

THE POTENTIAL USE OF MODIFIED SEED-BASED REGION GROWING TECHNIQUE FOR AUTOMATIC DETECTION OF BREAST MICROCALCIFICATIONS AND TUMOUR AREAS

NOR ASHIDI MAT ISA¹, SHAHRILL SABARUDDIN², UMI KALTHUM NGAH³,
KAMAL ZUHAI RI ZAMLI⁴ & MASRI AH MOHD NOOR⁵

Abstract. Breast cancer, a common female malignancy, does not show any symptoms in its early stage. Screening tests are therefore important to reduce the death rates. By far, mammography and ultrasound tests prove to be the tests for early detection of breast cancer. However, these tests have limitations such as blurriness, darkness and the existence of unwanted noise on the breast image, which can obscure breast tumours and microcalcifications; the salient features detected in cases of abnormality. Many image processing techniques have been introduced in order to detect the edges or segment these breast cancer morphologies including the seed based region growing (SBRG) algorithm. However, two parameters, namely the seed point and the threshold value of the conventional SBRG algorithm need to be determined manually. This paper attempts to automatically find these parameters by proposing the modified version of the SBRG algorithm. The proposed algorithm which is called the modified seed based region growing (MSBRG) algorithm has been tested on mammogram and ultrasound images to detect the breast tumour as well as microcalcification. The results show that the proposed method successfully detects and distinguish breast tumour and microcalcifications from the background and unwanted noises. The borders (edges) of regions found by the MSBRG algorithm are perfectly thin and connected. Hence, the size and shape of the regions will not be corrupted. These will certainly assist doctors in the breast cancer screening process.

Keyword: Modified seed based region growing, mammogram images, ultrasound images, microcalcification, breast tumour

Abstrak. Barah payu dara iaitu barah paling biasa di kalangan wanita, tidak menunjukkan sebarang simptom pada peringkat awal. Oleh itu, ujian saringan awal adalah penting untuk mengesan penyakit ini pada peringkat awal yang seterusnya mampu mengurangkan kadar kematian. Walau bagaimanapun, ujian-ujian ini mempunyai limitasi dari segi imej payu dara yang diperolehi seperti kabur, gelap dan dipengaruhi hingar. Ini akan menutup kawasan tumor dan mikrokalsifikasi. Banyak teknik pemprosesan imej yang diperkenalkan untuk mengesan pinggir atau meruas morfologi barah payu dara ini. Ini termasuk algoritma pertumbuhan kawasan secara titik benih (SBRG). Walau bagaimanapun, 2 parameter iaitu titik benih dan nilai ambang algoritma SBRG perlu

^{1,2,3,4&5} Pusat Pengajian Kejuruteraan Elektrik dan Elektronik, Universiti Sains Malaysia, Kampus Kejuruteraan, 14300 Nibong Tebal, Pulau Pinang, Malaysia

¹ Tel: +604-5995999 ext 6051, Fax: +604-5941023, Email: ashidi@eng.usm.my

³ Tel: +604-5995999 ext 6022, Fax: +604-5941023, Email: umi@eng.usm.my

⁴ Tel: +604-5995999 ext 6079, Fax: +604-5941023, Email: eekamal@eng.usm.my

ditentukan secara manual oleh pengguna. Penyelidikan ini mencadangkan penentuan kedua-dua parameter ini secara automatik dengan melakukan pengubahsuaian terhadap algoritma SBRG. Algoritma yang dinamakan pertumbuhan kawasan secara titik benih terubahsuai (MSBRG) telah diuji ke atas imej mammogram dan ultra bunyi untuk mengesan tumor payu dara dan juga mikrokalsifikasi. Keputusan yang diperoleh menunjukkan bahawa teknik yang dicadangkan telah berjaya mengesan dan mengasingkan tumor dan mikrokalsifikasi daripada latar belakang dan hingar. Pinggir kawasan yang diperoleh oleh algoritma MSBRG adalah halus dan bersambung penuh. Ini memastikan bahawa saiz dan bentuk kawasan yang dikesan tidak terganggu. Ini mampu menolong doktor dalam proses pengesanan barah payu dara.

Kata kunci: Pertumbuhan kawasan secara titik benih terubahsuai, imej mammogram, imej ultra bunyi, mikrokalsifikasi, tumor payu dara

1.0 INTRODUCTION

To date, cancer is the second leading cause of death in Malaysia [1]. According to MAKNA, an estimated 40 000 new cancer cases were reported in Malaysia yearly [1]. Generally, about 20% of population or one in four may be at risk of cancer in their lifetime [1]. Breast cancer is the most common malignancy affecting women in Malaysia. It ranks second to lung cancer as the cause of deaths in Malaysia.

