ISSN 2319 - 2720

Volume 1, No.1, July – August 2012
International Journal of Computing, Communications and Networking
Available Online at http://warse.org/pdfs/ijccn02112012.pdf

Cell Image Analysis using Fuzzy Clustering and Morphology based computation

C. K Nath[1], Jyotismita Talukdar[2]
[1] phtassam@gmail.com
[2] jyotismita4@gmail.com

ABSTRACT

Fuzzy c-Means algorithm and morphological analysis is used to study the microscopic images obtained from Pap smear test of cervical region. The clustering process is validated with two robust validation factors. The clustering process help in distinguishing the two primary parts of a cell: nucleus and cytoplasm. Morphological study helps the heath personals to gauge the extent of any normal or abnormal biological change in the cells.

1. INTRODUCTION

Fuzzy clustering is a strong tool handle ambiguities present the data. A medical image consisting of millions of pixels in a multidimensional space presents us a good example of such ambiguous data. Fuzzy approach can help to overcome the challenges presented by such data. Segmentation of medical images is task which requires high degree of fuzziness. Also a computer assisted method is less prone to human error.

Colour information is one of the most obvious and readily available information regarding an image. The pixel value of the images can be considered as input for an intelligent method like clustering. To achieve isolation of non-coherent objects morphological investigation is required.

In our study, through fuzzy clustering segmentation of the Pap smear image is done. Morphological analysis helps in detecting the contour of the objects that is the cell nuclei.

2. FUZZY CLUSTERING

Fuzzy clustering is a branch of cluster analysis. It is widely used in pattern recognition field. Fuzzy C-mean (FCM) algorithm is based on Euclidean distance. Suppose $X = \{x_1, x_2,....x_n\}$ is a sample set where n is the number of image pixels to determine their membership. u_{ik} is the membership degree of x_i belonging to i class and $U = \{u_{ik}\}$ is a fuzzy classification matrix which denotes the fuzzy membership of pixel k in cluster i. u_{ik} must satisfy the following condition:

$$\sum_{i=1}^{c} u_{ik} = 1, \forall k = 1,2,....n$$

the fuzzy membership value. The objective function used in traditional FCM algorithm is

$$J(U,V) = \sum_{k=1}^{n}\sum_{i=1}^{c} (u_{ik})^m \, (d_{ik})^2$$

'c' is the number of cluster, m is the fuzzifier, m>1, which controls the fuzziness of the method. u_{ik} is the membership value of the pixels and v_i is the cluster center in the subset i of the feature space. U is the fuzzy partition. The term $(d_{ik})^2$ is the Euclidean distance between a pixel and cluster center. FCM algorithm based on minimizing the dispersion between the cluster center i and pixel k to determine pixel membership using an objective function.

We constitute the membership values for three colour components R, G and B and these membership values are iteratively evaluated to obtain the optimal clustering result. Once the pixels are regrouped, the cluster centers need to be recomputed to minimize *J(V)*. For the new cluster centers, we can again regroup the pixels to reduce *J(V)*. The process can be repeated until J cannot be reduced any further. In summary, the c-means clustering procedure consists of the following steps:

I. Determine the number of clusters c.
II. Partition the input samples into c clusters based on an approximation. If no rule of approximation exists, the samples can be partitioned randomly.
III. Compute the cluster centers.
IV. Assign each input sample to the class of the closet cluster center.
V. Repeat steps 3 and 4 until J cannot be reduced any further.

3. CLUSTER VALIDATION

A cluster validity method evaluates the quality of the clustering result. We have used two validation parameters: Partition Index and Partition Entropy. The partition index of U, denoted by *P(U)*, gives an average of the squared values of the membership grades found in the partition matrix

$$P(U) = \frac{1}{N}\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^2$$

Partition index quantifies the fuzziness of the partition matrices so that they can be marked and select the one with the lowest fuzziness. Partition Entropy is defined as

$$H(U) = \frac{1}{N}\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}\ln(u_{ik})$$

The values of the partition entropy range from 0 to *ln(N)*. The highest value is obtained when there is a uniform distribution of membership grades.

