

Masses Detection in Digital Mammogram by Gray Level Reduction using Texture coding Method

Al Mutaz .M. Abdalla
Faculty of Computer Science
and Information System
University Technology
Malaysia
Johor, Malaysia

Safaai Dress
Faculty of Computer Science
and Information System
University Technology
Malaysia
Johor, Malaysia

Nazar Zaki
Collage of Information
Technology
UAE University
Al Ain, UAE

ABSTRACT

Breast cancer is the most common cancer in women around the world. Various countries including the UAE offer asymptomatic screening for the disease. The interpretation of mammograms is a very challenges task and is subject to human error. Computer-aided detection and diagnosis have been proposed as a second reader for helping radiologists perform this difficult task. Texture features have been widely used as classification of masses in digital mammogram. In this paper we proposed a method for automatic detection of masses in digital mammogram. The proposed method uses the coding technique achieved good accuracy with Linear Discriminant Analysis (LDA) classification. The classification accuracy by using the coded images is improved much compared to one that obtained from the original image.

Keywords

Mammogram, LDA, GLCM, ANN

1. INTRODUCTION

Breast cancer is the second causes of death for women around the world. Any average woman has one chance in eight (or about 12%) of developing breast cancer during her life. Early detection of breast cancer by means of screening mammography has been established as an effective way to reduce the mortality rate resulting from breast cancer (Smith, 1995; Tabar, 1995) [1,2]. Despite significant recent progress, the recognition of suspicious abnormalities in digital mammograms still remains a difficult task. There are at least several reasons for that. First, mammography provides relatively low contrast images, especially in the case of dense or heavy breasts. Second, symptoms of abnormal tissue may remain quite subtle. For example, speculated masses that may indicate a malignant tissue within the breast are often difficult to detect, especially at the early stage of development [3].

The recent use of textural features and machine learning (ML) classifiers has established a new research direction to detect breast cancer. Texture features have been widely used as classification of masses in digital mammogram. Texture features can be classified into three classes based on what they are derived from gray level co-occurrence matrices (GLCM), gray level difference statistics (GLDS), and run length statistics (RLS) [4]. Many studies have been focused on statistical texture features [5] used the three types of statistical texture feature for masses detection and classification, [6] used the RLS to extract the texture information from the region of interest. [7] Used

texture features in detection of masses in mammogram. [8] Used texture analysis for the classification of mammographic masses. [9] Used statistical texture features for early detection of masses in digitized mammogram. [10] Used second order statistics and ANN for detection of masses in digital mammogram. [11, 12] have been focused on general issue of textural analysis on mammographic images, in the context of detection of the boundary of tumors and micro-calcifications. Texture features is considered to be one of the widely used tool in masses detection in digital mammogram. The Haralick features and co-occurrence matrix are the most methods of textural analysis for feature extraction which are used to classify the regions of interest into benign and malignant masses. Most of the texture analysis methods are applied directly on the original or filtering images. The computational of the co-occurrence matrix to the whole image is required large computing time. However decreasing the size of the co-occurrence matrix by reducing the number of gray levels in the image without losing the information's is must. The main drawback of the rank coding method is the gray level of the image which does not exceed 16 levels while the window must be of the size 3X3 pixels. In order to solve this problem we propose to increase the gray levels to 64 bit and selecting the proper window size to cover all the local information in the border of the selected pixel. The gray levels of the digital image are replaced to codes to perform a coded mammographic image. The method has been tested using a sample of 60 mammograms from Digital Database for Screening Mammography (DDSM) [13] the data consisting of 30 Benign and 30 Malignant. This work is divided into three sections. Section one describes the texture features extracted from the region of interest. Section two describes the texture coding method. Section three describes the proposed method of texture coding applied to the digital mammogram in order to detect the suspicious region. In the last section results and conclusion are shown and the application of the proposed method is discussed.

2. FEATURE EXTRACTION AND SELECTION

Statistical texture features have been proven to be powerful in classifying masses and normal breast tissues [14]. The implementation of feature extraction procedure relies on the quality of the texture, which is the main descriptor for all the mammograms. In this work, we concentrate on gray level co-occurrence matrix (GLCM).

Co-occurrence matrices describe the level of gray level configuration and are used to explain the matrices of relative frequencies. Therefore the GLCM is an advanced method of using different combinations of pixel brightness values which are the gray levels in the image.

