An Automated Mean-shift based Segmentation for Pigmented Skin Lesions

Zhao Liu[1], Jiuai Sun[1], Melvyn Smith[1], Lyndon Smith[1], Rob Warr[2]

[1]Machine Vision Laboratory, University of the West of England, Bristol, UK

[2]Department of Plastic Surgery, Frenchay Hosptital, Bristol, UK

Abstract

This paper presents an unsupervised segmentation scheme to isolate pigmented skin lesion from surrounding normal skin. An adaptive mean-shift algorithm combined with maximal similarity based region merging is applied with a colour-spatial feature space to improve the efficiency and robustness of the segmentation approach. Upon comparison, the proposed method demonstrates good performance in achieving an automatic segmentation on various real skin data collected by ourselves and those downloaded from public dataset.

1. Introduction

Melanoma becomes one of the most common skin cancers in the UK. Most melanoma originates from irregular spreading of melanocyte cells which are responsible for producing the pigment melanin that colours the skin. As such melanoma usually has unique features of the colour and shape. Detection of a malignant tumor in its early stage not only reduces the mortality rate, but is helpful in reducing the expense associated with treatment. Measurement of features for diagnosis from images initially requires the detection and localization of the pigmented lesion area. Therefore image segmentation is considered to be the first step for achieving diagnosis in the skin cancers.

In order to accurately segment the pigmented area, Xu et al. [1] proposed a heuristic method using double thresholds to isolate skin from lesion through a few selected border points. Schmid [2] introduced a fuzzy c-means based lesion segmentation method which required centres of skin and tumor areas as a prior knowledge. In these supervised approaches, segmentation results are dependent on the initial selection of normal skin and the suspicious lesion areas. Iyatomi et al. [3] automatically extracted a dermatologist-like lesion region by combining pixel-based and region-based algorithms which rely on the approximation of the colour distributions of normal skin and pigmented lesion. However, colour information alone proved insufficient for a reliable automated segmentation of lesion [4]. Cluster overlaps in the colour feature space caused by additive noise as well as intrinsic artefacts usually results in poor skin-lesion separation.

In this paper, an adaptive mean-shift and maximal similarity based region merging method is used to achieve an automatic skin lesion segmentation. By appending the 2D coordinates to RGB colour feature space, a 5D feature space is achieved to improve the segmentation result. The experiments validated that the proposed approach can automatically and accurately separate pigmented lesion and the surrounding normal skin on the data acquired from various imaging devices.

2. Methodology

Before proceeding segmentation task, a 2D anisotropic diffusion algorithm [5] is first applied with skin image to reduce the noise while preserving the significant features; then a contrast limited adaptive histogram equalization [6] is used to deal with the large variation in the natural

skin pigmentation across the population; finally the image pixel values in RGB channels are normalized and stretched to the same range [0 255] to obtain a similar dynamic range as well as to reduce the sensitivity to lighting conditions.

There are essentially three parts in the entire segmentation framework. An adaptive mean-shift algorithm [7] is first applied to a compact of 5D feature vector, which includes colour and spatial information of each pixel. This step outputs an initial set of clusters. A further iterative region merging stage is followed up to prune the number of clusters by grouping the clusters with maximal similarity in the colour histogram [8]. Finally a weighted kernel *k*-means [9] is introduced to assign the remaining clusters to normal skin or pigmented lesion.

