Segmentation of Stretched Pap Smear Cytology Images using Clustering Algorithm

R. Ghafar[1], N.A. Mat Isa[1], U.K. Ngah[1], M.Y. Mashor[1], N.H. Othman[2]
[1] School of Electrical & Electronic Engineering, Universiti Sains Malaysia, Engineering Campus, Nibong Tebal, Penang, Malaysia
[2] School of Medical Science, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

Abstract
Papanicolaou test or better known as Pap test is the most popular and effective screening test for cervical cancer. At time, however, the detection of abnormal or cancerous cervical cells can be missed due to technical and human errors. In certain cases, the Pap smear images are blurred and highly affected by unwanted noises. These factors can hide and obscure the important cervical cells morphologies thus increasing the rate of false diagnosis rate. In this study a segmentation technique for Pap smear cytology images is proposed. The proposed technique begins by applying stretching process to enhance the contrast of Pap smear cytology images. The stretched Pap smear cytology images will then be segmented into three cervical cell's structures; nucleus, cytoplasm and background. The results show that the proposed segmentation technique produced better segmentation performance compared to the conventional clustering algorithms. The size and shape of cervical cells are also maintained through the proposed technique. This effort will assist pathologists for easier and better cervical cells morphological extraction.

Keywords
Clustering algorithm, Pap smear image, segmentation, stretching process

Introduction
Cervical cancer is the second common malignancy among woman after breast cancer. Without any symptom in its early stage, screening process is very important to detect any abnormal or cancerous cervical cells. Papanicolaou test or Pap test is the most popular and effective screening examination for cervical precancerous stage (Breen *et al.*, 2001, Framer, 2001, Adami *et al.*, 1994). A computer model analysis has found that the proportion of cervical cancer reduction is approximately 30% for a smear taken once every 10 years, 80% for one done every five years, 90% for every three years, 91% for every two years and 92% for an annual smear (Kuie, 1996).

At time, however, the determination of abnormal or cancerous cervical cells can be missed due to technical or human errors. Studies by Othman *et al.* (1997) and Hislop *et al.* (1994) have shown that some Pap smear cytology images are blurred and highly affected by unwanted noises, such as blood, air artifact, vagina discharge etc. These problems can hide and obscure the important morphologies of cervical cells. Thus, the Pap smear cytology image will be referred as inadequate samples for cervical cancer screening process.

In this work, a segmentation technique for Pap smear cytology images is proposed. The technique consists of stretching and clustering process. The proposed segmentation technique will segment the Pap smear images into two cervical cells structures, comprising nucleus and cytoplasm, and background. The segmented Pap

smear cytology images will assist pathologists for easier cervical cells morphological extraction, such as shape and size of nucleus and cytoplasm. Through this effort, pathologists can be assisted in performing better cervical cancer screening process.

Methodology

As mentioned, the proposed segmentation technique combines the stretching and clustering process. The block diagram of the proposed method is shown in Figure 1.

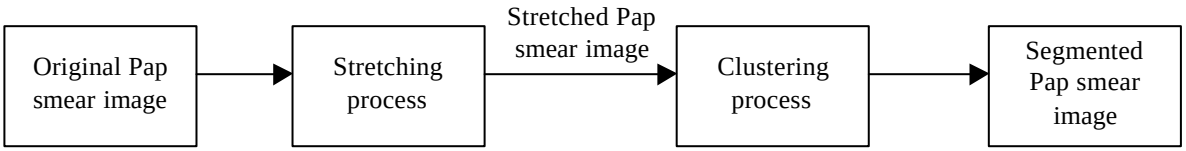


Figure 1: The proposed segmentation process for Pap smear cytology images.

In the stretching process, distribution of grey level of the original Pap smear images will be stretched. Ideally, the stretching process will result in an increase in the dynamic range of grey levels thereby enhancing the contrast of the Pap smear cytology image.

In the segmentation process, the stretched Pap smear images will be segmented into two cervical cell structures, which are the nucleus and cytoplasm, and the other is the background. This segmented Pap smear image will benefit pathologists in determining the shape and size of nucleus and cytoplasm during the morphological extraction process. In this study, two commonly used clustering algorithms were applied namely the non-adaptive *k*-means and fuzzy *c*-means clustering (Ghafar *et al.*, 2002, Mat-Isa *et al.*, 2003).

Results and Discussion

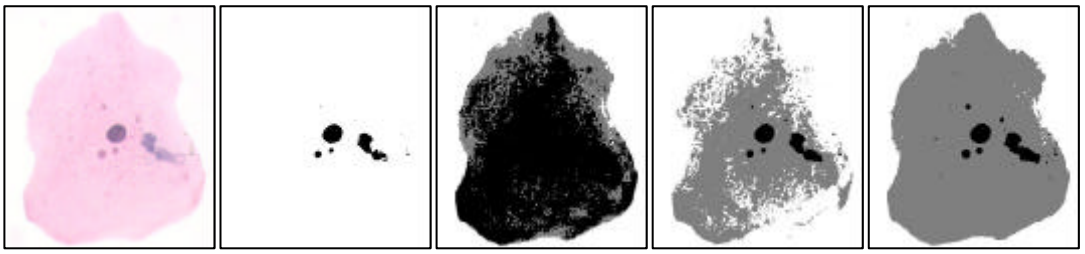
The proposed segmentation technique has been tested on three Pap smear cytology images; i.e. Pap1, Pap2 and Pap3. Figures 2, 3 and 4 show the results of the segmentation process of Pap1, Pap2 and Pap3 images respectively. The performance of the proposed segmentation technique was compared to the conventional non-adaptive *k*-means and fuzzy *c*-means clustering algorithm. For all of the figures, image (a) represents the original Pap smear cytology image, image (b) and (c) show the segmentation results for conventional non-adaptive *k*-means and fuzzy *c*-means clustering algorithm respectively, while image (d) and (e) represent the results for the proposed segmentation technique using the non-adaptive *k*-means and fuzzy *c*-means clustering respectively.

