International Journal of Soft Computing 10 (2): 163-168, 2015
ISSN: 1816-9503

Combined Approach on Cervical Cytology Image Segmentation

[1]S. Anantha Sivaprakasam and [2]E.R. Naganathan
[1]Department of MCA, P.S.R. Engineering College, Sivakasi, India
[2]Department of Computer Science and Engineering, Hindustan University, Chennai, India

Abstract: Medical image segmentation is a complex and challenging task due to the inherent nature of the image. A number of different segmentation algorithms for these images have been developed. The single arithmetic of colour image segmentation has some deficiencies and defects, so we can use combine different algorithms according to the actual situation for segmentation. Here, researchers propose a new and novel approach called colour image segmentation based on JSEG and Robust k-means algorithms applied on cytological image. This algorithm segments the cytological image properly without manual parameter adjustment for each image and simplifies texture and colour. In this, the segmentation consists of two major stages. In the first stage, the cervical colour image is quantized to represent several representing classes which can be used to differentiate regions in the image. In the second stage, hit rate regions with similar Colour Regions Merging algorithm, i.e., Robust k-means algorithm is applied. Experiment results show that this method provides good segmentation results on different cytological images.

Key words: Cytological image segmentation, cluster, region growing, hit ratio region, k-means

INTRODUCTION

Image segmentation is one of the key concepts in image processing and used in many applications. Colour image segmentation plays a vital role in many areas, especially in medical field. From the segmentation results, it is possible to identify the region of interest and objects needed for the analysis in the image. Recently, many techniques are employed in image segmentation. For example, stochastic model based approaches (Belongie *et al*., 1998), morphological watershed based region growing (Shafarenko *et al*., 1997), k-means clustering (Naganathan *et al*., 2013; Meskine and Bahloul, 2012), graph partitioning, etc. (Shi and Malik, 1997).

Segmentation can be carried out on various types of images. For example, satellite image processing in the application of remote sensing, the brain MR image analysis in the application of medicine, the plates of illegal vehicle region segmentation in the traffic image analysis, the image region of interest extraction in the object-oriented image compression and content-based image retrieval (Belongie *et al*., 1998). Quantitative evaluation methods have also been suggested (Borsotti *et al*., 1998)

In these applications, image segmentation is usually used for image analysis, identification and compress code, etc. As far as colour image segmentation on cervical cytology image is concerned most of proposed Colour Image Segmentation Methods are the combination of the existing gray scale image segmentation methods.

Commonly used cervical colour image segmentation methods are histogram threshold, feature space clustering, region-based approach based on Edge Detection Methods, Fuzzy Methods, artificial neural network approach (Lulio *et al*., 2011) based on Physical Model Methods, Stochastic Model based approaches (Delignon *et al*., 1997; Ma and Manjunath, 1997), etc.

Colour images with homogenous regions are segmented with an algorithm is to generate clusters in the colour space for segmenting the colour images with homogeneous region (Comaniciu and Meer, 1997). The spatial arrangement of pixels using a region-growing is one of the techniques to segment images with texture whereby a homogeneity mode is defined with pixels grouped in the segmented region. In order to segment texture images, one must consider different scales of images. Here, an unsupervised Colour-Texture Regions Segmentation algorithm (Panjwani and Healey, 1995) is ideal. It tests the homogeneity of a given colour-texture pattern which is computationally more feasible than model parameter estimation. It deals with the following assumptions for the acquired image:

- Image containing homogeneous colour-texture regions
- Colour information is represented by quantized colours
- Colours between two neighbouring regions are distinguishable

Corresponding Author: S. Anantha Sivaprakasam, Department of MCA, P.S.R. Engineering College, Sivakasi, India

In this study, we deal with cervical image segmentation based on J-seg and Robust k-means algorithm. The segmentation of cytological images is properly performed without manual parameter adjustment for each image and simplifies texture and colour. This method consists of two phase. In the first phase, colour quantization is carried out. In the second phase, spatial segmentation is carried out.

