Detection of Microcalcification Clusters in Mammograms Using a Difference of Optimized Gaussian Filters

Samuel Oporto-Díaz[1], Rolando Hernández-Cisneros[2], and Hugo Terashima-Marín[2]

[1] Universidad Nacional de Ingeniería, FIIS
Av. Túpac Amaru 210 Rímac, Lima, Peru
soporto@aurigacorp.com.pe

[2] Tecnológico de Monterrey, Campus Monterrey, Centro de Sistemas Inteligentes,
Av. Eugenio Garza Sada 2501, Monterrey, N.L. Mexico
{A00766380, terashima}@itesm.mx

Abstract. Since microcalcification clusters are primary indicators of malignant types of breast cancer, its detection is important to prevent and treat the disease. This paper proposes a method for detection of microcalcification clusters in mammograms using sequential Difference of Gaussian filters (DoG). In a first stage, fifteen DoG filters are applied sequentially to extract the potential regions, and later, these regions are classified using the following features: absolute contrast, standard deviation of the gray level of the microcalcification and a moment of contour sequence (asymmetry coefficient). Once the microcalcifications are detected, two approaches for clustering are compared. In the first one, several microcalcification clusters are detected in each mammogram. In the other, all microcalcifications are considered in a single cluster. We demonstrate that the diagnosis based on the detection of several microcalcification clusters in a mammogram is more efficient than considering a single cluster including all the microcalcifications in the image.

1 Introduction

Breast cancer is one of the main causes of death in women and early diagnosis is an important means to reduce the mortality rate. Mammography is one of the most common techniques for breast cancer diagnosis, and microcalcifications are one type of objects that can be detected in a mammogram. Microcalcifications are calcium accumulations of 0.1 mm to 2 mm wide, and they are indicators of the presence of breast cancer. Microcalcification clusters are groups of three or more microcalcifications that may appear in areas smaller than 1 cm^2, and have a high probability of becoming a malignant lesion.

Nevertheless, the predictive value of mammograms is relatively low, compared to biopsy. The causes of this low sensitivity [5] are the low contrast between the cancerous tissue and the normal parenchymal tissue, the small size of microcalcifications and possible deficiencies in the image digitalization process. The sensitivity may be improved having each mammogram be checked by two or more

M. Kamel and A. Campilho (Eds.): ICIAR 2005, LNCS 3656, pp. 998–1005, 2005.

radiologists, with the consequence of making the process inefficient by reducing the individual productivity of each specialist. A viable alternative is replacing one of the radiologists by a computer system, giving a second opinion [2], [13].

Several methods have been proposed for detection of microcalcifications in mammograms, like wavelets, fractal models, support vector machines, mathematical morphology, bayesian image analysis models, high order statistic, fuzzy logic, etc. The use of a Difference of Gaussian Filters (DoG) for detection of potential microcalcifications has been addressed by Dengler et al. [4] and Ochoa [11]. In this work, we developed a procedure that applies a sequence of Difference of Gaussian Filters (DoG), in order to maximize the amount of detected probable individual microcalcifications in the mammogram (signals), which are later classified by an artificial neural network (ANN) in order to detect real microcalcifications. Later, microcalcification clusters are identified. Additionally, the hypothesis to be tested states that the diagnosis accuracy of the mammograms is higher if it is based on the diagnosis of every microcalcification cluster detected in the mammogram, instead of considering of all the microcalcifications in the mammogram as a single cluster.

The rest of this document is organized as follows: in the second section, the proposed procedure is described in detail. In the third section, the experimental results are shown and discussed, and finally, in the fourth section, the conclusions are presented, and some comments about future work are also mentioned.

2 Methods

The mammographic images used in this project were provided by The Mammographic Image Analysis Society [12]. The MIAS database contains 322 images, and only 25 of them contain microcalcifications. Among these 25 images, 13 cases are diagnosed as malignant and 12 as benign. The size of all images is 1024x1024 pixels, digitized at 8 bits. Several related works have used this same database [3], [6], [7], [10].

