3D surface roughness measurement for scaliness scoring of psoriasis lesions

Article *in* Computers in Biology and Medicine · August 2013

DOI: 10.1016/j.compbiomed.2013.08.009 · Source: PubMed

6 authors, including:

Esa Prakasa
Indonesian Institute of Sciences
21 PUBLICATIONS **34** CITATIONS
SEE PROFILE

Vijanth Sagayan Asirvadam
Universiti Teknologi PETRONAS
185 PUBLICATIONS **336** CITATIONS
SEE PROFILE

Hermawan Nugroho
SEGi University College
32 PUBLICATIONS **142** CITATIONS
SEE PROFILE

Azura Mohd Affandi
Hospital Kuala Lumpur
23 PUBLICATIONS **102** CITATIONS
SEE PROFILE

Computers in Biology and Medicine 43 (2013) 1987–2000

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

3D surface roughness measurement for scaliness scoring of psoriasis lesions

M. Hani Ahmad Fadzil [a,*], Esa Prakasa [a], Vijanth Sagayan Asirvadam [a], Hermawan Nugroho [a], Azura Mohd Affandi [b], Suraiya Hani Hussein [c]

[a] Centre for Intelligent Signal and Imaging Research, Department of Electrical & Electronic Engineering, Universiti Teknologi PETRONAS, Perak, Malaysia
[b] Department of Dermatology, Hospital Kuala Lumpur, Malaysia
[c] KPJ Damansara Specialist Hospital, Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 23 March 2012
Accepted 15 August 2013

Keywords:
Skin surface roughness
Polynomial surface fitting
k-means clustering
Fuzzy *c*-means clustering
Agreement analysis

ABSTRACT

Psoriasis is an incurable skin disorder affecting 2–3% of the world population. The scaliness of psoriasis is a key assessment parameter of the Psoriasis Area and Severity Index (PASI). Dermatologists typically use visual and tactile senses in PASI scaliness assessment. However, the assessment can be subjective resulting in inter- and intra-rater variability in the scores. This paper proposes an assessment method that incorporates 3D surface roughness with standard clustering techniques to objectively determine the PASI scaliness score for psoriasis lesions. A surface roughness algorithm using structured light projection has been applied to 1999 3D psoriasis lesion surfaces. The algorithm has been validated with an accuracy of 94.12%. Clustering algorithms were used to classify the surface roughness measured using the proposed assessment method for PASI scaliness scoring. The reliability of the developed PASI scaliness algorithm was high with kappa coefficients > 0.84 (almost perfect agreement).

1. Introduction

Psoriasis is an incurable but treatable skin disorder. The disease is caused by the immune system that fails to signal the skin to grow new skin cells correctly. Instead, this condition accelerates the growth cycle of the skin cells, from the normal 28 days to 4 days [1]. Psoriasis is characterised as a reddish lesion on the skin surface. The lesion gets thicker with coarse white scales at higher severity levels. The psoriasis disease makes the patient feel distress with itchiness and coldness, and the conditions can prevail for long durations [2]. These conditions can significantly impact quality of life and cause social and psychological problems [3]. Fig. 1 shows examples of plaque psoriasis lesions on several body regions.

In this paper, the problems of skin roughness measurement for psoriasis assessment have been investigated in order to develop a practical solution for use in clinical practice. Psoriasis is a chronic skin disease that affects about 2–3% of the world population [4,5]. Neimann et al. reported that psoriasis prevalences from epidemiological studies around the world conducted from 1964 to 2005, varied from 0.6% to 4.8% [6]. In a recent review, Chandran et al. also reported the variation of psoriasis prevalences from several countries: 0.60–6.50% in Europe, 0.70–3.15% in North and South America and 0.08–4.00% in Africa. For countries in the Asian and Australian regions, the prevalences were reported to be 0.05–5.30% and 2.30–2.57%, respectively [7]. The Dermatological Society of Malaysia reported a psoriasis prevalence of 3% in Malaysia [8]. Periodical medical treatment for a psoriatic patient is important as the disease cannot be cured.

The Psoriasis Area and Severity Index (PASI) scoring method is considered by dermatologists as the gold standard for the severity assessment and evaluation of treatment efficacy [9]. The PASI assessment is, however, not applied in daily practice because it is laborious and time consuming; it takes nearly 30 min to assess the scores completely for a patient. It is also known that the scoring can be subjective resulting in inter and intra-rater variability. To perform PASI scaliness scoring, representative lesions at four body regions representing the head, trunk, upper limbs, and lower limbs are selected. Both visual and tactile senses are used to determine the score.

This paper focuses on the PASI scaliness assessment based on an objective roughness measurement of 3D images of psoriasis lesions. The studies for other PASI parameters, such as lesion area, erythema, and thickness have been reported by the authors in [10–14].

The remainder of the paper is organised as follows: Section 2 elaborates the related work on the measurement of skin surface roughness and the research problem formulation. Section 3 describes the development of the surface roughness algorithm for skin surfaces. Section 4 provides the agreement analysis on the PASI scaliness scoring and Section 5 concludes the paper.

* Corresponding author. Tel.: +605 368 8586.
E-mail address: fadzmo@petronas.com.my (M.H. Ahmad Fadzil).

0010-4825/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compbiomed.2013.08.009

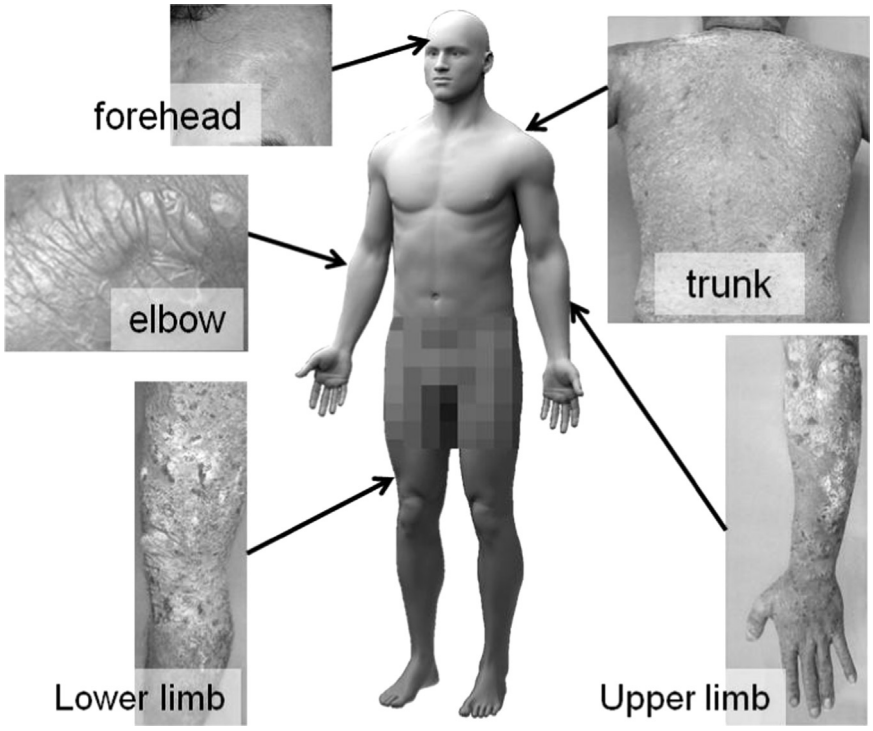

Fig. 1. Plaque psoriasis lesions.

2. Related work and problem formulation

2.1. Measurement of skin surface roughness

Alternatively, it is possible to assess psoriasis scaliness by characterising the lesion surface. Surface characterisation has been widely used in diverse application fields, such as road monitoring, material industry, remote sensing, and medical engineering. Skin surface roughness is a common parameter in skin condition assessment. From the evaluation of skin surface parameters (skin roughness, scaliness and *stratum corneum* hydration), it has been found that roughness varies with different anatomic areas and age groups [15]. Initially, skin roughness is evaluated by using a stylus profilometer on a negative impression of the skin surface that is made from silicone [16]. Skin roughness measurement can be categorised into two methods, invasive and non-invasive methods [17]. Invasive methods include scanning electron microscopy on skin tissue samples [21], shadow measuring [18–21] and 3D stylus profilometry on a skin surface replica [22,23]. In invasive methods, sample skin tissues from patients are taken. In non-invasive methods replicas of the skin are made instead. An example of a non-invasive method for skin roughness measurement is the tribo-acoustical system [24,25]. This system applies an acoustic pressure to the skin surface, which is then converted into roughness information. A precise and automatic system is required for controlling the normal load and sliding velocity of the roughness sensor. Laser profilometry has been used to determine *in vivo* skin surface roughness but its drawback is that it cannot provide fast scanning [26,27].

The digital image analysis has been used for measuring the surface roughness parameters of skin [28]. Surface roughness can be measured from the speckle texture of the surface. The image of the speckled texture is projected from the scattering of coherent light on rough surfaces. The contrast of the speckle has been found to be related to the surface roughness [29]. The reflected illumination is obtained from two coherent plane waves that are emitted from two slightly different angles [30]. Texture analysis methods, such as co-occurrence matrix [31] and fractal analysis [32] have been applied to determine roughness from speckled images. In a recent research, Tchvialeva et al. [33] proposed the speckle contrast method to determine surface roughness. The results have been validated with the results obtained from the fringe projection method. The stripe projection method has been applied in the non-contact 3D surface measurement of the skin roughness [34] of wrinkled skin [35] and skin with acne scarring [36]. In this method, a parallel stripe pattern is projected onto the skin surface.

The image of the projected surface is acquired using a 2D camera. A 3D surface image can be constructed from a series of images with varied stripe sizes. The stripe projection method enables a fast and *in vivo* measurement of the skin surface. For accurate measurement, the camera lens has to be positioned parallel to the skin surface. However, measurement inaccuracies can occur for deep-pitted skin surfaces. Hoppe and Sauermann applied the Discrete Fourier Transform (DFT) to skin profiles to investigate differences in the skin roughness of untreated and treated skin [37]. Although it was found that there was no differentiation in the spectra of the untreated and treated skins, the power spectrum of the amplitude of the skin profile shows the treatment progression of the skin. The treatment progression is interpreted from the decreasing of the power spectrum values. In assessing the scar surface roughness of burnt skin, the S_a parameter has been found more reliable than the S_z (average of 5 highest peaks and 5 deepest valleys of measured profile) and peak count (number of peaks per 0.1 mm length) parameters [38]. High interclass correlation coefficients (>0.90) of S_a were obtained ;in intra- and inter-rater variability evaluations.

