Segmentation of Acne Lesion using Fuzzy C-means Technique with Intelligent Selection of the Desired Cluster

Javed Khan, Aamir Saeed Malik, Nidal Kamel, Sarat Chandra Dass, Azura Mohd Affandi

***Abstract*— Segmentation is the basic and important step for digital image analysis and understanding. Segmentation of acne lesions in the visual spectrum of light is very challenging due to factors such as varying skin tones due to ethnicity, camera calibration and the lighting conditions. In this approach the color image is transformed into various color spaces. The image is decomposed into the specified number of homogeneous regions based on the similarity of color using fuzzy C-means clustering technique. Features are extracted for each cluster and average values of these features are calculated. A new objective function is defined that selects the cluster holding the lesion pixels based on the average value of cluster features. In this study segmentation results are generated in four color spaces (RGB, rgb, YIQ, I1I2I3) and two individual color components (I3, Q). The number of clusters is varied from 2 to 6. The experiment was carried out on fifty images of acne patients. The performance of the proposed technique is measured in terms of the three mostly used metrics; sensitivity, specificity, and accuracy. Best results were obtained for Q and I3 color components of YIQ and I1I2I3 color spaces with the number of clusters equal to three. These color components show robustness against non-uniform illumination and maximize the gap between the lesion and skin color.**

I. INTRODUCTION

Acne is a common skin disorder that has influenced people of all ages and especially adolescent population. About 85% of the adolescent population has this skin pathology [1]. Its pathological evolution involves chronic disorder of pilosebaceous units, follicular epidermal hyper-proliferation and p-acne (propionic bacteria) activity [2]. It affects those areas of human body where hair follicles are comparatively in large number for example, face, chest and upper part of the back [3]. A huge amount of money is spent on the treatment, medication and laser therapy of acne throughout the world. In the United States, approximately \$100 million is invested on products related to acne treatment [4]. Similarly in Malaysia, \$100 is spent per patient per year.

Acne can be divided into many types based on its pathological evolution such as acne conglobata, acne cosmetica and acne vulgaris. Among the different types of acne, acne vulgaris is the most common one. It comprises about 99% of all acne cases [5]. Acne vulgaris can be subdivided into five main types; comedone, papule, pustule, nodule and cyst as shown in Fig.1. Comedone is further grouped into two types; white-head comedone and black-head comedone. This categorization of acne lesions is based on the size, shape, color, depth and inflammation of the lesion. The severity of acne lesion increases from comedone to cyst gradually. Acne lesions mostly appear on face. It causes inflammation and may leave permanent scars on face. Acne affects the patients not only physically but it also has emotional effects on patients [6]. Acne patient avoids social gathering and hesitates from interacting with other people.

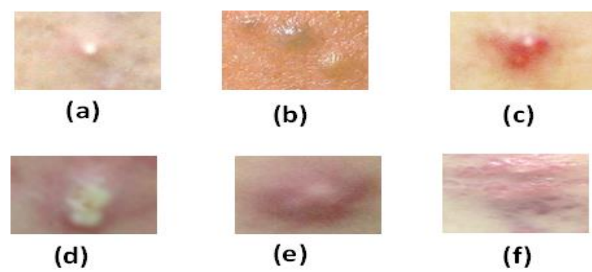

Figure 1. Type of acne vulgaris lesion (a) white-head comedone (b) black-head comedone (c) papule (d) pustule (e) nodule (f) cyst

Two methods have been used for the assessment of acne Vulgaris lesions; lesion counting method and comparison of acne patient's images with reference photographs. In lesion counting method, the facial skin of a patient is examined; the type of acne lesion is identified and total number of lesions is calculated. In photographic method, a dermatologist looks at the images of acne patients against the reference photographs and approximates the severity of acne patient based on visual inspection. The lesion counting method is very time consuming, tedious and subjective. In photographic method, the severity of acne patient is evaluated subjectively. The involvement of human experts makes both methods subjective. In both methods, intra-rater and inter-rater variability have been observed. In intra-rater variability, the severity of the same acne patient is differently rated by the same expert at different times while in case of inter-rater variability; the severity of an acne patient is evaluated differently by different experts. In order to make acne assessment objective, accurate and faster, an automated computer vision based system is needed.

