

CHALMERS



UNIVERSITY OF GOTHENBURG

PREPRINT 2012:18

On Efficiency of Combined Daubechies Wavelets and Statistical Parameters Applied in Mammography

M. ASADZADEH
E. HASHEMI
A. KOZAKEVICIUS

*Department of Mathematical Sciences
Division of Mathematics*

CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg Sweden 2012

Preprint 2012:18

**On Efficiency of Combined Daubechies Wavelets
and Statistical Parameters Applied
in Mammography**

M. Asadzadeh, E. Hashemi and A. Kozakevicius

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg, Sweden
Gothenburg, September 2012

Preprint 2012:18
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2012

ON EFFICIENCY OF COMBINED DAUBECHIES WAVELETS AND STATISTICAL PARAMETERS APPLIED IN MAMMOGRAPHY

M. ASADZADEH^{1,†}, E. HASHEMI^{1,2} AND A. KOZAKEVICIUS^{1,3,‡,*}

ABSTRACT. This note is a study of combined Daubechies (Db2) wavelet transform and the statistical parameters skewness and kurtosis applied to detection of microcalcification in mammography. We have succinctly introduced the concept of vanishing moments and derived scaling and wavelet functions using generating functions. The efficiency of the discrete algorithm is heavily relied on the order of performing wavelet approximation and the statistical procedure. The vital importance of both wavelet and statistical parameter approaches as well as the ordering issue in performing the analysis are justified through implementing numerical examples for some clinical data.

Keywords: Daubechies wavelet, vanishing moments, mammography, microcalcification, skewness, kurtosis.

AMS Subject Classification: 65T60, 68U10, 97K80, 97M60.

1. INTRODUCTION

Mammography is a common expression for specific type of imaging that uses a low-dose x-ray (photon-beam) therapy to examine breast tissue and a mammography exam is called a mammogram [13]. One of the main indicators of breast cancer searched in mammograms is a set of clusters formed by microcalcifications, these are tiny calcium deposits in breast tissue, that appear as small bright spots in the imaging [13].

In the last 15 years several mammography/image processing methods have been developed in order to help radiologists with the task of detecting microcalcifications. Among them wavelet based methods have been designed also in association with different statistical measurements [11, 12]. According to [12], microcalcifications in mammograms correspond to high frequency coefficients of the image spectrum, and a possible procedure to detect and extract these calcifications is simply to decompose the mammography image by wavelet transforms [2], suppress the low frequency subband (scaling coefficients block), and reconstruct a new image considering only the high frequency wavelet coefficients. The draw-back in this approach is that such procedure leads to a high number of false positive results, since no distinction is done among the high intensity wavelet coefficients.

To filter out false positive results and identify the regions of interest a useful tool is to employ the statistical parameters, called *skewness* and *kurtosis*. This is in the sense that, in a region containing microcalcifications, the symmetry of the Gaussian distribution of wavelet coefficients is deteriorated and hence the tails of their distribution are heavier, and/or their heights are

[‡] The corresponding author.

¹Chalmers University of Technology, Sweden, ² Polytechnique de Torino, Italy, ³Universidade Federal de Santa Maria- UFSM, Brazil,

e-mail:mohammad@chalmers.se, eliza.hashemi@gmail.com, alicek@ufsm.br

[†] The research of this author is supported by the *Swedish Foundation of Strategic Research (SSF) in Gothenburg Mathematical Modeling Center (GMMC)* and Swedish Research Council (VR).

The author* would like to thank the National Counsel of Technological and Scientific Development- CNPq (Brazil) under grant No 201457/2010-5 and the Mathematical Sciences Department of Chalmers University of Technology and University of Gothenburg.

Manuscript received xx.

peaked or flat [11]. The statistical parameters measuring these deformation effects in Gaussian distributions are the third and fourth order correlation parameters, called skewness and kurtosis, respectively. Therefore their computation for the wavelet coefficients may be used to characterize regions most similar to microcalcifications clusters, as the regions of interest. The analysis of these parameters, in connection with wavelet transforms goes back to [12], where the first analysis were done considering undecimated wavelet transforms. The ideas initially presented in [12] were then revisited in [11].

