad (14)$$

The iteration step τ is chosen as a value smaller than $0.5/N_d$ [37] where N_d is the number of signal dimensions. Applying separately on the tilt/slant pattern feature means, N_d is equal to 1, so τ is chosen as a value smaller than 0.5. The central finite difference is used to evaluate the partial differential equation of Eq. 14.

2.3 ABCD features

2.3.1 Asymmetry

In order to determine the asymmetry of a lesion, we have to find the centre of the lesion region, which is defined through moments (see Appendix 1). The two principal centroidal axes, which are 90° apart, are used to approximate the best axis of symmetry. Reflecting the lesion area by the two axes will result two non-overlapping area differences, which are equal to zero if the lesion is perfectly symmetrical and nonzero (as in most cases) if the lesion is asymmetrical. The least of the two differences $\Delta S_{\min}$ is used to calculate the asymmetry index (AI), which is defined as the ratio of the non-overlapping area to the original lesion area.

$$\text{AI} = \frac{\Delta S_{\min}}{S_l} \times 100 \quad (15)$$

2.3.2 Border

The border irregularity index (BI) is defined as the roundness ratio [29] as

$$\text{BI} = \frac{P^2}{4\pi S_l} \quad (16)$$

where P and S_l denote the perimeter and the area of the lesion, respectively. If x_i, $i = 1, 2,\ldots, N$ are sample points of the boundary, then the perimeter is given by

$$P = \sum_{i=1}^{N-1} \|x_{i+1} - x_i\| + \|x_N - x_1\| \quad (17)$$

where $\|.\|$ denotes the Euclidean distance. In a digital image, the area of the lesion can be evaluated by counting the number of pixels within the lesion. The ratio is smallest when the border profile is a circle, while it gets larger as the shape of the border deviates from a circle to indicate the increasing irregularities of the border.

2.3.3 Colour

With the possibility of reducing variations caused by different people and other environmental effects, "relative colour" instead of "absolute colour" is also used here [36]. It is defined as the normalised value of a colour component within the lesion subtracted from the normalised value of that colour component in the background skin.

$$\begin{bmatrix} r' \\ g' \\ b' \end{bmatrix} = \begin{bmatrix} r_{\text{lesion}} - r_{\text{skin}} \\ g_{\text{lesion}} - g_{\text{skin}} \\ b_{\text{lesion}} - b_{\text{skin}} \end{bmatrix} \quad (18)$$

where (r', g', b') denote the relative colours in red, green and blue colour components, $(r_{\text{lesion}}, g_{\text{lesion}}, b_{\text{lesion}})$ represent the existing lesion colours, $(r_{\text{skin}}, g_{\text{skin}}, b_{\text{skin}})$ represent the average colour values of the surrounding skin computed based on [8]. Variegated colours within the lesion imply high variances in the respective red (R), green (G) and blue (B) colour components. So three colour features are selected as the standard deviations σ_r, σ_g, σ_b in the red, green and blue relative colour spaces.

2.3.4 Diameter

The diameter of the lesion in pixel is defined as the longest distance between two sample points on the lesion boundary while the line between the sample points must pass through the centre of the mass (m_c, n_c).

$$D = \max\left\{\left\|x_i - x_j\right\| \text{where } (m_c, n_c) \in \overline{x_i x_j}\right\} \qquad (19)$$

where x_i and x_j denote the two boundary sample points, and the line segment between x_i and x_i must pass through the centre of the mass. Given the knowledge of an image's magnification specification, the scale calculated by Eq. 19 can be converted from pixels to millimetres.

2.4 Feature selection

The 10 features used for the combination study will be the asymmetry index (AI), the border irregularity index (BI), the standard deviations $\sigma_r, \sigma_g, \sigma_b$ in the relative red, relative green and relative blue colour spaces, the diameter D and the proposed tilt pattern features including OT and MT and slant pattern features including OS and MS.

Although a large number of independent features are available for lesion classification, not all of these features contributed equally well to solve the classification problem. Sometimes the best classification result is not determined by the complete set of the input features $\{x(1), x(2),\ldots, x(M)\}$ where M is the number of features, and it is decided only by a subset of them $\{x(1), x(2),\ldots, x(m)\}$ where $m < M$. The purpose of the feature selection scheme is to select the optimal combination of features, which gives the best classification results. One way is to exhaustively evaluate all possible combinations of the input features. However, computational cost of this exhaustive search scheme is prohibitively high.

One commonly used feature selection scheme is forward selection. The forward feature selection procedure begins by evaluating the classification performances of all feature subsets that consist of only one input feature so that we can find the best individual feature, $X(1)$. Next, it finds the best subset consisting of two features: the winner of one input feature, $X(1)$, and one other feature from the remaining $(M - 1)$ input features. So there are a total of $(M - 1)$ pairs. After that, the input subsets with three and more features are evaluated. According to forward selection, the best subset with m features is the *m*-tuple consisting of $X(1), X(2),\ldots, X(m)$, while overall the best feature set is the winner out of all the M steps.