2.2. Problem formulation

Lesion area, erythema, scaliness and thickness are the four parameters associated with the PASI score and they are assessed separately using specific procedures. The scores from each region are weighted and totalled accordingly to provide a PASI score ranging from 0 to 72. At least a 75% reduction in the PASI score is

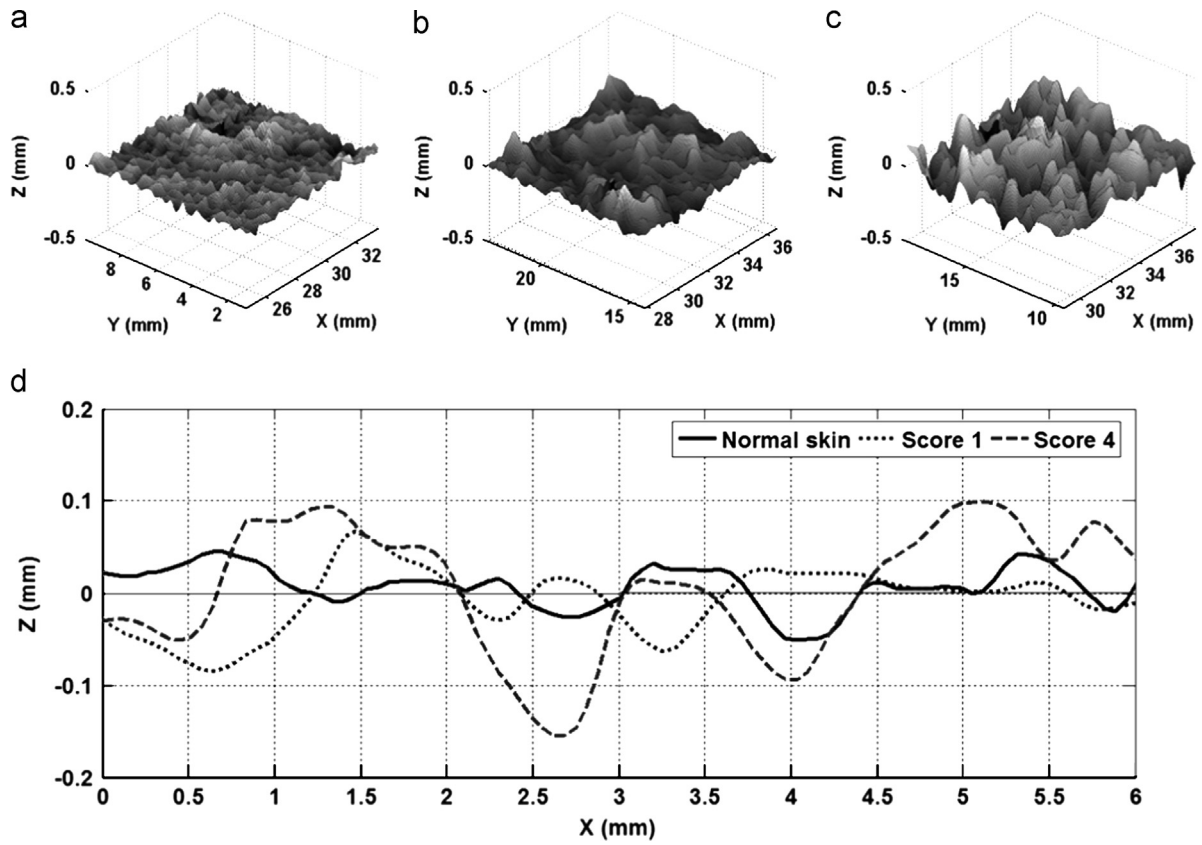

Fig. 2. 3D surfaces of (a) normal skin, psoriasis lesion, (b) score 1, and (c) score 4; (d) rough profiles of normal skin (solid line), lesion score 1 (dotted line), and lesion score 4 (dashed line).

required for the treatment efficacy to be considered a clinically meaningful improvement.

In this paper, the study has been limited to PASI scaliness assessment whereas the studies for the other PASI parameters such as lesion area, erythema, and thickness have been reported in [10–14]. Here, a digital image analysis method has been developed to objectively assess the PASI scaliness of psoriasis lesions by defining the PASI scaliness visual descriptors in terms of surface roughness. The surface roughness of psoriasis lesions on the human body is a feature that is measurable using 3D imaging approaches and can be used to differentiate PASI scaliness scores. Dermatologists assess the PASI scaliness severity based on the visual and tactile appearances of the lesions. The scaliness severity is scaled into four scores, from score 1 to score 4. The higher the score, the more severe the condition is. For the proposed system, the surface roughness feature has been used instead to represent the scaliness severity. It has been observed that as a psoriasis lesion gets more severe, the lesion surface becomes rougher due to the appearance of coarser scales.

Rougher surfaces, which are caused by irregular stacks of dead skin cells, are found with psoriasis lesions of higher severities. A 3D optical scanner is now available to perform a fast surface scan at high resolution. Fig. 2(a)–(c) describes the 3D correspondence between skin surface roughness and scaliness severities (scores). Fig. 2(d) illustrates the roughness profiles of normal skin; i.e., lesions with score 1 and score 4. In lesions with score 4, large vertical deviations have been observed that are caused by excessive coarse scales irregularly stacked on the lesion surfaces.

The average roughness (S_a) was adopted because it represented the actual vertical deviations of a surface thus minimising the impulse noise. Surface profile frequency was not considered as a roughness parameter because the frequency did not represent the roughness of the skin surface. Standard clustering methods were applied in the classification of the surface roughness of the psoriasis lesions in order to develop a complete psoriasis scoring system that would be both reliable and objective.

3. Description of PASI scaliness scoring method

3.1. Surface roughness algorithm

In this work, the psoriasis lesion surface was scanned using the stripe projection method. As specified by ISO 14460-1, the surface profile is classified into three components namely, roughness, waviness and form [39]. The surface wavelength or peak-to-peak spacing is used to differentiate the surface profile components [39–41]. The roughness is identified as the finer ridges' profile compared to waviness, which is smoother than roughness. The form is a curved surface without the fine ridges and is characterised as a profile with very long-range deviation [41]. The vertical deviations are calculated by subtraction between the lesion surface and its estimated waviness surface as shown in Fig. 3(a).

Here, the surface roughness is determined by averaging the vertical deviations of the lesion surface. In other implementations, several statistical parameters were used to determine the surface roughness, such as average roughness (S_a), root mean square roughness (S_q), maximum profile peak height (S_p), minimum profile valley depth (S_v), maximum height of the profile (S_t) etc. [40]. The vertical or horizontal distance of the roughness profile was extracted as the surface data to determine the roughness parameters [42]. Al-Kindi and Shirinzadeh found that the S_a and S_q parameters were accurate in evaluating surface measurement [43].

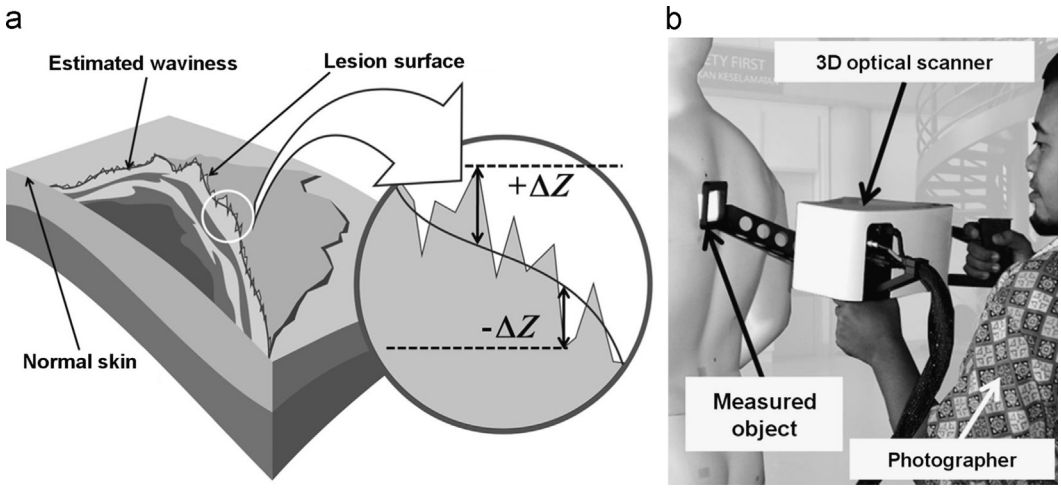

Fig. 3. (a) Cross-sectional view of a psoriasis lesion and estimated waviness surfaces. (b) A scanning process for acquiring scaliness data.

The other amplitude parameters, such as S_p, S_v, and S_t were less accurate because these parameters were influenced by the local maximum peak or valley of the surface profile.

A 3D optical scanner, Phase shift Rapid *In vivo* Measurement of Skin (PRIMOS) portable, was used to acquire the lesion surface images (Fig. 3(b)). The scanner is designed for 3D *in vivo* measurements of skin surface structure [44]. In the scanner system, a 3D surface image is constructed by applying the structured light projection method. This method has some advantages, such as non-contact measurement, high-speed scan (<63 ms) and high-resolution 3D surface. The scanner is able to image a 3D surface area of 40 mm $\times$ 30 mm. The highest spatial and depth resolution achieved are 0.0630 mm and 0.0040 mm, respectively.

As the lesion surface is a curved surface, a polynomial surface fitting is used to fit the lesion surface. Second and third order polynomials are applied to estimate the 3D waviness of the surfaces from the rough lesion surface. The best fit value obtained from among the polynomial orders is used to select the best estimated waviness. Higher polynomial orders (higher than third order) are not selected in estimating the waviness surface. This is to avoid the over-fitting of the roughness profile of the lesion surface [45]. In an over-fitting scenario, the estimated waviness tends to follow the actual profile of the lesion surface rather than the curved skin surface. Therefore, vertical deviations will not be obtained accurately. The optimum sample area for the surface roughness determination has been presented by the authors in [46].

Some other filtering methods are also available for the extraction of waviness and/or form profiles. These methods include Gaussian filtering [47], Fourier transform [48], spline filter [49] and wavelet filter [50]. In order to obtain the reference surface for determining vertical deviations, the Gaussian filter was applied to the surface profile. However, the filter was not able to provide the accurate reference line estimation at the edge regions of the image because the filtered data points were not available in these regions. The line was deflected from the surface curvature. It occurred if the sample size of the profile was small and the form component was more dominant than the roughness and waviness components [40,51].

The Fourier transform has been used to identify a wavelength threshold for roughness–waviness differentiation. However, the input signal was assumed to be comprised of sinusoid signals which did not represent the actual characteristic of the natural surface [40]. Spline filtering can solve problems associated with the Gaussian filter, such as the edge distortion and poor performance of a less undulated surface [40]. The disadvantage of this method is the computation cost for the calculations along the x- and y-coordinates [49]. The wavelet filter can decompose the surface profile into different frequency components (roughness, waviness and form) allowing a roughness analysis to be performed on the surface components separately [50]. As with the spline filter, the wavelet filter also requires intensive computation [40] and large sample points to increase the precision in the representation.

As reported in an earlier work [52], the polynomial surface fitting algorithm was applied to determine the reference surface for the surface roughness assessment of a curved surface. This algorithm is suitable for small areas with minimal outliers. Quadratic polynomial surface fitting is adequate for most nominally curved surfaces, such as spherical and cylindrical surfaces with less computation and resource allocation costs. Vertical deviations were determined by subtracting the rough lesion surface from the estimated waviness surface to obtain the deviation surface. These polynomial based surface fittings are suitable for small surface areas [52] as found in psoriasis lesion assessment. In small areas, the vertical undulation of the estimated waviness is less and can be accurately fitted with the second or third order polynomials. The general form of polynomials can be written by the following equations [53]:

The second order polynomial

$$z_2(x,y)=(a_1x^2+a_2x+a_3)y^2+(a_4x^2+a_5x+a_6)y+(a_7x^2+a_8x+a_9) \tag{1}$$

The third order polynomial

$$z_3(x,y)=(a_1x^3+a_2x^2+a_3x+a_4)y^3+(a_5x^3+a_6x^2+a_7x+a_8)y^2 \times(a_9x^3+a_{10}x^2+a_{11}x+a_{12})y^2+(a_{13}x^3+a_{14}x^2+a_{15}x+a_{16}) \tag{2}$$

To create a surface based on a polynomial equation, polynomial coefficients are determined for the second or third polynomials. The coefficient of determination (R^2) is used to measure the fittingness of the polynomial fitting [54]. A good fit can be obtained if R^2 is above 0.9. The best fitting result of either of these two polynomial orders (second or third) is selected based on the highest R^2. The equation of R^2 is expressed in the following equation:

$$R^2=1-\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(z(x_i,y_j)-w_k(x_i,y_j))^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}(z(x_i,y_j)-\bar{z})^2} \tag{3}$$

Variable $z(x_i, y_j)$ represents the elevation of the lesion surface, $\bar{z}$ represents the elevation average and $w_k(x_i, y_j)$ is the fitted value at (x_i, y_j) using the kth order polynomial. The surface roughness is determined by using the average roughness (S_a) equation [42]. In this equation, the surface roughness is calculated by averaging the absolute vertical deviation of all data points. The average roughness, S_a, is defined as the equation below. Variable $e(x_i, y_j)$ denotes the vertical deviation of the lesion surface at location (x_i, y_j).

$$S_a = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|z(x_i,y_j) - w_k(x_i,y_j)| = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|e(x_i,y_j)| \quad (4)$$

The surface roughness measurement method reported by present authors in their previous work [55] found to cause over-fitting on a large lesion area. To avoid over-fitting, the surface roughness method is applied to smaller areas by dividing the sample area into four subdivided surfaces [56]. The surface roughness algorithm is described by the flow chart shown in Fig. 4. In this flow chart, the fitting process of the second and the third orders are performed separately in order to get the best fit based on R^2. A 3D lesion surface matrix, $Z_0(x,y)$, with size $M \times N$, is assigned to the algorithm. The coordinates, x, y, and its value, $z(x,y)$ are then used to calculate some polynomial coefficients. The number of data points in the lesion surface is $M \times N$.

