Identification of Abnormalities in Mammograms Images Using Methods in the Spatial Domain

B. Luna-Benoso

Instituto Politécnico Nacional
Escuela Superior de Cómputo
Av. Juan de Dios Batíz, esq. Miguel Othón de Mendizábal
Ciudad de México 07320, México
blunabenoso@yahoo.com

J. C. Martínez Perales

Instituto Politécnico Nacional
Escuela Superior de Cómputo
Av. Juan de Dios Batíz, esq. Miguel Othón de Mendizábal
Ciudad de México 07320, México
pepe_cruzm@hotmail.com

Rolando Flores Carapia

Instituto Politécnico Nacional
Centro de Innovación y Desarrollo Tecnológico en Cómputo
Av. Juan de Dios Batíz, esq. Miguel Othón de Mendizábal
Ciudad de México 07700, México
rfcarapia@yahoo.com

Victor Manuel Silva García

Instituto Politécnico Nacional
Centro de Innovación y Desarrollo Tecnológico en Cómputo
Av. Juan de Dios Batíz, esq. Miguel Othón de Mendizábal
Ciudad de México 07700, México
Abstract

This paper presents a fast and efficient automated method for segmentation abnormalities present in mammogram images. For this, it uses methods in the spatial domain such as high-boost filter, binarization by thresholding and mainly related components. The method was applied to images obtained from the database mini-MIAS data base (Mammogram Image Analysis Society database (UK)), for this, it has used images that have abnormalities regardless if they are benign or malignant.
Identification of abnormalities in mammograms images

1 INTRODUCTION

Each year are diagnosed 1,151,000 new cases of breast cancer worldwide [1], becoming a health problem that is increasing, being the leading cause of death among women. Early detection of cancer can reduce mortality, for this, women must be performed clinical tests such as mammograms [2]. Mammography is a clinical examination that uses x-rays to detect breast abnormalities, abnormalities can be detected from nodules and changes in breast tissue that can later develop into cancer and that neither the woman nor a specialist can detecting physical examination [3, 4, 5]. Previously, the results of the mammograms were recorded on photographic plates, however, the problem of radiation exposure is an important factor, so, in 2000, the Food and Drug Administration (FDA for short in English: Food and Drug Administration) approved a new method for mammography: digital mammograms [6, 7, 8, 9], and this will reduce the radiation exposure. Digital mammograms breast image stored on a computer, where the specialist conducts the necessary studies.

Based on digital images of mammograms, this paper proposes a methodology to carry out the segmentation process and get the extraction of abnormalities present using spatial domain methods, ie, by processes that operate directly on the pixels.

2 PROPOSED MODEL

Before getting into detail with the proposed model, it is worth mentioning the definition of what is a digital image. A digital image is a two-dimensional function \( f(x, y) \) of the light intensity (brightness) at a point in space, where \( (x, y) \) coordinates of that point considering the origin at the top left of the image [10]. Since a digital image is a function \( f(x, y) \) discretized in both spatial coordinates and in brightness, often commonly represented as a two dimensional matrix \( F_{ij} = (f_{ij})_{H \times W} \) where \( H \) and \( W \) represent the image size (\( H \) and \( W \) referring to height and width of the image respectively) with \( f_{ij} = f(x_i, x_j) \) (Figure ??).

The method described in this work is carried out as follows: First consider digital mammography images, which are grayscale as shown in (Figure ??).
Given a digital image of a mammogram the segmentation method consists of 5 steps: (1) application of high-boost filter, which makes use of the high-pass filter, (2) difference between the original image and that produced by the passage (1), (3) binarization by thresholding, (4) removing the connected component corresponding to the pectoral muscle to be displayed in a mammogram image, and (5) elimination of the noise by the location of the related segment of greater area. Promptly immediately presents each step [10]:

2.1 High-boost filter

The goal of high-boost filter is to enhance the fine details of an image, in this case used to enhance abnormalities presented in mammography images. The high-boost filter is defined as:

\[
High - boost(f_{ij}) = (A - 1)_{(f_{ij})} + FPA(f_{ij})
\]

where \(F_{ij} = (f_{ij})_{H \times W}\) represents the original image, the variable \(A\) is an amplification factor and FPA is the high pass filter that is defined immediately. The highpass filter uses a 3 x 3 mask as shown in Figure ??.

\[
\begin{array}{ccc}
-1 & -1 & -1 \\
-1 & 0 & -1 \\
-1 & -1 & -1 \\
\end{array}
\]

Figure 3: Grayscale image obtained from a digital mammogram.

The mask is traversed for each pixel of the image, such that if \(f(x, y)\) represents the image that will apply the high-pass filter, then the new image \(h(x, y)\) is given by:

\[
h_{ij} = \begin{cases} 
255 & \text{if } FPA(f_{ij}) > 255 \\
0 & \text{if } FPA(f_{ij}) < 0 \\
f_{ij} & \text{otherwise}
\end{cases}
\]
where

\[ FPA(f_{ij}) = -f_{i-1,j-1} - f_{i-1,j} - f_{i,j-1} - f_{i+1,j-1} + c f_{i,j} - f_{i,j+1} - f_{i+1,j} - f_{i+1,j+1} \]

Figure ?? shows the application of high-boost filter applied to two images of the mini-MIAS repertoire with application factor \( A = 2.5 \).