The proposed solution model is shown in Figure 1. The general procedure receives a digital mammogram (I) as an input, and it is conformed by five stages: pre-processing, detection of potential microcalcifications (signals), classification of signals into real microcalcifications, detection of microcalcification clusters and classification of microcalcification clusters into benigns and malignants.

2.1 Pre-processing

The main objective of this stage is to eliminate those elements in the image that could interfere in the process of identifying microcalcifications. A secondary goal is to reduce the work area only to the relevant region that exactly contains the breast. The procedure receives the original image as an input. First, a median filter is applied in order to eliminate the background noise; second, a binary image is created from the filtered image, where each pixel represents a 16x16 window centered in the corresponding pixel from the original image. If the gray

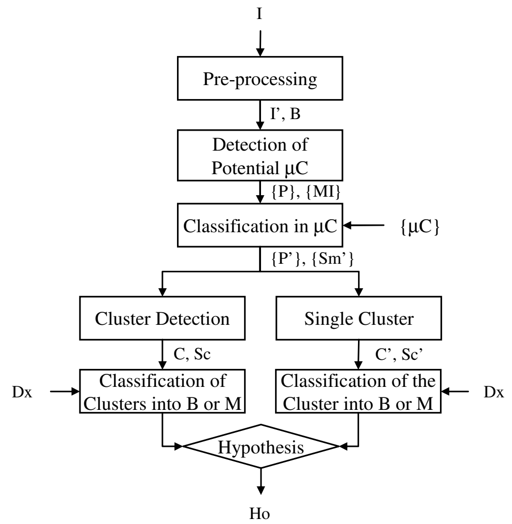

Fig. 1. Proposed Model for Hypothesis Testing

average level of the window is below certain threshold (established empirically, after visually analyzing the histograms of average gray level for several window sizes), a zero value is placed in the binary image; otherwise, a unitary value is placed. Third, an automatic cropping procedure is applied in order to delete the background marks and the isolated regions, so the image will contain only the region of interest. The result is a smaller image, with less noise.

2.2 Detection of Potential Microcalcifications (Signals)

This stage has the aim of detecting the mass centers of the potential microcalcifications in the image (signals). The pre-processed image of the previous stage is the input of this procedure. The optimized difference of two gaussian filters (DoG) is used for enhancing those regions containing bright points. A gaussian filter is obtained from a gaussian distribution, and when it is applied to an image, eliminates high frequency noise, acting like a smoothening filter. A DoG filter is built from two simple gaussian filters. These two smoothening filters must have different variances. When two obtained images after applying each filter separately are subtracted, an image containing only the desired frequency range is obtained. The DoG filter is obtained from the difference of two gaussian functions, as follows: it is shown in equation 1, where x and y are the coordinates of a pixel in the image, k is the height of the function and σ_1 and σ_2 are the standard deviations of the two gaussian filters that construct the DoG filter.

$$DoG(x,y) = k_1 e^{(x^2+y^2)/2\sigma_1^2} - k_2 e^{(x^2+y^2)/2\sigma_2^2} \tag{1}$$

The resultant image after applying a DoG filter is globally binarized, using a threshold. In Figure 2, an example of the application of a DoG filter is shown.

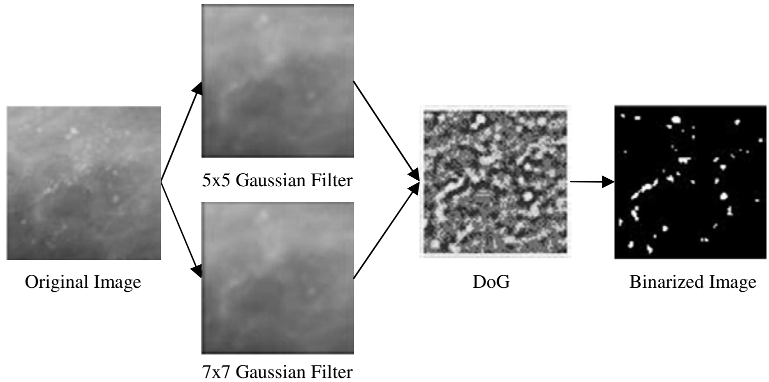

Fig. 2. Example of application of a DoG filter (5x5, 7x7)