II. RELATED WORK

The segmentation of acne lesion from skin is a very challenging and difficult task. There is a variation in skin tones due to ethnicity and also the appearance of color changes when images are taken under slightly different lighting conditions [7]. As different light sources have different color temperatures and contribute to different colors. Similarly variation can be observed in lesion color as well. In order to cope with such problems, color space is needed which has two important properties; robustness against illumination and increase the gap between the skin

Research supported by Universiti Teknologi PETRONAS, Malaysia.
Javed Khan, Aamir Saeed Malik, Nidal Kamel, Sarat Chandra Das are with Centre for Intelligent Signal and Imaging Research (CISIR), Department of Electrical & Electronic Engineering, Universiti Teknologi PETRONAS, Perak, Malaysia (email: aamir_saeed@petronas.com.my).
Azura Mohd Affandi is with the Department of Dermatology, Hospital Kuala Lumpur, Malaysia.

and lesion color. In order to find such color space or a component of a color space, four color spaces (RGB, normalized RGB, I1I2I3, YIQ) have been used.

A. *Color spaces*

Color has been used as an important feature in the segmentation of color images [8]. Many color space have been developed to represent the color information of color images. The default color space for most of the image capturing and storing device is RGB (Red, Green and Blue). In computer vision and image processing, RGB color space is converted into other color spaces through either linear or non-linear transformation. RGB color space is very sensitive to illumination variations [9]. Secondly the channels of this color space are highly correlated and represent redundant information. To minimize the dependency of RGB color space on illumination, it is transformed into normalized RGB color space denoted by small letters—rgb. Equations (1)-(3) depict the normalization of the individual channel.

$$r = \frac{R}{R+G+B} \tag{1}$$

$$g = \frac{G}{R+G+B} \tag{2}$$

$$b = \frac{B}{R+G+B} \tag{3}$$

Ohta [10] developed a color space from RGB color space through linear transformation according to (4) through (6). I1 represents the luminance information while the other two components I2 and I3 represent chrominance information.

$$I1 = \frac{(R+B+C)}{3} \tag{4}$$

$$I2 = (R-B) \tag{5}$$

$$I3 = \frac{(2G-R-B)}{2} \tag{6}$$

Another very important and mostly used color space is YIQ. It separates the luminance (Y) and chrominance information (I, Q) of RGB color space. The illumination effects can be eliminated by using only the color components (I, Q). YIQ color space is obtained from RGB color space using (7).

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 1.0000 & 1.0000 & 1.0000 \\ 0.9560 & -0.2720 & -01.1060 \\ 0.6210 & -0.6470 & 1.70000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{7}$$

B. *Fuzzy C-means (FCM)*

FCM is an iterative clustering technique [11] in which the association of a pixel to a cluster is measured in terms of membership degree having values in the range [0 1]. In this technique a pixel can belong to more than one cluster with different values of association degree. An image is decomposed into the specified number of clusters by maximizing the objective function shown in (8).

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij}^{m} \left\| x_i - q_j \right\|^2 , \ 1 \le m < \infty \tag{8}$$

Where N is the total number of pixel to be clustered, C is the specified number of clusters and μ represents membership degree. In each iteration, it is updated according to (9). The exponential parameter *m* is called fuzzier and appropriate value for this is 2. In (8), q_j holds the centroids of j^{th} cluster.

$$\mu_{ij} = \sum_{k=1}^{C} \left(\frac{\|x_i - q_j\|}{\|x_i - q_k\|} \right)^{\frac{-2}{m-1}} \tag{9}$$

Like the membership parameter, the centroids of the clusters are recalculated according to (10).

$$q_j = \frac{\sum_{i=1}^{N} \mu_{ij}^{m} . x_i}{\sum_{i=1}^{N} \mu_{ij}^{m}} \tag{10}$$

The algorithm stops when either maximum number of iteration is reached or change in the centroid values of clusters in two consecutive iterations is less than the specified threshold, that is $\left| \mu_{ij}^{(k+1)} - \mu_{ij}^{k} \right| < \varepsilon$, where ε represents the threshold value. Apart from this conventional FCM, many variants of this technique are available [12-14].