2.5 Lesion classification

Because feature selection is the main focus of this paper, we did not choose a very complex classification system

such as ensemble classifiers [31] (in fact, the design of ensemble classifiers [21] should be another topic to be discussed). Instead, a single support vector machine (SVM) classifier [4], which has the advantages of simplicity and efficiency, while giving good classification power is used in this combination study. Specifically, the nonlinear SVM with a multilayer perceptron kernel function is chosen. The theory of SVM requires the input vectors to be nonlinearly mapped to a very high-dimension feature space, typically much higher than the original feature space. In this feature space, data from the two classes can always be separated by a hyperplane. The support vectors are those transformed training vectors that are equally close to the hyperplane and therefore are the most informative for defining the optimal separating hyperplane for the classification task and the most difficult patterns to classify. Among many hyperplanes that might classify the data, only the hyperplane that maximises the margin between the two classes is used for classification. Therefore, learning is formulated as an optimisation problem with the target of maximising the distance from the hyperplane to the support vectors, or equivalently maximising the nearest distance between a point in one separated hyperplane and a point in the other separated hyperplane.

3 Results

A total of 46 lesion subjects were collected over a period of 2 years at the collaborating dermatological clinics using the Skin Analyser. Consent forms were signed by participating patients involved in the study. For confidentiality reasons, each subject lesion collected was assigned a unique number and kept anonymously. The research ethics committee of the NHS (UK) approved our methods of using the clinical subjects in this work. Of the total 46 lesions, 12 are MMs and 34 are from nine other types of benign lesions. The 34 benign lesions include both non-melanocytic lesions such as four dermatofibromas (DFs), five intra-dermal naevi (IN), three hyperkeratotic squamous papillomas (HSP), eight seborrhoeic keratoses (SKs) and also melanocytic lesions such as two dysplastic naevi (DN), eight compound naevi (CN), two congenital naevi (CGN), one junctional naevus (JN) and one blue naevus (BLN). Inclusion of pigmented non-melanocytic lesions in lesion classification has largely been ignored by previous computer-based diagnosis systems. However, some pigmented non-melanocytic lesions can even be mistaken for melanocytic lesions even by experienced specialists [28]. Therefore, they should be included to test the accuracy of classification systems. All the lesions acquired are used for the classification experiments, and we did not artificially choose skin lesions for classification.

Table 1 Classification performance ranking in percentage [listed as sensitivity/specificity (overall classification rate)] for the 4 sample groups and overall samples for the 10 features using a nonlinear SVM classifier

Rank	Feature	Group 1 (%)	Group 2 (%)	Group 3 (%)	Group 4 (%)	Overall (%)
1	OS	**88.9/75.0 (81.0)**	62.6/75.0 (70.0)	77.8/75.0 (76.2)	**100.0/75.0 (85.0)**	**82.3/75.0 (78.0)**
2	MT	**100.0/66.7 (81.0)**	**75.0/75.0 (75.0)**	88.9/66.7 (76.2)	87.5/66.7 (75.0)	87.8/68.8 (76.8)
3	MS	**88.9/75.0 (81.0)**	**75.0/75.0 (75.0)**	77.8/75.0 (76.2)	87.5/66.7 (75.0)	82.3/72.9 (76.8)
4	Red	100.0/41.7 (66.7)	50.0/75.0 (65.0)	66.7/75.0 (71.4)	87.5/75.0 (80.0)	76.0/66.7 (70.8)
5	OT	88.9/50.0 (66.7)	75.0/50.0 (60.0)	66.7/75.0 (71.4)	100.0/50.0 (70.0)	82.6/56.3 (67.0)
6	Asymmetry (AI)	66.7/41.7 (52.4)	25.0/41.7 (35.0)	**88.9/83. (85.7)**	87.5/58.3 (70.0)	67.0/56.3 (60.8)
7	Border (BI)	77.8/25.0 (47.6)	62.5/58.3 (60.0)	77.8/41.7 (57.1)	75.0/66.7 (70.0)	73.3/47.9 (58.7)
8	Blue	100.0/33.3 (61.9)	87.5/16.7 (45.0)	66.7/41.7 (52.4)	100.0/50.0 (70.0)	88.5/35.4 (57.3)
9	Diameter (D)	77.8/50.0 (61.9)	87.5/16.7 (45.0)	100.0/33.3 (61.9)	62.5/50.0 (55.0)	81.9/37.5 (56.0)
10	Green	100.0/33.3 (61.9)	87.5/16.7 (45.0)	33.3/16.7 (23.8)	75.0/41.7 (55.0)	74.0/27.1 (46.4)

The best classification performance(s) for each lesion group is(are) highlighted in bold

It is understood that the sample size is relatively small due to patients' attendance rate at the collaborating clinics; therefore, we need to employ a convincing method to test the classification result. Leave-one-out cross-validation (LOOCV) is chosen because it can give an unbiased classification performance for each feature or subset feature. In a LOOCV scheme, a classifier will be tested on the one sample but trained on all but the one testing sample. So training and test samples are independent with each other. The LOOCV performance for the feature or subset feature is evaluated as the average classification result of all samples.