MATERIALS AND METHODS

First, colours in the image are coarsely quantized without significantly degrading the colour quality. The purpose is to extract a few representing colours which can be used to differentiating neighboring regions in the image. A good colour quantization technique is important for the segmentation process and quantized image is segmented using J-seg and Robust k-means algorithm.

Colour quantization: In the first stage, the colour space is quantized with little perceptual degradation by using the Quantization algorithm (Deng *et al*., 1999a) with minimum colouring. Consider each colour is associated with a class. The original image pixels are replaced by classes to form the class maps (Duda and Hart, 1970) (texture composition) for the next stage. Before performing the hit rate regions, the J-image (Deng *et al*., 1999b) a class map for each windowed colour region, the edges and textures of the processing image is represented as positive and negative values must be created with pixel values used as a similarity algorithm for the hit rate region. These values are called J-values that are calculated from a window placed on the quantized image where the J-value belongs (Deng *et al*., 1999b). Figure 1 shows the image segmentation steps.

Here, unsupervised segmentation method is used. This segmentation (Lulio *et al*., 2011) method is to extract representative colours differentiating neighbouring regions in the acquired image. To smooth the image and existing noise, the colour quantization using peer group filtering is applied through perceptual weighting on individual pixels. Then, new values indicating the smoothness of the local areas are obtained and a weight is assigned to each pixel, prioritizing textured areas to smooth areas. These areas are identified with a quantization vector to the pixel colours based on General Lloyd Algorithm (GLA) which the perceptually uniform L×u×v colour space is adopted, presenting the overall distortion D:

$$D = \sum_i D_i = \sum_i \sum_n v(n) \| x(n)-c_i \|^2 \to x(n)C_i \qquad (1)$$

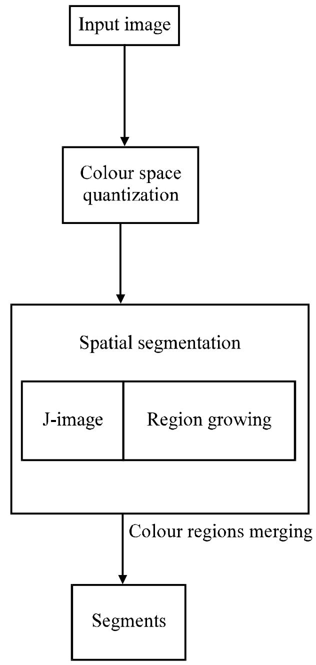

Fig. 1: The proposed image segmentation steps

And it is derived for:

$$c_i = \frac{\sum v(n)x(n)}{\sum v(n)} \to x(n)C_i \qquad (2)$$

Where:

c_i = The centroid of cluster C_i

$x(n)$ and $v(n)$ = The colour vector and the perceptual weight for pixel n

D_i = The total distortation for C_i

After cluster merging for colour quantization, a label is assigned for each quantized colour, representing a colour class for image pixels quantized to the same colour:

```
Segmented 1          Segmented 2

+++++++++            +#+ #+#+#+
+++++++++            -+- +-+ - +-
+++++++++            +#+#+# +#+
_____________        _____________
---------            - +-+-+- +-
---------            +#+#+#+#+
---------            -+-+- +- +-
_____________        _____________
#########             +#+#+#+#+
#########             -+-+-+ -+-
#########            +#+#+#+#+
```

Two different class-map representing three distinct classes of data points. In Fig. 1, class-map 1 indicates three regions containing a single class of data points for segmentation process and class-map 2 is not segmented indicating colour uniformity. The symbols (+, -, #) denotes the label values (J-value) for three distinct data points.