2.1 Adaptive Mean-shift Clustering

The mean-shift algorithm is a nonparametric clustering technique which does not require prior knowledge of the number of clusters, and does not constrain the shape of the clusters [7]. Let $p_i \in O^d, \mathrm{i} = 1,2,3,\cdots$ be the set of feature vectors in a d-dimensional feature space. The implication of the mean-shift property is that the iterative procedure

$$p^{(j+1)} = \frac{\sum_{i=1}^{n} \frac{1}{h_i^{d+2}} x_i g\left(\left\|\frac{p^{(j)}-p_i}{h_i}\right\|^2\right)}{\sum_{i=1}^{n} \frac{1}{h_i^{d+2}} g\left(\left\|\frac{p^{(j)}-p_i}{h_i}\right\|^2\right)} \qquad j = 1,2,3,\cdots \tag{1}$$

is a hill climbing process to the nearest stationary density point, which guarantees the convergence to the local maximum after a few iterations. Here function $g(p)$, is the profile of the associated kernel $G(p) = \mathrm{c}_{g,d} g(\|p\|^2)$, h_i is the window size determining the range of influence of the kernel located in p_i, and j is the iteration number.

Constant mean-shift using a fixed window size h instead of h_i for each feature vector might result in clusters over splitting due to its small value or unexpected excessive merging from the selection of a large window. Therefore an adaptive value of h_i is required at each feature point p_i. This so called adaptive mean-shift (AMS) [10] jointed with colour-spatial feature space forms the basis of our segmentation scheme.

Taking $h_i = \|p_i - p_{i,k}\|$, where h_i is the L_1 distance between p_i and its *k*-nearest-neighbour $p_{i,k}$. The window size h_i is the only numerical parameter in the AMS, thus the choice of k will have a significant influence on the initial clustering from the AMS and will further affect the final segmentation result. The experimental in section 3.1 will demonstrate that the selection of k can be flexible in a large interval without greatly influence in the segmentation accuracy. An example of malignant melanoma is shown in Fig.1(a). Fig.1(b) gives the clustering map output from the AMS.

2.2 Iterative Cluster Pruning

After the initial AMS, there are still hundreds or thousands clusters left. Therefore a maximal similarity based region merging algorithm (MSRM) is carried out on the analysis of colour histogram. This method is adaptive to the content of the input image and avoids the problem of a preset threshold in advance [8]. So it is appropriate for a large variation of natural skin pigment across various skin tumors.

We quantize each RGB channel into 16 bins and therefore obtain a colour histogram of $16 * 16 * 16 = 4096$ bins for each region. Then Bhattacharyya coefficient $\rho(\mathrm{R},\mathrm{Q}) = \sum_{\mu=1}^{4096} \sqrt{\mathrm{Hist}_{\mathrm{R}}^{\mu} \cdot \mathrm{Hist}_{\mathrm{Q}}^{\mu}}$ is used to describe the similarity between regions R and Q, where $\mathrm{Hist}_{\mathrm{R}}$ and $\mathrm{Hist}_{\mathrm{Q}}$ are the

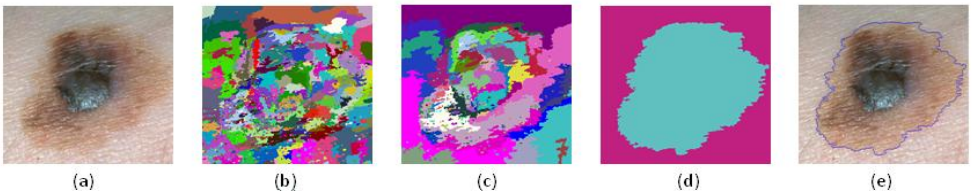

Fig.1: An example of the segmentation of real skin image. (a) Input Image. (b) 1st segmentation map after AMS. (c) 2nd segmentation map after MSRM. (d) Final segmentation map. (e) Segmented borders overlaid on the original image.

quantized colour histograms, and the superscript μ represents the μ^{th}element inside them. Two similar regions therefore result in similar colour histograms, and hence have a higher Bhattacharyya coefficient ρ between them.