Some lesion classes have only a very small number of samples, including BLN, CGN, JN and HSP; therefore, the classification results may not be representative of the true discriminating power of the feature, and the 34 benign lesions are subsequently split into four sample groups. Sample group 1 includes only melanocytic lesions: one JN and eight CN based on the fact that both lesions are typically small, smooth and slightly raised. Sample group 2 includes only non-melanocytic lesions: eight SKs which are among the most common classes of benign lesions and have a distinct appearance from other benign lesions. Lesions in this class can vary significantly in visual appearance including size, shape, colour and texture. Sample group 3 includes only non-melanocytic lesions (five IN and four DFs) based on the fact that they both have raised and nodular shape. Sample group 4 is made up of a combination of non-melanocytic lesions of three HSP with melanocytic lesions of two DN, two CGN and one BLN.

Regarding each feature's discriminating capability, its classification performance with regard to each sample group using a nonlinear SVM classifier with LOOCV is listed in Table 1. Among the 10 features, the best classification performance is achieved by MT for group 1, MT and MS for group 2, asymmetry for group 3 and OS for group

4. If judged by the overall classification performance for all lesion samples, the OS feature has demonstrated to be the best one amongst all the 10 features. The reason why the asymmetry feature has shown the best performances for group 3 and the relative red colour feature has shown good result for group 4 is not surprising as group 3 includes IN and DF mainly have a round or nodular shape, while group 4 contains the BLN and the CGN mainly have a uniform or even colour distribution.

For single feature, the proposed 3D skin surface texture features have provided the best classification results for group 1, 2 and 4. At the same time, their classification results are also comparable to that of the asymmetry for group 3. If judged by the overall performances, the proposed OS has demonstrated the best classification results among all the 10 features. In general, the 3D features have shown better classification results than the 2D ABCD features, a finding consistent with [38] which indicates that the 3D features are better than the well-established border [22] and colour features using single classifier systems. The authors acknowledge that the 2D features used in this study are classic but rather simplistic compared to the 3D features, so using another set of 2D features might give different classification results. However, since each feature will only focus on one property of pigmented lesions, the purpose of this paper is not to select a single gold feature but to assess the discriminating power of the combined features, so the conclusions drawn from the single-feature classification experiment should be seen as indicative not conclusive.

Regarding the possible combinations of feature subset, the forward scheme as mentioned in Sect. 2.4 is used to select the optimal feature subset for (1, 2,…, m) features. Because OS is the best feature in the one-feature experiment, it is used in the subsequent two-feature

Table 2 Classification performances in percentage [listed as sensitivity/specificity (overall classification rate)] of the 4 sample groups and overall using a nonlinear SVM classifier for the best feature subsets (with feature size 2–6) selected by the forward feature selection scheme

No. of features	Index	Feature combination	Group 1 (%)	Group 2 (%)	Group 3 (%)	Group 4 (%)	Overall (%)
2	a	OS + Green	**88.9/83.3 (85.7)**	75.0/66.7 (70.0)	66.7/75.0 (71.4)	100.0/83.3 (90.0)	82.6/77.1 (79.3)
	b	Red + AI	88.9/58.3 (71.4)	50.0/66.7 (60.0)	88.9/75.0 (81.0)	**100.0/100.0 (100.0)**	81.9/75.0 (78.1)
	c	OS + MS	**88.9/83.3 (85.7)**	75.0/66.7 (70.0)	66.7/75.0 (71.4)	100.0/83.3 (90.0)	82.6/77.1 (79.3)
3	a	OS Green + AI	**88.9/83.3 (85.7)**	**75.0/83.3 (80.0)**	88.9/75.0 (81.0)	100.0/83.3 (90.0)	88.2/81.3 (84.2)
		Red + AI + BI	66.7/50.0 (57.1)	75.0/66.7 (70.0)	**100.0/83.3 (90.5)**	87.5/100.0 (95.0)	82.3/75.0 (78.2)
4	a	OS + Green + AI + Blue	**88.9/83.3 (85.7)**	75.0/75.0 (75.0)	**100.0/83.3 (90.5)**	100.0/91.7 (95.0)	91.0/83.3 (86.5)
	b	Red + AI + BI + Blue	77.8/58.3 (66.7)	62.5/66.7 (65.0)	**100.0/83.3 (90.5)**	**100.0/100.0 (100.0)**	85.1/77.1 (80.5)
5	a	OS + Green + AI + Blue + Red	**88.9/83.3 (85.7)**	**87.5/75.0 (80.0)**	**100.0/83.3 (90.5)**	100.0/91.7 (95.0)	**94.1/83.3 (87.8)**
	b	Red + AI + BI + Blue + Green	100.0/50.0 (71.4)	**100.0/66.7 (80.0)**	88.9/75.0 (81.0)	**100.0/100.0 (100.0)**	97.2/72.9 (83.1)
6	a	OS + Green + AI + Blue + Red + OT	**88.9/83.3 (85.7)**	**87.5/75.0 (80.0)**	88.9/83.3 (85.7)	100.0/83.3 (90.0)	91.3/81.3 (85.4)
	b	Red + AI + BI + Blue + Green + D	100.0/66.7 (81.0)	75.0/66.7 (70.0)	77.8/66.7 (71.4)	**100.0/100.0 (100.0)**	88.2/75.0 (80.6)

