A Hybrid System Based on a Filter Bank and a Successive Approximations Threshold for Microcalcifications Detection

Humberto Ochoa[1], Osslan Vergara[1], Vianey Cruz[1,2], Efrén Gútierrez[1]
[1]Universidad Autónoma de Ciudad Juárez, Chih, México
[2] Centro Nacional de Investigación y Desarrollo Tecnológico (cenidet), Cuernavaca; México
Emails: hochoa@uacj.mx, *overgara@uacj.mx, vianey@cenidet.edu.mx, egutierre@uacj.mx

***Abstract*— In this paper, we propose a new hybrid system for microcalcifications detection in digital mammograms, using the combination of the CDF 9/7 filter bank and a successive approximations threshold. Microcalcifications are low contrast samples and only have a few pixels in diameter which are difficult to detect. We shown that microcalcifications have not only support in high frequency regions, but also along the entire frequency spectrum. The digital mammograms are analyzed and the lowest frequency subband dropped. After recovering the image a successive threshold is calculated to keep the samples with higher amplitudes. Results show that the proposed method reveals accurately the small injuries in digital mammograms.**

***Index Terms*— Breast cancer, Microcalcifications detection, Filter bank, DCF 9/7.**

I. Introduction

Mammography is a method that uses low dose of x-rays to produce a picture of the breast. This method is also known as screen-film mammography or simply mammogram and is the most common widely used technique to determine the existence of breast cancer [1]. Breast cancer is the most common cause of death in middle–aged women [2]. In addition, from 30 to 50% of the tissue surrounding malignant tumors of the breast, contains groups of microcalcifications [3], [4].

Microcalcifications are small deposits of calcium in the breast that cannot be felt but only can be seen on a mammogram. These specks of calcium may be benign or malignant and could be a first cue of cancer. Clusters of microcalcifications have diameters from some μm up to approximately 200 μm [5]. On a digital mammogram, microcalfications appear as a group from one up to few number of high intensity samples, usually considered regions of high frequency on a digital mammogram.

Among the methods used to detect microcalcifications on digital mammograms are those that use wavelet transform implemented as filter banks [6], [7], [8]. Wavelets are mainly used because of their dilation and translation properties, suitable for non stationary signals [9].

In this paper we use a Cohen-Daubechies-Feauveau (CDF) 9/7 filter bank [10], [11] to separate the mammogram using 3 levels of wavelet decomposition. Decomposition is followed by discarding the lowest frequency subband. After reconstruction, a threshold based on successive approximations was applied to recover the regions of highest intensity.

II. The Frequency Support of a Microcalcification

In this section, the frequency support of microcalcifications is analyzed. Thus, their energy across the frequency spectrum is considered. More levels of wavelet decomposition are used to increase the energy of a microcalcification in order to detect them.

A. Experimental Analysis of the Energy of Microcalcifications via DCT

In order to investigate the frequency support of a microcalcification, twelve microcalcifications of 8 x 8 samples each were taken from different digital mammograms [12]. The mean of each microcalcification was subtracted and the 2D Discrete Cosine Transform (DCT) [13] was applied to change from the spatial domain into the frequency domain. Fig. 1 shows the image of a collection of several microcalcifications used for the experimental analysis.

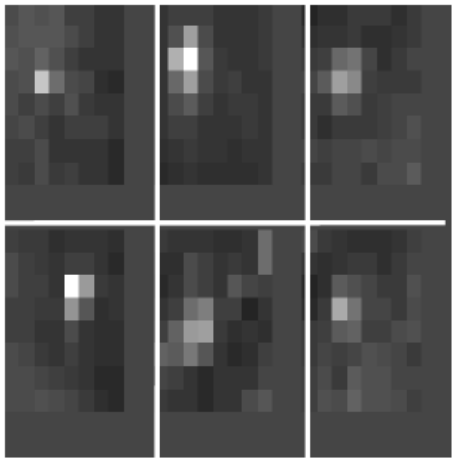


Figure 1. Microcalcifications obtained from different digital mammograms.