The main differences between the proposed approach in here and the methodology presented in [11] is that, now the discrete Daubechies wavelet transform with two vanishing moments, Db2, is employed, leading to simpler analysis algorithms, with the advantage of minimizing the boundary distortions caused by long sized filter wavelet functions, as for example wavelets with a higher number of vanishing moments. With the Db2 filters it is possible to detect and precisely localize continuous piecewise linear signals. A property that is carried out also in the two dimensional context, i.e. to detect and localize subregions inside the image corresponding to linear data variations. For such a signal or image Db2 would yield zero wavelet coefficients (once the linear information is entirely represented by the scaling coefficients due to the two vanishing moments) at the linearity subregions. Nevertheless at the site of singularities and some close neighborhoods the wavelet coefficients would capture the higher variations, which due to the type of singularity, can be spread out for many levels of the transformation.

Another difference of the proposed algorithm and the previous studies is that now there is no need to consider any overlapping subregions. This change implies in a major simplification of the microcalcification detection procedure, since no strategy for selecting and dealing with the subregions analysis is necessary. Instead, in our algorithm, three subbands with the corresponding wavelet coefficients are considered, where on each row and column of these subbands, the statistical parameters skewness and kurtosis are computed. The vectors containing these statistical quantities, computed by rows and columns, are then thresholded, keeping only the significant values, those which are higher than a chosen threshold parameter (here 80% of the maximum of the details of each band). The crossing of rows and columns associated to the significant values determine candidate regions of microcalcifications clusters.

An outline of this work is as follows. In Section 2 a swift framework to the Daubechies wavelet (Db2) construction is summarized, highlighting the importance of the vanishing moments property. We have also introduced some aspects of the statistical parameters: skewness and kurtosis. The analysis of the algorithm is presented in Section 3. Also a series of tests are designed in order to validate the methodology and to stress issues related to the relevance of the combination of the two chosen parameters with the wavelet transform. Numerical results, presented in Section 4, are based on the analysis of mammography from the web page of the American cancer society. The main advantage of considering data from this data base is that annotations from specialists are provided with data, which is primordial for the validation of the developed algorithms. Finally, Section 5 is devoted to comments and concluding remarks.

2. WAVELET FRAMEWORK AND STATISTICAL PARAMETERS

2.1. The Daubechies wavelet DB2. In this section we wish to motivate the analytical aspects considered in the choice of the Daubechies Wavelet with two vanishing moments (Db2) in our study of mammography. A vast class of wavelet families suffers from some major drawbacks which can lead to some inconveniences during the analysis. Haar wavelets have compact support but are discontinuous. Linear spline wavelets are continuous, but the orthogonal scaling function and associated wavelet have infinite support, *they do, however, decay rapidly at infinity*, [6].

The hierarchy of Daubechies wavelet family contains the Haar wavelet, which is the only discontinuous one. The other orthonormal Daubechies wavelets are both compactly supported and continuous. The degree of smoothness of Daubechies wavelets increase with raising their

hierarchy number, associated to the capacity the scaling functions have to span polynomials from degree 0 until a certain degree $n - 1$. The number of vanishing moments of the wavelet family is then said to be n , since the wavelet coefficients for those n smooth functions are all zero, for all levels of decomposition (by the orthogonality between wavelets and scaling functions) [5].

To detect a biological disorder in human tissue, we are dealing with certain biochemical models in a bounded three dimensional domain with particular singularity aspects. On the other hand a raw model for human tissue/organ would involve differential equations and hence requires to assume a certain degree of smoothness. The main computational objective of this study is two-fold: i) To model a problem setting using optimally small number of degrees of freedom in order to keep the cost down. ii) To single out the healthy tissue regions from the regions of interest in order to help the oncologist in radiation treatment of malignant tumor. Therefore, here Db2 fits well for the purpose of our study, specially because the considered data is discrete and the way boundaries are treated matters now for the final analysis.