Like all other organs of the body, the breast is made of many types of cells. Normally, cells divide to produce more cells in order to support the needs of the body or to replace dead cells. Although not needed, rapidly dividing cells may form mass of tissue called growth or tumours, which can be benign or malignant [2]. When abnormal cell changes persist over time and become severe, these cells can develop into cancer cells. Over a long period of time, breast cancer cells can invade nearby tissues and may spread through the bloodstream and lymphatic system to other parts of the body [2].

The early stages of breast cancer may be asymptomatic. Therefore, screening is recommended. The symptoms will only be experienced by the patient as the breast tumour grows in size. The symptoms include lump or thickening in the breast or underarm, a change in size or shape of the breast, nipple discharge or nipple retraction, redness or scaling of the skin or nipple and ridges or pitting of the breast skin [2, 3]. Thus, early detection is vital to improve breast cancer prognosis. Prognosis of breast cancers has improved, which is believed to be due to improve treatment and early detection.

Mammography is by far, the proven method of early detection of breast cancer [1]. Mammography is a low-dose x-ray system for examination of the breasts. The images of the breasts are viewed on an x-ray film. Based on the images, radiologists identify any abnormalities if present in the breast, such as cysts, microcalcifications and tumours.

Sometimes, particularly if a doctor thinks that a patient may have a cyst or a patient is young and has dense breasts, the patient may be referred for an ultrasound examination. An ultrasound image is obtained using sound wave emissions. High-

frequency sound waves are passed through breast tissues and then converted into images on a viewing screen. Ultrasound complements other tests. If an abnormality is seen on mammography or felt by physical exam, ultrasound is the best way to find out if the abnormality is solid (such as a benign fibroadenoma or malignant lesion) or fluid-filled (such as a benign cyst). However, both mammography and ultrasound tests cannot conclude that any abnormalities occur in the breast is cancerous. Thus, further examination called biopsy is recommended. A biopsy test is the only way to know for sure if a patient has cancer because it allows doctor to obtain cells that can be examined under a microscope [1, 3].

Although both mammography and ultrasound tests are vital to improve breast cancer prognosis, these tests are not without their problems. Most mammogram and ultrasound images are of low contrast and are affected by unwanted noises. Furthermore, features indicative of abnormalities (e.g. the presence of microcalcifications which are tiny calcium deposits) are very subtle and minute [4]. Difficulties arise in the interpretation of the images by radiologists. This study attempts to apply the edge detection and region segmentation techniques on mammogram and ultrasound images in order to detect microcalcification as well as tumour areas. Automatic edge detection technique, namely modified seed-based region growing is proposed to ascertain the abnormalities in a region of interest in mammogram and ultrasound images.

2.0 AUTOMATIC EDGE DETECTION TECHNIQUE

In image processing field, many edge detection techniques with promising performance have been introduced [5 - 7]. The seed-based region growing (SBRG) becomes one of the well-known edge detection techniques. Studies by [5, 7, 8] proved that unlike gradient and second derivative methods; such as Laplacian, Sobel and Kirsch, the borders (edges) of regions found by SBRG algorithm are perfectly thin and connected. Thus, the size and shape of the regions will not be corrupted. Furthermore, the SBRG algorithm is more stable with respect to noise.

However, based on conventional implementation of SBRG algorithm described in [7, 8], the algorithm faces limitation where two parameters of the SBRG algorithm, namely the threshold value and the seed point, which must be determined manually by users. These are done on a trial and error basis and must be repeated until satisfactory results are obtained. This leads to time-consuming issue. Furthermore, the results obtained from the processes are highly subjective to the user. Based on [7, 8], the general concept of the conventional SBRG algorithm is to find the edges of the regions of interest by using region growing technique from a seed point. However, once the seed point does not fulfil certain conditions, as stated by [7, 8], the region growing will be stopped even though the whole image is not considered to be grown yet. The setback of the SBRG algorithm is the problem of a trapped seed point, which can cause incomplete edge detection process.

This study proposes the modified seed based region growing algorithm, whereby the threshold value and the seed point are automatically determined. The modified seed based region growing algorithm involves two stages. In the first stage, clustering algorithm is applied to the image to find the threshold value. The number of clusters depend on the number of region to be segmented on the image. The cluster values found by the clustering algorithm will be used as threshold values for each region of interest. In the second stage, this threshold values will be used in the conventional SBRG algorithm. Modifications are made to the conventional SBRG algorithm where the seed point is automatically determined. Section 2.1 will discuss the implementation of moving k -means clustering algorithm to find the threshold value while Section 2.2 will illustrate the MSBRG algorithm.