DCT is an orthogonal transform and completely decorrelates a signal. Besides, DCT can be applied to matrices of small dimension. Therefore, each 8 x 8 samples of microcalcification was transformed into an 8 x 8 different non-overlapping frequency components. We obtained as a result a matrix of coefficients arranged in zig-zag as shown in Fig. 2, from the DC component (top

left corner) to the highest frequency component (bottom right corner) [13]. The coefficients fulfill the condition shown in Eq. 1.

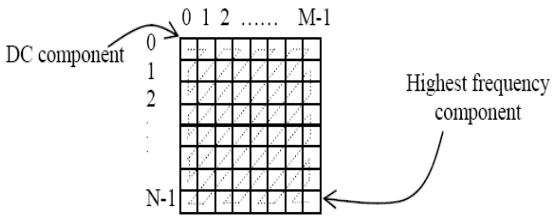


Figure 2. Arrangement of DCT coefficients.

$$En[x]=\frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}|x(n,m)|^2=\frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1}|X(k,l)|^2=En[X] \quad (1)$$

Where $x(m,n)$, are the samples of the microcalcification in the spatial domain, $X(k,l)$ the transformed coefficients, N the number of rows, M the number of columns in the sample domain, $En[x]$ the energy of the samples in the spatial domain, and $En[X]$ the energy of the coefficients in the transformed domain. The microcalcifications shown in Fig. 1 correspond to the 3D plots of Fig. 3. Peaks represent the specks of calcium and are surrounded by noise. Thus, the microcalcifications have a noise component $\eta(m,n)$ added to it. Therefore, the total effect can be expressed with Eq. 2.

$$\hat{x}(m,n)=x(m,n)+\eta(m,n) \quad (2)$$

Where $\hat{x}(m,n)$ is the microcalcification plus a noise component. According to the microcalcification definition and, in order for a microcalcification to be detected, two conditions must be observed: $x(m,n)$ must be compactly supported and $En[x]>En[\eta]$, otherwise the injury is not a microcalcification.

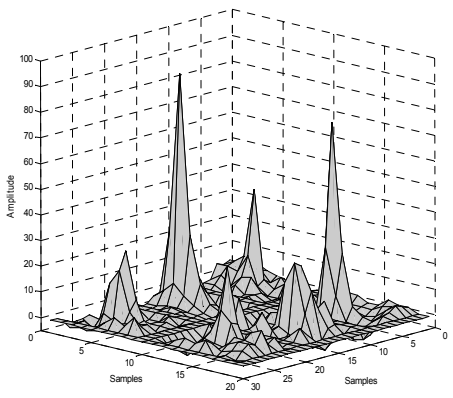


Figure 3. Amplitudes of 6 microcalcifications shown in Fig. 1 contaminated with noise.

B. Zonal Filter

In order to discard (set to zero) from the highest to the lowest frequency coefficients, seven zonal filters were applied to the transformed coefficients. Fig. 4 shows the masks applied to the DCT coefficients. The dark area corresponds to the discarded coefficients and white squares are set to one.

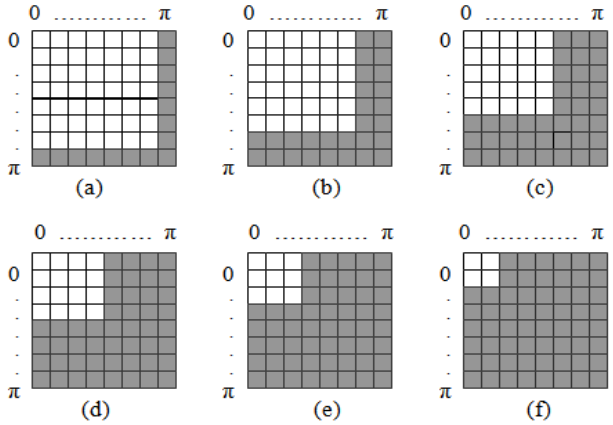


Figure 4. Zonal filters for discarding a) 15, b) 28, c) 39, d) 40, e) 55, and f) 60 coefficients.