Below we give a brief mathematical background for the construction of Db2. In this setting, to introduce Db2 wavelets we state the properties that guarantee the pointwise and L_2 convergence of scaling function defined by a polynomial $p(\xi) = P(e^{-i\xi})$, (with $P(e^{-i\xi}) = \frac{1+e^{-i\xi}}{2} = e^{-i\xi/2} \cos(\xi/2)$). Obviously $p(\xi)$ satisfies the properties

$$\begin{cases} p(0) = 1 \\ |p(\xi)|^2 + |p(\xi + \pi)|^2 = 1 \\ |p(\xi)| > 0, \quad \text{for } |\xi| \leq \pi/2. \end{cases} \quad (1)$$

We recall the scaling equation

$$\phi(x) = \sum_k p_k \phi(2x - k),$$

wher the coefficients p_k are given by

$$p_k = 2 \int_{-\infty}^{\infty} \phi(x) \overline{\phi(2x - k)} dx.$$

Now let $\phi_0(x)$ be the Haar scaling function and define

$$\phi_n(x) = \sum_{k \in Z} p_k \phi_{n-1}(2x - k), \quad \text{for } n \geq 1, \quad (2)$$

where, Z is the set of integers. Then the sequence $\{\phi_n\}$ converges pointwise and in L_2 to a function $\tilde{\phi}$, satisfying the orthogonality condition

$$\int_{-\infty}^{\infty} \tilde{\phi}(x - n) \tilde{\phi}(x - m) dx = \delta_{nm}, \quad (3)$$

and the scaling equation $\tilde{\phi}(x) = \sum_k p_k \tilde{\phi}(2x - k)$. Obviously $\tilde{\phi} \equiv \phi$. The function

$$P(e^{-i\xi}) = \frac{1 + e^{-i\xi}}{2} = e^{-i\xi/2} \cos(\xi/2),$$

is called the generating polynomial and is an example of a polynomial satisfying (1). The iterative procedure above yields the Daubechies scaling function ϕ , and the associated wavelet is given by

$$\psi(x) = \sum_{k \in Z} (-1)^k p_{1-k} \phi(2x - k). \quad (4)$$

A general form for the generating polynomial is given by

$$P(z) = \frac{1}{2} \sum_{k \in Z} p_k z^k, \quad (5)$$

where z is a complex number with $|z| = 1$. A common truncation is of the form $P_N(z) = \frac{1}{2} \sum_{k=0}^{2N-1} p_k z^k$, where in the two simplest cases with $N = 1$ (Haar) and $N = 2$ (Daubechies), the polynomials are

$$P_1(z) = \frac{1}{2}(1+z), \quad \text{and} \quad P_2(z) = \left(\frac{1+\sqrt{3}}{8} + \frac{1-\sqrt{3}}{8}z \right) (1+z)^2. \quad (6)$$

In general we have the following result.

Proposition 2.1.1 (Daubechies). *For every N , there will be $2N$ nonzero, real scaling filter coefficients $p_0, p_1, \dots, p_{2N-1}$, resulting in a scaling function and wavelet that are supported on the interval $0 \leq t \leq 2N-1$. They are chosen so that the corresponding degree $2N-1$ polynomial $P_N(z) = \frac{1}{2} \sum_{k=0}^{2N-1} p_k z^k$ has the factorization*

$$P_N(z) = (z+1)^N \tilde{P}_N(z), \quad (7)$$

where the degree of \tilde{P}_N is $N-1$ and $\tilde{P}_N(-1) \neq 0$.

This guarantees that the associated wavelets will have precisely N vanishing moments.