The lesion surface is then divided into 2×2 subdivided surfaces to give four subdivided surfaces namely, D_1, D_2, D_3 and D_4. Half division is applied symmetrically to each side of the lesion area, thus the size of each subdivided surface becomes $(M/2)(N/2) = L$. Fig. 5 displays the division process of a 3D lesion surface into four subdivided surfaces.

The k variable represents the total number of loops. This variable is used to count the subdivided surface that has been measured. Here, the maximum number of loops is 4. For each subdivided surface, D_1–

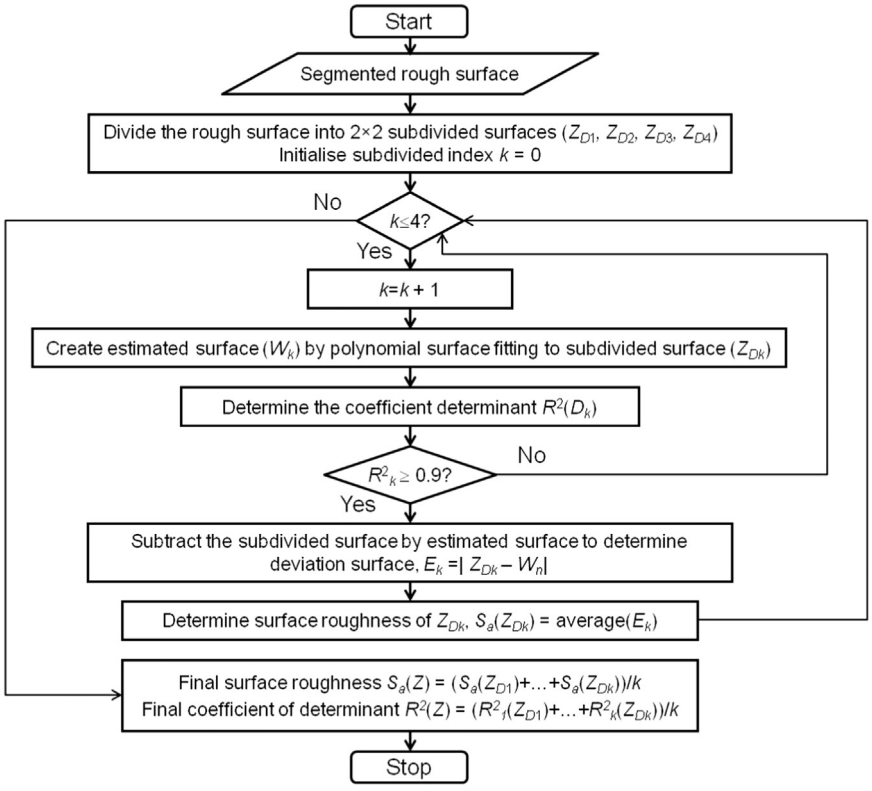

Fig. 4. Flow chart of the surface roughness algorithm.

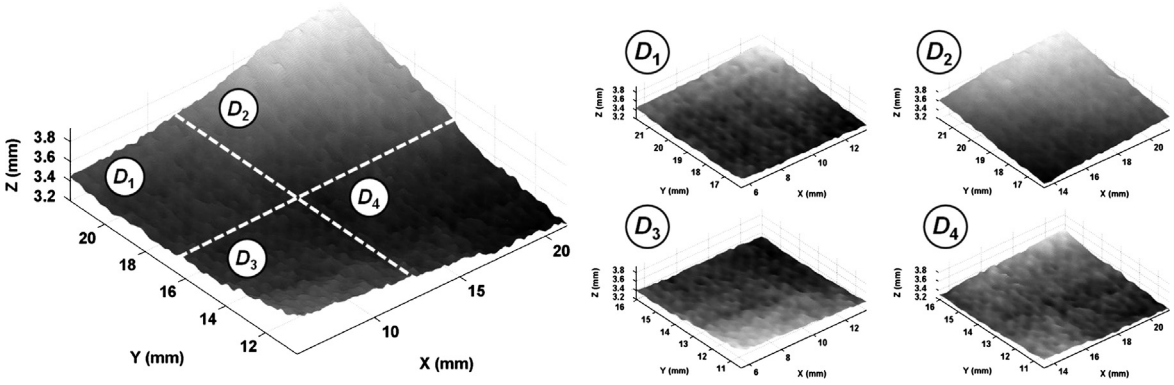

Fig. 5. The segmented rough surface of the lesion model and its 4×4 subdivided surfaces.

D_4, an estimated waviness needs to be constructed. To create an estimated waviness based on the high order polynomial, the polynomial coefficients are determined from coordinates x, y, and $z(x, y)$ through a matrix inversion. This inversion is performed separately for all subdivided surfaces as follows:

$$VA = Z \quad A = V^{-1}Z \tag{5}$$

where matrix V contains the elements of the polynomial equation, the coefficients of the polynomial equation are stored in matrix A, and matrix Z represents the subdivided lesion surface.

Once the coefficients A have been determined, the estimated waviness surface can be constructed. The waviness is obtained by applying calculated polynomial equation with polynomial coefficients A at the evaluated points, (x, y). The matrix equation of (5) can be written as $V_2A_2 = Z_{D_k}$ to denote a second order polynomial fitting at subdivided surfaces D_k, $(k = 1, 2, 3, 4)$. The elements of matrices V_2, A_2 and Z_{D_k} are given by the following forms:

$$V_2 = \begin{bmatrix} x_1^2y_1^2 & x_1y_1^2 & y_1^2 & \cdots & 1 \\ x_2^2y_2^2 & x_2y_2^2 & y_2^2 & \cdots & 1 \\ x_3^2y_3^2 & x_3y_3^2 & y_3^2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_L^2y_L^2 & x_Ly_L^2 & y_L^2 & \cdots & 1 \end{bmatrix} \tag{6}$$

Here, variable L refers to the total number of data points at a subdivided surface:

$$A_2 = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_8 & a_9 \end{bmatrix}^T \tag{7}$$

$$Z_{D_1} = [z_{D_1}(x_1, y_1) \quad z_{D_1}(x_2, y_2) \quad \cdots \quad z_{D_1}(x_L, y_L)]^T \tag{8}$$

The matrix of polynomial coefficients A_2 is unknown and the inversion $A_2 = V_2^{-1}Z_{D_k}$ is applied to determine its values. The determination of polynomial coefficients A_2 can be written as

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_9 \end{bmatrix} = \begin{bmatrix} x_1^2y_1^2 & x_1y_1^2 & y_1^2 & \cdots & 1 \\ x_2^2y_2^2 & x_2y_2^2 & y_2^2 & \cdots & 1 \\ x_3^2y_3^2 & x_3y_3^2 & y_3^2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_L^2y_L^2 & x_Ly_L^2 & y_L^2 & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} z_{D_1}(x_1, y_1) \\ z_{D_1}(x_2, y_2) \\ z_{D_1}(x_3, y_3) \\ \vdots \\ z_{D_1}(x_L, y_L) \end{bmatrix} \tag{9}$$

Fig. 6 shows various surface profiles, e.g., lesion surface, estimated waviness surface, and deviation surface. The waviness surface is first estimated by applying the second order polynomial surface fitting with coefficients A_2. In this figure, the waviness surface is overlaid by a transparent lesion surface, $Z_{D_k}(x, y)$, to provide an impression of the fitting process. The estimation is limited for (x, y) coordinates at the D_k area. An equation for determining $W_2(x, y)$ is shown by following expressions:

$$W_2 = V_2A_2 \tag{10}$$

$$\begin{bmatrix} w_2(x_1, y_1) \\ w_2(x_2, y_2) \\ w_2(x_3, y_3) \\ \vdots \\ w_2(x_L, y_L) \end{bmatrix} = \begin{bmatrix} x_1^2y_1^2 & x_1y_1^2 & y_1^2 & \cdots & 1 \\ x_2^2y_2^2 & x_2y_2^2 & y_2^2 & \cdots & 1 \\ x_3^2y_3^2 & x_3y_3^2 & y_3^2 & \vdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_L^2y_L^2 & x_Ly_L^2 & y_L^2 & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_L \end{bmatrix} \tag{11}$$

The coefficient of determination R^2 is calculated to evaluate the fitting result. Here, $W_2(x, y)$ is accepted if R^2 is within [0.9, 1.0]. The equation for determining R^2 is provided in (3). The notation $R_2^2(D_k)$ is used to denote the R^2 of the second order polynomial surface fitting at the subdivided surface, D_k. Therefore, (3) can be rewritten as

$$R_2^2(D_k) = 1 - \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(z_{D_k}(x_i, y_j) - w_k(x_i, y_j))^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}(z_{D_k}(x_i, y_j) - \bar{z}_{D_k})^2}, \quad k = 1, 2, 3, 4 \tag{12}$$

If $R_2^2(D_k)$ is computed ≥ 0.9 then the surface roughness, S_a, is determined at the subdivided surfaces, D_k. The deviation surface, $E_2(x, y)$, as shown in Fig. 6(c), is determined by subtracting the estimated waviness, $W_2(x, y)$, from the lesion surfaces, $Z_{D_k}(x, y)$, as seen in Fig. 6(a). The input variables for this equation are deviation surface, $E_2(x, y)$, with matrix size $(M/2)(N/2)$.

$$S_{a,2}(D_k) = \frac{1}{((M/2)(N/2))}\sum_{i=1}^{M/2}\sum_{j=1}^{N/2}|z_{D_k}(x_i, y_j) - w_2(x_i, y_j)|, \quad k = 1, 2, 3, 4 \tag{13}$$

Notation $S_{a,2}(D_k)$ is used to represent the surface roughness of the subdivided surfaces, D_k, that is determined by the second order polynomial surface fitting. These computations are performed separately for each subdivided surface. The overall surface roughness of the lesion surface is obtained by averaging the surface roughness of the subdivided surfaces. The surface roughness of a subdivided surface will not be included in the final calculation if its R^2 is not within an acceptable interval [0.9, 1.0]. Thus, the overall surface roughness of the lesion surface can be expressed as

$$\overline{S_{a,2}} = \frac{\sum_{k=1}^{N_D} S_{a,2}(D_k)}{N_D} \text{ if } 0.9 \leq R_2^2(D_k) \leq 1.0, \quad N_D \leq 4 \tag{14}$$

The overall coefficient of determination R^2 is computed as well by applying the equation below:

$$\overline{R_2^2} = \frac{\sum_{k=1}^{N_D} R_2^2(D_k)}{N_D} \text{ if } 0.9 \leq R_2^2(D_k) \leq 1.0, \quad N_D \leq 4 \tag{15}$$

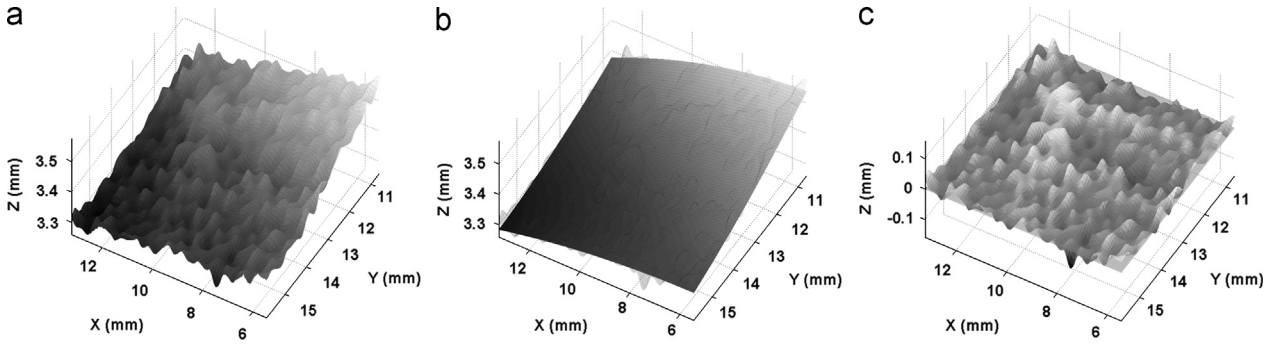

Fig. 6. The 3D surfaces involved in the surface roughness determination: (a) lesion surface, (b) estimated waviness, and (c) deviation surface ($S_{a,2}(D_1)$=0.012 mm and R_2^2=0.917).