Zonal filters throw-outs intervals from a) $\left[\frac{7}{8}\pi,\pi\right]^2$, b) $\left[\frac{3}{4}\pi,\pi\right]^2$, c) $\left[\frac{5}{8}\pi,\pi\right]^2$, d) $\left[\frac{1}{2}\pi,\pi\right]^2$, e) $\left[\frac{3}{8}\pi,\pi\right]^2$, and f) $\left[\frac{1}{4}\pi,\pi\right]^2$ rads/sample. Notice that the filters remove high frequency intervals, keeping the low pass intervals.

In order to investigate the effect of the discarded coefficients, the energy of the interval kept was calculated. The strength of the microcalcification is determined by the percent of retained energy after the zonal filter (see Table I).

TABLE I.
PERCENT OF RETAINED ENERGY AFTER ZONAL FILTER

	% of Total Energy	Zonal Filter (a)	Zonal Filter (b)	Zonal Filter (c)	Zonal Filter (d)	Zonal Filter (e)	Zonal Filter (f)
		% of retained energy					
Micro 1	100	85.24	74.85	66.10	60.11	47.09	35.40
Micro 2	100	89.72	77.89	67.40	55.76	21.20	13.98
Micro 3	100	98.28	94.11	77.89	72.27	31.86	6.45
Micro 4	100	83.39	72.05	55.71	48.48	14.35	3.20
Micro 5	100	97.14	92.41	83.46	77.42	47.26	5.65
Micro 6	100	93.64	83.55	77.26	67.32	30.51	11.36

C. Frequency Support of Microcalcifications

In Table I we reported the percentage of retained energy after zonal filters. It should be noted that filters (b) and (d) retain frequencies in the interval $\left[0,\frac{3}{4}\pi\right)^2$ and $\left[0,\frac{1}{2}\pi\right)^2$ respectively.

However, when filter (b) is applied; more than 82% on average of the total energy is retained and when filter (d) is used more than the 50% of the total energy is retained. Nevertheless, one could expect the energy to drop more drastically when the filter affects the high pass regions because we assume that a microcalcification is a high frequency signal.

Each microcalcification on Fig. 3 can be represented as in Eq. 2. Conversely, it is necessary to also look into the effect of the noise, so that we can have a better approximation of this signal. Thus, we set to zero some low magnitude samples around the microcalcification and leave a set of 4 x 5 samples representing the microcalcification as shown in Fig. 5. We did not apply any filtering because it was necessary to keep the shape and the actual amplitude of the surge. Without loss of generality, Eq. 2 can now be expressed as:

$$\hat{x}(m,n) \cong x(m,n) \tag{3}$$

In Table II we report the percentage of retained energy after removing the noise and applying the zonal filters of Fig. 4. It should be noted that all frequency bands are affected but the energy content of the noise is much less energy than the energy of the microcalcification. Also we notice that the percent of retained energy is still high for high frequencies.

TABLE II
PERCENT OF RETAINED ENERGY AFTER ZONAL FILTER.

	% of Total Energy	Zonal Filter (a)	Zonal Filter (b)	Zonal Filter (c)	Zonal Filter (d)	Zonal Filter (e)	Zonal Filter (f)
		% of retained energy					
Micro 1	100	80.50	74.09	60.97	39.34	33.82	5.61
Micro 2	100	87.53	81.02	67.08	43.49	34.68	5.56
Micro 3	100	98.75	94.21	70.64	59.42	26.24	5.22
Micro 4	100	80.41	67.34	55.27	25.56	21.20	2.28
Micro 5	100	98.03	92.50	84.19	54.67	45.41	5.93
Micro 6	100	93.26	79.25	59.767	44.988	23.95	3.48

Fig. 6 (a) shows the spatial support of one of the microcalcifications of Fig. 3; and Fig. 6 (b) shows the frequency response. The frequency content is distributed in the entire frequency spectrum with approximately the same magnitude.