J-value: All necessary segmentation information, after colour quantization is extracted and relocated to a class-map (Belongie *et al*., 1998). A specific region contains pixels from a colour class set which is distributed in image regions. These regions, forming each one, a class-map has distributed points in all spatial data segments (Deng *et al*., 1999b), corresponding a two-dimensional plane and represents the cartesian position vector (x, y). In order to calculate the J-value, Z is defined as the set of all points of quantized image then z = (x, y) with z∈Z and being m the average in all Z elements. C is the number of classes obtained in the quantization. Then, Z is classified into C classes, Z_i are the elements of Z belonging to class i where i = 1, ..., C and mi are the element averages in Z_i.

$$m = \frac{1}{N}\sum_{z \in Z} Z \tag{3}$$

$$m_i = \frac{1}{N_i}\sum_{z \in Z} Z \tag{4}$$

The J-values are as follow:

$$J = \frac{SB}{SW} = \frac{ST\text{-}Sw}{Sw} \tag{5}$$

Where:

$$S_T = \sum_{z \in Z} \|z\text{-}m\|^2 \tag{6}$$

$$S_w = \sum_{i=1}^{c} \times \sum_{z \in Z} \|z\text{-}mi\|^2 \tag{7}$$

The parameter S_T represents the sum of quantized image points within the average in all Z elements. Thereby, the relation between S_B and S_W, denotes the measures of distances of this class relation for arbitrary non-linear class distribution (Duda and Hart, 1970). J for higher values indicates with homogeneous colour regions. The distance and consequently, the J-value, decrease for images with uniformly colour classes. Each segmented region could be recalculated, instead of the entire class-map with new parameters adjustment for J average. J_K represents J calculated over region k, M_k is the number of points in region k, N is the total number of points in the class-map with all regions in class-map summation:

$$J = \frac{1}{N}\sum_{k} M_K \; J_K \tag{8}$$

For a fixed number of regions, a criterion for J is intended for lower values.

Spatial Segmentation algorithm: The characteristics of the J-images allow us to use a Region-Growing Method to segment the image. Here, Robust k-means algorithm is used.

Robust K-Means algorithm (RKM): This algorithm is an extension of k-means algorithm that removes outlier. This algorithm (Karsin, 2008) uses the Information Bottleneck Method as a foundation for its solution to geometric clustering problems. Clustering, from an information theory standpoint is a for lossy compression based on the ratio of datum to clusters. The main goal is to retain as must relevant information about the location of the data points while compressing the data point into the clusters. The RKM regulates this compression with the lagrange parameter λ. The clustering criterion of this algorithm is:

$$\max[I(x, c)\text{-}\lambda I(c, i)\; p(c|1)] \tag{9}$$

Where:

I = Data index
c = Cluster index
x = Location of the datum

The goal is to retain the maximum amount of relevant information using the λ. Figure 2 shows a flow chart of the steps in the Spatial Segmentation algorithm. Consider the original image as one initial region. The algorithm starts segments all the regions in the image at an initial large scale. It then repeats the same process on the newly segmented regions at the next smaller scale until the minimum specified scale is reached.

Valley determination: A set of small initial areas are determined as the base for region growing. These areas have the lowest local J-values simply called as valleys. A heuristics for the valley determination presupposed a condition for initial regions to be determined as the pattern growing. As follows:

- Calculate the average and the standard deviation of the local J-values in the region, denoted by σ_J and μ_J, respectively
- Set a threshold T_J at

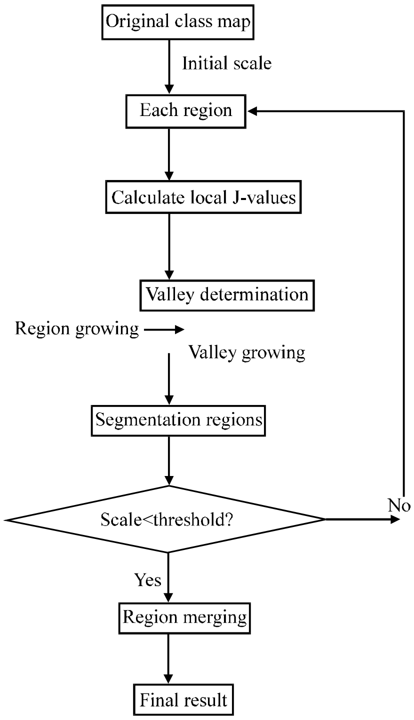

Fig. 2: Flow-chart of the steps in spatial segmentation

$$T_J = \mu_J + \alpha\sigma_J \tag{10}$$

The condition to candidate valley points for pixels with local J-values is determined $T_J > J$. Connect the points based on the 4 connectivity and obtain the valleys:

- For candidate valleys smaller than the spatial segmentation relation between scale and image size they are denoted as valleys
- A preset parameter values [-0.6, -0.4, -0.2, 0, 0.2, 0.4] is given for variable α which gives the most number of valley

Valley growing: The new regions are then grown from the valleys. It is slow to grow the valleys pixel by pixel. A faster approach is used in the implementation:

- Remove holes in the valleys
- Average the local J-values in the remaining unsegmented part of the region and connect pixels below the average to form growing areas. If a growing area is adjacent to one and only one valley, it is assigned to that valley
- Calculate local J-values for the remaining pixels at the next smaller scale to more accurately locate the boundaries
- Grow the remaining pixels one by one at the smallest scale. Unclassified pixels at the valley boundaries are stored in a buffer. Each time, the pixel with the minimum local J-value is assigned to its adjacent valley and the buffer is updated till all the pixels are classified

Region merge: An initial segmentation of the image is obtained by applying region growing. It often has over segmented regions. These regions are merged based on their colour similarity. The quantized colours are naturally colour histogram bins. The colour histogram features for each region are extracted and the distances between these features can be calculated. Since, the colours are very coarsely quantizes in this algorithm it is assumed that there are no correlation between the quantized colours. Therefore, a Euclidean distance measure is applied directly. First, distances between two neighbouring regions are calculated and stored in a distance table. The pair of regions with the minimum distance is merged together. The colour feature vector for the new region is calculated and the distance table is updated. The process continues until a maximum threshold for the distance is reached. After merging, the final segmentation results are obtained.

Algorithm:

a) Read the input image;
b) Convert RGB to L×a×b
c) Apply Colour Quantization algorithm
d) Find J-value
e) Apply Robust k-means algorithm
   i) Perform region growing and region merging process
f) Display the segmented image

RESULTS AND DISCUSSION

This algorithm is tested on a variety of cytological images. In this algorithm, the user gives the three parameters. The first parameter is a threshold for the colour quantization process. It determines the minimum distance between two quantized colours. The second one is the number of scales desired for the image. The last one is a threshold for region merging. These parameters are necessary because of the varying image characteristics in different applications. The algorithm works well on a variety of images using a fixed set of parameter values. This algorithm also provides good results when applied to gray scale images where intensity values are quantized the same way as the colours. Segmented images are dimmed to show boundaries (Fig. 3). It can be seen that the results are quite good. This system is developed in Matlab 7.5 Version.

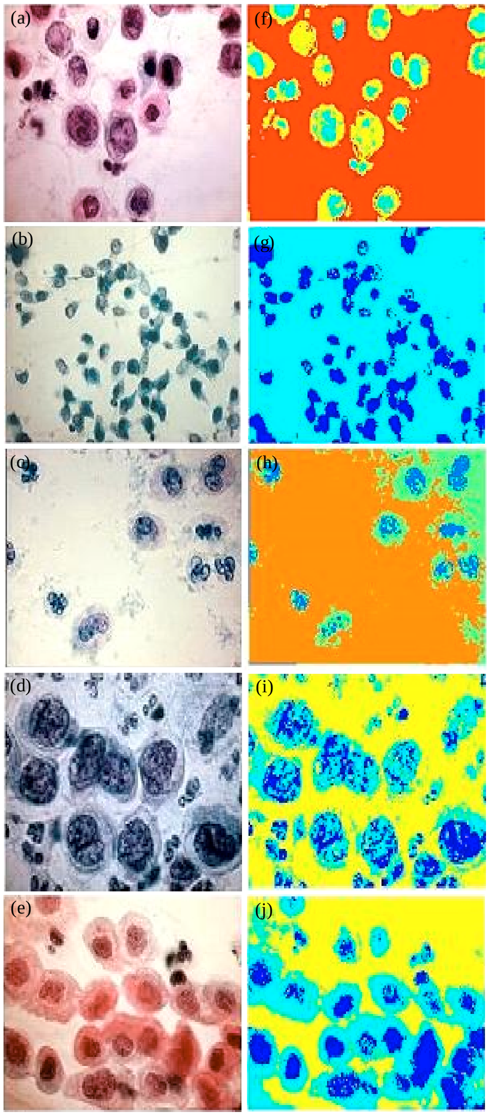

Fig. 3: a-d) The left images are original images and e-k) right images are segmented image after applying J-seg and RKM