3.2. Validation of surface roughness algorithm

A validation study of the surface roughness algorithm has been conducted in order to verify the assessment approach on human like curved surfaces. To model a skin lesion, a surgical tape that has a regular and uniform texture on its surface was used. It was made from an elastic material in order to adapt to the curvature of any surface. To preserve its texture characteristics, the tape had to be pasted onto a smooth surface only. The average surface roughness of 33 lesion models, $\overline{S_a}$, was used as the reference. These reference models were pasted on a hard paperboard flat surface (Fig. 7). By applying the lesion models on the flat surface, vertical undulations due to a curved surface would not occur. Thus, the vertical deviations of the lesion model only contributed to the roughness calculation.

In this validation, the lesion model was scanned three times by applying a series of scans on the scanner mode. The scanning process of the 33 lesion models resulted in 99 3D surface images. The surface roughness of a lesion model was represented by the average of the surface roughness from these three consecutive scans. The average surface roughness, $\overline{S_a}$, and its standard deviation were found to be 0.0122 ± 0.0011 mm ($\bar{x}_{Ref} \pm \sigma_{Ref}$). To validate the surface roughness algorithm on the curved surfaces of human skin, the lesion models were placed on the surfaces of specially fabricated mannequins for medical purposes. The validation on the mannequin used a similar tape material. The number of lesions for the body region were 6 (head), 87 (upper limb), 159 (trunk), and 138 (lower limb). The validation results presented in this study improved from those previous studies that were reported by present authors [55,56]. In the previous study [55], surface locations could not be imaged accurately due to rigidity of human model skin surface. In the successive study [56], mannequin surface with elastic material is used, to enable the 3D optical scanner to obtain a focused image. This paper involved more lesion models (in form of 2D image and 3D surface) using mannequin which are shown in Fig. 8. Here, lesion models were distributed and pasted onto life-size mannequin to represent the lesions.

Three hundred and ninety lesion models were pasted and distributed onto several locations of the mannequin's body. The lesion models were placed in the centre point of the grid arrangement with a size of 40×30 mm^2. The centre point was selected to provide an accurate image because a highly focused area is normally obtained at this point. The regions of the lesion models were defined following four regions (head, upper limb, trunk, and lower limb) for PASI scoring. Since each model required a certain rectangular space, the number of lesions was limited by the available areas for each body region.

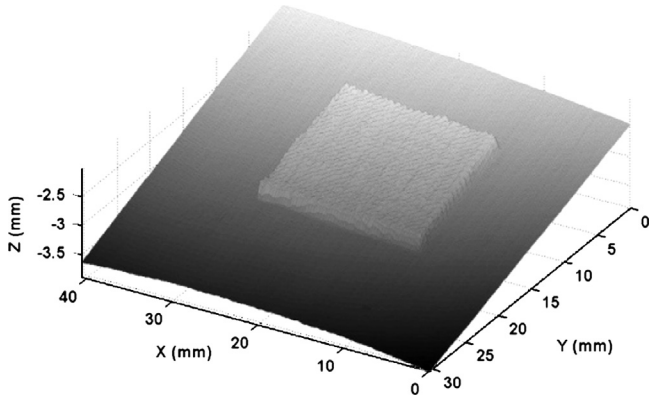

Fig. 7. A 3D surface of lesion model of surgical tape is pasted on the flat surface of a paperboard.

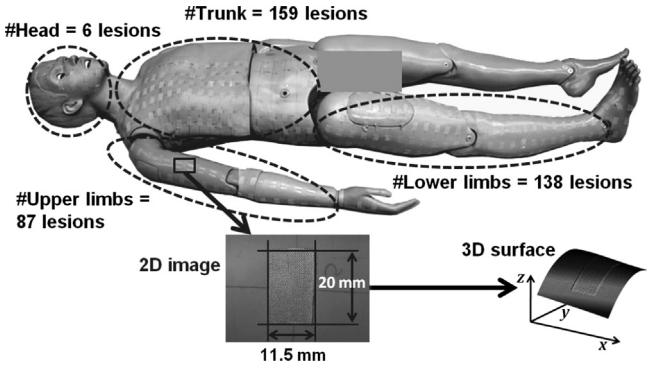

Fig. 8. Lesion models on life-size mannequin.

The validation process determined the reliability of the roughness measurement on a 3D curved skin surface. The algorithm was considered valid if the measured surface roughness of the surgical tape (lesion model) was constant at any location on the skin surfaces. In this study, the surface of a mannequin's skin was used to simulate the human skin surfaces. The average surface roughness, $\overline{S_a}$, and the standard deviation of the lesion models were found to be 0.0130 ± 0.0013 mm ($\bar{x}_{Curve} \pm \sigma_{Curve}$).

To determine the accuracy of the surface roughness measurement, the average S_a of the 33 lesion models on a flat surface was used as a reference ($S_{a(Flat)}$). The $S_{a(Flat)}$ value was found to be 0.0122 ± 0.0011 mm ($\bar{x}_{Ref} \pm \sigma_{Ref}$) and the S_a of the lesion models on the mannequin surface ($S_{a(Curve)}$) was found to be 0.0130 ± 0.0013 mm ($\bar{x}_{Curve} \pm \sigma_{Curve}$). The surface roughness of the lesion models on flat and mannequin surfaces are shown in histograms in Fig. 9(a) and (b). The mean of the curved surface histogram was shifted to the right compared to the flat surface histogram. It indicated that surface curvature of the mannequin surface introduced a positive error to the surface roughness calculation.

To analyse the measurement reliability of the surface roughness algorithm, the following equations were applied to the statistical values of the surface roughness of the flat and curved surfaces. The calculation results of these equations are described in Table 1.

$$Error = |\bar{x}_{Curve} - \bar{x}_{Ref}| \tag{16}$$

$$\sigma_{Total} = \sqrt{\sigma_{Ref}^2 + \sigma_{Curve}^2} \tag{17}$$

$$Accuracy = \left[1 - \frac{|\bar{x}_{Curve} - \bar{x}_{Ref}|}{\bar{x}_{Ref}}\right] \times 100\% \tag{18}$$

The computed accuracy was the minimum accuracy which implied that the algorithm accuracy was not less than 94.12%. Based on the calculated standard deviation ($\sigma_{Total} = 0.0017$ mm), the 95% confidence interval ($Z_{\alpha/2} = 1.96$) for the calculated values were in the range of ± 0.0019 mm for the sample provided. This measurement range showed that it lay within the depth resolution of the PRIMOS camera (0.0040 mm) [57]. The algorithm accuracy of this work is better than the accuracy presented in the previous works [55,56]. The accuracy has been improved from $\geq 80\%$, $\geq 89.30\%$, to $\geq 94.12\%$.

3.3. Classification methods for PASI scaliness scores

In previous works [55], the proposed algorithm was not implemented on the actual psoriasis lesions. A repeatability study of surface roughness measurement on psoriasis lesions has been conducted by present authors in [56]. This paper implements surface roughness algorithm on the actual psoriasis lesions with built in classification

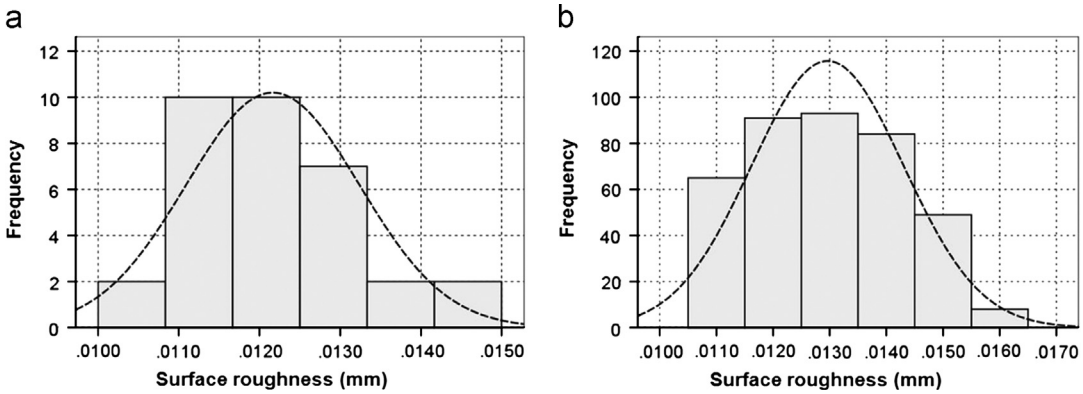

Fig. 9. The histograms of the lesion models pasted on (a) flat and (b) curved mannequin surfaces.

Table 1
Performance evaluation of the surface roughness algorithm.

Surfaces	Surface roughness parameter		
	$\bar{x}$ (mm)	σ (mm)	Accuracy (%)
Reference	0.0122	0.0011	
Curve	0.0130	0.0013	
Absolute Error	$\lvert 0.0130-0.0122\rvert = 0.0008$	$\sqrt{0.0011^2+0.0013^2} = 0.0017$	$\left(1-\frac{0.0008}{0.0122}\right)\times 100\% = 94.1\%$

system for scoring PASI scaliness. This 3D surface roughness measurement and scoring techniques was adopted in a real clinical study that was conducted in hospital with presence of dermatologists.

The dataset of the surface roughness measurement was 1999 psoriasis lesion images obtained from 204 registered psoriasis patients at Hospital Kuala Lumpur, Malaysia. There are two types of assessments performed by the developed system. The first type namely, the single assessment, is performed only once for each lesion sample. The measurement of the second type namely, the double assessment, is performed two times for each lesion sample.

The psoriasis lesion images were divided into two groups: training and test datasets. The training dataset consisting of 1351 images was used to train the clustering algorithm. Most of these images (93.3%) were obtained from the single assessment of the psoriasis lesions. There were several images obtained from the double assessment. However, these images were not used as training data if either the first or second assessment gave $R^2 < 0.9$ or the difference between the first and second assessment was higher than the two standard deviations of the assessment difference ($> 2\sigma_{Diff}$). The assessment difference was computed from the surface roughness difference between the first and second assessments. It was ensured that there were no duplications in the training dataset and all of the lesions were fitted with $R^2 \geq 0.9$. The test dataset consisted of 648 scanned images acquired from the double assessments on 324 lesions. 1892 lesions have been scored 1, 2, 3 or 4 by dermatologists. From the dermatologist's assessment, the score distributions were: 1266 (66.9%) score 1, 409 (21.6%) score 2, 153 (8.1%) score 3 and 64 (3.4%) score 4. It can be seen that the proportion of lesions with low scores was greater than the high scores due to the treatment received.