Suppose N is a group set of clustering regions output from the AMS step, a set of the adjacent regions $\hat{S}_R = \{Z_i\}, i = 1,2,\ldots,u$ (u is the number of adjacent regions of R) is formed for each $R \in N$. And for each Z_i, another set of adjacent regions $\hat{S}_{Z_i} = \left\{S_j^{Z_i}\right\}, j = 1,2,\ldots,v$ (v is the number of adjacent regions of Z_i) is constructed. The similarity between region Z_i and its adjacent regions can be calculated according to the following merging rule:

$$\rho(R, Z_i) = \begin{cases} true & if\ \rho(R, Z_i) = \max_{j=1,2,\ldots,v}\left(Z_i, S_j^{Z_i}\right) \\ false & otherwise \end{cases} \tag{2}$$

which means that the selected region R will be merged with its adjacent region Z_i only if the similarity $\rho(R, Z_i)$ is the maximal one among all the similarities within $\rho(Z_i, \hat{S}_{Z_i})$. Fig.1(c) shows the 2nd segmentation map after the MSRM.

2.3 Kernel K-means Clustering in Colour Feature Space

The weighted kernel *k*-means [8] is introduced as the last step to assign the remaining clusters to normal skin or pigmented lesion according to their RGB colour values. The objective function is defined as:

$$E\left(\{\pi_c\}_{c=1}^{l}\right) = \sum_{c=1}^{l} \sum_{X \in \pi_c} w(X)\, \|\Phi(X) - m_c\|^2$$

$$m_c = \frac{\sum_{Y \in \pi_c} w(Y)\Phi(Y)}{\sum_{Y \in \pi_c} w(Y)}, \qquad c = 1,2,..l \tag{3}$$

where X and Y are the colour vectors for two different clusters output from MSRM, and these two clusters are assigned to the same class c in kernel *k*-means step. $w(X)$ and $w(Y)$ are the weights standing for the relative portion of the total number of points inside cluster X and cluster Y respectively. Fig.1(d) shows the final segmentation result after clustering and Fig.1(e) outlines the border over the original image.

3. Experimental Results

The proposed segmentation framework is validated on 113 sets of real skin data: 74 from our own dataset captured at Frenchay hospital and 39 sets of public dataset downloaded from Dermatology Image Atlas [9]. Throughout the experiments, manual segmentations given by dermatologists are used as ground truth for the performance evaluation. The Tanimoto coefficient $TC(i) = \frac{N_{AB}^i}{N_A^i + N_B^i - N_{AB}^i} \times 100\%$ is used to qualify the accuracy, where i is the cluster index represents

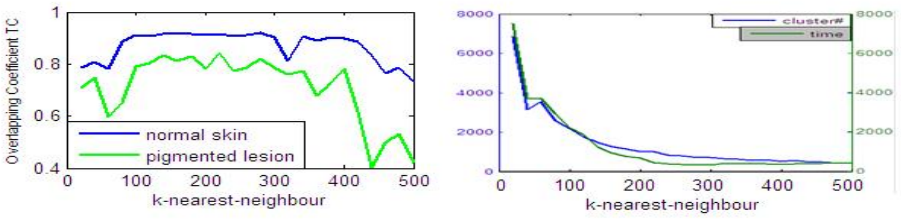

Fig.2: Sensitivity of k. (a) Overlapping Coefficient TC of k range from [20 500] for normal skin (blue) and pigmented lesion (green). (b) Number of remaining modes (blue) and computation time (green) plot together. Vertical axes in the left and right represent mode numbers and seconds.

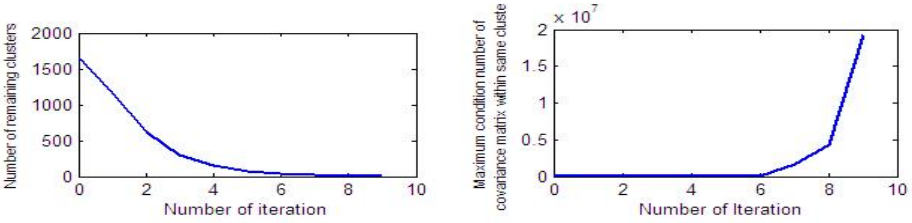

Fig.3: (a) Change of number of clusters with the increase of iteration number. (b) Change of maximum colour variation with the increase of iteration number.

skin/lesion, N_{AB}^{i} denotes the number of pixels assigned to skin/lesion by ground truth and automated segmentation simultaneously. N_A^i and N_B^i are the numbers of pixels of skin/lesion for ground truth and computed segmentation respectively.