2.1 Moving K-Means Clustering Algorithm

As mentioned previously, in the modified seed based region growing algorithm, the clustering algorithm is implemented on the mammogram or ultrasound image to automatically find the threshold value for classifying two regions of clusters, i.e. the object of interest (the breast tumour or the microcalcification) and the background. Most of the previous studies on digital images used k -means and fuzzy c -means (FCM) clustering algorithms as clustering techniques [9, 10]. However, both clustering algorithms did not always produce good performance due to dead centre, centre redundancy and trapped centre at local minima problems [11]. In the year 2000, Mashor [11] proposed the moving k -means clustering algorithm (a modified version of k -means clustering). The moving k -means clustering algorithm has successfully reduced the problems stated. For this reason, the moving k -means clustering algorithm is proposed in this study to find the cluster values.

Consider one mammogram or ultrasound image with $N_b \times N_l$ pixels (where b and l are the number of row and column of the image respectively) to be clustered into n_c regions. Let $p(x,y)$ be a pixel to be clustered and C_j is the j -th cluster (centre) ($x = 1, 2, \dots, N_b, y = 1, 2, \dots, N_l$ and $j = 1, 2, \dots, n_c$). Based on these considerations, the implementation of the moving k -means clustering algorithm [11] in finding the cluster values of the medical image is as follows:

- (i) Initialise the centres and α_0 , and set $\alpha_a = \alpha_b = \alpha_0$ (where α_0 is a small constant value, $0 < \alpha_0 < 1/3$ and should be chosen to be inversely proportional to the number of centres. α_a and α_b are small constants).
- (ii) Consider the first pixel, $p(0,0)$. Calculate the Euclidean distance between the pixel and all the centres. Assign the pixel as the member of the centre with the smallest Euclidean distance.
- (iii) Repeat Step (ii) for all pixels in the image.
(Note: After this step, each pixel will be a member of one and only one of the centres with the nearest Euclidean distance value)

- (iv) Calculate the new centre positions for each centre using:

$$C_j = \frac{1}{n_j} \sum_{y \in c_j} \sum_{x \in c_j} p(x, y); \quad (1)$$

- (v) Check the fitness of each centre using equation:

$$f(C_j) = \sum_{y \in c_j} \sum_{x \in c_j} (\|p(x, y) - C_j\|)^2; \quad (2)$$

where $j = 1, 2, \dots, n_c$
 $x = 1, 2, \dots, N_c$
 $y = 1, 2, \dots, N_l$

- (vi) Find C_s and C_l , the centre that has the smallest and the largest value of $f(\bullet)$.
(vii) If $f(C_s) > \alpha_a f(C_l)$,
(a) Assign the members (pixels) of C_l to C_s if $p(x, y) < C_l$, where $x, y \in C_l$, and leave the rest of the members (pixels) to C_l .
(b) Recalculate the positions of C_s and C_l according to:

$$\left. \begin{aligned} C_s &= \frac{1}{n_s} \sum_{y \in c_s} \sum_{x \in c_s} p(x, y); \\ C_l &= \frac{1}{n_l} \sum_{y \in c_l} \sum_{x \in c_l} p(x, y); \end{aligned} \right\} \quad (3)$$

Note: C_s will give up its members (pixels) before Step vii(a) and, n_s and n_l in Equation (3) are the number of the new members (pixels) of C_s and C_l respectively, after the reassigning process in Step vii(b).

- (viii) Update α_a according to $\alpha_a = \alpha_a - \alpha_a/n_c$ and repeat Steps (vi) and (vii) until $f(C_s) \geq \alpha_a f(C_l)$.
(ix) Reassign all pixels to the nearest centre and recalculate the centre positions using Equation (1).
(x) Update α_a and α_b according to $\alpha_a = \alpha_0$ and $\alpha_b = \alpha_b - \alpha_b/n_c$ respectively, and repeat Steps (v) to (ix) until $f(C_s) \geq \alpha_b f(C_l)$.
(xi) Sort the centres in ascending order where $C_1 < C_2 < \dots < C_{n_c}$.

(Note: In this study, two cluster values, C_1 and C_2 , were determined where each cluster value represents the tumour or microcalcification and background regions respectively.)

2.2 The Modified Seed Based Region Growing Algorithm

Having found the threshold value from the moving k -means clustering algorithm, the modified seed based region growing (MSBRG) algorithm would then be applied on mammogram or ultrasound images for the following main purposes:

- (i) This algorithm should be able to detect the edges of the tumour or microcalcification and clearly outline the edges of the tumour and microcalcification so that the edges are distinguishable and easily seen from the whole image.
- (ii) This algorithm should be able to segment the tumour or microcalcification from the rest of the image and clearly distinguish the tumour or microcalcification from the whole image effectively. This is done by changing all the pixels surrounded by the edges of the tumour or microcalcification with one grey level value (i.e. in this case, the value 0 is used to form a black region).

For the same consideration in Section 2.1, the algorithm of the MSBRG can be implemented as:

- (i) Implement three pre-processing algorithms to the image, namely median filter, histogram normalization and histogram equalization.
- (ii) Choose $N \times N$ neighbourhood as shown in Figure 1 (for $N = 5$ where N must be an odd number equal to or greater than 3).