Two standard clustering methods, *k*-means and Fuzzy *c*-Means Clustering (FCM) algorithms, were applied for the scaliness clustering. These clustering algorithms were considered because there was no ground truth available to implement supervised clustering. By applying the *k*-means clustering [58,59] as an unsupervised clustering algorithm to the surface roughness dataset, the psoriasis lesions were classified into four score groups. The boundary value between two adjacent score groups was determined by finding the middle point between the calculated centroids of the adjacent groups. The centroids of the PASI scores for the dataset were found to be 0.024 mm (Score 1), 0.040 mm (Score 2), 0.061 mm (Score 3) and 0.106 mm (Score 4). The boundary value between Score 1 and Score 2 was determined by finding the middle point between the Score 1 and Score 2 centroids. This method was also applied to determine the next boundary levels. The boundary levels of the surface roughness for the PASI scaliness scores are given in Table 2. The PASI scaliness scores were then determined by applying the boundary levels to the surface roughness of the psoriasis lesions.

Table 2
Boundary levels of the surface roughness for the PASI scaliness scores.

Score	Boundary levels
1	$S_a \leq 0.032$ mm
2	$0.032 < S_a \leq 0.051$ mm
3	$0.051 < S_a \leq 0.084$ mm
4	$S_a > 0.084$ mm

Table 3
Score centroids of subsets 1 and 2.

PASI scaliness score	Subset 1 (N=676)		Subset 2 (N=675)		Centroid difference (mm)
	Score centroid (mm)	N	Score centroid (mm)	N	
1	0.024	226	0.024	227	0.001
2	0.038	247	0.038	264	0.000
3	0.057	169	0.059	155	0.004
4	0.104	34	0.104	29	0.000

The ability of the *k*-means clustering algorithm to classify psoriasis lesions based on surface roughness has been validated. Validation was conducted by applying the *k*-means clustering algorithm to the divided datasets. The dataset was randomly partitioned

into two equal-sized subsets. The consistency of the score centroids for both subsets showed that the scores were properly clustered (maximally separated). The score centroids were stable since the sample size of each score was statistically acceptable ($N \geq 30$). Table 3 shows the stability of the score centroids of the partitioned subsets. The centroid differences of each score between Subsets 1 and 2 were smaller than 0.004 mm. To observe the centroid consistency on a smaller sample size (Table 4), the dataset is randomly partitioned into 3 subsets (Subsets 1–3). The centroid consistency of the subsets can be observed in two and three partitions. The centroid differences were not more than 0.005 mm except for the centroid difference of score 4 at three partitions which gave 0.008 mm. The centroid difference was greater due to the size of the clustered dataset being less than 30 samples.

The adaptation of the FCM clustering [60] with membership degrees of the training dataset were obtained, clustered and scattered according to the existing clusters as depicted in Fig. 10.

From the FCM clustering results, it was found that the membership degrees were not calculated at the surface roughness values outside the values of the training dataset. Therefore, a membership function was required to determine the membership degrees or probabilities of the surface roughness at any value. The membership degree or probability of a particular roughness (S_a) belonging to a scaliness score was determined by applying the membership functions. Gaussian functions were applied to fit each of the score clusters. The general form of the Gaussian equation for representing the membership function was given by (19). The subscript n on the variable $P_n(S_a)$ represented the group of the scaliness score. Eq. (19) was constructed from the summation of two Gaussian fitting functions, which had different means and standard deviations. For the first Gaussian function, coefficient a_1 is the maximum height of the distribution, b_1 is the mean of the Gaussian distribution, and c_1 is the standard deviation. The same meanings are also applied for the second Gaussian function.

$$P_n(S_a) = a_1 \exp\left(-\left(\frac{S_a - b_1}{c_1}\right)^2\right) + a_2 \exp\left(-\left(\frac{S_a - b_2}{c_2}\right)^2\right) \quad (19)$$

Table 5 shows the coefficients of the membership functions that are represented as Gaussian functions for the roughness classification in the PASI scaliness scoring. All of the membership functions were fitted with $R^2 \approx 1$. The membership functions of the PASI scaliness scores are shown in Fig. 11.

The PASI scaliness score was determined by applying the rules of scaliness score as formulated in (20). An input surface roughness was classified into a particular score if the input gave the highest probability among the scaliness scores. Several lesion surfaces that have been scored by FCM algorithm are shown in Fig. 12.

$$S(S_a) = \begin{cases} 1, & \max(P_1(S_a), P_2(S_a), P_3(S_a), P_4(S_a)) = P_1(S_a) \\ 2, & \max(P_1(S_a), P_2(S_a), P_3(S_a), P_4(S_a)) = P_2(S_a) \\ 3, & \max(P_1(S_a), P_2(S_a), P_3(S_a), P_4(S_a)) = P_3(S_a) \\ 4, & \max(P_1(S_a), P_2(S_a), P_3(S_a), P_4(S_a)) = P_4(S_a) \end{cases} \quad (20)$$

Given 324 data points used as the test dataset, 289 (89.2%) data points were successfully classified as the same cluster in the first and second assessments. However, 35 (10.8%) data points of the double assessments were classified in different clusters. The misclassifications were analysed and were able to be categorised into three types based on the location of the occurrences.

The first type was the misclassification in determining either score 1 or score 2. Score 1 might have been found in the first assessment, and then followed by score 2 in the second assessment. The second type was the classification uncertainty of score 2 or 3. The last type was the scoring vagueness between score 3 and 4. Two different scores were able to be obtained in the second and the third types as exemplified by the first type. These measurement inconsistencies occurred when the surface roughness data points were located too close to the score boundary levels.

Table 4
Score centroids of subsets 1–3.

PASI scaliness score	Subset 1 (N=451)		Subset 2 (N=450)		Subset 3 (N=450)		Centroid difference (mm)
	Score centroid (mm)	N	Score centroid (mm)	N	Score centroid (mm)	N	
1	0.024	156	0.023	164	0.025	158	0.001
2	0.040	180	0.038	155	0.039	164	0.001
3	0.061	90	0.058	117	0.058	96	0.002
4	0.102	25	0.109	14	0.097	32	0.008

Table 5
Coefficients of the Gaussian functions for the surface roughness classification of the PASI scaliness scores.

PASI scaliness score	a_1	b_1	c_1	a_2	b_2	c_2	R^2
Score 1 (S_1)	0.609	0.029	0.007	0.888	0.018	0.011	0.994
Score 2 (S_2)	0.514	0.038	0.006	0.858	0.048	0.010	0.992
Score 3 (S_3)	0.635	0.064	0.010	0.784	0.081	0.018	0.995
Score 4 (S_4)	0.708	0.118	0.023	0.792	0.160	0.039	0.997

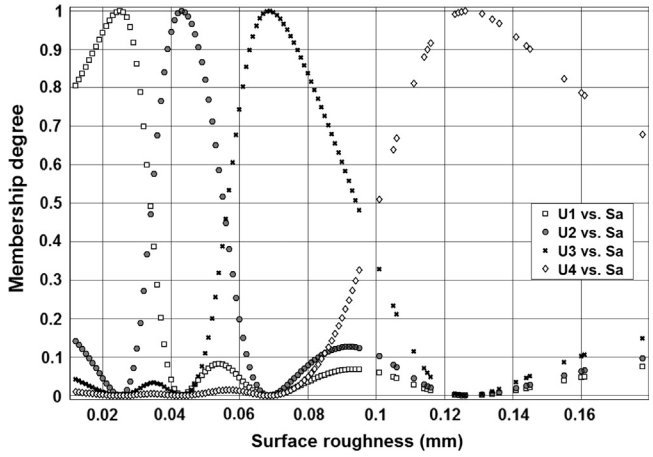

Fig. 10. Membership degrees scattering of the clustered dataset.

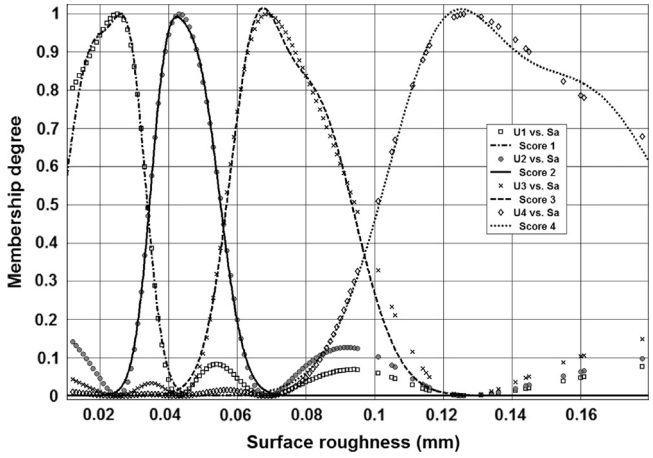

Fig. 11. Membership functions of the PASI scaliness scores.

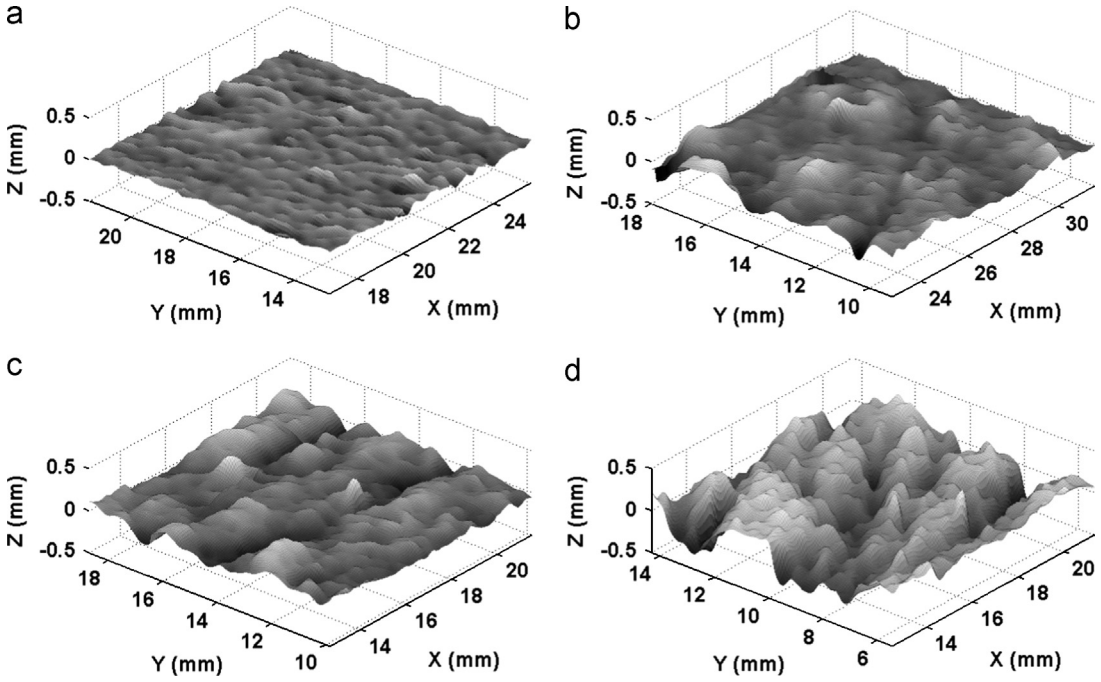

Fig. 12. Deviation surfaces of psoriasis lesions: (a) score 1, (b) score 2, (c) score 3, and (d) score 4, determined by the surface roughness algorithm and classified by the FCM algorithm.