Thus, both polynomials $P_1(z)$ and $P_2(z)$ in (6) have the factorization (7) with $\tilde{P}_1(z) = \frac{1}{2}$ and $\tilde{P}_2(z) = \frac{1+\sqrt{3}}{8} + \frac{1-\sqrt{3}}{8}z$. Since P_N are real, $\overline{P_n(-z)} = P_n(-\bar{z})$, and one can easily show that the scaling function ϕ_N and wavelet ψ_N have Fourier transforms given by

$$\hat{\phi}_N(\xi) = \frac{1}{\sqrt{2\pi}} \prod_{j=1}^{\infty} P_N(e^{i\xi/2^j}), \quad (8)$$

$$\hat{\psi}_N(\xi) = e^{-i\xi/2} P_N(e^{i\xi/2}) \hat{\phi}_N(\xi/2). \quad (9)$$

Then, we have that

$$\hat{\psi}_N^{(k)}(0) = \begin{cases} 0 & k = 0, \dots, N-1, \\ -N!(-i/2)^N \tilde{P}_N(-1)/\sqrt{2\pi} \neq 0, & k = N. \end{cases} \quad (10)$$

This yields the following result by Daubechies:

Proposition 2.1.2 (Daubechies). *The wavelet function ψ_N possesses exactly N vanishing moments, i.e., we have*

$$\int_{-\infty}^{\infty} x^k \psi_N(x) dx = \begin{cases} 0 & k = 0, \dots, N-1, \\ -(2^{-N} N!/\sqrt{2\pi}) \tilde{P}_N(-1), & k = N. \end{cases} \quad (11)$$

Daubechies wavelets are classified according to the number of vanishing moments they have. The smoothness of the scaling function and wavelet increases with their number of vanishing moments. As we mentioned above, the case $N = 1$ is the same as the Haar case, where both scaling function and wavelet are discontinuous. For $N = 2$, the Daubechies scaling function and wavelet are continuous but certainly do not have smooth derivatives. In the $N = 3$ case, both scaling and wavelets functions are continuously differentiable. For more details see, e.g. [3, 1].

Why is it useful to have vanishing moments? The short answer is that vanishing moments are a key factor in many applications-compression, noise removal, singularity detection, for example. According to expression (11) for the moments computation, in the case $N = 2$ (Db2), the first two moments ($k = 0, 1$) vanish, and no information is represented in terms of the wavelet, just with respect to the scaling function. For the numerical applications this implies that only half of the discrete data, exactly the scaling coefficients, is necessary to represent the original signal, since all wavelets coefficients will be zero, as indicated by (11).

For polynomial of degree $k = 2$ considering Db2, $-(2^{-1}/\sqrt{2\pi}) \tilde{P}_2(-1) = -1/(2\sqrt{2\pi}) \cdot \sqrt{3}/4$ is the corresponding computed moment. Thus for the third moment we have

$$\int_{-\infty}^{\infty} x^2 \psi_2(x) dx = -\frac{1}{8} \sqrt{\frac{3}{2\pi}}. \quad (12)$$

We can use these moments to approximate and estimate the wavelet coefficients for smooth signals and show that these coefficients will be small when the level j is high. If f is a smooth, twice continuously differentiable signal, then for sufficiently large j , using Taylor expansion, its j, k -wavelet coefficient is approximated as,

$$\begin{aligned} d_k^j &= \int_{-\infty}^{\infty} f(x) 2^{j/2} \psi_2(2^j x - k) dx = \int_0^{3 \cdot 2^{-j}} f(x + 2^{-j} k) 2^{j/2} \psi_2(2^j x) dx \\ &\approx \int_0^{3 \cdot 2^{-j}} \left(f(2^{-j} k) + x f'(2^{-j} k) + \frac{1}{2} x^2 f''(2^{-j} k) \right) 2^{j/2} \psi_2(2^j x) dx. \end{aligned} \quad (13)$$

Now, since the first two moments vanish and the third is given in (12), we can approximate the wavelet coefficients for any level j and any position k as

$$d_k^j \approx -\frac{1}{16} \sqrt{\frac{3}{2\pi}} 2^{-5j/2} f''(2^{-j} k). \quad (14)$$

Note that the formula (14) is exact in any region where the signal (the function f) is a polynomial of degree ≤ 2 in x . As an application of this formula, we find a point where an otherwise smooth function has a discontinuity in the derivative. This application is called *singularity detection*, and this process can be used, among other things, to the purposes of interest in this work: the detection of microcalcification regions in mammogram images.