3.1 Parameter Selection

3.1.1 Sensitivity of k-Nearest-Neighbour

We randomly selected images and change the variable k ranging from 20 to 500 to evaluate the algorithm sensitivity to it. Here the image shown in Fig.1(a) is used as an example because the colour inside the pigmented lesion greatly varied; and the contrast between skin and part of the pigmented lesion areas is low. These properties can be the excellent factors for the evaluation of the selection of k.

Fig.2(a) plots the coefficient TC corresponding to each k value for both normal skin and pigmented lesions. It can be observed that both TC give high values and stay constant when k is in the range of [100 300]. A significant decrease occurs thereafter, especially for k larger than 400. When the large number of k-nearest-neighbour is introduced the computed window size h_i in the AMS also increases, which may cause unacceptable smoothing in the low contrast region; whereas when k is in the range [20 80], the segmentation results are also not very good due to the sensitivity to the tiny artefacts caused by small h_i in applying the AMS. From Fig.2(b) it is evident that the running time greatly decreases with the increase of k, because fewer remaining clusters need to be merged in the cluster pruning step. The k nearest-neighbour is set to 240 in our work by comprehensively considering the accuracy as well as computation efficiency.

3.1.2 MSRM and Kernel k-means

In theory, a maximal similarity based region merging and weighted kernel k-means algorithm could be individually applied after AMS until the desired clustering number is achieved.

However, we apply both methods in order to avoid the drawbacks of each technique, and make them complement each other.

The maximal similarity based region merging algorithm utilizes the "max" operator which is sensitive to outliers and therefore may result in excessive merging with the increase in the number of iteration; whereas kernel k-means is unsuitable for the large images, as the kernel matrix makes the method inefficient for a standard PC due to the memory limitation. Taking the same image in Fig.1 as an example, Fig.3 plots the cluster number and the largest condition number of the covariance matrix within the same cluster in MSRM for each iteration. With an increase in iterations, the number of remaining clusters decreases while the largest condition number increases exponentially. From 7th iteration, the growth of largest condition number starts to change at a significant pace between two successive iterations, which means excessive merging might take place. In order to prevent the risk of overly merging, region merging stops when the number of remaining clusters reaches or is less than 0.25 times the number of clusters obtained from the AMS.

3.2 Performance Evaluation

In this experiment, we investigate the performance of the proposed approach by comparing it with four state-of-art segmentation techniques, including double threshold, fuzzy c-means, N-cut algorithm and active contour [1][2][10][11]. The resultant segmentations for eight randomly selected images are shown in Fig.4, and the statistics for the whole 113 sets of images are given in Table.1. It can be observed that our segmentation scheme gives the highest average coefficient TC and lowest standard deviation of lesion. This demonstrates that the proposed segmentation scheme is more accurate as well as consistent in the segmentation of skin lesions. Moreover, the proposed approach performs well on both our own and public datasets. But we also noticed that the only image to fail using our method is associated with very strong specular artefacts inside, and none of the other methods could provide reliable segmentation either.

