International Journal of Natural and Engineering Sciences 4 (2): 47-50, 2010
ISSN: 1307-1149, E-ISSN: 2146-0086, www.nobel.gen.tr

A Pilot Study on Image Analysis Techniques for Extracting Early Uterine Cervix Cancer Cell Features

Babak Sokouti

Faculty of BioMedical Engineering, Islamic Azad University-Tabriz Branch, Tabriz, Iran

Corresponding author
Email:b.sokouti@gmail.com

Received : March 10, 2010
Accepted : May 15, 2010

Abstract

The second most common and preventable form of cancer among women worldwide is cervical cancer in which the signs for this disease can be detected in the early Pap smear screening of cervical cells. To improve the efficiency of expert diagnosis, we will need to automate the feature extraction of cervical cancer cells by the means of image processing techniques. This article employs image processing techniques to get the special features of normal, precancerous and cancerous cell images. We extract spectral features for cervical cancer cell detection. This article uses the noise decrease filters, OTSU threshold to make it ready for processing through 2-D Fourier and logarithmic transforms. By drawing the linear plot, we will be able to extract the feature of normal, precancerous and cancerous cells according to the texture and morphology automatically. These linear plots will be unique which can classify the cells in three groups normal, precancerous and cancerous cells.

The experiment shows that extracted unique features for each cell will provide evidences for diagnoses even in cytopathology images in which the nucleus and cytoplasm segmentation algorithms suffer from complex overlaying cells.

Keywords: Cytopathological cell images, Pap smear, cervical cancer, medical image analysis

INTRODUCTION

Cervical cancer is a major issue in women's health today, especially in developing countries. It takes years for an abnormal cell to grow, and we need the earliest possible sign of abnormality to let for the earliest treatment [1, 2, 3]. ThinPrep Pap smear screening can be useful as an early diagnostic process which can diagnose the cervical cancer [4, 5]. Finding abnormal cells in ThinPrep Pap smears is an error-prone and difficult problem for pathologists. Therefore, the need for an automated screening tool is desirable [6 ,7, 8]. Most of the researches done on automatic cervical screening try to segment the nucleus and cytoplasm accurately to detect the abnormal cells. Even with 100% segmentation accuracy, the presence of blood, inflammatory cells or complex overlaying cells, cell shape features may fail to show the differences between normal and abnormal cells [9, 10, 11, 12]. In this paper, we propose a new approach from known techniques to distinguish normal, precancerous and cancerous cells. This takes advantages of spectral property and avoids segmentation difficulties. First, each cell image cropped from a cytopathological ThinPrep image is used as an input in the image processing algorithm. Then the output will be three unique linear plots which present different patterns in normal, precancerous and cancerous cells.

MATERIALS and METHODS

The images of uterine cervical cells are captured by a high resolution digital camera mounted on the microscope placed in ALZAHRA HOSPITAL of Tabriz. A Cell includes a nucleus surrounded by cytoplasm. As a traditional way, a pathologist evaluates the cytoplasm and the background of slide. The abnormality features are as below:

- Size: There is an increased size of the nucleus compared to the cytoplasm

- Shape: Smooth, circular, oval outline belongs to normal nuclei
- Texture: Rough textures belongs to abnormal nuclei
- Chromaticity: Abnormal nuclei are darker than normal ones

The digitized images in size of 1024*768 are acquired and saved as JPEG or JPG image file format. Then the pathologist crops a single specific cell from the image to extract its grey-scaled properties from the ThinPrep Pap smear screening images.

The image processing block diagram of ThinPrep images of cervical cells is shown in Figure 1. In the proposed block diagram, we are supposed to find a way to demonstrate a plot that shows the nuclei and cytoplasm rate distribution. The median filter as a spatial filter, whose response is based on ranking the pixels contained in the image, is a noise reducer of certain types of noises. It replaces the value of a pixel by the median of the grey levels in the neighbourhood pixel. A 2-D median filter of 5*5 neighbourhood is applied to the input pixel so the output pixel will contain the median value of this process. In the third step, sharpening images includes subtracting a blurred version of an image from an image itself to distinguish between nuclei, cytoplasm and other ingredients which can be expressed as:

$$f_s(x,y) = f(x,y) - \bar{f}(x,y) \quad (1)$$

In which $f_s(x,y)$ shows the sharpened image gained from unsharp masking and $\bar{f}(x,y)$ shows a blurred version of $f(x,y)$. It uses a Gaussian blur low pass filter with a standard deviation of 10 pixels with a total size of 15*15 filter and the scaling factor of 0.9.