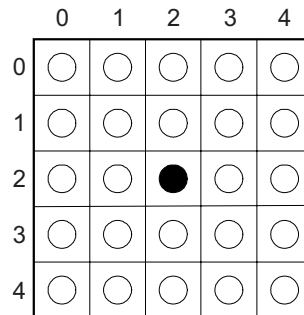


Figure 1 Location of the seed pixel and its 5×5 neighbourhood

- (iii) Take the value of tumour or microcalcification cluster as a threshold value, β , which is determined by using the moving k -means clustering algorithm in the earlier process.
- (iv) Examine all the pixels in the image. Set the first pixel with grey level value higher than β as the initial seed location, $p_0(x, y)$.

Note: The initial seed location must be located at the centre of all its $N \times N$ neighbours, as shown in Figure 1.

- (v) Calculate the mean grey level, \bar{x} (which is known as region mean) and the standard deviation, σ , for $N \times N$ neighbourhood according to Equations (4) and (5) respectively.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad (4)$$

$$\sigma = +\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (5)$$

where x_i is the grey level value for i -th pixel and n is the total of pixels in $N \times N$ neighbourhood.

- (vi) Compare each neighbour pixel with the initial seed pixel. Add a pixel to a region if it qualified for the region through either one of the two conditions listed below [8]:
- If the gradient of the neighbour pixel is less than 95% of the equalized histogram and its grey level value is more than or equal to the preselected threshold, b .
 - If the gradient of the neighbour pixel is more than or equal to 95% of the equalized histogram and the grey level of the pixel is not more than or equal to one standard deviation away from the region mean.
- (vii) Set the neighbour pixel, which is added to the region, as the new seed location.
- (viii) Repeat Steps (v) to (vii) until the region cannot be grown or all the pixels have been considered.
- (ix) Change the grey level of the pixel that cannot be grown with the value of 0. (Note: After this step, the edge of breast tumour or microcalcification will be detected as black edge.)
- (x) Set next pixel with grey level more than β as the new initial seed location, $p_0(x, y)$ if the pixel is not been grown yet. Repeat Steps (v) to (ix) until all the pixels of the image have been considered.

For region growing procedure, there are three possible ways for a seed point to grow as shown in Figure 2.

3.0 RESULTS

The proposed method was applied on three case studies. In the first case study, the proposed method was used to detect tumour region on mammogram images. Three mammogram images, namely Mammo001, Mammo002, and Mammo003 were

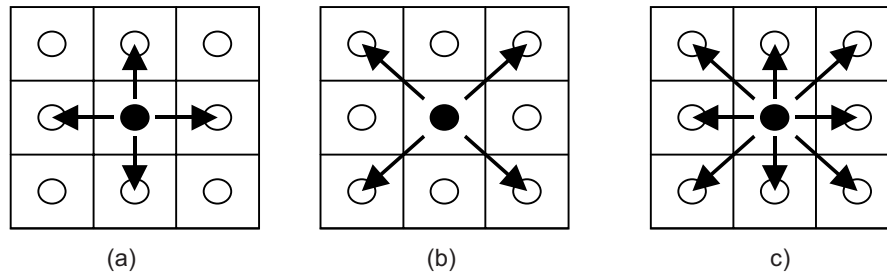


Figure 2 Three possible ways for seed pixel to grow: (a) 4 adjacent neighbours, (b) 4 diagonal neighbours, (c) 8 surrounding neighbours

examined for this case study. In the second case study, the proposed method was applied on three ultrasound images, namely UltraS001, UltraS002 and UltraS003, in order to detect breast tumour. For the third case study, two mammogram images, namely Mammo004 and Mammo005, were used in order to detect microcalcifications. The results for the first, second and third case studies are as shown in Sections 3.1, 3.2 and 3.3 respectively.

3.1 Results for Detection of Tumour Region on Mammogram Images

Figures 3, 4 and 5 show the results of edge detection of tumour region for image Mammo001, Mammo002 and Mammo003 respectively. For each figure, image (a) shows the original mammogram image. Image (b) shows the resultant image after applying the proposed method.