As described in Table 2, three boundary levels were defined. These boundary levels distinguished a score cluster from the other clusters. The levels labelled as B_1, B_2, and B_3. B_1 was used to separate score 1 from score 2. The second boundary was B_2 that divided score 2 from score 3. The last boundary was B_3. This boundary split score 3 from score 4. The values of B_1, B_2, and B_3 were 0.032 mm, 0.051 mm, and 0.084, respectively. The closeness of the measured surface roughness to each boundary level was computed by applying the Euclidean distance equation. The equation is given by

$$D_{B_i}(S_a) = \sqrt{(S_{a_1} - B_i)^2 + (S_{a_2} - B_i)^2}, \quad i = 1, 2, 3 \tag{21}$$

S_{a_1} and S_{a2} were the measured surface roughness from the first and the second assessments. B_i was used to represent the boundary level at i as mentioned in the previous paragraph. The Euclidean distance equation was applied to the test dataset that had been misclassified in the score clustering. In this work, there were 21 misclassification cases of type 1 (boundary of scores 1 and 2), 10 cases of type 2 (boundary of scores 2 and 3), and only 4 cases for the type 3 (boundary of scores 3 and 4). The closeness of the measured surface roughness to the existing boundary levels were determined by applying (21). Furthermore, an average of the Euclidean distance was calculated to indicate the closeness of the misclassification type with a certain boundary level.

The closeness measurement results are given in Table 6. The distances were determined from the misclassification cases on the test dataset. Misclassification type 1 occurred when the measured surface roughness values were close to the boundary level B_1. This closeness was shown by the $\overline{D_{B_1}}$ value (0.003 mm) which was 11% smaller than D_{B_2} (0.028 mm) and D_{B_3} (0.075 mm). Misclassification type 2 was indicated by the smallest value of D_{B_2} (0.003 mm) compared to D_{B_1} (0.026 mm) and D_{B_3} (0.048 mm). For type 3, the trend followed the previous types. The smallest Euclidean distance was obtained at the D_{B_3} (0.004 mm) value. D_{B_1} (0.072 mm) and D_{B_2} (0.046 mm) were found to be much larger than D_{B_3}.

A small Euclidean distance at a certain boundary level was only found in the misclassification cases. Small Euclidean distances could not be found in the data points that were correctly classified. All of the surface roughness data points were located beyond from the boundary levels. Thus, the misclassification cases could be minimised. The average of the Euclidean distance was much higher than 0.005 mm for all of score groups, as listed in Table 7.

Table 6
Average of the Euclidean distances of the surface roughness and boundary levels of the *k*-means and FCM classification algorithms.

Misclassification type	*k*-means				FCM			
	$\overline{D_{B_1}}$ (mm)	$\overline{D_{B_2}}$ (mm)	$\overline{D_{B_3}}$ (mm)	*N*	$\overline{D_{B_1}}$ (mm)	$\overline{D_{B_2}}$ (mm)	$\overline{D_{B_3}}$ (mm)	*N*
Type 1, scores 1 and 2	0.003	0.028	0.075	21	0.003	0.030	0.088	17
Type 2, scores 2 and 3	0.026	0.003	0.048	10	0.033	0.004	0.056	8
Type 3, scores 3 and 4	0.072	0.046	0.004	4	0.091	0.060	0.004	4

To avoid misclassification, it is recommended that more than a single data point should be collected if the surface roughness measured is near the boundary levels. The surface roughness input is then determined by averaging the surface roughness values of the multiple measurements. The surface roughness ranges for the multiple measurements are defined in Table 8.

The second clustering method reported in this paper is FCM clustering. This algorithm aims to improve the classification reliability especially at a border region between two adjacent clusters. Rigidity at the border region in the *k*-means clustering is reduced by implementing the fuzziness of FCM clustering.

An example that shows the *k*-means rigidity was when the results of two consecutive measurements were classified into two different clusters. It could occur when the boundary level was located in the middle area of the first and the second assessments. The first measurement was classified as belonging to the first cluster whereas the second measurement was considered to be in the second cluster. Here, the same dataset was used by FCM clustering. The clustering iteration gave the membership degrees of each data point. For this

Table 7
Average of the Euclidean distances of the surface roughness and boundary levels determined from the correct classification cases of the *k*-means and FCM algorithms.

Score group	*k*-means				FCM			
	$\overline{D_{B_1}}$ (mm)	$\overline{D_{B_2}}$ (mm)	$\overline{D_{B_3}}$ (mm)	*N*	$\overline{D_{B_1}}$ (mm)	$\overline{D_{B_2}}$ (mm)	$\overline{D_{B_3}}$ (mm)	*N*
Score 1	0.012	0.039	0.085	79	0.012	0.043	0.101	104
Score 2	0.011	0.016	0.063	110	0.014	0.018	0.076	114
Score 3	0.046	0.019	0.028	70	0.053	0.022	0.036	55
Score 4	0.120	0.093	0.046	30	0.128	0.097	0.039	22

Table 8
The surface roughness ranges near the boundary levels of the *k*-means and FCM classification algorithms.

Misclassification type	Surface roughness interval (mm) $S_a-3\sigma_{Total} \le B_i \le S_a+3\sigma_{Total}$	
	k-means algorithm	FCM algorithm
Type 1, scores 1 and 2	$0.027 \le S_a \le 0.037$	$0.029 \le S_a \le 0.084$
Type 2, scores 2 and 3	$0.046 \le S_a \le 0.056$	$0.051 \le S_a \le 0.106$
Type 3, scores 3 and 4	$0.079 \le S_a \le 0.089$	$0.092 \le S_a \le 0.147$

dataset, 31 iteration steps were required to achieve stability at an objective function value of 0.0503. The objective function determined the summation of the cluster member aggregation and the separability of the existing clusters. Ideally, the value of the final objective function should be nearly zero. This condition was achieved when all of the cluster members were located close to its cluster centroids and when there were no intersections among the membership functions, as found in hard clustering.

The same test dataset was applied to FCM algorithm. From the 324 data points tested, 295 (91.0%) data points were correctly classified in the double assessment sessions. Misclassified data points were found at 29 (9.0%) data points. As observed in the *k*-means algorithm, similar misclassifications also happened in the FCM implementation. It occurred in the locations near the boundary levels of the score groups. Here, the boundary levels of the FCM algorithm were not considered in the classification process. These boundary levels were determined in order to find the transition boundary from a certain score to another score group.

A boundary level is defined as an intersection point of two overlapping membership functions. Fig. 13 shows an example of a boundary level between score 1 and score 2. The boundary level is depicted as a green circle. For this data point, both overlapping membership functions (score 1 and score 2) gave the same membership degree for a certain surface roughness. This membership degree relation can be expressed by $P_1(0.034)=P_2(0.034)=0.4564$. An incorrect classification would occur if the measured surface roughness values were found to be in the left and the right sides of the boundary level. The first assessment, S_{a1}, was classified as score 1 because the membership function of score 1 gives the largest membership degree, $P_1(0.030)=0.8684$. Conversely, in the second assessment, the highest membership degree, $P_2(0.036)=0.6781$, was given by the membership function of score 2. Thus, the measured surface roughness S_{a2} was considered as score 2.

Finally, all of the boundary levels of the FCM algorithm were obtained from the intersection point of the membership functions. The boundary levels of misclassification types 1, 2, and 3 were 0.034 mm, 0.056 mm, and 0.097 mm, respectively. These levels were not equal to the boundary levels of the *k*-means algorithm. In the FCM algorithm, the levels depend on the curve shape of the membership functions. Therefore, the levels were not exactly

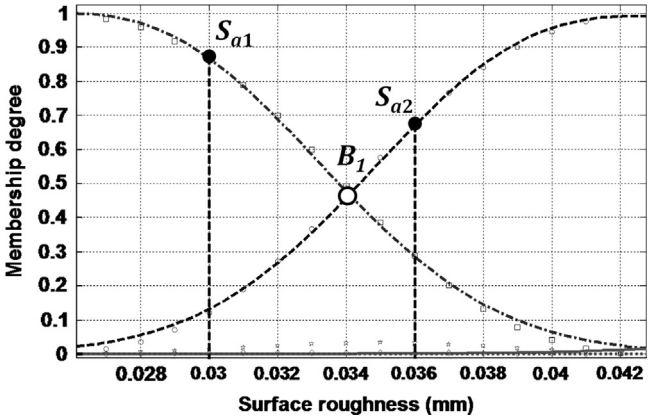

Fig. 13. A boundary level (white circle) splits the membership functions of score 1 (dot-dashed line) and score 2 (dashed line).

located in the middle point of the two cluster centroids, as was found in the *k*-means algorithm.

As conducted in the previous section, the average of the Euclidean distance was computed to find the closeness of the measured surface roughness with the boundary levels. Eq. (21) was applied to compute the Euclidean distance for each misclassified data point. The closeness of the measured data to the boundary levels of the FCM algorithm are listed in Table 6. The data was summarised from the misclassified data points. Misclassification type 1 existed when the measurement results were located near to the first boundary level, B_1. The average distance, $\overline{D_{B_1}}$ (0.003 mm), was small and not more than 10% of D_{B_2} (0.030 mm) and D_{B_3} (0.088 mm). Misclassification types 2 and 3 were also indicated by the closeness of the measured surface roughness to the boundary levels. A small value of D_{B_2} (0.004 mm) shows the closeness of the measured surface roughness to the boundary level of misclassification type 2. The occurrence of misclassification type 3 was explained by a small distance value of D_{B_3} (0.004 mm).

These small distances did not occur for the cases with the correct classifications. The measured surface roughness values were located far from the boundary levels. Therefore, misclassification cases could be avoided in these measurements. The distances for all of the score groups are summarised in Table 7. Two methods have been recommended to minimise the misclassification cases. The first method was performed by acquiring more data points if the surface roughness lay close to the boundary levels. An average value of this multiple measurement was then applied to the score of the surface roughness. This first method is also suggested to resolve the boundary problems of the *k*-means algorithm. Table 8 summarises the surface roughness ranges that needed to be considered for performing multiple measurements.

In the second method, the score group was determined by comparing the membership degrees of the decided score groups. As illustrated in Fig. 13, the final score could not be determined due to two different scores from two measurement sessions. The final score could be obtained by comparing the maximum membership degrees of the first ($P_1(0.030)=0.8684$) and the second ($P_1(0.036)=0.6781$) assessments. From the comparison, the result was $0.8684 > 0.6781$ and thus the final score for this data point was score 1. A mathematical expression was formulated to determine the final score,$S(S_a)$, from the *n*-times measurements. Let $m_i(S_i, P_{\max,i})$ represent the *i*th measurement which has a final score, S_i, at the maximum membership degree $P_{\max,i}$. The equation of the final score can be given by

$$S(S_a) = S_k,\ \max(P_{\max,1}, P_{\max,k}, \ldots, P_{\max,n}) = P_{\max,k}(S_a) \quad (22)$$

The coverage area of the algorithm was limited by the scanning frame size of 40×30 mm^2 of the 3D optical scanner. Lesions could

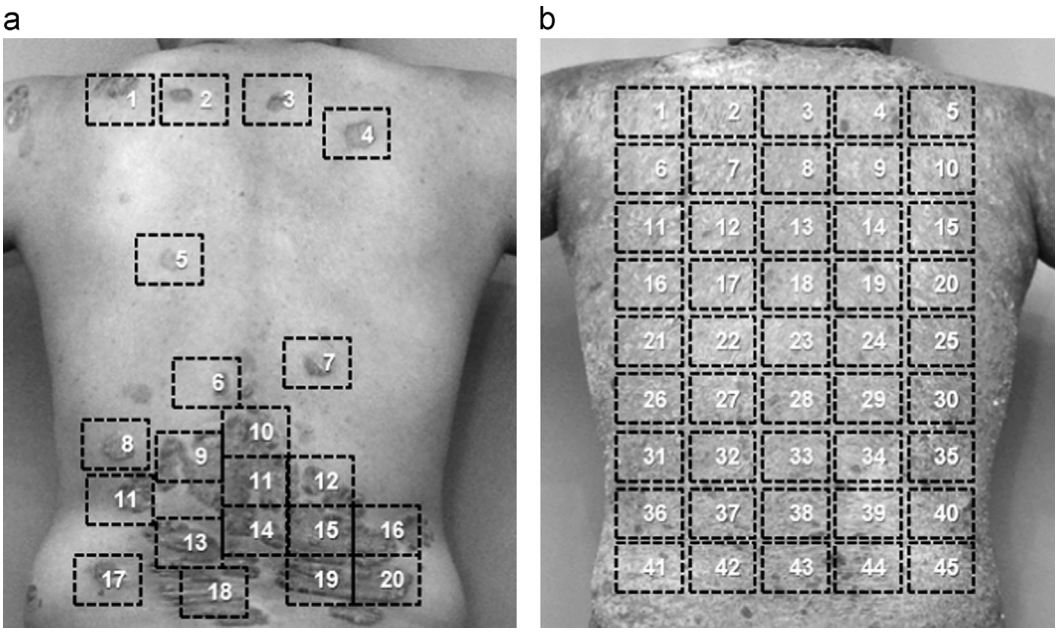

Fig. 14. Multiple scans for a single large lesion: (a) large lesion covering part of the region and (b) single large lesion covering the whole region.

be widely distributed on the body regions or appear as a single lesion covering a large area. In order to measure a larger area ($>40\times30$ mm^2), multiple scans and assessments on adjacent areas could be performed. As shown by the examples in Fig. 14, lesions can be represented by several sampled areas. The average surface roughness for the assessed region would then be determined by averaging the surface roughness of the sampled areas.