The rows indexed as "a" indicate that the features are selected from the combinations of the 2D and 3D features. The rows indexed as "b" indicate that they are selected only from the 2D features. The row indexed as "c" indicates that it is selected only from the 3D features; the result in this category only displays the combination including two 3D features (OS and MS). This is because adding more 3D features cannot further improve the classification result in this category. The best classification performance(s) for each lesion group is(are) highlighted in bold

subset selection steps. Then, the best two features are kept for selecting the best three features. The procedure repeats until the combination of m features (here $m = 6$) has been computed. Table 2 lists the classification performances for both the combination with 3D features (indexed as "a") and the combinations of only the 2D features (indexed as "b"). For the former, the combinations beyond six features are not listed because the forward selection cannot select any new ones different to the existing six features. Also, inclusion of more features cannot improve the classification results indicating data redundancy. It can be seen that the classification results improve as the number of features increases until five features then starts to degrade with six features.

For combined features, the best classification result is achieved as the combination of the five features, which gives a very promising classification result of 87.8 %. It is a substantial improvement over (1) overall 78.0 % achieved by the best single feature, (2) overall 83.1 % achieved by the best combination of the 2D features and (3) overall 79.3 % achieved by the best combination of the 3D features. Looking at the results for individual group, the best result by the 2D and 3D combination is also better than the best result by only the 2D combination in two out four groups, and both combinations match each other in one group. In the other group (group 4), the latter only performs slightly better in specificity, and the former's result of 91.7 % specificity is also a very good result. Both combinations share the same good 100 % sensitivity in this group.

Comparing the classification results when the number of features is 2, 3, 4, 5 and 6, the combinations of the 2D

and 3D features selected by forward selection scheme have shown better performances than those of only the 2D features. The best 2D and 3D combination has five features including OS, AI and three colour components but without the BI and the diameter feature. The best 2D combination also has five features, including AI, BI and all three colour component features but without the diameter feature. This suggests that (1) using all the ABCD features in either combination does not give the best classification result and (2) the 3D skin surface texture features can improve the 2D ABCD-based classification if used in combination with colour and asymmetry features. Therefore, it can provide complementary, useful and very discriminating information. Since ABCD features represent different properties of a lesion, it is interesting to see whether the two 2D features (diameter and BI) not selected by the feature selection can have some values in improving individual group's classification results in the 2D and 3D feature combination.

The best result of five or six combined 2D and 3D features including the diameter is listed in Table 3 a[1*] and a[1**], respectively. Some interesting observations can be made here: the best result with five features performs better overall than the one with six features. However, the latter achieved 100 % sensitivity and very high specificity (83.3, 83.3 and 91.7 %, respectively) for 3 out of 4 groups. The only group it performs worse than both the former and the solely 2D features is group 2 which consists of SKs. However, SKs belong to a benign lesion class, that has a distinct visual appearance, different to other benign lesions and its size tends to be larger

Table 3 Classification performances in percentage (listed as sensitivity/specificity (overall classification rate)) between 2D + 3D feature combinations and only 2D feature combinations using a nonlinear SVM classifier

No. of features	Index	Feature combination	Group 1 (%)	Group 2 (%)	Group 3 (%)	Group 4 (%)	Overall (%)
5	a^{1*}	OS + AI + Red + Blue + D	100.0/75.0 (85.7)	50.0/66.7 (65.0)	**88.9/91.7 (90.5)**	100.0/91.7 (95.0)	84.7/83.3 (84.0)
	a^{2*}	OS + AI + Red + Blue + BI	77.8/66.7 (71.4)	**100.0/83.3 (90.0)**	100.0/83.3 (90.5)	100.0/91.7 (95.0)	**94.4/81.3 (86.7)**
	b	Red + AI + BI + Blue + Green	100.0/50.0 (71.4)	100.0/66.7 (80.0)	88.9/75.0 (81.0)	**100.0/100.0 (100.0)**	97.2/72.9 (83.1)
6	a^{1**}	OS + AI + Red + Blue + Green + D	**100.0/83.3 (90.5)**	50.0/58.3 (55.0)	**100.0/83.3 (90.5)**	100.0/91.7 (95.0)	87.5/79.2 (82.7)
	a^{2**}	OS + AI + Red + Blue + Green + BI	55.6/58.3 (57.1)	**100/83.3 (90.0)**	100.0/83.3 (90.5)	**100.0/100.0 (100.0)**	88.9/81.3 (84.4)
	b	Red + AI + BI + Blue + Green + D	100.0/66.7 (81.0)	75.0/66.7 (70.0)	77.8/66.7 (71.4)	**100.0/100.0 (100.0)**	88.2/75.0 (80.6)