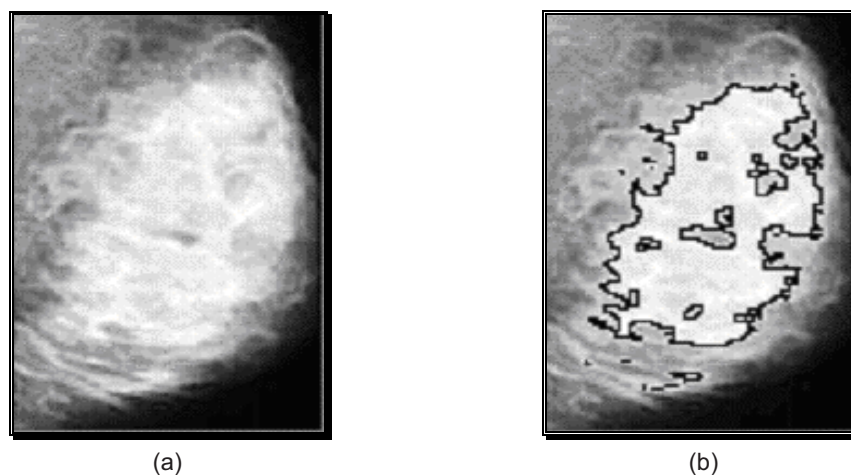


Figure 3 (a) Original image of Mammo001 (b) The result for tumour edge detection

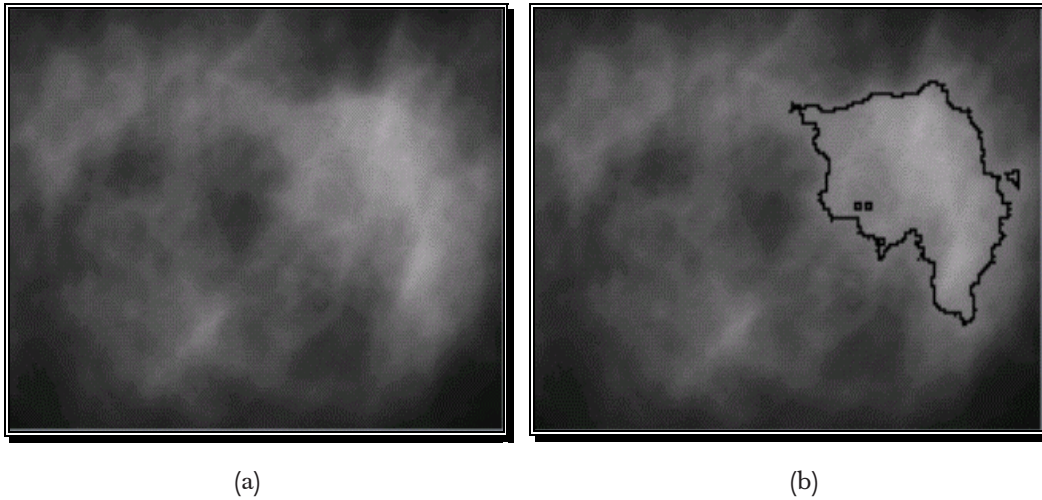


Figure 4 (a) Original image of Mammo002 (b) The result for tumour edge detection

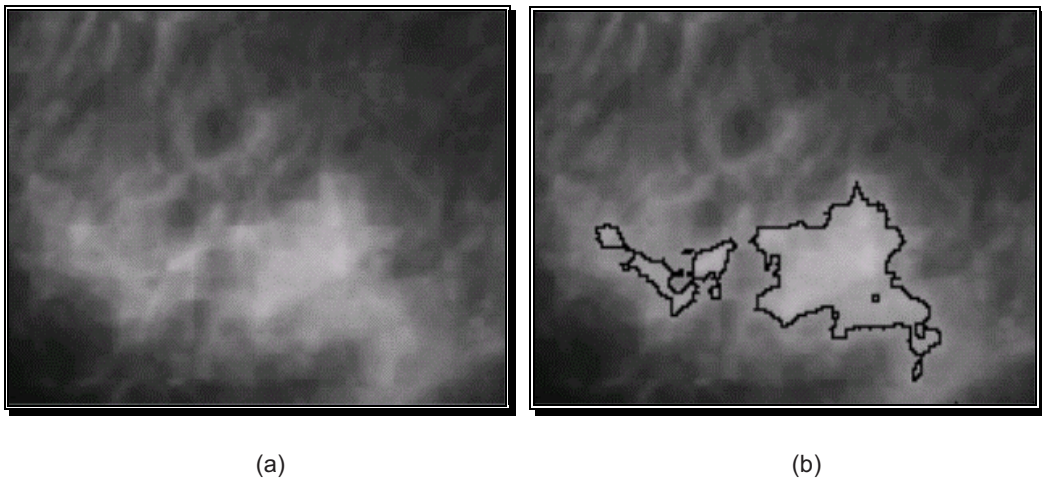


Figure 5 (a) Original image of Mammo003 (b) The result for tumour edge detection

3.2 Results for Detection of Tumour Region on Ultrasound Images

The results for the second case study are as shown in Figures 6, 7 and 8. Figures 6, 7 and 8 show the results of edge detection of tumour region on image UltraS001, UltraS002 and UltraS003 respectively. For each figure, image (a) shows the original ultrasound image, while image (b) shows the resultant image after applying the proposed method.