III. Material and Method

A. *Dataset*

The results are generated with fifty color image of acne patients. The images were captured at Hospital Kuala Lumpur Malaysia under the supervision of dermatologist. Five images were taken from facial regions (nose, forehead, left cheek, right cheek and chin) of each patient after signing the consent form. All images were taken at constant of four feet to avoid variation in acne lesion size due to relative distance.

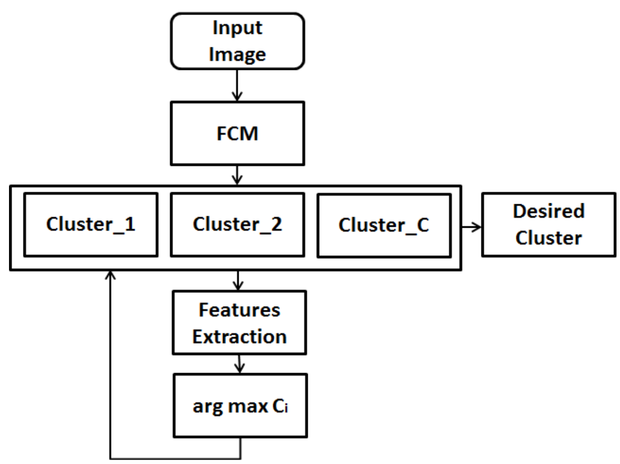

Figure 2. Flow chart of the proposed approach

B. *Proposed Method*

Initially highlights are removed using method developed in [15] and then the RGB color image is transformed into the desired color space. This colour image is decomposed into the specified number of homogeneous regions using fuzzy c-means clustering technique. The color space of each cluster is converted into I1I2I3 color space where I3 component is used for determining the cluster holding the lesion pixels. Average values are calculated for each cluster as in (11).

$$f_i = {}^{1}\!/_{n_i} \sum_{i \in C} x_i \tag{11}$$

Where n_i is the number of pixels in i^{th} cluster, *i= 1, 2,…C* , is the cluster number. The average value calculated for each cluster is passed to the objective function defined as in (12).

$$z = arg \max_{i \in C} f_i \tag{12}$$

The parameter z holds the maximum value of the objective function and index of corresponding cluster. The cluster corresponding to the maximum value of z holds acne lesions and represents the segmentation of the color image.

IV. RESULTS AND DISCUSSION

The results are produced with the proposed method using four color spaces; RGB, rgb, I1I2I3 and YIQ and two individual color components I3 and Q of I1I2I3 and YIQ respectively and presented in table 1-to-table 6.

A. *Performance Analysis*

Three metrics (Specificity, Sensitivity and Accuracy) are selected to measure the performance of segmentation results. These metrics are calculated from true positive (TP), true negative (TN), false positive (FP) and false negative (FN) values. True positive means that a segmentation technique correctly identifies a lesion pixel while true negative indicates that a skin pixel is correctly identified. When a skin pixel is identified as a lesion pixel, it is called false positive and when a lesion pixel is identified as a pixel lesion, it is called false negative. Specificity is the fraction which measures how many skin lesions are correctly identified out of total skin pixel according to (13).

$$Specificity\ (\%) = \frac{TN}{TN+FP} \times 100 \tag{13}$$

Similarly sensitivity measures that how many lesion pixels are correctly identified out of the total lesion pixels. It is calculated as (14).

$$Sensitivity\ (\%) = \frac{TP}{TP+FN} \times 100 \tag{14}$$

Finally the accuracy represents the ratio of correct decision made to the total made decision. It is calculated as in (15)

$$Accuracy(\%) = \frac{TP+TN}{TP+FN+TN+FP} \times 100 \tag{15}$$

TABLE 1. SEGMENTATION RESULTS FOR RGB COLOR SPACE

No. of Clusters	Sensitivity (%)	Specificity (%)	Accuracy (%)
2	82.83	44.05	47.28
3	71.35	68.26	68.28
4	58.99	78.74	76.87
5	49.64	82.96	79.8
6	44.70	85.76	82.14

TABLE 2. SEGMENTATION RESULTS FOR NORMALIZED RGB COLOR SPACE

No. of Clusters	Sensitivity (%)	Specificity (%)	Accuracy (%)
2	85.15	52.22	54.22
3	71.94	73.19	72.74
4	63.82	87.05	84.64
5	53.46	93.48	89.89
6	46.93	95.37	90.74