As an example, consider a piecewise linear function defined in a compact interval, with a slope change at a single interior point in the interval. Since $f'' \equiv 0$ where f is linear, hence (14) implies that the only nonzero wavelet coefficient will come from a small region near the corner point where the slope changes.

Example 2.1.3. Consider the piecewise linear signal

$$f(x) = \begin{cases} 0.5x + 1, & 2 \leq x \leq 3 \\ 1.5(x - 1), & 3 < x \leq 4 \end{cases}$$

After sampling the signal at 128 equally spaced points, we perform one level of decomposition using the Db2 ($N = 2$ Daubechies) wavelet with $2N$ scaling filters. Thus our starting finest level of representation is indexed by $j = 7$, and $j = 6$ is the level immediately coarser obtained by the transform. The only two considerable non-zero wavelet coefficients are d_{30}^j and d_{31}^j corresponding to a wavelet supported on the interval $[3 - \alpha h_l, 3 + \alpha h_l]$, where $\alpha = 2N$ is the size of the scaling filters and $h_l = lh_0$ is the corresponding spacing according to the level of decomposition, here $l = 1$ and the finest grid spacing for the original grid is $h_0 = 1/64$. Consequently, the singularity is supported in this interval.

2.2. Skewness and Kurtosis. A thorough wavelet study would yield a final imaging that does not miss any abnormalities. This would include both malignant tumors as well as healthy bins. To detect the true positives, the sort of false positives are heavily relied on the match between the two statistical measuring parameters: *skewness* and *kurtosis*. More specifically, the need to study skewness and kurtosis is due to the fact that the normal curve often fails to give an adequate representation for a considered data.

These quantities, which are playing a crucial role in many applications with data interpretations, are not usually described mathematically. Below we give a short description of the concepts of skewness and kurtosis and refer the reader to standard statistical texts for further properties and some detailed descriptions. The need to study skewness and kurtosis is due to the fact that the normal curve often fails to give an adequate representation for a considered data. While skewness is a measure of the asymmetry of the probability distribution function, the kurtosis can be employed as a measure of its peakedness. In Pearson's curve system it has been suggested that to model the data selected on the base of observed third (skewness) and fourth (kurtosis) statistical moments. This is elaborated in some details in Moors [9]. Then, the skewness of a

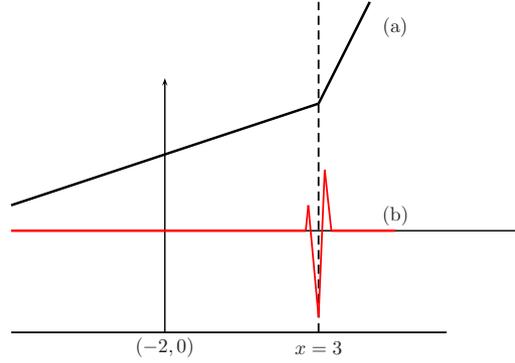


Figure 1: (a) A piecewise linear signal with a singularity at $x = 3$. (b) The wavelet component.

random variable X is often measured by the standardized third moment $\gamma_1 = \mu_3/\sigma^3$, and the fourth standardized moment is defined as $\gamma_2 = \mu_4/\sigma^4$, where $\mu_i := \frac{1}{M} \sum_{k=1}^M (x_k - \bar{x})^i$, $i = 3, 4$. The value of these moments may become arbitrarily large, and therefore hard to interpret.