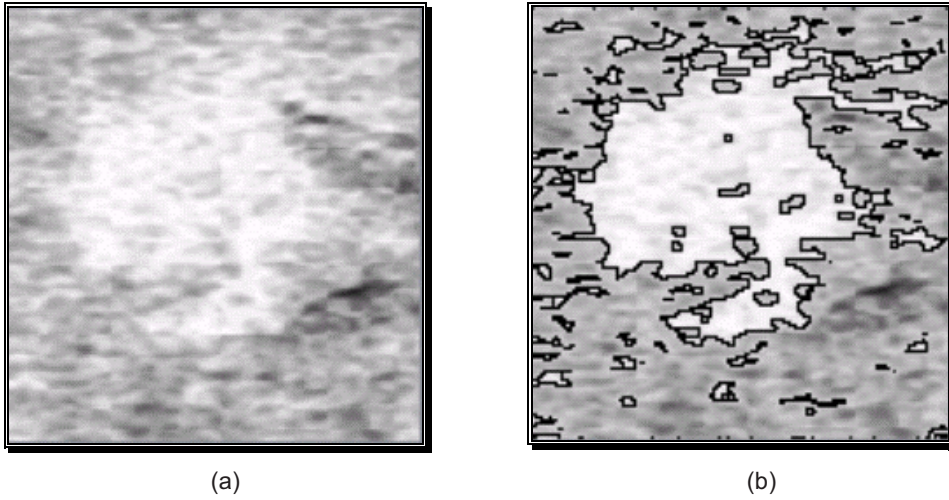


Figure 6 (a) Original image of UltraS001 (b) The result for tumour edge detection

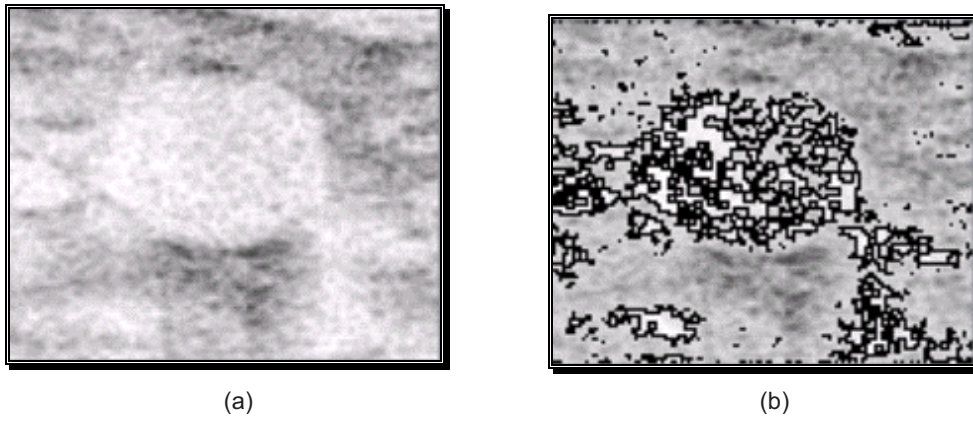


Figure 7 (a) Original image of UltraS002 (b) The result for tumour edge detection

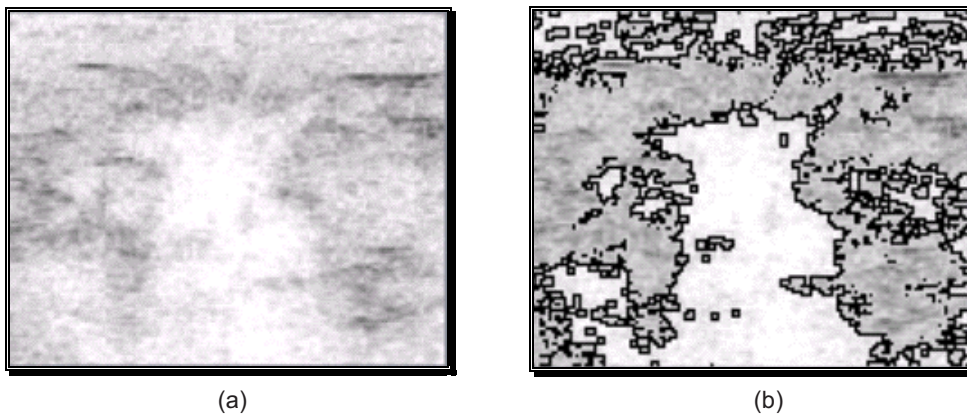


Figure 8 (a) Original image of UltraS003 (b) The result for tumour edge detection

3.3 Results for Detection of Microcalcifications on Mammogram Images

The results for the third case study are as shown in Figures 9 and 10. Figures 9 and 10 show the results of edge detection of microcalcification on image Mammo004 and Mammo005 respectively. For each figure, image (a) shows the original mammogram image, while image (b) shows the resultant image after applying the proposed method.