TABLE 3. SEGMENTATION RESULTS OF OHTA'S COLOR SPACE

No. of Clusters	Sensitivity (%)	Specificity (%)	Accuracy (%)
2	81.33	43.57	46.74
3	73.38	68.87	69.16
4	65.14	77.66	76.33
5	60.42	80.84	81.85
6	52.87	86.33	83.30

TABLE 4. SEGMENTATION RESULTS FOR YIQ COLOR SPACE

No. of Clusters	Sensitivity (%)	Specificity (%)	Accuracy (%)
2	83.97	45.69	48.90
3	73.87	69.41	69.63
4	63.14	80.49	78.77
5	52.91	84.78	81.94
6	46.46	87.25	83.73

TABLE 5. SEGMENTATION RESULTS FOR Q COMPONENT OF YIQ COLOR SPACE

No. of Clusters	Sensitivity (%)	Specificity (%)	Accuracy (%)
2	98.17	65.42	67.62
3	89.67	93.19	92.63
4	80.19	97.12	95.20
5	65.16	98.81	95.49
6	52.36	99.38	94.90

TABLE 6. SEGMENTATION RESULTS FOR I3 COMPONENT OF I1I2I3 COLOR SPACE

No. of Clusters	Sensitivity (%)	Specificity (%)	Accuracy (%)
2	98.26	65.05	67.26
3	89.54	91.62	91.05
4	77.40	97.02	94.86
5	62.30	98.81	95.35
6	62.30	98.81	95.35

Higher values for sensitivity, specificity, and accuracy show good segmentation results. It is possible to increase the specificity at the cost of decreased sensitivity and similarly the sensitivity can be increased at the cost of decreased value of specificity. Within each of the six tables, the segmentation results vary by selecting different number of clusters and comparatively good results are obtained when the image is decomposed into three clusters.

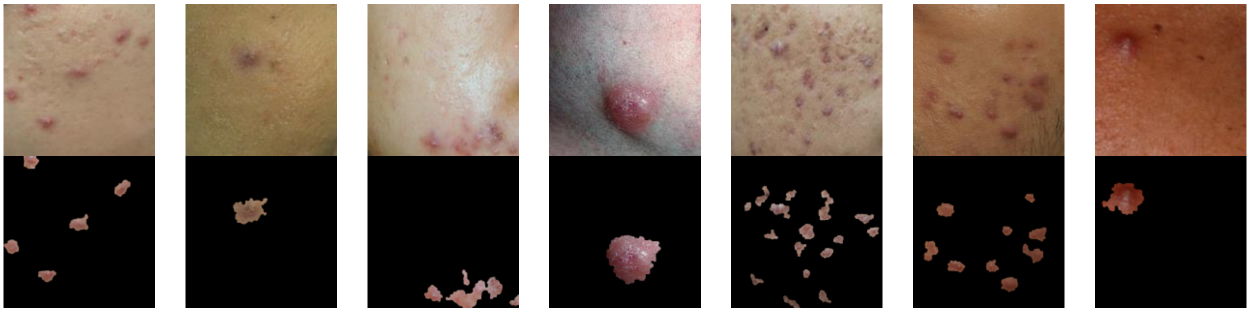

Figure 3. Segmentation results for Q component of YIQ color space with number of cluster equal to three.

The results also suggest that using luminance component or the color space which is highly vulnerable to illumination can adversely affect the segmentation. The second rows (corresponding to three clusters) of table 5 and table 6 show higher values for all the three metrics (specificity, sensitivity and accuracy) used. Table 5 represents the segmentation results of using only the Q component of the YIQ color space. Table 6 corresponds to the segmentation results of using only the I3 component of I1I2I3 color space. The results of these tables are very similar. Such results are expected because the way these two components (I3, Q) are obtained from the RGB color space, is similar. For example, Q=0.6210R+1.7000B-0.6470G, and similarly I3= (2G-R-B)/2; in deriving these two components from RGB color space; red, green and blue channel values are multiplied with slightly different coefficients. Secondly subtracting G from the sum of R and B components as in deriving of Q or subtracting sum of R and B from weighted G as in deriving I3, can obtain quite similarly values but with opposite signs (positive, negative). These components show robustness against illumination and maximize the gap between the skin and lesion. Table 1 suggests that RGB color space performs very poor in segmentation. Table 2 demonstrates that normalized RGB color space is better than RGB color. Table 3 and 4 indicate that using all components of I1I2I3 and YIQ are not useful as the luminance components I1 and Y are highly affected from illumination and their higher values dominates the overall segmentation results.