(a) (b) (c) (d) (e)

Figure 2: Results for segmentation of Pap1 image using various segmentation techniques.

(a) (b) (c) (d) (e)

Figure 3: Results for segmentation of Pap2 image using various segmentation techniques.



(a) (b) (c) (d) (e)

Figure 4: Results for segmentation of Pap3 image using various segmentation techniques.

Let us consider the segmentation performance produced by conventional non-adaptive *k*-means and fuzzy *c*-means clustering algorithms. For all of the Pap smear cytology images, the results show that both conventional clustering algorithms produced poor segmentation performance as shown in image (b) and (c) for non-adaptive *k*-means and fuzzy *c*-means clustering respectively. The reason is because both conventional clustering algorithms cannot avoid dead centre, centre redundancy and trapped centre in the local minimum problems, which result in improper segmentation process of those Pap smear cytology images (Mat-Isa *et al.*, 2003). The results also show that the shape and size of the nucleus and cytoplasm of those Pap smear cytology images are corrupted.

However, the non-adaptive *k*-means and fuzzy *c*-means clustering algorithms are able to reduce those problems when they were applied on stretched Pap smear cytology images as shown in Figure 2(d) and (e), Figure 3(d) and (e), and Figure 4(d) and (e) for Pap1, Pap2 and Pap3 images respectively. Both proposed techniques produced better performance as shown in image (d) and (e) using the non-adaptive *k*-means and fuzzy *c*-means clustering algorithms respectively. Overall, the proposed segmentation technique using fuzzy *c*-means clustering algorithm produced the best segmentation performance compared to other techniques. The shape and size of nucleus and cytoplasm regions were maintained. This ensures a proper segmented Pap smear image to assist the doctor for easier and better extraction of morphologies of cervical cells.

Conclusion

This study proposed segmentation process using non-adaptive *k*-means and fuzzy *c*-means clustering algorithms on stretched Pap smear cytology images. The results in the previous section showed that the proposed segmentation techniques produced better segmentation performance compared to conventional non-adaptive *k*-means and fuzzy *c*-means clustering algorithms. The proposed techniques have the capability to segment the Pap smear cytology images into nucleus, cytoplasm and background regions. The shape and size of nucleus and cytoplasm of cervical cells were also maintained. This can provide a clear segmented Pap smear cytology images for easier and better extraction process of morphologies of cervical cells by pathologists.

References

[1] Adami, H. O., Ponten, J., Sparen, P., Bergstrom, R., Gustafsson, L. & Friberg, L. G. 1994. Survival Trend After Invasive Cervical Cancer Diagnosis in Sweden Before and After Cytologic Screening. *Cancer*. **73**(1). 140-147.

[2] Breen, N., Wagener, D. K., Brown, M. L., Davis, W. W. and Barbash, R. B. 2001. Progress in Cancer Screening Over a Decade: Results of Cancer Screening From the 1987, 1992 and 1998 National Health Interview Surveys. *Journal of the National Cancer Institute.* **93**(22). 1704-1713.

[3] Framer, P. S. 2001. Screening for Cancer: Progress but More Can Be Done. *Journal of the National Cancer Institute.* **93**(22). 1676-1677.

[4] Ghafar, R., Mashor, M. Y. & Othman, N. H. 2002. Segmentation of Pap Smear Slides Images Using K-Means Clustering. *In Proceedings of Kuala Lumpur International Conference on Biomedical Engineering.* 41-43. Kuala Lumpur, Malaysia.

[5] Hislop, T. G., Band, P. R., Deschamps, M., Clarke, H. F., Smith, J. M. & Ng, V. T. Y. 1994. Cervical Cancer Screening in Canadian Native Women: Adequacy of The Papanicolaou Smear. *The Journal of Clinical Cytology and Cytopathology*. **38**(1). 29-32.

[6] Kuie, T. S. 1996. *Cervical Cancer: Its Causes and Prevention*. Singapore: Times Book Int.

[7] Mat-Isa, N. A., Mashor, M. Y. & Othman, N. H. 2003. Comparison of Segmentation Performance of Clustering Algorithms for Pap Smear Images. *Proceedings of International Conference on Robotics, Vision, Information and Signal Processing (ROVISP2003)*. 118–125. Penang, Malaysia. 22-24 January.

[8] Othman, N. H., Ayub, M. C., Aziz, W. A. A., Muda, M., Wahid, R. & Selvarajan, S. (1997). Pap Smears – Is It An Effective Screening Methods for Cervical Cancer Neoplasia? – An Experience with 2289 Cases. *The Malaysian Journal of Medical Sciences*. **4**(1). 45-50.