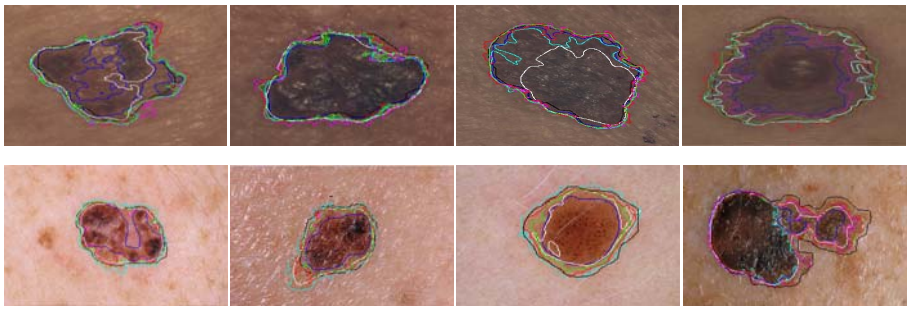

Fig.4: Eight example images: top row shows the images in our dataset, bottom row are the images from public dataset. Ground truth (black), double threshold (blue), fuzzy c-means (cyan), N-cut (white), active contour (magenta), our method with colour space only (green), and our method with colour-spatial space (red).

Algorithms	Ave. TC of lesion	STD of TC	Worst TC of lesion	Failed NO.
Double threshold	0.7426	0.1634	0.3781	5
Fuzzy C-means	0.8018	0.1117	0.4932	5
N-Cut	0.7614	0.1574	0.3290	8
Active contour	0.8216	0.1084	0.5138	3
Colour space only	0.8112	0.1138	0.4754	4
Spatial-Colour space	0.8523	0.0513	0.7451	1

Table 1: Average and standard deviation of skin lesion for 113 test images.

4. Conclusions

This paper presents an automated skin lesion segmentation approach to separate pigmented lesions from normal skin. In the comparison with four other state-of-art algorithms, the approach proposed gives the highest average overlapping coefficient TC with lowest standard deviation. Moreover, as the adaptive mean-shift associates the spatial-colour coherence and groups neighbouring pixels to the close cluster in spite of the large local colour variations, spatial information proves helpful in improving the segmentation by increasing the accuracy from 81.12% to 85.23% as well as halving the deviation from 11.38% to 5.13%.

5. References

[1] Xu, L., Jackowski, M., Goshtasby, A.: Segmentation of skin cancer images. Image Vision and Computering. 17(1), 65—74, 1999.

[2] Schmid, P.: Segmentation of digitized dermatoscopic images by two-dimensional color clustering. IEEE Trans. Med. Imaging, 18(2), 164—171, 1999.

[3] Ivatomi, H. Oka, H. Saito, M.: Quantitative assessment of tumour area extraction from dermoscopy images and evaluation of the compute-based methods for automatic melanoma diagnostic system. Melanoma Research 16(2), 183—190, 2006.

[4] Zhou H, Chen M, Zou L et al. Spatially Constrained Segmentation of Dermoscopy Images. Proc. of the IEEE Int. Symposium on Biomedical Imaging, pp. 800-803, 2008.

[5] Perona, P., Malik, J.: Scale-Space and Edge Detection Using Anisotropic Diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7), 629—639, 1990.

[6] Zuiderveld, K., “Contrast limited adaptive histogram equalization”, Graphics gems IV, Academic Press Professional, Inc., pp: 474—485, 1994.

[7] Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(5), 603—619, 2002.

[8] Ning, J., Zhang, L., Zhang, D., Wu, C.: Interactive image segmentation by maximal similarity based region merging. Pattern Recognition, 43, 445—456, 2010.

[9] Dhillon, I. S., Guan, Y., Kulis, B.: Kernel k-means, spectral clustering and normalized cuts. Proc. 10th ACM Knowledge Discover and Data Mining, pp: 551—556, 2004.

[10] Mayer, A., Greenspan. H.: An Adaptive Mean-Shift Framework for MRI Brain Segmentation. IEEE Transactions on Medical Imaging, 28(8): 1238—1250, 2009.

[11] Shi, J., Malik, J.: Normalized Cuts and Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8): 888—905, 2000.

[12] Houhou, N., Thiran, J., Bresson, X.: Fast Texture Segmentation based on Semi-Local Region Descriptor and Active Contour. Numerical Mathematics: Theory, Methods and Applications, 2(4): 445—468, 2009.