V. Conclusion

It can be concluded from the result and discussion section that the proposed method can generate good results. However the number of clusters that produce comparatively better results may vary, depending on the application and content of the images. In this study, segmentation results are generated with number of clusters varying from 2 to 6. The results suggest that three is the optimal number of clusters. Secondly RGB color space is not suitable to be used for segmentation when there is non-uniform illumination in the images. In case of the color spaces in which the luminance and chrominance parts are separate, the luminance component adversely affects and dominates the overall segmentation results. In this study two components (I3, Q) are discovered which have good capability to separate lesion from skin and show robustness against non-uniform illumination. The best result obtained in this study is specificity = 93.19%, sensitivity = 89.67% and accuracy = 92.63% for Q component of YIQ color space with the number of clusters equal to three. Similar results (specificity = 91.62%, sensitivity = 89.54%, accuracy = 91.05%) are also obtained for I3 component of I1I2I3 color space with number of clusters equal to three. As mentioned in the results and discussion section, the way these two color components are derived is similar. In future study, it is expected that the results can be improved by considering textural and shape features along with color features.

References

[1] W. D. James, "Acne," New England Journal of Medicine, vol. 352, pp. 1463-1472, 2005.

[2] N. B. Simpson and W. J. Cunliffe, "Disorders of the sebaceous glands," Rook's Textbook of Dermatology, Seventh Edition, pp. 2121-2196, 2004.

[3] J. Tan, J. Tang, K. Fung, A. K. Gupta, D. R. Thomas, S. Sapra, et al., "Prevalence and severity of facial and truncal acne in a referral cohort," Journal of drugs in dermatology: JDD, vol. 7, pp. 551-556, 2008.

[4] L. Cordain, S. Lindeberg, M. Hurtado, K. Hill, S. B. Eaton, and J. Brand-Miller, "Acne vulgaris: a disease of Western civilization," Archives of dermatology, vol. 138, pp. 1584-1590, 2002.

[5] R. Ramli, A. S. Malik, A. F. M. Hani, and A. Jamil, "Acne analysis, grading and computational assessment methods: an overview," Skin research and technology, vol. 18, pp. 1-14, 2012.

[6] R. G. Fried and A. Wechsler, "Psychological problems in the acne patient," Dermatologic therapy, vol. 19, pp. 237-240, 2006.

[7] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods," Pattern recognition, vol. 40, pp. 1106-1122, 2007.

[8] S. L. Phung, A. Bouzerdoum, and D. Chai Sr, "Skin segmentation using color pixel classification: analysis and comparison," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 27, pp. 148-154, 2005.

[9] H.-D. Cheng, X. Jiang, Y. Sun, and J. Wang, "Color image segmentation: advances and prospects," Pattern recognition, vol. 34, pp. 2259-2281, 2001.

[10] Y.-I. Ohta, T. Kanade, and T. Sakai, "Color information for region segmentation," Computer graphics and image processing, vol. 13, pp. 222-241, 1980.

[11] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," Computers & Geosciences, vol. 10, pp. 191-203, 1984.

[12] J. Kang, L. Min, Q. Luan, X. Li, and J. Liu, "Novel modified fuzzy c-means algorithm with applications," Digital Signal Processing, vol. 19, pp. 309-319, 2009.

[13] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," Image Processing, IEEE Transactions on, vol. 19, pp. 1328-1337, 2010.

[14] K. S. Tan, W. H. Lim, and N. A. M. Isa, "Novel initialization scheme for Fuzzy C-Means algorithm on color image segmentation," Applied Soft Computing, vol. 13, pp. 1832-1852, 2013.

[15] J. Khan, A. S. Malik, N. Kamel, S. C. Das, and A. M. Affandi, "Recovery of color loss due to specular illumination in Acne patient images," in Biomedical Engineering and Sciences (IECBES), 2014 IEEE Conference on, 2014, pp. 740-743.