ACTA STEREOL 1996; 15/2: 125-130
ORIGINAL SCIENTIFIC PAPER

A MULTISTAGE METHODOLOGY FOR DETECTION OF MICROCALCIFICATIONS IN DIGITAL MAMMOGRAPHY

Hervé Guillemet, Habib Benali, Edmond Kahn, Robert Di Paola

INSERM U66, CHU Pitié Salpetrière, 91 Bd de l'Hôpital
75013 Paris, France

ABSTRACT

This paper describes a coarse-to-fine approach for detection and segmentation of clustered microcalcifications. Starting from digitized images of whole mammographic films, the method involves three main stages. First, the breast is segmented. Then a texture analyzer based on the fractional Brownian motion model is applied to select those regions that may contain a cluster. The last stage computes the isophote map of the selected regions and considers the relative positions of isophotes to recognize the microcalcifications.

The method was evaluated using a database of 150 mammographic images of patients who underwent a surgical biopsy for a cluster of microcalcifications. In 95% of mammograms, the clusters are correctly localized if we accept an average 1,5 false positive cluster per mammography.

Key words: image analysis, fractal, isophote, microcalcifications, mammogram.

INTRODUCTION

In the last few years, breast cancer has become the leading cause of cancer-related death in women. Recent studies in developed countries show that one woman over ten will develop breast cancer during her lifetime (Basset and Gold, 1987). It is well known thatearly detection is the most effective means of decreasing mortality (Feig, 1988). With the increased concern in mass screening programs based on X-ray examination (mammography), more breast cancers are now detected in earlier stage. Unfortunately interpretation of mammographic films is difficult, requesting special training from radiologists, and lacks specificity. When dealing with minimal breast cancers, the most frequent abnormality is the presence of an accumulation of small calcium deposits in tissue, appearing as a cluster of small luminous spots, known as *microcalcifications* (MC). MC are also the hardest abnormalities to deal with for they can be hardly distinguishable and be benign as well as malignant. The aim of our work is to assist radiologists by supplying a system for detecting and interpreting clusters of MC. This system provides the following functionalities:

1. Detection and segmentation of MC: this step extracts the regions of the digitized image containing clusters of MC, and determines the contour of each MC.
2. Measure of shape parameters: quantitative shape features of the objects described at the previous step are supplied to the radiologist for a more relevant diagnosis.
3. Malignancy estimation: Each cluster is given an automatic malignancy grading.

Only the first step, detection and segmentation of MC, is presented in this paper.

MATERIAL

A database of 150 mammographic files were selected. All women underwent a surgical biopsy for isolated clustered microcalcifications. One film per file was digitized, using a 12bits scanner at a pixel size of 80 μm. All clusters (usually 1 in a mammogram) were localized by radiologists. After processing by the system, the results were read and compared to their expertise.

METHODS

Microcalcifications are small luminous objects with variable shapes and variable contrasts. Even the most expert radiologist could wonder whether a small feature he observes on a mammography is truly a MC. This suggests that the automatic method we need to accurately recognize MC must achieve a high degree of complexity. Unfortunately, the average size of MC (200 μm) requires a very fine sampling of the film, and thus leading to great amount of data. The approach we adopted reduces step by step the amount of data by using different methods from a coarse scale to a fine one.

Pre-processing steps

A mammographic image consists in three kinds of pixels, depending on the material crossed by X rays: white for lead, grey for tissue and black for air. These three classes are determined by an automatic thresholding method based on histogram segmentation. The breast is recognized as the biggest object made of grey pixels. This simple operation rejects 75% of the whole processed image.

Another essential preprocessing step is noise equalization. The screen-film system converts X-ray exposure to optical density in a non-linear way. Our digitizer also introduce distortion when converting optical density to gray-levels. Thus the information is not uniformly distributed along the grey scale, leading to sensitivity variation with the global intensity of a region. We rectify this non-uniform distribution by applying an adapted requantification (Karssmeijer and van Earning, 1991).