Here, "a^{1*}" and "a^{2*}" denote the best classification performances of the feature combination including the diameter feature and border irregularity feature, respectively, along four other different 2D and 3D features. "a^{1**}" and "a^{2**}" denote the best classification performances of six 2D and 3D features including the five features of "a^{1*}" and "a^{2*}", respectively, plus the green component. The rows indexed as "b" indicate that they are the best results selected by the forward selected only from the 2D features. The best classification performance(s) for each lesion group is(are) highlighted in bold

compared with many other benign lesions. In fact, the single-diameter feature performed poorly in group 2 (with accuracy below 50 % as listed in Table 1). It is this weakness that decreases the overall performance of the combination. However, in the clinics, a majority of SKs can be identified successfully by trained dermatologists. Therefore, if they can be excluded in the first place manually, the classification result by this 2D and 3D feature combination is significant in the context that it can help doctors to distinguish more difficult lesions from group 1, 3 and 4. This again proves that including the 3D feature (OS) in the combination can improve the 2D ABCD classification.

In the next experiment, we assessed whether BI can be useful in the 2D and 3D combination. The best results by the five and six features including BI are listed in Table 3 a^{2*} and a^{2**}, respectively. In both results, the combination with BI has shown very promising result for group 2, 3 and 4 (all with 100 % sensitivity and high specificity). It is significant that it is able to classify group 2, which has not been satisfactorily classified by all the other combinations so far. The group that this combination does not perform well is group 1. Group 1 that consists of junctional and compound naevi that are typically small; therefore, they are more likely to be classified correctly with the assistance of the diameter feature. However, inclusion of the diameter would likely to have difficulties in classifying other benign lesions such as SKs which tend to be large in size and more likely to be larger than 6 mm and therefore classified as MMs. On the other hand, shape-based features such as border irregularity are likely to suffer more from noise effects for small lesions than for larger lesions, therefore preventing it from making the correct classification. Indeed, in Table 1, the single border feature performed poorly for group 1 (overall accuracy

below 50 %); it is the only group that this feature is not so capable of correct differentiation. In the most likely cases, many CN and JN are probably estimated as having large border irregularities due to noise effects, therefore preventing the distinction from many MMs.

4 Discussion

Regarding different features or feature combinations, their strengths and weaknesses are discussed in this section.

4.1 Asymmetry

Asymmetry has demonstrated itself to be a useful feature, particularly in discriminating round and nodular lesions, i.e. IN and DFs of group 3, from MM. On the other hand, the asymmetry feature has shown poor performance in discriminating the benign lesions in group 4, which include CGN and DN that are even considered as difficult by the dermatologists at the collaborating clinic. Reasonably good classification is achieved for sample group 4. Overall, the asymmetry feature has demonstrated the second best classification performance among the ABCD features next to relative red.

4.2 Border

Border irregularity has performed reasonably well for sample group 4 while behaves poor results for the others. Therefore, the simple border feature used in this paper is not sufficient for the differentiation between MM and benign lesions. Although more sophisticated border features

[1, 22, 27] may lead to better classification results, they still suffer from the drawbacks below: firstly, although most MMs would have irregular border profiles, many benign lesions would also have large border irregularity indices [9]. Secondly, the border feature is very sensitive to imaging noise [19, 33, 34], this is particularly true for small lesions where the signal to noise ratio is low. Thirdly, the ground-truth border profile drawn by the dermatologists may even be different from person to person [9], seriously affecting the subsequent lesion analysis and classification. A recent finding [38] indicates that border features of [22] demonstrated inferior classification performances using a single nonlinear classifier than both the 3D features and the 2D colour features.

4.3 Colour

Colour has also shown the best classification performance for sample group 4. This is understandable as group 4 contains BLN and CGN (benign lesions with mainly uniform and even colour), which can be easily distinguished from most variegated-coloured MMs. Overall, the standard deviation of the relative red colour has demonstrated the best overall classification performance amongst all the colour features, which is consistent with other research on relative colour features [6]. It is also the best feature amongst all the ABCD features.