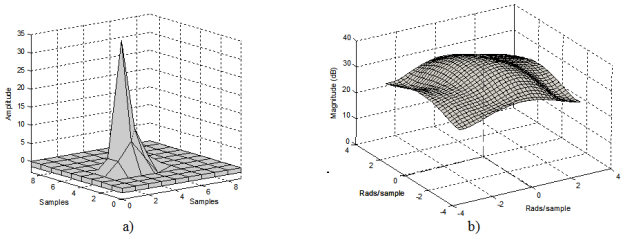


Figure 6. Microcalcification. a) Spatial support and (b) its frequency response.

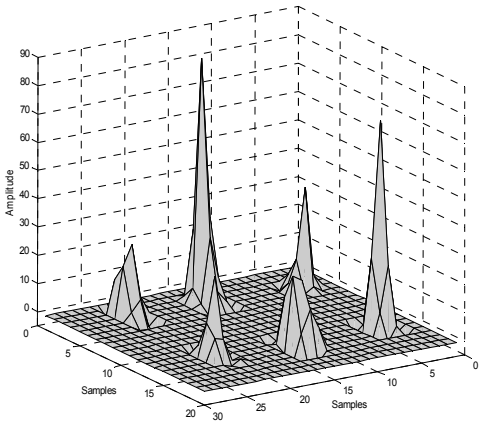


Figure 5. Amplitudes of microcalcifications after removing noise.

III. THE WAVELET TRANSFORM

Wavelet transform decomposes a signal onto a set of bases functions called "*wavelets*". The Discrete Wavelet Transform (DWT) is used for discrete signals and it is closely related to the multirate signal processing technique [9]. The basis functions are obtained from a single "*wavelet mother*" $\psi(n)$ by dilations and contractions (scaling) $\psi_{a,b}(2^a n - b)$, $a,b \in Z$.

The particular wavelet decomposition relates to filter banks. Wavelet-based image decomposition is a filtering process. For a given image $\boldsymbol{X}$ of size N x M, subband decomposition can be performed as follows: $h1(n)$ and $h2(n)$ are low and high pass wavelet filters respectively with frequency response $H1(\omega)$ and $H2(\omega)$ respectively. With low pass filtering we obtain the background and with high pass filtering the details of the image $\boldsymbol{X}$.

Filtering along rows is followed by downsampling the columns, then filtering along columns is followed by downsampling the rows (see Fig. 7). Since we downsample by a factor of two in both directions the size of the resulting subbands is $N/2$ x $M/2$ and can be expressed as $H1H1X$, $H1H2X$, $H2H1X$, and $H2H2X$. The subband $H1H1X$ contains the lowest frequency coefficients or smooth information and background intensity of the image and $H1H2X$, $H2H1X$, and $H2H2X$ contain the detail information. $H1H2X$ gives the horizontal high frequencies (vertical edges), $H2H1X$ vertical high frequencies (horizontal edges), and $H2H2X$ high frequencies in both directions (diagonal edges).

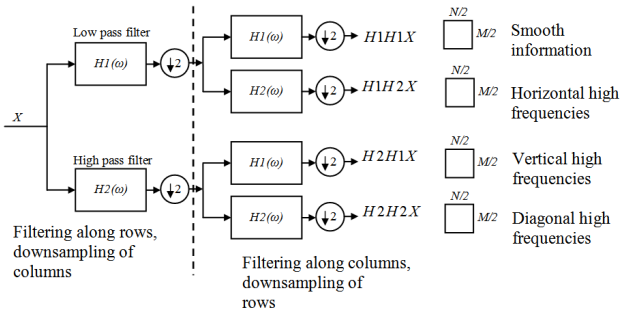


Figure 7. One level of 2-D subband decomposition.

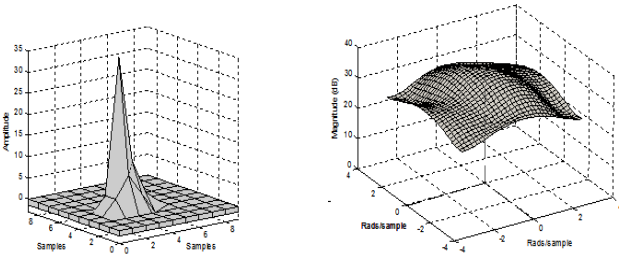
When methods such as Discrete Wavelet Transform (DWT) [6], [7], [14] are used to detect microcalcifications, the lowest frequency subband is discarded and the digital image mammogram recovered. If one level of DWT is used, the interval of discarded frequencies is $\left[0, \frac{\pi}{2}\right)^2$, and corresponds to ¼ of the total number of coefficients. This corresponds to apply the zonal filter of Fig. 4 (d). From Table II we can notice that the energy kept on average is 44.57 % only. As the number of decomposition levels increases, less energy is discarded. Therefore, the retained energy increases, increasing the energy of a microcalcification as well as the noise energy. However, the noise amplitude is shorter