4. Agreement analysis on the PASI scaliness scores

Inter-rater variation of two independent observers can be analysed by using Cohen's Kappa (κ) coefficient. The Landis inter-pretation on the Kappa coefficient values can be summarised as shown in Table 9 [61]. The highest agreement namely, almost perfect agreement, can be obtained at $0.81\le\kappa\le0.99$.

The Kappa coefficient value of PASI scaliness scores on 1283 lesions was determined to evaluate the agreement between two dermatologists. The Kappa coefficient value between Dermatologist 1 and Dermatologist 2 was 0.55 and thus was categorised as a moderate agreement. The Kappa coefficient of 0.55 indicated that the dermatologists achieved a 55% agreement of total assessment, and 45% agreement was by chance. A perfect agreement can be achieved if the Kappa coefficient is greater than 0.80. As a result, the dermatologists' scores could not be considered as the ground truth for evaluating the algorithm performance since the Kappa coefficient among the dermatologists was only 0.55.

To evaluate the performance of the PASI scaliness algorithm, several lesion samples were imaged in two successive scans. The PASI scaliness scores of the lesion images were then analysed by the same user in a separate calculation. The number of samples was 324 lesions (648 images). The number of tested lesions was less than the number of the scored lesions by the dermatologists because not all of lesions were scanned twice. The Kappa coefficients between the first assessment and second assessment obtained from the *k*-means and FCM clustering algorithms were then evaluated. Table 10 shows the summary of the scoring agreement performed by the dermatologists and the clustering algorithms – *k*-means clustering and FCM. Table 11 lists the comparisons of the Kappa coefficients obtained from three scoring methods. The Kappa coefficients were computed in the same way as exemplified in the previous section. It can be seen that the FCM could provide a better Kappa coefficient agreement compared to the *k*-means clustering.

Table 9
Agreement interpretation of Kappa coefficient.

Score	Agreement interpretation
<0	Less than chance agreement
0.01–0.20	Slight agreement
0.21–0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–0.99	Almost perfect agreement

With Kappa coefficients values of not less than 0.84, perfect agreement between the first assessment and second assessment has been achieved. The performance of the PASI scaliness algorithm was therefore more consistent as compared to the assessments between the two dermatologists. The FCM performance (0.8708) of the surface roughness classification was slightly better than the *k*-means (0.8473) because the FCM technique could solve the boundary problems better. Fig. 15 depicts an example of the double measurement result. These measurements were taken from lesions at an upper limb region (lesion numbers: 1600 and 1601). The first and second measurements were $S_a=0.030$ mm and 0.033 mm, respectively. In the *k*-means clustering, the measurement results were classified as the different score clusters (score 1 and score 2) because the boundary level of score 1 and score 2 was located in the middle of the measurement results. These measurement results were sharply classified into different scores by the *k*-means clustering; even the images were acquired from the same lesion. However, FCM was able to accommodate this condition by giving the membership degrees or the probabilities of the measured data to clusters 1 and 2. Although the membership degrees changed, the final decision was the same; both measurements were classified into score 1.

5. Conclusion

The surface roughness of the psoriasis lesion can be used as a scaliness feature since the roughness caused by the scale's

Table 10
Summary of the scoring agreement performed by the dermatologists and clustering algorithms (*k*-means and FCM).

First assessment vs. second assessment	Dermatologist					*k*-means					FCM				
	Second assessment					Second assessment					Second assessment				
	1	2	3	4	*N*	1	2	3	4	*N*	1	2	3	4	*N*
First Assessment															
1	634	200	4	0	838	79	10	0	0	89	104	12	0	0	116
2	37	204	39	0	280	11	110	7	0	128	5	108	4	0	117
3	4	24	75	5	108	0	3	70	2	75	0	6	59	1	66
4	1	0	16	40	57	0	0	2	30	32	0	0	1	24	25

Table 11
Comparison of the Kappa coefficients of the dermatologist, *k*-means, and FCM.

Classification method	Kappa coefficient	*N*
Performed by dermatologist	0.5500	1283
k-means	0.8473	324
FCM	0.8708	324

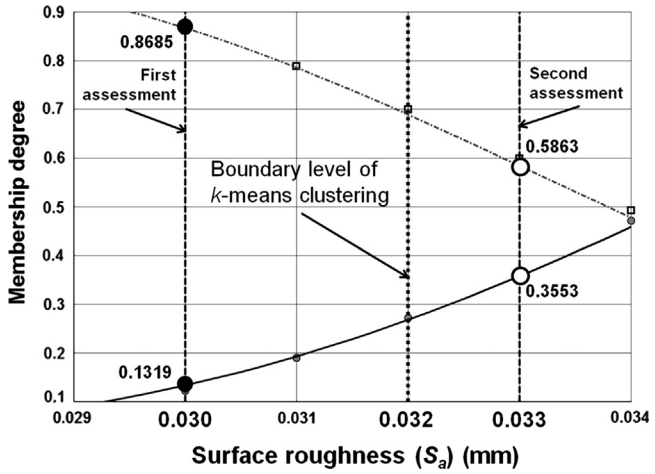

Fig. 15. Clustering of the *k*-means and FCM at the boundary of two clusters – score 1 and score 2. A boundary level of the *k*-means clustering (dotted-line) crosses the membership functions of score 1 (dot-dashed line) and score 2 (solid line).

appearance is related with the scaliness severity. The developed surface roughness algorithm for 3D curved surfaces for measuring the surface roughness of the psoriasis lesion has been validated with an accuracy of 94.12%. The surface roughness algorithm has been applied to 1999 psoriasis lesion images. The result gives the roughness of the psoriasis lesion of within 0.010–0.187 mm. Unsupervised clustering algorithms (*k*-means and FCM) were used to classify the surface roughness for the PASI Scaliness scoring. The centroid consistency of the subsets was observed in two and three partitions. The dermatologists' assessments could not be considered as truly ground truth since the dermatologists' agreement (Kappa coefficient) was found to be 0.55.

The reliability of the developed PASI scaliness algorithm was high as shown by the perfect agreement values of the kappa coefficients (*k*-means – 0.8473 and FCM – 0.8708) and thus alleviated the issues of the intra- and inter-rater variability. The surface roughness intervals at the cluster boundaries have been defined for both clustering algorithms. FCM clustering slightly improved the agreement between two successive measurements of *k*-means clustering from 0.8473 to 0.8708. This was due to the FCM capability of resolving scoring problems at cluster boundaries.

Conflict of interest statement

None declared.

Acknowledgements

The research has been a collaborative effort between the Centre for Intelligent Signal and Imaging Research (CISIR), Universiti Teknologi PETRONAS and the Dermatology Department, Hospital Kuala Lumpur. The work was funded by the Ministry of Science, Technology and Innovation, Malaysia under the Techno Fund Grant TF 0308C041. The Clinical Study Proposal (NMRR-09-1098-4863) was approved by the Clinical Research Centre, Ministry of Health, Malaysia.

Appendix A. Supplementary material.

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.compbiomed.2013.08.009.

References

[1] Lippincott Williams, Wilkins, Pathophysiology: A 2-in-1 Reference for Nurses, Lippincott Williams & Wilkins, Philadelphia, 2005.
[2] Nili N. Alai, MD, FAAD, Psoriasis, 2011.
[3] M.J. Bhosle, A. Kulkarni, F. SR, R. Balkrishnan, S.R. Feldman, Quality of life in patients with psoriasis, Health Qual. Life Outcomes 4 (2006) 35.
[4] National Psoriasis Foundation, Statistics, National Psoriasis Foundation, 2012.
[5] International Federation of Psoriasis Associations, Facts About Psoriasis, International Federation of Psoriasis Associations, 2012.
[6] A.L. Neimann, S.B. Porter, J.M. Gelfand, The epidemiology of psoriasis, Expert Rev. Dermatol. 1 (2006) 63–76.
[7] V. Chandran, S.P. Raychaudhuri, Geoepidemiology and environmental factors of psoriasis and psoriatic arthritis, J. Autoimmun. 34 (2010) J314–J321.
[8] P.D. Malaysia, Overview of Psoriasis in Malaysia, 2011.
[9] S.R. Feldman, G.G. Krueger, Psoriasis assessment tools in clinical trials, Ann. Rheum. Dis. 64 (2005) 65–68.
[10] A.F. Mohamad Hani, E. Prakasa, V.S. Asirvadam, H. Nugroho, C.H. Chong, A.M. Affandi, et al., Development of body surface area measurement using multi-view imaging for psoriasis area assessment, Australas. Phys. Eng. Sci. Med. 34 (2011) 559–637.
[11] A.F.M. Hani, E. Prakasa, H. Nugroho, A.M. Affandi, S.H. Hussein, Body surface area measurement and soft clustering for PASI area assessment, in: Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, IEEE, San Diego, USA, 2012, pp. 4398–4401.
[12] A.F.M. Hani, H. Nugroho, E. Prakasa, V. Asirvadam, A.M. Affandi, S.H. Hussein, Soft clustering of lesion erythema for psoriasis assessment, J. Invest. Dermatol. 132 (2012).
[13] A.F.M. Hani, E. Prakasa, H. Nugroho, V.S. Asirvadam, Implementation of fuzzy *c*-means clustering for psoriasis assessment on lesion erythema, in: 2012 IEEE Symposium on Industrial Electronics and Applications (ISIEA), IEEE, Kuala Lumpur, Malaysia, 2012.
[14] A.F.M. Hani, H. Fitriyah, E. Prakasa, V.S. Asirvadam, S.H. Hussein, M.A. Azura, *In vivo* 3D thickness measurement of skin lesion, in: 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), IEEE, Kuala Lumpur, Malaysia, 2010, pp. 155–160.
[15] W. Manuskiatti, D.A. Schwindt, H.I. Maibach, Influence of age, anatomic site and race on skin roughness and scaliness, Dermatology 196 (1998) 401–407.