4.4 Diameter

Being the simplest and the most straightforward feature among the ABCD features, diameter shows some promises in classifying group 1 and group 3. In particular, results for group 1 are more understandable as it consists of CN and JN, which are typically small compared to many MMs. However, diameter alone is not capable of differentiating between MMs and benign lesions. This is because some benign lesions may be in variable size such as SKs in group 2, IN and DFs in group 3 and HSP in group 4, making it difficult to give the correct classification results.

4.5 Single feature versus combined features

Comparing Tables 1 and 2, all the five combinations of the 2D and 3D features selected by the forward selection scheme outperform any single feature's overall classification result. Here, a sign test [13] is used to validate our claim that the former is superior to the latter in classification performance. The null hypothesis is that their classification performances are equivalent. By counting the number of wins or losses or ties, the former has won all the five cases. This gives a p value of $1/2^5 \begin{pmatrix} 5 \\ 0 \end{pmatrix} = 0.031$, which is enough to reject the hypothesis. Therefore, the

classification performances are different between the former and the latter. Judging on the classification results, the former is indeed better than the latter. Therefore, it is fair to say that each single feature has its own limitations and shortcomings for lesion classification, so there is not a gold feature that can give the best classification between MMs and benign lesions without assistance from other features.

4.6 2D feature combination versus 2D and 3D feature combination

Judging on the performances in Table 2, the 2D and 3D feature combinations outperform their solely 2D combination counterparts when the number of features in the combination is from 2 to 6, respectively. A sign test again is used to validate our claim that the former is better than the latter. Here, the null hypothesis H_0 is: the two types of combinations are equivalent in performance. The alternative hypothesis H_1 is: one type of combination is better in classification performance than the other. As this is a sign test, a straightforward way to compare the overall performance. In all the five feature combinations, the one with the 3D features has shown better overall performances than without. So the corresponding in this sign test p value is $1/2^5 \begin{pmatrix} 5 \\ 0 \end{pmatrix} = 0.031$, which indicates that there is enough evidence to reject the null hypothesis. Therefore, the alternative hypothesis is valid indicating one combination is better than the other. Judging from the classification performances in Table 2, this further validates our claim that the combinations with the 3D features are better than the combinations with only the 2D features.

4.7 3D feature combination only

The combinations of only 3D features are also studied here for comparisons. The best classification in this category ended with two features (OS + MS), and adding more 3D features cannot improve the classification result. Adding the most disrupted feature in the slant pattern (MS) is able to improve both sensitivity and specificity slightly than using only the overall disrupted feature (OS). This is also a promising result as both features reflect surface variations (disruptions) in the z-axis (slant pattern). Therefore, they are the features unique in 3D and can potentially reveal more complementary information in addition to the 2D features than the tilt pattern features (OT + MT), which reflects surface variations in the 2D x–y plane. Due to the number of 3D features available, we are unable to compare the classification performances with the other two-feature sets [i.e. (1) 2D features alone and (2) combined 2D and 3D features] beyond two features. Therefore, it remains to be seen whether more

useful features can be found in 3D to further improve its classification performance.

4.8 Values of border and diameter

From the results in Table 3, it can be seen that the other two features (1) border irregularity and (2) diameter not selected by the forward selection in Table 2 have also justified their values. If used separately in combination with the AI, colour and 3D features, they can be useful in improving the correct classifications of non-melanocytic benign lesions from SKs and melanocytic benign lesions (CN and JN) from MMs, respectively. If non-melanocytic SKs are considered as less difficult to be correctly identified by many trained dermatologists and can be excluded manually beforehand, then the inclusion of the diameter feature in the 2D and 3D combination has more significance in assisting dermatologists in recognising other more difficult melanocytic and non-melanocytic benign lesions. If classifying SKs automatically is also important to reduce the cost of human involvement, it seems that a more sophisticated classification system involving multistage ensemble design maybe a right way forward. In this case, the border feature can be used to exclude non-melanocytic SKs as malignant in the first stage, and the diameter feature can be used to exclude many small benign melanocytic lesions such as CN and JN.

5 Conclusion

A computer-assisted diagnosis system of malignant melanoma consists of three steps (1) data acquisition, (2) feature extraction/selection and (3) classification. Improvements on lesion classification can be made on all three steps. This paper is focused on the second step, feature extraction and selection. An experimental study is conducted on the many possible combinations of 3D features with the traditional 2D features in current use. Judging on classification performances using a single nonlinear SVM classifier, the many possible feature combinations have demonstrated that the 3D features are useful in improving existing classifications based purely on (1) single feature and (2) combinations of the 2D features.

Out of all the feature combinations, the one including both the 3D feature, the overall skin slant disruption and the 2D features, three colour channel features, and the asymmetry index feature has shown the best overall classification rate. However, the other two unselected ABCD features including border irregularity and diameter have also demonstrated their values. Inclusion of border in the 2D and 3D feature combination has shown promising results with 100 % sensitivity and high specificity for 3 out of 4 lesion groups. The exception is the

group with small compound and junction naevi whose noises are likely to hamper the correct estimation of the border irregularity feature. Inclusion of diameter in the combination has also shown very satisfactory results with 100 % sensitivity and high specificity for 3 out of 4 lesion groups. The exception is the group with SKs, which tend to be large in size, and therefore more likely to be mistaken with many MMs which are also large in size.