A region-labeling algorithm allows the identification of each one of the points. Then, a segmentation algorithm extracts small 9x9 windows, containing the region of interest whose centroid corresponds to the point centroid. The size of the windows is adequate for containing the signals, given that at the current resolution, their area is 5x5 pixels in average.

Three selection methods are applied in order to transform a point into a signal. The first one performs selection according to the region area, the second one according to the gray level and the third one according to the gray gradient. The result is a list of signals represented by their centroids. In order to detect the greater possible amount of points, six gaussian filters of sizes 5x5, 7x7, 9x9, 11x11, 13x13 and 15x15 are combined, two at a time, to construct 15 DoG filters that are applied sequentially. Each one of the 15 DoG filters was applied 51 times, varying the binarization threshold in the interval $[0, 5]$ in increments of 0.1. The points obtained by applying each filter are added to the points obtained by the previous one, deleting the repeated points. The same procedure is repeated with the points obtained by the remaining DoG filters. All of these points are passed later to the three selection procedures.

2.3 Classification of Signals into Real Microcalcifications

The objective of this stage is to identify if an obtained signal corresponds to an individual microcalcification or not. With this in mind, a set of features are extracted from the signal, related to their contrast and shape. From each signal, 47 features are extracted: seven related to contrast, seven related to background contrast, three related to relative contrast, 20 related to shape, six related to the moments of the contour sequence and the first four Hu invariants.

There is not an a priori criteria to determine what features should be used for classification purposes, so the features pass through two feature selection processes [8]: the first one attempts to delete the features that present high correlation with other features, and the second one uses a derivation of the forward sequential search algorithm, which is a sub-optimal search algorithm.

The algorithm decides what feature must be added depending of the information gain that it provides, finally resulting in a subset of features that minimize the error of the classificator. After the two selection processes, only three features were selected and used for classification (absolute contrast, standard deviation of the gray level and the third moment of contour sequence).

A back-propagation neural network is used to classify each signal and to obtain those signals that correspond to real microcalcifications. The number of inputs for the neural network is three, equal to the number of selected features. According to Kolmogorov's theorem [9], and considering the number of inputs as $n = 6$, three layers were considered. The hidden layer contains $2n + 1 = 7$ neurons, and the output layer has only one neuron. The transfer function of each neuron is the sigmoid hyperbolic tangent function, and the error is measured with the mean square error function.

Even before the detection of microcalcification clusters, the global performance of the neural network at the classification of individual microcalcifications was 85%, confirmed by experts, and related to the application of the DoG filters in sequence. After the individual microcalcifications were detected, they had to be grouped in clusters. Two clustering procedures were proposed for comparison: a) Detection of microcalcification groups that can form clusters. In a single mammogram, one or more clusters can be identified; and b) Consideration of all the microcalcifications in a mammogram as part of a single cluster [1].

2.4 Detection of Microcalcification Clusters

During this stage, the microcalcification clusters are identified. The algorithm tries to locate those microcalcification clusters occupying regions where the quantity of microcalcifications per cm^2 is higher. The microcalcifications forming a cluster are later labeled. From each cluster, a cluster feature set is extracted.

There is an additional clustering procedure that considers all the microcalcifications identified in the mammogram as members of a unique cluster. From each cluster, 30 features are extracted: six related to the shape of the cluster, six related to the area of microcalcifications and 10 related to the contrast of the microcalcifications. The same two feature selection procedures mentioned earlier are also included in this stage. Only three cluster features were selected for the classification process (minimum diameter, minimum radius and average radius).