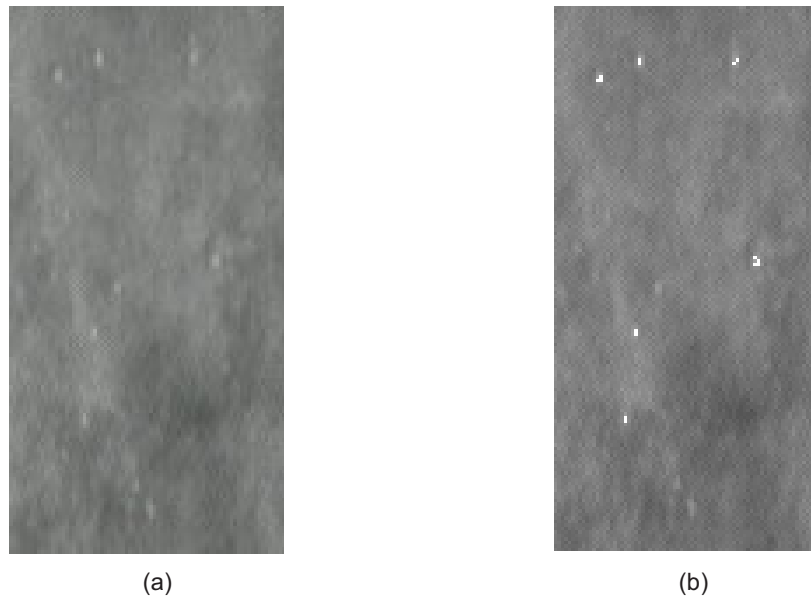


Figure 9 (a) Original image of Mammo004 (b) The result for microcalcification detection

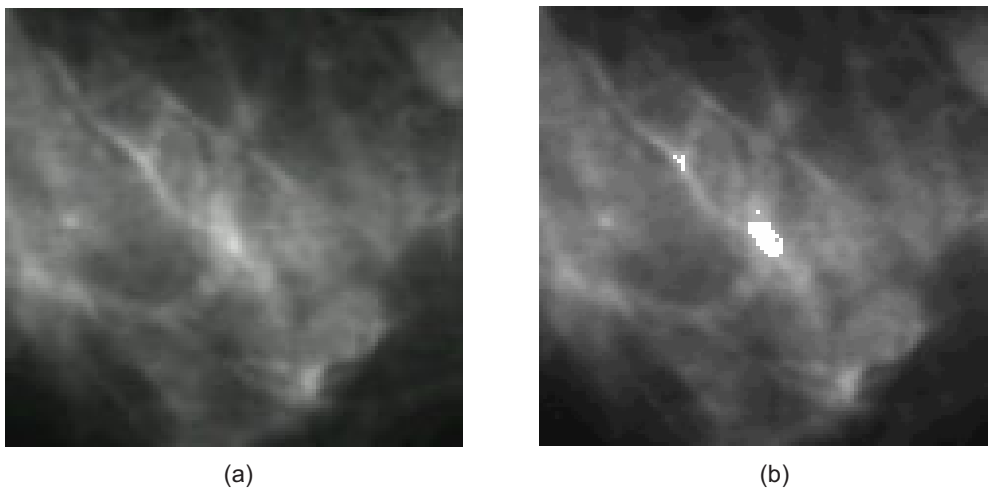


Figure 10 (a) Original image of Mammo005 (b) The result for microcalcification detection

4.0 DISCUSSIONS

Section 3 shows the results of the resultant medical images after applying the proposed MSBRG algorithm. For detection of breast tumour, the MSBRG algorithm is tested with mammogram and ultrasound images as presented in Sections 3.1 and 3.2 respectively. For each type of medical image, three case studies are used.

From Section 3.1, the original mammogram images are considered as dark images as shown in Figures 3(a), 4(a) and 5(a) for image Mammo001, Mammo002 and Mammo003 respectively. Manual screening for the size and shape of breast tumour could be difficult because of blurriness and low contrast of breast tumour regions. After applying the proposed MSBRG algorithm, the resultant images are more meaningful for interpretation. The MSBRG algorithm successfully detected the breast tumour and distinguished them from the background. The edges of breast tumours are fully connected, thus sustaining their shapes.

Promising resultant images are also obtained when the proposed MSBRG algorithm is implemented on ultrasound images as shown in Figures 6, 7 and 8 for image UltraS001, UltraS002 and UltraS003 respectively. From Figures 6(a), 7(a) and 8(a), the original ultrasound images are highly affected by unwanted noises. However, without smooth texture, the MSBRG algorithm is capable to detect the tumour regions and successfully sustain its shape as shown in Figure 6(b) for image UltraS001. However, the results obtained in Figures 7(b) and 8(b) show that the resultant images are still affected by unwanted noises. However, the MSBRG algorithm is still capable of outlining the edges of the tumour regions which could provide basic knowledge on the location, shape and size of the tumours.