Future work can also be carried out to improve the third step, classification by using more sophisticated classifiers such as multistage ensemble classifiers with each stage designed to exclude either SKs or CN/JN, respectively. Another reason for an ensemble design is that based on the current study, a classifier can perform well against a class of lesions if it was trained with the same class lesions or classes that have similar appearances. In a clinical trial where no lesion class is known beforehand, a confidence vote of multiclassifiers seems reasonable where each classifier is trained against different lesion classes and a confidence score is collected to determine the final result. Also, we acknowledge that the experimental data used in this study are relatively small compared to others used in the literature. Therefore, a larger data set is desirable to arrive at more reliable results in future studies. Nevertheless, based on this study, the 3D features have demonstrated clearly its additive values in improving the existing 2D ABCD-based computer-assisted diagnosis of malignant melanoma.

Appendix 1: Locating a lesion's centre of mass and principal axis using moment

The moment of order $(p + q)$ for an $M \times N$ digital image is given by

$$m_{pq} = \sum_{m=1}^{M} \sum_{n=1}^{N} m^p n^q S(m, n) \qquad (20)$$

The centralised moments are given by

$$m_{pq}^c = \sum_{m=1}^{M} \sum_{n=1}^{N} (m - m_c)^p (n - n_c)^q S(m, n) \qquad (21)$$

where (m_c, n_c) is the centre of the mass, which is defined as

$$m_c = \frac{m_{10}}{m_{00}} \quad \text{and} \quad n_c = \frac{m_{01}}{m_{00}} \tag{22}$$

and $S(m, n)$ is a binary image generated as

$$S(m, n) = \begin{cases} 1 & \text{if } (m, n) \in S_l \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

where S_l denotes the lesion region. Then, the direction of the principle axis of a lesion is given by

$$\omega_0 = \frac{1}{2} \tan^{-1} \left(\frac{2m_{11}^c}{m_{20}^c - m_{02}^c} \right) \tag{24}$$