[16] C. Edwards, R. Heggie, R. Marks, A study of differences in surface roughness between sun-exposed and unexposed skin with age, Photodermatol. Photoimmunol. Photomed. 19 (2003) 169–174.
[17] P. Corcuff, G.E. Pierard, Skin imaging: state of the art at the dawn of the year 2000, Curr. Probl. Dermatol. 26 (1998) 1.
[18] P. Corcuff, J. de Rigal, J.L. Leveque, S. Makki, P. Agache, Skin relief and aging, J. Soc. Cosmet. Chemists 34 (1983) 177–189.
[19] C.F. Corcuff, P. J. Leveque, A fully automated system to study skin surface patterns, Int. J. Cosmet. Sci. (1984) 167–176.
[20] P. Corcuff, F. Chatenay, A. Brun, Evaluation of anti-wrinkle effects on humans, Int. J. Cosmet. Sci. 7 (1985) 117–126.
[21] P. Corcuff, J.L. Leveque, G.L. Grove, A.M. Kligman, The impact of aging on the microrelief of peri-orbital and leg skin, J. Soc. Cosmet. Chemists 38 (1987) 145–152.
[22] T.G. Mathia, P. Pawlus, M. Wieczorowski, Recent trends in surface metrology, Wear 271 (3-4) (2011) 494–508.
[23] P. Nardin, D. Nita, J. Mignot, Automation of a series of cutaneous topography measurements from silicon rubber replicas, Skin Res. Technol. 8 (2002) 112–117.
[24] C. Pailler-Matti, H. Zahouani, Study of adhesion forces and mechanical properties of human skin *in vivo*, J. Adhes. Sci. Technol. 18 (2004) 1739–1758.
[25] H. Zahouani, R. Vargiolu, G. Boyer, C. Pailler-Mattei, L. Laquieze, A. Mavon, Friction noise of human skin *in vivo*, Wear 267 (2009) 1274.
[26] M. Breuer, W.Voss, Proving for Efficacy: Laser Profilometry, Dermatest GmbH, Munster [available from Dermatest].
[27] Z. Maor, S. Yehuda, W. Voss, Skin smoothing effects of Dead Sea minerals: comparative profilometric evaluation of skin surface, Int. J. Cosm. Sci. 19 (1997) 105–110.
[28] M. Loden, Atopic dermatitis and other skin diseases, in: K.P. Wilhelm (Ed.), Bioengineering of the Skin: Skin Imaging and Analysis, Informa Healthcare USA, Inc., New York, 2007.
[29] H. Fujii, T. Asakura, Effect of surface roughness on the statistical distribution of image speckle intensity, Opt. Commun. 11 (1974) 35–38.
[30] F. Dalmases, R. Cibrián, M. Buendía, C. Romero, R. Salvador, J. Montilla, Speckle correlation technique to determine roughness in the dermatologic interval, Phys. Med. Biol. 33 (1988) 913–922.
[31] Lu Rong-Sheng, Gui-Yun Tian, Duke Gledhill, Steve Ward, Grinding surface roughness measurement based on the co-occurrence matrix of speckle pattern texture, Appl. Opt. 45 (2006) 8839–8847.
[32] Z. Li, H. Li, Y. Qiu, Fractal analysis of laser speckle for measuring roughness [6027-64], in: Proceedings—SPIE The International Society for Optical Engineering, 2006, p. 60271S.
[33] L. Tchvialeva, H. Zeng, I. Markhvida, D.I. McLean, H. Lui, T.K. Lee, Skin roughness assessment, New Dev. Biomed. Eng. (2010) 341–358.
[34] G. Frankowski, M. Chen, Optical 3D *in vivo* measurement of human skin surfaces with PRIMOS, in: Proceedings of the Fringe, 2001.
[35] Ute Jacobi, Mai Chen, Gottfried Frankowski, Ronald Sinkgraven, Martina Hund, Berthold Rzany, Wolfram Sterry, Jürgen Lademann, et al., *In vivo* determination of skin surface topography using an optical 3D device, Skin Res. Technol. 10 (2004) 207–214.
[36] P.M. Friedman, G.R. Skover, G. Payonk, A.N.B. Kauvar, R.G. Geronemus, 3D in-vivo optical skin imaging for topographical quantitative assessment of non-ablative laser technology, Dermatol. Surg. Official Publ. Am. Soc. Dermatol. Surg. et al 28 (3) (2002) 199–204.
[37] U.D.O. Hoppe, G. Sauermann, Quantitative analysis of the skin's surface by means of digital signal processing, Analysis 123 (2000) 105–123.
[38] M.C.T. Bloemen, M.S. van Gerven, M.B. a van der Wal, P.D.H.M. Verhaegen, E. Middelkoop, An objective device for measuring surface roughness of skin and scars, J. Am. Acad. Dermatol. 64 (2011) 706–715.
[39] L. Blunt, X. Jiang, Advanced Techniques for Assessment Surface Topography: Development of a Basis for 3D Surface Texture Standards "Surfstand", Kogan Page Science, London; Sterling, VA, 2003.
[40] B. Muralikrishnan, J. Raja, Computational Surface and Roundness Metrology, 1st ed, Springer, London, 2009.
[41] D.J. Whitehouse, Handbook of Surface Metrology, Institute of Physics Publishing, Bristol; Philadelphia, 1994.
[42] K. Carneiro, J. Garnaes, C. Jensen, J. Jørgensen, L. Nielsen, O. Jusko, G. Wilkening, G. Barbato, G. Picotto, G. Gori, S. Livi, G. Hughes, H. McQuoid, Scanning Tunnelling Microscopy Methods for Roughness and Micro Hardness Measurements (1994).
[43] G.A. Al-Kindi, B. Shirinzadeh, Feasibility assessment of vision-based surface roughness parameters acquisition for different types of machined specimens, Image Vision Comput. 27 (2009) 444–458.
[44] GFMesstechnik, User Manual: PRIMOS Optical 3D Skin Measuring Device, 2008.
[45] Y. Lee, S. Lee, I. Ivrissimtzis, H.-P. Seidel, Overfitting control for surface reconstruction, in: ACM International Conference Proceeding Series, Citeseer, 2006, pp. 231–234.
[46] A.F.M Hani, E. Prakasa, H. Nugroho, A.A.M. Affandi, S.S.H. Hussein, A.F.M. Hani, Sample area for surface roughness determination of skin surfaces, in: 2012 4th International Conference on Intelligent and Advanced Systems (ICIAS), vol. 1, Universiti Teknologi PETRONAS, Kuala Lumpur, 2012. pp. 328–332.
[47] J. Dariusz, Gaussian filters with profile extrapolation, Precis. Eng. 35 (2011) 602–606.
[48] E.K. Regis Hoffman, R. Hoffman, E. Krotkov, Terrain roughness measurement from elevation maps, Mob. Robots IV (2003) 104–114.
[49] H. Zhang, Y. Yuan, W. Piao, A universal spline filter for surface metrology, Measurement 43 (2010) 1575–1582.
[50] X. Jiang, P. Scott, D. Whitehouse, Wavelets and their applications for surface metrology, CIRP Ann.: Manuf. Technol. 57, (1), (2008) 555–558, ISSN 0007–8506, 〈http://dx.doi.org/10.1016/j.cirp.2008.03.110〉.
[51] K. Michael, Form filtering by splines, Measurement 18 (1996) 9–15.
[52] W.P. Dong, E. Mainsah, K.J. Stout, Reference planes for the assessment of surface roughness in three dimensions, Int. J. Mach. Tools Manuf. 35 (1995) 263–271.
[53] J.S. Yoon, C. Ryu, J.H. Lee, Developable polynomial surface approximation to smooth surfaces for fabrication parameters of a large curved shell plate by Differential Evolution, Comput.-Aided Des. 40 (2008) 905–915.
[54] G.J. Borradaile, Statistics of Earth Science Data: Their Distribution in Time, Space, and Orientation, Springer, Berlin; New York, 2003.
[55] M.A. Fadzil, E. Prakasa, H. Fitriyah, H. Nugroho, A. Affandi, S. Hussein, Validation on 3D surface roughness algorithm for measuring roughness of psoriasis lesion, Int. J. Biol. Life Sci. 8 (2012) 205.
[56] A. Hani, E. Prakasa, H. Fitriyah, H. Nugroho, A. Affandi, S. Hussein, High order polynomial surface fitting for measuring roughness of psoriasis lesion, in: H. Badioze Zaman, P. Robinson, M. Petrou, P. Olivier, T. Shih, S. Velastin, et al., (Eds.), Visual Informatics: Sustaining Research and Innovations, Springer, Berlin, Heidelberg, 2011, pp. 341–351.
[57] GFMesstechnik, Technical Data of 3D PRIMOS Portable, 2008, p. 2.
[58] J.V. De Oliveira, W. Pedrycz, Advances in Fuzzy Clustering and its Applications, John Wiley and Sons Ltd., The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2007.
[59] R.O. Duda, D.G. Stork, P.E. Hart, Pattern Classification and Scene Analysis. Part 1, Pattern Classification, Wiley, New York, Chichester, 2000.
[60] R. Xu, D.C. Wunsch, Clustering algorithms in biomedical research: a review, IEEE Rev. Biomed. Eng. 3 (2010) 120–154.
[61] A.J. Viera, J.M. Garrett, Research series – understanding interobserver agreement: the Kappa statistic, Fam. Med. 37 (2005) 360.

Ahmad Fadzil Mohamad Hani received B.Sc. (1st Class Hons.) in Electronics Engineering, M.Sc. in telematics and Ph.D. in image processing from University of Essex, UK in 1983, 1984 and 1991, respectively. Currently, he is a Professor at the Department of Electrical and Electronic Engineering and heads the Centre for Intelligent Signal & Imaging Research, Universiti Teknologi PETRONAS. His research interests are digital signal and image processing, biomedical image analysis, computer vision, vision inspection systems, image compression/coding, remote sensing, digital elevation models, fractal analysis, blind deconvolution, and independent component analysis. He has authored and co-authored over 180 research articles on his research interests and holds several patents.

Esa Prakasa received B.Eng. in Nuclear Engineering from Universitas Gadjah Mada, Indonesia in 1998 and M.Eng. in Electrical Engineering from Universitas Gadjah Mada, Indonesia in 2001. Currently, he is a Ph.D. candidate in the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia. His research interests are 3D medical imaging, computer vision, and pattern recognition. He is on leave from the Research Center for Informatics, Indonesian Institute of Sciences, Bandung, Indonesia.

Vijanth Sagayan Asirvadam received B.Sc. (Hons.) in Statistics from University of Putra Malaysia in 1997. He received the M.Sc. with Distinction and Ph.D. from Queen's University of Belfast, Ireland. Currently, he is an Associate Professor at the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia. His research interests include signal modelling, system identification, adaptive/recursive and sequential learning. He is also the current secretary of IEEE Malaysian section of Signal Processing chapter.

Hermawan Nugroho received B.Sc. in Electronics Engineering from Institut Teknologi Bandung, Indonesia in 2005 and M.Sc. in Electrical Engineering from Universiti Teknologi PETRONAS, Malaysia in 2007. Currently, he is a Ph.D. candidate in the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia. His research interests are medical imaging, seismic imaging, intelligent system and signal processing.

Azura Mohd Affandi received M.B.Ch.B. (Bachelor of Medicine) from the University of Manchester, United Kingdom in 1999. She underwent physician training and received membership from the Royal College of Physician (UK) in 2005. She then pursued further training in dermatology and obtained Advanced Master of Dermatology from University Kebangsaan Malaysia in 2009. In 2011, she was elected as a fellow in Onco-Dermatology and completed 9 months fellowship in Sydney, Australia. She is currently working as a dermatologist in the Department of Dermatology, Hospital Kuala Lumpur. Her research interests are cutaneous malignancies and medical imaging in psoriasis.

Suraiya Hani Hussein has served as consultant dermatologist and head of department at Department of Dermatology, Hospital Kuala Lumpur. She was President of the Dermatological Society of Malaysia (1996–1998). She is currently working as a consultant dermatologist in the Damansara Specialist Hospital, Kuala Lumpur.