Figure 4: (a) and (b) show digital mammography image, and (c) and (d) show the application of high-boost filter to Images (a) and (b) respectively with \( A = 2.5 \).

2.2 Difference between images

Given two images \( f(x, y) \) and \( g(x, y) \), the difference between images \( (f - g)(x, y) \) is defined as follows:

\[
(f - g)_{ij} = \begin{cases} 
255 & \text{if } f_{ij} - g_{ij} > 255 \\
0 & \text{if } f_{ij} - g_{ij} < 0 \\
 f_{ij} & \text{otherwise}
\end{cases}
\]

Figure ?? shows the result after the difference between the original image and the image after applying the high-boost filter of the figures shown in Figure ??.
Figure 5: Result of applying difference between the original image and the one obtained after applying high-boost filter.

2.3 Binarization by thresholding

Given a an image \( f(x,y) \) and a variable \( v \), the binarization by thresholding used is defined as follows:

\[
bin_{ij} = \begin{cases} 
255 & \text{if } f_{ij} \leq v \\
0 & \text{if } f_{ij} > v 
\end{cases}
\]

Figure ?? shows the result of applying binarization by thresholding with \( v = 120 \) to the images shown in figure ??.

Figure 6: Binarization by thresholding.

2.4 Removing connected component corresponding to the pectoral muscle.

Shows that each digital mammography image shows a section corresponding to the pectoral muscle of the person, this section is always on top of the image, either right or left, as shown in Figure ??.
To delete the image corresponding to the pectoral muscle, we consider that the image is in black and white, then we place a pixel on the front lines that belonged to his chest and we label it, and we use the 8-connectivity to label equally to all those pixels that belong to this connected set of pixels. Figure ?? shows the result of removing the section for the pectoral muscle of the images shown in Figure ??.

Figure 7: Removing the connected component corresponding to the person’s chest.

### 2.5 Elimination of noise by locating connected segment larger area

Once the mammogram image has obtained, result after the series of steps (1) to (4) of the proposed method, images are observed as in Figure ???. These figures show type noise salt and pepper. To eliminate this noise will seek the connected component with the greatest area, and the rest is eliminated, so that with this process, the resulting image is segmentation of abnormality present in mammography as shown in Figure ??.

Figure 8: Segmentation of abnormalities found in mammograms using the proposed methodology.

Figure ?? Shows the result of the segmented image superimposed on the original image. The blue section is segmented section of the steps defined in the methodology.
3 EXPERIMENTS AND RESULTS

In this study we used digital mammography images obtained from "The mini-MIAS database of mammograms" [11], which are of size 1024 x 1024. The repertoire of mini-MIAS images, only used the images that present any abnormality (101 images), regardless whether it is benign or malignant.

Given an image of the bank of images, was applied high-boost filter with amplification factor $A = 2.5$, was applied immediately the difference between the original image and the one obtained when applied to high-boost filter for the binarization by thresholding is consider the threshold value of $v = 120$. Until this series of steps, the images obtained are as shown in Figure 6, these figures are unusual in displaying an image section corresponding to the pectoral muscle, was observed in each of the images this section always appears at the upper left or right of the image, to remove this section was scanned pixels in the top of the image to locate one that corresponded to the section of the pectoral muscle, which is what we want to eliminate, once located one pixel, we searched all the pixels that were connected to it, that is, connected component sought to correspond to the section to be removed. After removing the pectoral muscle, section images were obtained as shown in Figure 7, which clearly shows the presence of noise in the form of salt and pepper. To remove noise and stay with the binary image corresponding to the abnormality, connected component looked with the largest area, which corresponds to the abnormality, and the rest was removed from the image. Finally achievement segment the image and display only the presence of the abnormality, and to show it, then overcame the segmented image with the original image.

Figure ?? shows the experimental corresponding to another pair of images of the image bank of mini-MIAS.
Identification of abnormalities in mammograms images

Figure 10: (a) and (c) original images, (b) and (d) a segmented image, (c) and (e) segmented images superimposed on the original.

4 CONCLUSION
In this paper we presented a fast and efficient automated method for segmentation of digital images abnormalities in mammograms. The method was applied to image bank provided by "The mini-MIAS database of mammograms". The proposed segmentation method can be a useful aid for early diagnosis of breast cancer, or for those who wish to carry out a computer system which addresses the problem of pattern classification.

5 ACKNOWLEDGMENT
The authors would like to thank the Instituto Politécnico Nacional (Secretaría Académica, COFAA, SIP, ESCOM, CIDETEC) and CONACYT for their economical support to develop this work.
References


Received: September 25, 2013