After applying the median and unsharp mask filters, we segment the images into nuclei and background (including the cytoplasm and the image background) by OTSU threshold. We then use the texture and chromaticity features to evaluate the normality and abnormality (precancerous and cancerous) of the nuclei as the texture and big dark nuclei on its own is a reliable indicator of nuclear normality of abnormality. Also these features can be outstanding in comparison to the outcome of image segmentation. In the next step, we use threshold to get the binary version of the image. Threshold is a useful mean for separating objects from the background. If we choose T as a global threshold parameter, the threshold function will be achieved by the below formula:

$$g(x,y) = \begin{cases} 1 \\ 0 \end{cases} \quad (2)$$

By applying the threshold to the image, we can represent the cytoplasm as white and nuclei as black. The threshold function calculations are done by Otsu's method. Each $M \times N$ cell image can be represented as a function $f(x,y)$ $x = 0,..M-1, y = 0,..,N-1$. The 2-D Fourier Transform of f shown by $F(u,v)$ is given according to the equation below:

$$F(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(\frac{\boldsymbol{u}}{M}+\frac{\boldsymbol{y}}{N})} \quad (3)$$

For $u = 0,..M-1, v = 0,..,N-1$, we can expand this with sines and cosines in the frequency domain which is determined by u and v. This Transform is used to convert the threshold images in to spatial frequency domain. We need to place the zero frequency unit at the frequency space centre, we are able to shift the image before applying the transformation. The general form of log transform is illustrated as below which will be applied to the Fourier magnitude result as the whole result will be complex:

$$s = c\log(r+1) \quad (4)$$

In which c is a constant and we assume $r \geq 0$. A narrow range of grey level values in the input image will be mapped into wider range of output level through this transformation. This will expand the values of dark pixels in an image while compressing the higher level values which will reduce the DC value of all pixel values to present the details more than before. In this stage the properties of three types of cells will be detectable.

A mean filter is a spatial filtering that reduces the noise at the presence of noise in current images. This filter will calculate the average value of corrupted image $f(x,y)$ in an area defined by S_{xy} (represent set of coordinates in a rectangular subimage window of size $m \times n$, centred at point) showed by formula 5. As a result, a 5×5 mean filter will lessen the noise and by twice performing this filter the image will be smoothed. As by applying once, the noise reduction is not enough for automatic diagnosis.

$$\hat{g}(x,y) = \frac{1}{\boldsymbol{m}} \sum_{(s,t)\in S_{xy}} f(s,t) \quad (5)$$

The two dimensional images are needed to be analyzed by one dimensional image processing techniques. By drawing the linear plot of the image, we extract the centre row values of the images because of the high intensity of the values and plot them against their positions for further consideration.

RESULTS and DISCUSSION

The proposed bottom-up approach to cervical cancer detection in ThinPrep Pap smear images have been evaluated on a database containing 250 single cell images

randomly obtained from 20 patients (84 normal, 83 LSIL, 83 HSIL). First, all images are preprocessed to remove the existing noises and normalize the intensity. Then for classifying each cell as normal, LSIL and HSIL, after applying the noise reduction filters and transformations to spatial frequency domain, the algorithm extracts the linear plot of the cell images which are unique regarding their types. The result images for each processing step are also presented in Figure 2. As it is shown in Figure 3, a local minimum can be seen before the symmetry point and a local maximum after the symmetry point. This shows that the normal cell features are repeated regularly, but this feature is not seen in abnormal cells.

Similarly, according to the linear plot shown in Figure 4, the ascending graph property without a minimum or local maximum in the plot can be seen in the fully cancerous images.

The linear plot shown in Figure 5, shows a condition between the previous two cases (LSIL), unlike the normal cells. There is no local maximum, indicating that the cells have the ability of HSIL risk and are therefore preventable.

The extracted features in Figure 3,4 and 5 are related to 15 cells. The normal, LSIL, HSIL cells are included and indicates that each cell type whether a normal, LSIL or HSIL cell, it will have a unique frequency chart with 100% accuracy in comparison with other methods. This can specially be compared with nuclei and cytoplasm separation methods (with accuracy of 80% - 95%), this result can be regarded as a very accurate method in cervical cancer diagnosis field. The reason of why it is been said the accuracy is 100% without any statistical plots is provided by the means of the linear plots to manually classify the cells based on the plot structures. Further the plots can be used as an input for an intelligent neural network system for automatic classification.