2.5 Classification of Microcalcification Clusters into Benigns and Malignants

This stage has the objective of classifying each cluster in one of two classes: benign or malignant. This information is provided by the MIAS database. The classificator used in this stage is also a backpropagation neural network with three layers, again considering Kolgomorov's Theorem [9]. The performance measure for this classificator is the success rate.

Finally, the performances provided by both classification processes (detection of microcalcification clusters and single clustering) are compared.

3 Results

In order to demonstrate the proposed hypothesis in this work, an experiment was prepared for evaluating two treatments applied to two datasets, and identify if these treatments have some influence in the results or the variations are random.

Two data groups were prepared. From each data group, the following features were extracted: minimum diameter, minimum radius, and average radius. The first data group (G_{MC}) corresponds to 40 identified microcalcification clusters, using the density technique and a radius of 100 pixels. Each microcalcification cluster has a diagnosis provided by the MIAS database. The second data group (G_{SC}) corresponds to the 22 unique, single clusters obtained by considering all the microcalcifications in a mammogram as members of a single cluster. Each single cluster (mammogram) has a diagnosis provided by the MIAS database.

Table 1 presents the obtained performances (proportion of correctly diagnosed mammograms) after 25 runs of each neural network (treatment). We can observe that both means (0.91 for G_{MC} and 0.89 for G_{SC}) indicate very good and similar performances by both treatments, and a statistical test should be applied in order to know if there is a significant difference between them. If the mean of G_{MC} is significantly greater than the mean of G_{SC}, it would mean that considering one or more microcalcification clusters in a mammogram leads to a more accurate diagnosis than considering all the microcalcifications as members of a single cluster, thus confirmating the hypothesis of this work. The following hypotheses were formulated:

1. H_1: There is a significant difference between the means of G_{MC} and G_{SC}.
2. H_0: There is no significant difference between the means of G_{MC} and G_{SC}.

The F test was used to validate or discard the hypothesis H_0, and the probability is 95%. $F_{calculated}$ is 4.82, and $F_{table(0.48)}$ is 3.01. $F_{calculated}$ is greater than $F_{table(0.48)}$, so H_0 is rejected and we can conclude that there is a significant difference between the means of G_{MC} and G_{SC}, and the hypothesis of this research is confirmed.

Table 1. Obtained performances (proportion of correctly diagnosed mammograms) after 25 runs of each neural network (treatment)

N	G_{MC}	G_{SC}
1	0.98	1.00
2	0.93	0.91
3	0.98	0.95
4	0.90	0.86
5	0.85	0.82
6	0.95	0.91
7	0.93	0.91
8	0.95	0.91
9	0.83	0.82

N	G_{MC}	G_{SC}
10	0.85	0.82
11	0.95	0.91
12	0.95	0.91
13	0.90	0.86
14	0.88	0.82
15	0.93	0.91
16	0.93	0.91
17	0.90	0.86
18	0.95	0.91

N	G_{MC}	G_{SC}
19	0.90	0.86
20	0.88	0.86
21	0.95	0.91
22	0.90	0.86
23	0.90	0.86
24	0.95	0.95
25	0.88	0.86
mean	0.91	0.89
STD	0.04	0.05

4 Conclusions and Future Work

It is not possible to use a single DoG filter and a unique binarization threshold that maximizes the number of potential microcalcifications detected in a mammogram, because it would identify only some frequency ranges. The use of multiple DoG filters with different relations σ_1/σ_2 and different binarization thresholds solves this problem, because more ranges of frequencies are analyzed. The global performance achieved at the classification of individual microcalcifications was 85%, confirmed by experts, and related to the application of the DoG filters in sequence.