than the microcalcifications amplitude. Then, the energy of a recovered microcalcification is higher as more energy is used to represent it. Therefore, the definition of a microcalcification is limited by the number of wavelet decomposition levels applied to the mammogram and by the inherent noise.

IV. Experimental Results

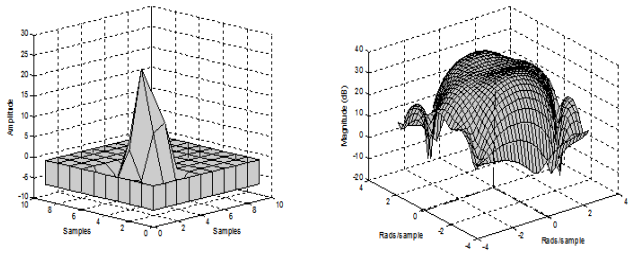
In this section, we discuss the results of tests carried out on several microcalcifications and the detected microcalcifications after using a successive approximations threshold. For experimental investigations, the Discrete Fourier Transform (DFT) was applied to injuries and the resulting frequency coefficients plotted to visualize the frequency support. Also in this section is shown the result of the detected injuries.

Figs. 8 and 9 (a) show the amplitudes of two different microcalcifications and Figs. 8 and 9 (b) show the frequency support after the application of the DFT. The spatial support of Fig. 8 (a) is less than that in Fig. 9 (a). Both microcalcifications have a large support in frequency domain. However, the injury of Fig. 8 (b) has an all pass characteristic while the one on Fig. 9 (b) has a low pass characteristic with a large extent in the band pass.



a) b)

Figure 8. a) Microcalcification with a large amplitude, short spatial support and b) its magnitude response.



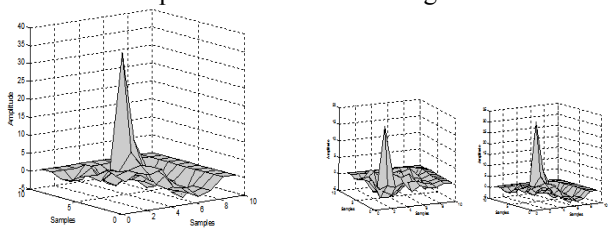
a) b)

Figure 9. a) Microcalcification with a short amplitude and b) its magnitude response.

This operation corresponds to applying the zonal filter of Fig. 4 (d). From Table 2 we can notice that the energy kept is, on average 48.46% approximately. As the number of decomposition levels increases, less energy is discarded and also the shape of a microcalcification is preserved. However, the energy of the surrounding noise also increases, but in less amount than that of the recovered microcalcification.

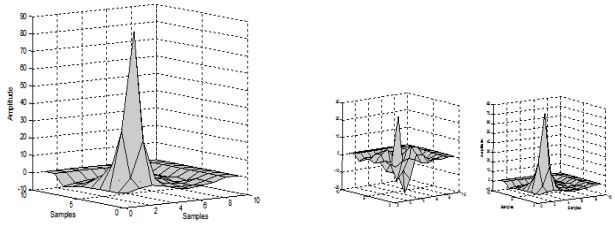
Regularly after IDWT, a threshold is optimized to separate the samples with larger magnitude from the noise. Thus, detection of a microcalcification is limited by the number of wavelet decomposition levels and the optimized threshold.