CONCLUSION

In this research, a new approach for colour image segmentation is applied on cytological images. The segmentation consists of colour quantization and spatial segmentation. A criterion for good segmentation is proposed. Applying the criterion to local image windows result in J-images which can be segmented using a Multi-Scale Region Growing Method. Result show that this algorithm provides good segmentation on a variety of cytological images. This shows that the segmentation results are much better than other existing approaches.

Future research work is to handle the varying shades of an object due to illumination and segregate, the boundaries of two neighbour regions clearly.

REFERENCES

Belongie, S., C. Carson, H. Greenspan and J. Malik, 1998. Color-and texture-based image segmentation using EM and its application to content-based image retrieval. Proceedings of the 6th International Conference on Computer Vision, Bombay, January 4-7, 1998, IEEE Computer Society, Washington, DC. USA., pp: 675-682.

Borsotti, M., P. Campadelli and R. Schettini, 1998. Quantitative evaluation of color image segmentation results Pattern Recogn. Lett., 19: 741-747.

Comaniciu, D. and P. Meer, 1997. Robust analysis of feature spaces: Color image segmentation. Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 17-19, 1997, San Juan, pp: 750-755.

Delignon, Y., A. Marzouki and W. Pieczynski, 1997. Estimation of generalized mixtures and its application in image segmentation. IEEE Trans. Image Process., 6: 1364-1376.

Deng, Y., C. Kenney, M.S. Moore and B.S. Manjunath, 1999a. Peer group filtering and perceptual color image quantization. Proceedings of the IEEE International Symposium on Circuits and Systems, Volume 4, May 30-June 2, 1999, Orlando, FL., pp: 21-25.

Deng, Y., B.S. Manjunath and H. Shin, 1999b. Color image segmentation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 2, June 23-25, 1999, Fort Collins, CO., pp: 446-451.

Duda, R.O. and P.E. Hart, 1970. Pattern Classification and Scene Analysis. John Wiley and Sons, New York.

Karsin, B., 2008. Review of Robust k-means algorithm and comparison with other clustering methods. http://www2.hawaii.edu/~karsin/RKM/RKMAnalysis.pdf.

Lulio, LC., M.L. Tronco and A.J.V. Porto, 2011. JSEG Algorithm and Statistical ANN Image Segmentation Techniques for Natural Scenes. In: Image Segmentation, Ho, P.G. (Ed.). InTech, Croatia, pp: 343-364.

Ma, W.Y. and B.S. Manjunath, 1997. Edge flow: A framework of boundary detection and image segmentation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 17-19, 1997, San Juan, pp: 744-749.

Meskine, F. and S.N. Bahloul, 2012. Privacy preserving k-means clustering: A survey research. Int. Arab J. Inform. Technol., 9: 194-200.

Naganathan, E.R., S.A. Sivaprakasam and V.S. Kumar, 2013. Enhanced colour image segmentation on cervical cytology image. Proceedings of the International Conference on Applied Mathematics and Theoretical Computer Science, January 24-25, 2013, Tamilnadu, India, pp: 215-218.

Panjwani, D.K. and G. Healey, 1995. Markov random field models for unsupervised segmentation of textured color images. IEEE Trans. Pattern Anal. Mach. Intell., 17: 939-954.

Shafarenko, L., M. Petrou and J. Kittler, 1997. Automatic watershed segmentation of randomly textured color images. IEEE Trans. Image Process., 6: 1530-1544.

Shi, J. and J. Malik, 1997. Normalized cuts and image segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 17-19, 1997, San Juan, Puerto Rico, Pages: 731.