Fractal texture classification

Mammary parenchyma has been proven to exhibit fractal properties (Caldwell et al., 1989) (Fefebvre et al., 1992). But in presence of accidents that disturbs the regularity of texture, like a particularly dense vessel, a scratch on the film or a calcification, these properties vanish (Fefebvre, 1993). In order to reduce the amount of data, we split the breast area into overlapping windows and only retain those windows that do not conform to a fractal model.

The model used is the fractional Brownian motion (fBm). fBm is a non stationary random process which can be described by its increment:

$$B_H(\vec{p}) - B_H(\vec{q}) \sim N\left(0, \sigma^2 |\vec{p} - \vec{q}|^{2H}\right) \tag{1}$$

Its covariance function is expressed as:

$$E\left(B_H(\vec{p}) B_H(\vec{q})\right) = \frac{1}{2}\sigma^2\left(|\vec{p}|^{2H} + |\vec{q}|^{2H} - |\vec{p} - \vec{q}|^{2H}\right) \tag{2}$$

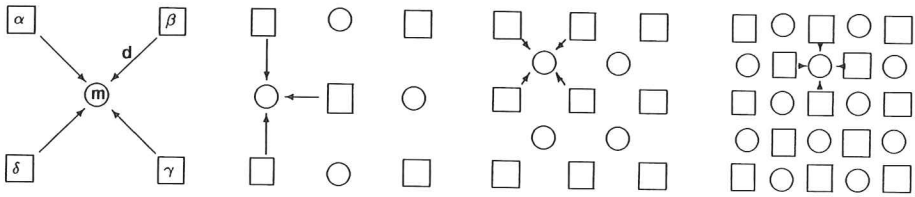

Fig. 1. Random midpoint displacement algorithm: formation of the 4 first generations of pixels. Circles and squares respectively denote new and old pixels.

The conformance to this model is quantified by considering an algorithm of fBm synthesis and calculating the likelihood that a sub-window is a realization of this algorithm. Several techniques for the synthesis of fBm has been developed. They use digital filtering of white noise sequences, summation of increments of discrete fractional Gaussian noise, or wavelets. The one we chose, called random midpoint displacement algorithm (RMDA), has been shown to have comparable accuracy with other techniques but be the most computationally efficient (Stoksik et al., 1995). Fig. 1 shows the formation process of an image by the RMDA. The four corners of the image, with known grey values α, β, γ and δ compose generation 0. A value m is given to the midpoint, which constitutes generation 1, by averaging the values of its neighbours and adding a random displacement, ε.

$$m = \frac{\alpha + \beta + \gamma + \delta}{4} + \varepsilon \tag{3}$$

If the resulting image is a fBm, α, β, γ, δ, and m must satisfy Eq. 1 and Eq. 2, which imply:

$$\begin{cases} E\left((m-\alpha)^2\right) = E\left((m-\beta)^2\right) = E\left((m-\gamma)^2\right) \\ \qquad\qquad = E\left((m-\delta)^2\right) = \sigma^2 d^{2H} \\ E\left((m-\alpha)(m-\beta)\right) = E\left((m-\beta)(m-\gamma)\right) = E\left((m-\gamma)(m-\delta)\right) \\ \qquad\qquad = E\left((m-\delta)(m-\alpha)\right) = \sigma^2 d^{2H}\left(1-2^{2H-1}\right) \\ E\left((m-\alpha)(m-\gamma)\right) = E\left((m-\beta)(m-\delta)\right) = \sigma^2 d^{2H}\left(1-2^{2H-1}\right) \end{cases} \tag{4}$$

Eq. 3 and Eq. 4 yield the expected variance s_d^2 of ε:

$$s_d^2 = E(\varepsilon^2) = \sigma^2 d^{2H}\left(1-2^{H-2}-2^{2H-3}\right) \tag{5}$$