The three features extracted from individual microcalcifications that maximize the rate of true positives and the success rate simultaneously are: absolute contrast, gray standard deviation of a microcalcification, and a moment of contour sequence (asymmetry coeficient). Contrast properties provide more information than shape properties for the classification of signals into microcalcifications. The performance achieved for classifying signals into microcalcifications are 70.8% for true-positives and 85.7% for all the examples. Despite the performance in the classification of individual microcalcifications is not commonly reported in literature, we consider that the performance obtained by the proposed method (sequence of DoG filters) is reasonably good. On the other hand, shape properties provide more information than contrast properties when microcalcification clusters are classified. In this case, the features that provide better results for maximizing the success rate of the classificator are minimum diameter, minimum radius, and average radius, all of them shape properties. The performance achieved at diagnosing a microcalcification cluster is 91%. After analyzing the result of the experiments, the main conclusion of this work is that diagnosing a mammogram based on one or more microcalcification clusters in the image provides better results than always considering all the microcalcifications in the image as a single and unique cluster.

Several subjects were not solved nor implemented for this research, and they are proposed as future work. It could be useful to use other mammography databases, and test how different resolutions could affect system effectiveness. The size of the gaussian filters could be adapted depending on the size of the microcalcifications to be detected and the resolution of images. The correspondence between the spatial frequency of the image and the relation σ_1/σ_2 has to be thoroughly studied. Different features could be extracted from the microcalcifications in the images and tested also. Enhancements on the architecture or training methods for the neural network, or even other approaches for classification could be proposed. Finally, it would be recommendable to study the obtained results using ROC curves, for comparison with other works.

Acknowledgments

This project was developed at the Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM), thanks to the support of the National Council of Science and Technology of Mexico (CONACYT).

References

1. Aghdasi, F., Ward, R. K., Palcic, B.: Classification of mammographic microcalcifications clusters. In Proc. of the CCECE, Vancouver, BC, Canada (2003) 1196–1199
2. Anttinen, I., Pamilo, M., Soiva, M., Roiha, M.: Double reading of mammography screening films: one radiologist or two? Clin. Radiol. **48** (1993) 414–421
3. Chandrasekhar, R., Attikiouzel, Y.: Digitization regime as a cause for variation in algorithm performance across two mammogram databases. Technical Report 99/05, Centre for Intelligent Information Processing Systems, Department of Electrical and Electronic Engineering, The University of Western Australia (1999)
4. Dengler, J., Behrens, S., Desaga, J. F.: Segmentation of microcalcifications in mammograms. IEEE Trans. Med. Imaging **12**(4) (1993) 634–642
5. Ganott, M. A., Harris, K. M., Klaman, H. M., Keeling, T. L.: Analysis of false-negative cancer cases identified with a mammography audit. The Breast Journal **5**(3) (1999) 166–175
6. Gulsrud, T. O.: Analysis of mammographic microcalcifications using a computationally efficient filter bank. Technical Report, Department of Electrical and Computer Engineering, Stavanger University College (2001)
7. Hong, B.-W., Brady, M.: Segmentation of mammograms in topographic approach. In IEE International Conference on Visual Information Engineering, Guildford, UK (2003)
8. Kozlov, A., Koller, D.: Nonuniform dynamic discretization in hybrid networks. In Proceedings of the 13th Annual Conference of Uncertainty in AI (UAI), Providence, Rhode Island (2003) 314–325
9. Kurkova, V.: Kolmogorov's theorem. In: Arbib, M. A. (ed.): The handbook of brain theory and neural networks. MIT Press, Cambridge, Massachusetts (1995) 501–502
10. Li, S., Hara, T., Hatanaka, Y., Fujita, H., Endo, T., Iwase, T.: Performance evaluation of a CAD system for detecting masses on mammograms by using the MIAS database. Medical Imaging and Information Science **18**(3) (2001) 144–153
11. Ochoa, E. M.: Clustered microcalcification detection using optimized difference of gaussians. Master Thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base (1996)
12. Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S., Taylor, P., Betal, D., Savage, J.: The Mammographic Images Analysis Society digital mammogram database. Exerpta Medica International Congress Series **1069** (1994) 375-378. http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html
13. Thurfjell, E. L., Lernevall, K. A., Taube, A. A. S.: Benefit of independent double reading in a population-based mammography screening program. Radiology **191** (1994) 241–244