Fig. 10, 11, and 12 (a) show microcalcifications with a short spatial support and high amplitude after subtracting the spatial mean. Figures 10, 11, and 12 (b) and (c), show the recovered microcalcifications after one and four DWT decomposition levels respectively. The shape and amplitude of the injury are drastically changed when one decomposition level is used. As the number of decomposition levels increases, the shape and amplitude approximates to the original microcalcification. Fig. 13 (a) shows a microcalcification with short amplitude. After one level of DWT decomposition (see Fig. 13 (b)) the shape and amplitude of the injury is totally missed. We also observe that the amplitude of the recovered injury is close to the amplitude of the surrounding noise.



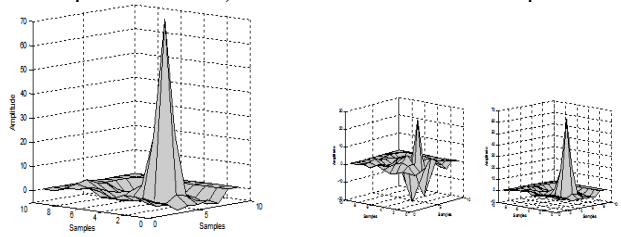
(a) (b) (c)

Figure 10. a) Microcalcification with a short spatial support and short amplitude, and recovered microcalcification after b) one level of DWT decomposition and c) four levels of DWT decomposition.



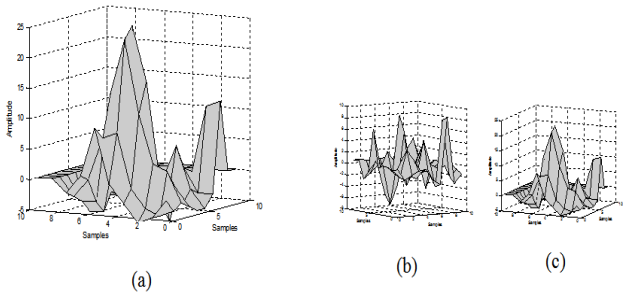
(a) (b) (c)

Figure 11. a) Microcalcification with a short spatial support and high amplitude, and recovered microcalcification after b) one level of DWT decomposition and c) four levels of DWT decomposition.



(a) (b) (c)

Figure 12. a) Microcalcification with a short spatial support and high amplitude, and recovered microcalcification after b) one level of DWT decomposition and c) four levels of DWT decomposition.



(a) (b) (c)

Figure 13. a) Microcalcification with a long spatial support and short amplitude, and recovered microcalcification after b) one level of DWT decomposition and c) four levels of DWT decomposition.

As a microcalcification has a large support along the frequency spectrum, they can be enhanced using a Discrete Wavelet Transform with orthogonal property and energy compaction of small details (impulses and thin lines) properties which are characteristics of the CDF 9/7 filter bank [9].

For experimental investigations, the CDF 9/7 filter bank is applied to the mammograms using 3 decomposition levels and discarding the lowest frequency subband before reconstruction. The three levels were applied to the image in order to keep most of the energy of small microcalcifications in the frequency domain.

To remove the inherent reconstructed noise a threshold based on successive approximations was applied to the recovered image. The initial threshold is calculated with equation 4.

$$Th = 2^{\lfloor log_2 \, max(c_{i,j}) \rfloor} \qquad c_{i,j} \in \hat{X} \qquad (4)$$

Where $\lfloor \ \rfloor$ is a floor operation, $\max(c_{i,j})$ is the maximum magnitude of the coefficients in the recovered image, and $\hat{X}$ is the recovered image. If the sample is greater or equal than the threshold then we say it is significant. In each pass, we look for significant samples in the entire image. If a sample is significant and it has not been significant previously, then it is fully recovered. After each pass, the threshold is halved and a new pass started. Consequently, higher magnitude samples are always recovered first, and we can decide the number of passes or bit planes to recover, in order to avoid noise.

Fig. 14 shows several detected microcalcification using 3 levels of wavelet decomposition and 3 bit planes recovered. Digital mammograms used to represent and measure microcalcifications in section I where not used in this section.