Displacement ε is then produced by realizing a Gaussian random variable with mean 0 and variance s_d^2. Generation 2 is computed as shown on Fig. 1 and the process is continued until the desired resolution is obtained. Given generation 0, the probability that an image is a realization of the RMDA is:

$$Pr(I) = Pr(I_1|I_0)Pr(I_2|I_1)...Pr(I_N|I_{N-1}) \quad (6)$$

Where $Pr(I_{i+1}|I_i)$ denotes the probability we obtain generation i + 1 knowing generation i:

$$\Pr(I_{i+1}|I_i) \propto \frac{1}{\sigma d^H} \prod_{p \in I_{i+1}} e^{-\frac{(v_p - m_p)^2}{2s_d^2}} \quad (7)$$

where v_p is the value of the midpoint and m_p is the average of its neighbours.

The index used to quantify the conformity of a window to fBm is deduced:

$$\xi = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_{I_i}}\sum_{p \in I_i}\frac{(v_p - m_p)}{\sigma^2 d^{2H}} \quad (8)$$

ξ linearly depends on *log(P(I))*. Its distribution can be shown to follow a χ^2 law.

Thus, the method for fractal texture classification consists in estimating H and σ on the whole parenchyma, computing ξ for each sub-window, and thresholding. For an average choice of the threshold, 90% of the image is eliminated this way. As a drawback, 10 clusters in the 150 mammograms of our database were incorrectly removed.

Microcalcification segmentation

Microcalcification recognition and segmentation are achieved through the creation of an isophote map. These maps, constituted of lines of equal grey level, are a way of describing a surface which carries a higher level information than that of a simple grey level lattice. They can contain the whole original information if the difference between elevation of 2 successive isophotes is chosen equal to 1, or achieve efficient data compression if it is chosen higher than 1.

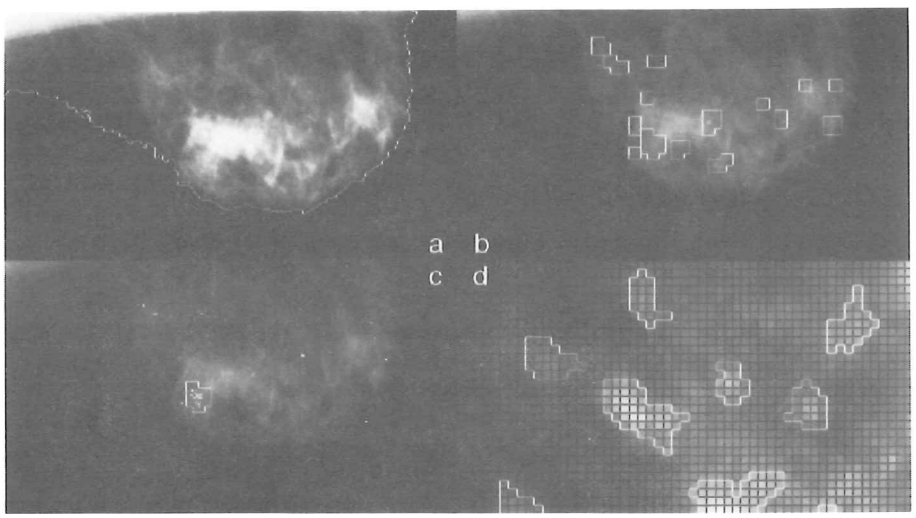

Fig.2. a) Original image and result of breast segmentation. b) Areas selected by fractal texture classification. c) Microcalcification detected in green and cluster retained in red. d) Isolines of several microcalcifications.

Detection and segmentation of MC consist in:

- Computing the isoline map for each sub-window selected in the previous stages
- Establishing inclusion relationships between isolines
- Selecting isolines that conform to a model of microcalcifications in term of isoline characteristics and relative positions.

An example of computed isoline is shown on Fig. 2. An isoline is considered *included* in another isoline if its bounding box is included in the other's. This is a rough approximation that we accept to save computation time. In this representation, MC appear like hills on a topographic map with imbricated circular isophotes. Our model considers the lentgth of isolines at the base of such hill and the way isoline lengths decrease when we approach the top of the hill.