GLCM are constructed by observing pairs of image cells distance *d* from each other and incrementing the matrix position corresponding to the gray level of both cells. This allows us to derive four matrices for each given distance and four different directions table 1. The image digitized at 256 gray levels which is leading to a co-occurrence matrix of size 256x256 [15]

Direction	Vector Representation for Direction
0°	(D,0)
45°	(D,D)
90°	(0,D)
135°	(-D,D)

Table 1: Description of texture feature derived from GLCM

Graphics	Top	In-between
Contrast	$\sum_{i,j=0}^{N-1} P_{ij} (i - j)^2$	Measures of intensity contrast between a pixel and its neighbor
Energy	$\sum_{i=1}^N \sum_{j=1}^N p(i, j)^2$	Is the sum of square elements in GLCM
Homogeneity	$\sum_{i,j=0}^{N-1} \frac{P_{ij}}{1 + (i - j)^2}$	Represent the closeness of the distribution of elements in GLCM to the GLCM diagonal
Correlation	$\frac{\sum \sum (ij)p(i, j)}{\sigma_x \sigma_y} - \mu_x \mu_y$	To see how correlated a pixel to its neighbor over the whole image

3. DECREASING THE GRAY LEVELS

The gray levels of an image are decreased from 256 gray levels to 16 levels (F, Eddaoudi, et al, 2011). The 256 gray levels are collected in 16 equal intervals by using the formula 1.

Table 2: Frequency distributions of the Ng gray levels

Gray levels	174	176	172	168	182	178	164	166	170	184	194
Frequency	8	3	3	2	2	2	1	1	1	1	1

(170, 184, and 194) are three gray levels to be eliminated

$$Y = \frac{16y}{256} \text{-----}1$$

Where Y represents the output of the image after decreasing the gray levels, y represents the original image. The rank coding method is already applied in satellite image processing [16]. In this paper we will apply it to the mammographic digital image for masses detection.

4. PROPOSED METHOD

The main objective of the proposed method is to locate the texture information presented in the mammograms by coding its textures and eliminating unnecessary information in order to reduce the data and computation time. The second objective is to keep the texture information in the coded image. In the image each gray levels pixel is assigned a code number. The original mammogram image is transferred to coded image and then the Haralick texture features from the co-occurrence matrix are extracted. Figure 1 showing the proposed method.

Five block sizes have been proposed (5x5, 7x7, 9x9, 11x11, and 13x13). The number of gray levels (Ng) is calculated for each block size as shown in the Figure 2 which contains 11 gray levels as shown in Table 2.

Let Ng be the number of gray levels in the initial image as shown in the Figure 2(a). From this figure we can find that the number of gray levels (Ng) is equal 11.

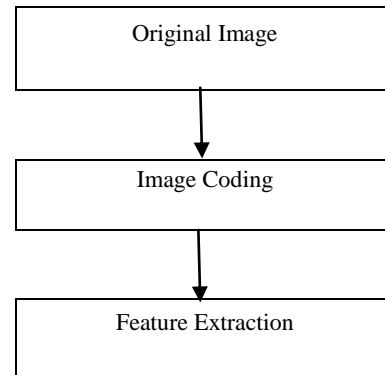


Fig 1: showing the proposed method using Mammogram coding

174 174 176 166 168
 174 194 178 172 178
 172 164 168 170 172
 174 176 174 174 184
 176 182 174 182 174

Figure 2(a): 11 number of gray levels:

174 174 176 166 168
 174 194 178 172 178
 172 164 168 170 172
 174 176 174 174 184
 176 182 174 182 174

F2. (b) Original image

174 174 176 166 168
 174 172 178 172 178
 172 164 168 176 172
 174 176 174 174 174
 176 182 174 182 174