V. Conclusions

The experiments presented in this paper indicate that the frequency content of microcalcifications is extended in most of the frequency spectrum, rather than in a high frequency region only. The approach presented in this paper was motivated by the ability of wavelets to decompose an image onto different scales and resolutions. This approach exploit the frequency selectivity and energy compaction property of the CDF 9/7 filter bank.

The properties give us the advantage to decompose the mammogram into more levels, to discard the least possible energy when low frequency subband is thrown away. This fact not only keeps more energy of microcalcification, but also more energy of the noise which is assumed to be less amplitude than microcalcifications. However, a threshold based on successive approximations detects the samples with more amplitude first.

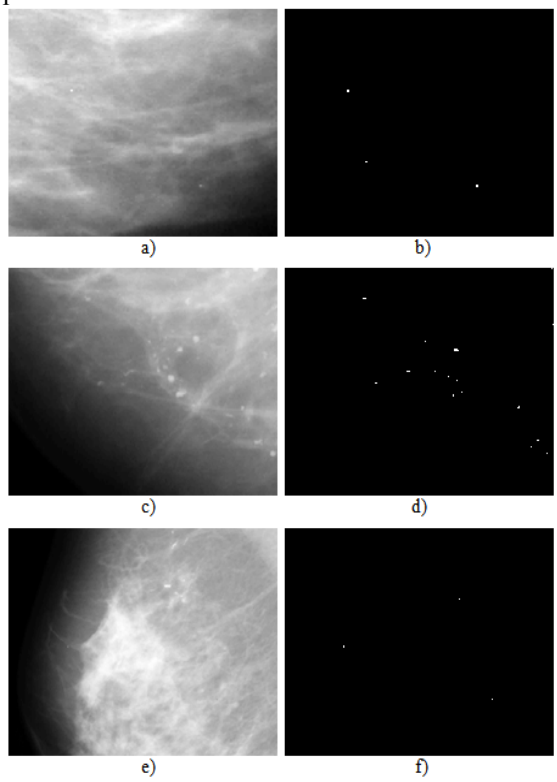


Figure 14. Digital mammograms from MIAS database (a) mdb210, (b) detected microcalcifications, (c) mdb245, (d) detected microcalcifications, (e) mdb249 and (f) detected microcalcifications.

References

[1] G. T. Barnes and G. D. Frey (eds), "Screen-film Mammography Imaging: Considerations and Medical Physics Responsibilities", *Medical Physics*, pp. 1– 46, 1991.

[2] R. A. Smith, Epidemiology of Breast Cancer, Categorical Course in Physics, Technical Aspects of Breast Imaging, Oak Brook, 1993.

[3] E. A. Sickles, "Mammographic Features of Early Breast Cancer", *American Journal of Roentgenology*, vol. 143, no. 3, pp. 461 – 464, 1984.

[4] E. A. Sickles, "Mammographic Features of 300 Consecutives Nonpalpable Breast Cancers", *American Journal of Roentgenology*, vol. 146, no. 4, pp. 661 – 663, 1986.

[5] J. K. Kook and H. P. Wook, "Statistical Textural Features for Detection of Microcalcifications in Digitized Mammograms", *IEEE Transactions on Medical Imaging*, vol. 18, no. 3, pp. 231 – 238, 1999.

[6] T. C. Wang and N. B. Karayiannis, "Detection of Microcalcifications in Digital Mammograms Using

Wavelets", *IEEE Transactions on Medical Imaging*, vol. 17, no. 4, pp. 498 - 509, 1998.
[7] R. Nakayama, Y. Uchiyama, K. Yamamoto, R. Watanabe and K. Namba, "Computer-Aided Diagnosis Scheme Using a Filter Bank for Detection of Microcalcification Clusters in Mammograms", *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 2, pp. 273 – 283, 2006.
[8] R. Nakayama, Y. Uchiyama, R. Watanabe, S. Katsuragawa, K. Namba and K. Doi, "Computer-Aided Diagnosis Scheme for Histological Classification of Clustered Microcalcifications on Magnification Mammograms", *Medical Physics*, vol. 31, no. 4, pp. 786 - 799, 2004.
[9] I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia, PA, SIAM, 1992.
[10] D. S. Taubman and M. W. Marcelin, *JPEG2000: Image Compression Fundamentals, Standards and Practice,* Boston, M.A, Kluwer, 2002.
[11] J. Maly and P. Rajmic, "Fast Lifting Wavelet Transform and its Implementation in Java", *International Federation for Information Processing (IFIP),* vol. 245, pp. 488 – 496, 2007.
[12] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, et al., "The Mammographic Image Analysis Society Digital Mammogram Database Experta Medica", *International Congress Series*, vol. 1069, pp. 375-378, 1994.
[13] K. R. Rao, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Boston, MA, Academic Press, 1990.
[14] G. Rezai-rad and S. Jamarani, "Detecting Microcalcification Clusters in Digital Mammograms Using Combination of Wavelet and Neural Network", *International Congress on Computer Graphics, Imaging and Vision: New Trends*, pp. 197–201, 2005.