Final processing

Sub-windows containing at least 3 aggregated MC are considered parts of a cluster. A classical method which detects high densities of points in a plane is applied to delineate the aggregates.

RESULTS

Results are summed up in Tbls. 1 and 2. Computation time of the whole process is 45 s CPU on a Sun/Sparc 10.

Table 1. Results of MC detection per cluster. The number of automatically detected regions is denoted by n. tp denotes the number of regions containing true clusters of MC and fp of those containing less than 3 MC. fn is the number of undetected clusters.

n	tp	fp	fn
344	181 (53%)	163 (47%)	30

Table 2. Results of MC detection per mammogram. N denotes the total number of mammograms. TP denotes the number of mammograms where at least 1 cluster was correctly localized, NFP those where no false-positive regions were found and FN those where nothing were found.

N	TP	NFP	FN
150	138 (92%)	74 (50%)	10 (6%)

DISCUSSION

Some issues must be developed in view of Tables 1 and 2:

- Among the 181 true positive regions, 11 area were ignored by the radiologists during his first expertise. This result could justify by itself the use of such Computer Aided Diagnonis when screening mammography. However, in most cases, these areas consisted in secondary clusters that the radiologist, blinded by the primary cluster which alone determined the diagnosis, did not see.

- 46 % of false positive regions are due to defects in the screen/film system which create small luminous spots similar to MC. 6 % are scratches on the film and 16% are areas containing macrocalfications of vascular calcifications, well known as benign lesions. 32% are other normal structures, like an overlap of two vessels. The first two sources of false-positive will vanish when digital mammographs will be used. Eliminating the other two sources would involve supplying knowledge about other calcifications, vessels and other structures to the system.
- The method and the database were realized in order to help radiologists for locating and contouring MC, and not for classifying patients for mass screening. Therefore results of Tbl. 1 are more relevant than results of Tbl. 2. The same methodology could however be applied in mass screening. Tbl. 2 suggests that in this case, half of the patients not carrying cluster would be correctly classified and 6% of carrying patients would be incorrectly classified as non-carrying.

CONCLUSION

We have presented a methodology for detection and segmentation of MC. An advantage of using isolines to recognize MC is that the description of the shape is immediately available for the characterization step. Current work is exploiting these results to establish an automatic malignancy grading of clusters. This classification is being compared to radiologists' and anatomopathologists' classifications.

ACKNOWLEDGMENTS

We would like to address special thanks to MECASERTO, IFSBM and *Région Ile de France* who support this work, and to Dr. Cherel and Dr. Frouge for their expert help.

REFERENCES

Basset L. W. and Gold R. H. G. Breast Cancer Detection. Grune and Stratton 1987.

Caldwell C. B., Stapleton S. J., Holdsworth D. W., Jong R. A., Weiser W. J., Cooke G., and Yaffe M. J. Characterization of mammographic parenchymal pattern by fractal dimension. In SPIE Proc. Medical Imaging 1989, volume 1092, pages 10-16. SPIE.

Feig S. A. Decreased breast cancer mortality through mammographic screening: results of clinical trials. Radiology 1988, 167:659-665.

Karssmeijer N. and van Earning L. Iso-precision scaling of digitized mammograms to facilitate image analysis. In Proc. SPIE 1991, volume 1444, pages 166-177.

Lefebvre F. Détection et Interprétation des Microcalcifications en Mammographie Numérique. PhD thesis 1993, Université Paris XI.

Lefebvre F., Benali H., and Kahn E. Fractal analysis of clustered microcalcifications in mammograms. Acta stereol 1992, 11:611-616.

Stoksik M., Lane R., and Nguyen D. Practical synthesis of accurate fractal images. Graphical Models and Image Processing 1995, 57(3):206-219.

Presented at the 9th International Congress for Stereology, Copenhagen, August 20th to 25th, 1995.