F2. (c) Image after gray levels reduction

Table 3: Frequency distributions of the image after gray levels reduction

Gray levels	174	176	172	168	178	182	164	166
Frequency	9	4	4	2	2	2	1	1

0 0 1 7 3
 0 2 4 2 4
 2 6 3 1 2
 0 1 0 0 0
 1 5 0 5 0

F2. (d) Coded image

Table 4: Substitution of the gray levels by codes

Gray levels	174	176	172	168	178	182	164	166
Codes	0	1	2	3	4	5	6	7

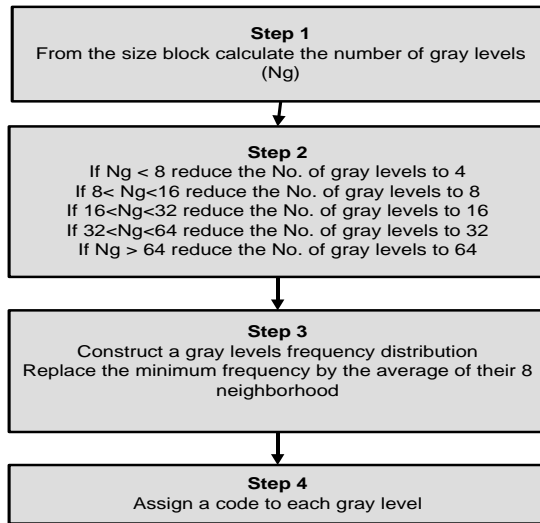


Figure 3: Steps for calculating the texture coding using gray levels reduction

5. EXPERIMENTAL RESULT

In our experiment we used a database of 60 mammograms, from the DDSM database (Digital Database for Screening Mammography). The database contains 30 benign (B) and 30 malignant (M). We evaluated the proposed coding method to those using the rank coding figure 4 showing the mammograms coded by the rank coding method. We used the linear Discriminant analysis (LDA) for the classification of the mammograms. Table 4 & Figure 5 summarize the percentages of classification rate. The classification accuracy shows that our proposed method achieved good result in classification of the masses and non- masses in digital mammogram. The coding method has strength of reducing the number of gray levels.

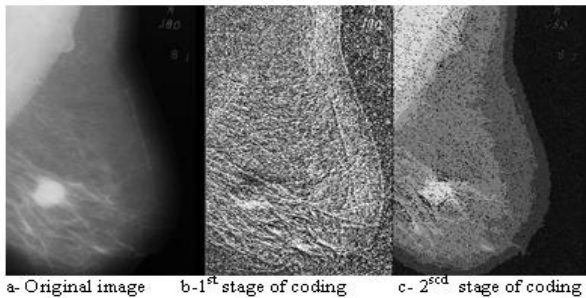
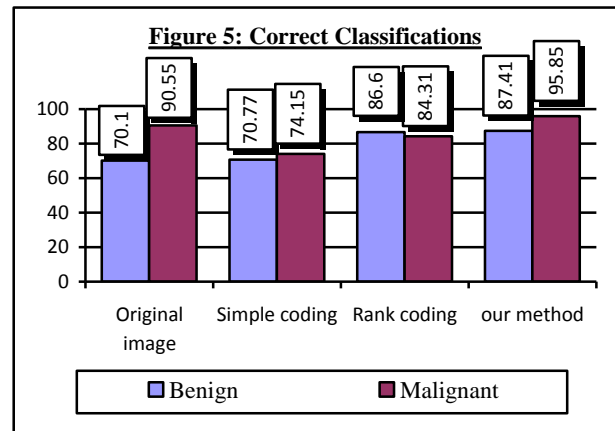
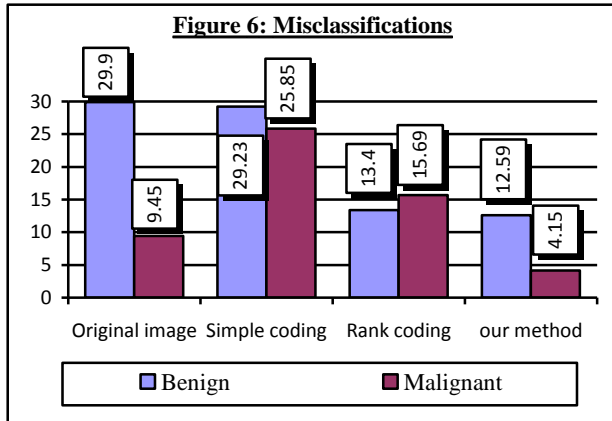


Figure 4: Mammogram coded by the rank coding method

Table 5: Frequency distributions of the Ng gray levels

Classification	Correct Classification		Misclassification	
	B	M	B	M
Original image	70.10%	90.55%	29.90%	9.45%
Simple coding	70.77%	74.15%	29.23%	25.85%
Rank coding	86.60%	84.31%	13.40%	15.69%
Our method of coding	87.41%	95.85%	12.59%	4.15%





6. CONCLUSION

In this paper we developed a new method for the detection and classification of masses in digital mammogram based on texture coding. In our proposed technique we replaced the gray levels of the pixels by assigning a code to each pixel in the image. The experimental results showed that our proposed method performed better than the original and ranked image. The classification accuracy using LDA applied to the coded images is 95.85%. The proposed method is reducing the number of values related to each pixel from 256 gray levels to 64 gray levels, this yields to reduce the computational time. Finally the proposed method of coding the image is promising method in reducing the information of the data without losing the image characteristics.