The last histopathological images shown in Figure 6 is related to the presence of several cell images in which we run the image processing process and has reached in two important facts:

- The proposed algorithm expressed without the need to separate nuclei and cytoplasm is able to detect the presence of abnormal cells.
- And if there is a cell overlap, it can still be able to detect the presence of abnormal cells.

CONCLUSION

The separation of nuclei and cytoplasm by the means of image analysis is so difficult. So many researches have been done on Pap smear classification. These researches are done for getting features from frequency domain of the images [8, 13, 14 ,15]. In this paper a novel method for analysing the cervical cell images by using features obtained from images and graphs of linear spectrum in Fourier domain is offered. Thus, images can be classified in 3 distinct groups of normal, LSIL and HSIL. It is shown that features obtained from frequency analysis and Fourier transform, can be used to classify single cell images. Also, in cases where there are overlapping cells or nuclei and cytoplasm separation is difficult, we are still able to classify these kinds of cells by the proposed algorithm.

For the future work, we are working on broader samples to use them for classifying the cells with artificial neural networks.

Conflicts of interests

The authors had no competing interests to declare in relation to this article.

REFERENCES

[1] N.H Othman, Cancer of the cervix——from bleak past to bright future. Kelantan, Malaysia: Pustaka Reka Publishing Company; 2003.

[2] V. Linasmita , Cervical cancer screening in Thailand –FHI-satellite meeting on the prevention and early detection of cervical cancer in the Asia and the Pacific region, 2006

[3] M. Bazoon, D.A. Stacey ,C. Chen, et al. A hierarchical artificial neural network system for the classification of cervical cells. IEEE World Congress on Computational Intelligence.,IEEE International Conference on Neural Networks, 1994;3526: 3525-3529.

[4] R.A. Kemp,C. MacAulay,D. Garner , et al. Detection of malignancy associated changes in cervical cell nuclei using feed-forward neural networks. Analytical Cellular Pathology, 1997; 14:31-40.

[5] B.J. Pernick, R.E Kopp, J. Lisa , et al. Screening of cervical cytological samples using coherent optical processing. Part 1 (ET). Appl. Opt. , 1978; 17:21.

[6] I.W. Ricketts. Cervical cell image inspection - a task for artificial neural networks. Network: Computation in Neural Systems , 1992; 3:15 – 18,.

[7] I.W. Ricketts, H. Banda-Gamboa,A.Y. Cairns , et al. Towards the automated prescreening of cervical smears.Applications of Image Processing in Mass Health Screening, IEE Colloquium on, 1992; 7/1-7/4

[8] S.J. McKenna ,I.W. Ricketts,A.Y Cairns, et al. A comparison of neural network architectures for cervical cell classification. Third International Conference on Artificial Neural Networks, 1993;105-109,

[9] R.F. Walker, P. Jackway, B. Lovell, et al. Classification of Cervical Cell Nuclei using Morphological Segmentation and Textural Feature Extraction. Second Australian and New Zealand Conference on Intelligent Information Systems,1994; 297-301.

[10] J.S. Hallinan, Evolving neural networks for the classification of malignancy associated changes. Intelligent Data Engineering and Automated Learning - IDEAL 2005.
[11] N.A. Mat-Isa. Automated detection technique for pap smear images using moving K-means clustering and modified seed based region growing algorithm. Int J Comput Internet Manage , 2005;13(3):45-59
[12] M.Y. Mashor Hybrid training algorithm for RBF network. Int J Comput Internet Manage , 2000;8(2):50-65.
[13] S.J. McKenna,I.W. Ricketts ,A.Y. Cairns, et al. Cascadecorrelation neural networks for the classification of cervical cells. Neural Networks for Image Processing Applications, IEE Colloquium on, 1992; 5/1-5/4 .
[14] N.A. Mat-Isa, M.Y. Mashor ,N.H. Othman NH, An automated cervical pre-cancerous diagnostic system, Artificial Intelligence in Medicine ,2008;42: 1-11.
[15] I.W. Ricketts,H. Banda-Gamboa ,A.Y. Cairns, et al. Automatic classification of cervical cells-using the frequency domain. Applications of Image Processing in Mass Health Screening, IEE Colloquium on,1992; 9/1-9/4.