4. MORPHLOGICAL ANALYSIS

Morphological analysis gives the idea about the shape of the object of study. Simple direction code or chain code can be used to trace the contour of an object. The chain code is chosen among 8 selected points on the boundary of the object (Figure 1). The angle between any consecutive lines connecting two consecutive pair of points is 45°. Any two opposite points can be taken as starting points of boundary. Suppose point 1 and point 5 are taken as the starting point of the boundary and let the respective boundaries be *b1* and *b5*. Then we can write mathematically

$b5 = b1 + p$

p is a real number.

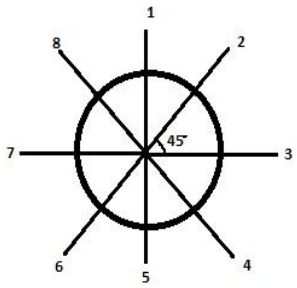

Figure 1: Direction code with 8 point

5. IMAGE COMPUTATION RESULT

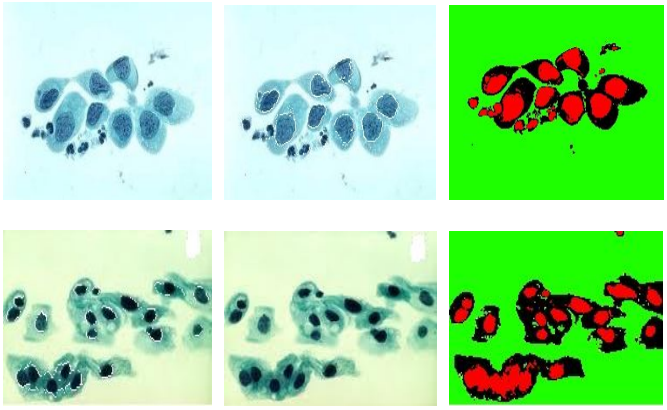

Figure 2: Result directional chain code and Fuzzy c-means algorithm

6. CONCLUSION

A simple yet worthy method of microscopic image analysis is presented in this paper. Colour information is passed as input to the conventional Fuzzy c-means algorithm for segmentation purpose. The cluster validation makes sure that the segmentation results are up to the mark. We have not tested images containing overlapping cells which may present some intricacies. The directional chain code traces the contour of the cell nuclei smoothly. The direction code can further be formulated to compute the shape deformation of the cells in presence of any disease. In our future work, we intend to achieve complete automation of the process by including a routine for choosing optimal cluster number. Other future works include reduction of the computational time, determination of shape deformation of the cells and classification of the cells as normal or abnormal based on some clinically defined parameters.

REFERENCES

1. A.K. Jain, M. N. Murthy and P. J. Flynn. **Data Clustering: A Review**, ACM Computing Surveys, Vol. 31, No. 3, September, 1999

2. Chin Wen et al. **Automatic Segmentation of abnormal cell nuclei from microscopic image analysis for Cervical Cancer screening**, Proc. IEEE 3rd International Conference on Neuro/Molecular Medicine and Engineering, 2009

3. George Dounias et al. **Automated identification of Cancerous smears using various competitive intelligent techniques**, Oncology Reports, 15, 2006

4. J. C. Bezdek. **Pattern recognition with Fuzzy Objective function algorithm**, Plenum Press, New York, 1981

5. J. C. Dunn. **A Fuzzy relative of the ISODATA process and its use in detecting compact, well-separated clusters**, J. Cybernets, Vol. 31. pp. 32-57, 1973

6. Jan Jantzen and George Dounias. **Analysis of PAP smear image data**, Proc. of Nature Inspired Smart Information System, NISIS 2006.

7. Kwang-Baek Kim, Sungshin Kim and Gwang-Ha Kim. **Nucleus Classification and Recognition of Uterine Cervical Pap-Smear Using FCM Clustering,**, ICANNGA 2007, Part II, LNCS 4432, pp. 290 – 299, Springer-Verlag Berlin Heidelberg, 2007

8. M. E. Plissiti et al. **Automated segmentation of cell nuclei in Pap smear images**, Proc. of IEEE International Special Topic Conference on Information Technology in Biomedicine, Greece, 2006

9. R. C. Gonzalez, R. E. Woods. **Digital Image Processing**, 2nd Edition, Pearson education India, New Delhi, 2003

10. Witold Pedrycz. **Knowledge-Based Clustering**, John Wiley & Sons, Inc. 2005.