7. REFERENCES

- [1] Smith,R.A., "Screening women aged 40-49: where are we today?" *J Natl Cancer Inst*, 1995, pp. 1198-1199.
- [2] Tabar,L., Fagerber,G., and Chen,R.H., "Efficacy of breast screening by age: new results from the swedish two country trial", *Cancer*, 1995, pp. 1412-1419.
- [3] Arodz,T., Kurdziel,M., Sevre,E.O.D., and Yuen,D.A., "Pattern Recognition Techniques for Automatic Detection of Suspicious-looking Anomalies in Mammograms", *Computer Methods and Programs in Biomedicine, Elsevier*, 2005, pp. 135-149.
- [4] H.D. Cheng, X.J. Shi, R. Min, L.M. Hu, X.P. Cai, H.N. Du (2006) "Approaches for automated detection and classification of masses in mammograms", *Pattern Recognition*, Vol. 39, pp. 646-668.
- [5] Mavroforakis,M.E.,Georgiou,H.V.,Dimitropoulos,N., Cavouras,D., and Theodoridis,S., "Mammographic masses characterization based on localized texture and dataset fractal analysis using linear, neural and support vector machine classifiers", *Artificial Intelligence in Medicine*, 2006, pp. 145—162.
- [6] Szekely, N, Toth, N. Pataki, B. (2004). A hybrid system for detecting masses in mammographic images. Instrumentation and measurement technology conference, 2004. IMTC 04. Proceeding of the 21st IEEE Vol3, 18-20 May 2004 pp2065-2070.
- [7] K. Bovis and S. Singh. Detection of masses in mammograms using texture features. 15th International Conference on Pattern Recognition (ICPR'00), 2:2267, 2000.
- [8] Mudigonda, N.R, Rangayyan, R. Desautels, J.E.L (2000); Gradients and texture analysis for the classification of mammographic masses Medical Image, IEEE Transaction on Vol 19, Issue 10, Oct. 2000 pp 1032-1043.
- [9] N. Youssry, F.E.Z. Abou-Chadi, and A.M. El-Sayad. Early detection of masses in digitized mammograms using texture features and neuro-fuzzy model. 4th Annual IEEE Conf on Information Technology Applications in Biomedicine, 2003.
- [10] Al Mutaz, M. A., Deris, S., Zaki, N. M. 2011. Detection of Masses in Digital Mammogram Using Second Order Statistics and Artificial Neural Network. *International Journal of Computer Science & Information Technology (IJCSIT)*. Vol.3 No.3
- [11] Lisboa,P.G., "A review of evidence of health benefit from artificial neural networks", *Neural Networks*, 2002, pp. 11-39.
- [12] Arivazhagan,S., and Ganesan,L., "Textural classification using wavelet transform", *Pat Rec Lett* , 2003, pp. 1513-21.
- [13] D. K. M. Heath, K.W. Bowyer, Current status of the digital database for screening mammography, in: *Proceeding of the Fourth International Workshop on Digital Mammography*, Kluwer Academic Publishers, 1998,pp.457-460.
- [14] Haralick,R.M., Shanmugam,K., Dinstein,I., "Textural features for image classification", *IEEE Trans Sys Man Cyb*, 1973, pp. 610—21.
- [15] Fatima Eddaoudi, Fakhita Regragui, Microcalcifications Detection in Mammographic Image Using Texture Coding. *Applied Mathematical Sciences*, Vol.5, 2011, no. 8, 381-393
- [16] M.Hanifi, F.Sedes, D.Aboutajdine, A.Lasfar. A new approach for coding Satellite. Images : Rank coding. *International Journal of Computational Science*. , 2009, Vol. 3, No. 1
- [17] G.Lohmann. Analysis and Synthesis of Textures: a Co-occurrence-Based Approach. *Computer and Graphics*, 1995, Vol. 19, No. 1, 29:36 J.Parkkinen, T.J. Selkainaho, Detecting texture periodicity from the Co-occurrence matrix. *Pattern Recognition Letters*, 1990, Vol. 11, 43:50