Humberto de Jesús Ochoa Domínguez received the B.S. degree in Industrial Engineering from The "Instituto Tecnológico de Veracruz", México, M.S. degree in Electronic Engineering from The "Instituto Tecnológico de Chihuahua", México, and Ph.D. degree in Electrical Engineering from "The University of Texas at Arlington", USA. In 1998 he was awarded the Chihuahua Prize for the "System to Classify Digital Mammograms into Normal and Abnormal using Texture Features and Microcalcifications Detection. He has conducted several tutorials/workshops on image/video coding/processing in Mexico, United States, Italy, Czech Republic and Republic of Malta. He has published in refereed journals and has been a consultant to industry and academia.

He worked for the Mexican Merchant Marine as Electronic Officer and for the Nokia Research Center in Irving, Texas. He currently serves as a professor at the Autonomous University of Ciudad Juárez (UACJ), Chihuahua, México.

Dr. Ochoa works at the Electrical and Computer Engineering Department. He is member of the IEEE Computer Society and the IEEE Circuits and Systems Society. He is a member of the National Research Systems.

Osslan Osiris Vergara Villegas was born in Cuernavaca Morelos, Mexico on July 3th, 1977. He received the B.S. degree in Computer Engineering from the "Instituto Tecnológico de Zacatepec (ITZ)," México in 2000, the M.Sc. in Computer Science at the "Center of Research and Technological Development (CENIDET)" in 2003, and the Ph. D. degree in computer Science from the National Center of Research and Technological Development (CENIDET), Morelos, México, in 2006. His research interests include pattern recognition, digital signal processing, computer vision, image compression and mechatronics.

He currently serves as a professor at the Autonomous University of Ciudad Juárez (UACJ), Chihuahua, México.

Dr. Vergara works for the Manufacturing an Engineering Department. He is member of the IEEE Computer Society. He is a member of the National Research Systems.

Vianey Guadalupe Cruz Sánchez was born in Cardenas Tabasco, México on September 14th, 1978. She received the B.S. degree in Computer Engineering from the "Instituto Tecnológico de Cerro Azul," México in 2000, the M.Sc. degree in Computer Science at the "Center of Research and Technological Development (CENIDET)" in 2004. She is a Ph. D. student of Computer Science at the National Center of Research and Technological Development (CENIDET), Morelos, México. Her research interests include hybrid systems, knowledge representation, neural networks, pattern recognition and digital image processing.

She currently serves as a professor at the Autonomous University of Ciudad Juarez (UACJ), Chihuahua, México. He works for the Electrical and Computer Engineering Department. Prof. Cruz is a member of the IEEE Computer Society.

Efrén Gutiérrez Casas received the B.S. degree in Industrial Engineering from The "Instituto Tecnológico de Ciudad Juárez", México, M.S. degree in Electrical and Computer Engineering from "The University of Texas at El Paso", USA, and the Ph.D. degree in Electrical and Computer Engineering from "The University of Texas at El Paso", USA. He has conducted seminars in Image Processing in México. He has published in refereed journals.

Currently he serves as a professor at the Autonomous University of Ciudad Juárez (UACJ), Chihuahua, México.

Dr. Gutiérrez works for the Electrical and Computer Engineering Department. He is member of the IEEE Computer Society.