References

1. Aribisala B, Claridge E (2005) A border irregularity measure using a modified conditional entropy method as a malignant melanoma predictor. In: International conference on image analysis and recognition, vol 3656, pp 914–921
2. Barata C, Figueiredo M, Celebi M, Marques J (2014) Color identification in dermoscopy images using gaussian mixture models. In: Proceedings of the IEEE international conference on acoustics, speech, and signal processing (ICASSP 2014), pp 3611–3615
3. Blum A, Luedtke H, Ellwanger U, Schwabe R, Rassner G, Garbe C (2004) Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. Br J Dermatol 151(5):1029–1038
4. Burges C (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Disc 2(2):121–167
5. Cancer Research UK (2015) CancerStats Key Facts—Skin Cancer. http://www.cancerresearchuk.org/cancer-info/cancerstats/keyfacts/skin-cancer/. Accessed 15 Jan 2015
6. Celebi M, Zornberg A (2014) Automated quantification of clinically significant colors in dermoscopy images and its application to skin lesion classification. IEEE Syst J 8(3):980–984
7. Celebi M, Kingravi H, Uddin B, Iyatomi H, Aslandogan Y, Stoecker W, Moss R (2007) A methodological approach to the classification of dermoscopy images. Comput Med Imaging Graph 31(6):362–373
8. Cheng Y, Swamisai R, Umbaugh S, Moss R, Stoecker W, Teegala S, Srinivasan S (2008) Skin lesion classification using relative color features. Skin Res Technol 14(1):53–64
9. Claridge E, Hall P, Keefe M, Allen J (1992) Shape analysis for classification of malignant melanoma. J Biomed Eng 14(3):229–234
10. Clawson K, Morrow P, Scotney B, McKenns D, Dolan O (2007) Computerised skin lesion surface analysis for pigment asymmetry quantification. In: International machine vision and image processing conference, pp 75–82
11. Clawson K, Morrow P, Scotney B, McKenns D, Dolan O (2007) Determination of optimal axes for skin lesion asymmetry quantification. In: IEEE international conference on image processing, vol 2, pp 453–456
12. D'Amico M, Ferri M, Stanganelli I (2004) Qualitative asymmetry measure for melanoma detection. In: IEEE international symposium on biomedical imaging: nano to macro, vol 2, pp 1155–1158
13. Demsar J (2006) Statistical comparisons of classifiers over multiple data sets. J Mach Learning Res 7:1–30
14. Ding Y, Smith L, Smith M, Warr R (2007) 3D skin texture analysis for early diagnosis for malignant melanoma. In: Proceedings of medical image understanding and analysis, pp 151–155
15. Ding Y, Smith L, Smith M, Sun J, Warr R (2008) Obtaining 3D malignant melanoma indicators through the analysis of skin tilt pattern and skin slant pattern. In: Proceedings of the MICCAI workshop on microscopic image analysis with application to biology (MIAAB)
16. Ding Y, Smith L, Smith M, Warr R, Sun J (2008) Enhancement of skin tilt pattern for lesion classification. In: IASTED conference on visualization, imaging and image processing, pp 1–6
17. Ding Y, Smith L, Smith M, Sun J, Warr R (2009) Obtaining malignant melanoma indicators through statistical analysis of 3D skin surface disruptions. Skin Res Technol 15(3):262–270
18. Ding Y, Smith L, Smith M, Sun J, Warr R (2010) A computer assisted diagnosis system for malignant melanoma using 3D skin surface texture features and artificial neural network. Int J Model Ident Control 90:370–381
19. Horn B (1986) Robot vision. MIT press, Cambridge
20. Iyatomi H, Oka H, Celebi M, Hashimoto M, Hagiwara M, Tanaka M, Ogawa K (2008) An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. Comput Med Imaging Graph 32(7):566–579
21. Kuncheva L, Whitaker C (2003) Measures of diversity in classifier ensembles. Mach Learn 51:181–207
22. Lee T, McLean D, Atkins M (2003) Irregularity index: a new border irregularity measure for cutaneous melanocytic lesions. Med Image Anal 7(1):47–64
23. Mazzarello V, Soggiu D, Masia D, Ena P, Rubino C (2006) Melanoma versus dysplastic naevi: microtopographic skin study with noninvasive method. J Plast Reconstr Aesthet Surg 59(7):700–705
24. Menzies S, Bischof L, Talbot H, Gutenev A, Avramidis M, Wong L, Lo S, Mackellar G, Skladnev V, McCartny W, Kelly J, Cranney B, Lye P, Rabinovitz H, Oliviero M, Blum A, Varol A, De'Ambrosis B, McCleod R, Koga H, Grin C, Braun R, Johr R (2005) The performance of solarscan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma. Arch Dermatol 141(11):1388–1396
25. Ng V, Benny Y, Fung M, Lee T (2005) Determining the asymmetry of skin lesion with fuzzy borders. Comput Biol Med 35(2):103–120
26. Pellacani G, Grana C, Seidenari S (2006) Algorithmic reproduction of asymmetry and border cut-off parameters according to the ABCD rule for dermoscopy. J Eur Acad Dermatol Venereol 20(10):1214–1219
27. Piantanelli A, Maponi P, Scalise L, Serresi S, Cialabrini A, Basso A (2005) Fractal characterisation of boundary irregularity in skin pigmented lesions. Med Biol Eng Compu 43(4):436–442
28. Rosado B, Menzies S, Herbauer A, Pehamberger H, Wolff K, Binder M, Kittler H, Corona R (2003) Accuracy of computer diagnosis of melanoma: a quantitative meta-analysis. Arch Dermatol 139(3):361–367
29. Rosenfeld A (1974) Compact figures in digital pictures. IEEE Trans Syst Man Cybern 4(2):221–223
30. Round A, Duller A, Fish P (2000) Lesion classification using skin patterning. Skin Res Technol 6(4):183–192
31. Sboner A, Eccher C, Blanzieri E, Bauer P, Cristofolini M, Zumiani G, Forti S (2003) A multiple classifier system for early melanoma diagnosis. Artif Intell Med 27(1):29–44
32. Stoecker W, Li W, Moss R (1992) Automatic detection of asymmetry in skin tumors. Comput Med Imaging Graph 16(3):191–197
33. Sun J, Smith M, Smith L, Coutts L, Dabis R, Harland C, Bamber J (2008) Reflectance of human skin using colour photometric

stereo: with particular application to pigmented lesion analysis. Skin Res Technol 14(2):173–179

34. Sun J, Liu Z, Ding Y, Smith M (2014) Recovering skin reflectance and geometry for diagnosis of melanoma. In: Scharcanski J, Celebi ME (eds) Computer vision techniques for the diagnosis of skin cancer. Springer, Berlin, pp 243–265

35. Tsai D, Chao S (2005) An anisotropic diffusion-based defect detection for sputtered surfaces with inhomogeneous textures. Image Vis Comput 23(3):325–338

36. Umbaugh S, Moss R, Stoecker W (1989) Automatic color segmentation of images with application to detection of variegated coloring in skin tumors. IEEE Eng Med Biol Mag 8(4):43–50

37. Weickert J (1998) Anisotropic diffusion in image processing. Teubner, Stuttgart

38. Zhou Y, Smith M, Smith L, Farooq A, Warr R (2011) Enhanced 3D curvature pattern and melanoma diagnosis. Comput Med Imaging Graph 35(2):155–165