The most challenging application for the MSBRG algorithm is determination of microcalcification in mammogram images. As most microcalcifications are presented as small white spots which may be equivalent to unwanted noises, the MSBRG algorithm should intelligently differentiate between them. The results obtained in Figures 9 and 10 show that the proposed MSBRG algorithm successfully segmented the microcalcifications and separated them from background as well as unwanted noises.

From Section 3, the results proved that the MSBRG algorithm is a reliable automatic technique to detect breast tumour from mammogram and ultrasound images. Besides that, the MSBRG algorithm is also able to detect microcalcification from mammogram image by successfully distinguishing them from unwanted noises and background. It proves that the MSBRG algorithm could successfully detect any size of region of interest (i.e. as big as tumour region and as small as microcalcification region).

In order to determine the capability of the proposed method to detect breast tumour and microcalcification, this study further employs comparison study with the conventional SBRG algorithm. We tested the proposed MSBRG and the conventional SBRG algorithm with 62 cases where 36 are normal (negative) cases

and the remaining 26 cases are positive. Similar performance was obtained for both algorithms where the percentage of accuracy is 93.55%. However, the results favour MSBRG algorithm because the MSBRG algorithm could determine the seed location and the threshold value automatically. It could save time in determining both parameters manually as for the conventional SBRG. Thus, the proposed MSBRG algorithm is not a user dependable technique.

5.0 CONCLUSIONS

A modified seed based region growing (MSBRG) algorithm, which is a modified version of conventional seed based region growing algorithm has been proposed in this paper. As compared to the conventional SBRG algorithm, the proposed method automatically determines the seed point and threshold value. The MSBRG algorithm is also proposed as an automatic edge detection tool for breast tumour and microcalcification. The results obtained are favourable as the proposed MSBRG algorithm provides more meaningful images. The regions of breast tumour and microcalcification are successfully detected. The size and shape of these regions have also been preserved.

Further to be pursued would be to improve the performance of the MSBRG algorithm. In addition, more case studies should be undertaken to test the MSBRG algorithm in order to establish its capability and reliability.

REFERENCES

- [1] National Cancer Council. 2005. Breast Cancer – Beat The Odds.... <http://www.makna.org.my/breastcancer.aspx>.
- [2] Cotran, R. S., V. Kumar, and S. L. Robbins. 1994. *Pathologic Basis of Disease*. Philadelphia, United State of America: W. B. Saunders Company.
- [3] Oncolink. 2005. Breast Cancer Information and Resources. <http://www.oncolink.org/types/article>.
- [4] Chan, H. P., K. Doi, C. J. Vyborny, K. L. Lam, and R. A. Schmidt. 1988. Computer Aided Detection of Microcalcifications in Mammogram – Methodology and Preliminary Clinical Study. *Investigative Radiology*. 23(9): 664-671.
- [5] Fan, J. P., D. K. Y. Yau, A. K. Elmagarmid, and W. G. Aref. 2001. Automatic Image Segmentation by Integrating Color-Edge Extraction and Seeded Region Growing. *IEEE Transaction on Image Processing*. 10(10): 1454-1466.
- [6] Rughooputh, S. D. D. V., and H. C. S. Rughooputh. 1999. Image Segmentation of Living Cells. Proceedings of International Conference on Robotics, Vision and Parallel Processing for Automation. Ipoh, Perak, Malaysia. 1: 8-13.
- [7] Ngah, U. K., T. H. Ooi, S. N. Sulaiman, and P. A. Venkatachalam. 2002. Embedded Enhancement Image Processing Techniques on a Demarcated Seed Based Grown Region. Proceedings of Kuala Lumpur International Conference on Biomedical Engineering. Kuala Lumpur, Malaysia. 170-172.
- [8] Romberg, J., W. Akram, and J. Gamiz. 1997. Image Segmentation Using Region Growing. <http://www.owlnet.rice.edu/elec539/projects97/WDEKknow/index>
- [9] Ghafar, R., M. Y. Mashor, and N. H. Othman. 2002. Segmentation of Pap Smear Slides Images Using K-Means Clustering. Proceedings of Kuala Lumpur International Conference on Biomedical Engineering. Kuala Lumpur, Malaysia. 41-43.

- [10] Razaee, M. R., P. M. J. Zwet, B. P. F. Lelieveldt, R. J. Geest, and J. H. C. Reiber. 2000. A Multiresolution Image Segmentation Technique Based on Pyramidal Segmentation and Fuzzy Clustering. *IEEE Trans. on Image Processing*. 9(7): 1238-1248.
- [11] Mashor, M. Y. 2000. Hybrid Training Algorithm for RBF Network. *International Journal of the Computer. The Internet and Management*. 8(2): 50-65.