For a robust measurement, skewness should satisfy the following property, (see [10]):

- (i) For any $a > 0$, and b , $\gamma_1(f) = \gamma_1(af + b)$.
- (ii) If $f(x)$ is symmetrically distributed, then $\gamma_1(f(x)) = 0$.
- (iii) $-\gamma_1(f(x)) = \gamma_1(-f(x))$

(iv) If F and G are cumulative distribution functions of $f(x)$ and $g(x)$, and $F <_c G$, then $\gamma_1(f(x)) \leq \gamma_1(g(x))$, where $<_c$ is the skewness-ordering among distributions. One aspect in our study would rely on justification of these properties for the probability density distribution of sampled data. This is a rather established test in mathematical statistics, see, e.g. [10]. However, there is an equally important statistical measurement quantity: namely the *kurtosis*, which either in combination with the skewness, or independently, may be employed for the analysis of statistical data.

Corresponding robustness conditions for kurtosis are not so easy to formulate and therefore are missing in the literature. Nevertheless, there are other suggestions characterising kurtosis, e.g. Moors [9] showed that kurtosis is easily interpreted as a measure of dispersion around two values $\mu \pm \sigma$. For justification of this we shall rely on the following definition:

Definition 2.2.1. For any distribution the kurtosis k is defined as the normalized fourth central moment (if it exists). Hence, for a random variable X with expectation $\mu := E(X)$, variance $\sigma^2 := V(X)$ and $E(X^4) < \infty$,

$$k = E(X - \mu)^4 / \sigma^4,$$

i.e., here μ_4 is defined as $\mu_4 := E(X - \mu)^4$.

Introducing a standardized variable $Z := (X - \mu)/\sigma$ yields

$$k = E(Z^4) = V(Z^2) + [E(Z^2)]^2 = V(Z^2) + 1. \quad (15)$$

Hence, k may be interpreted as a measure of the dispersion of Z^2 around its expectation 1, or that of the dispersion Z around ± 1 . Therefore k measures the dispersion of X around the two values $\mu \pm \sigma$. One can show that for any symmetric two-point distribution (15), i.e. k , attains the minimum value 1. This indicates that high kurtosis are possible in two cases: either

concentration of probability mass near μ , or concentration of probability mass in the tails of the distribution.

An alternative description is based on the following definition.

Definition 2.2.2. For a random variable X the *octiles* are defined by

$$P(X < E_i) \leq i/8, \quad P(X > E_i) \leq 1 - i/8, \quad i = 1, 2, \dots, 7. \quad (16)$$

For continuous X with distribution function F , the octiles are unique ($E_i = F^{-1}(i/8)$ for $i = 1, 2, \dots, 7$), and therefore (16) is simplified as

$$F(E_i) = i/8, \quad i = 1, 2, \dots, 7. \quad (17)$$

For $E_6 > E_2$ a conventional measure proposes an alternative and robust kurtosis measure defined by

$$\Lambda := (E_7 - E_5) + (E_3 - E_1)/(E_6 - E_2). \quad (18)$$

Moors justified this estimator on the ground that the two terms, $(E_7 - E_5)$ and $(E_3 - E_1)$, are large (small) if relatively little (much) probability mass is concentrated in the neighborhood of E_6 and E_2 , corresponding to large (small) dispersion around $\mu \pm \sigma$. For further studies we refer the reader to [9]. Hence, both definitions are indicating that the two approaches yield to the same characterization of kurtosis as being just *a measure of dispersion around $\mu \pm \sigma$* .

In our case we specifically deal with a sample of $M = 2^{m_0}$ measurements. Then, for simplicity, using a somewhat modified notation, the sample skewness S and kurtosis K are given by

$$S := \gamma_1 = \frac{\frac{1}{M} \sum_{k=1}^M (x_k - \bar{x})^3}{\left(\frac{1}{M} \sum_{k=1}^M (x_k - \bar{x})^2\right)^{3/2}}, \quad \text{and} \quad K := \gamma_2 = \frac{\frac{1}{M} \sum_{k=1}^M (x_k - \bar{x})^4}{\left(\frac{1}{M} \sum_{k=1}^M (x_k - \bar{x})^2\right)^2} - 3, \quad (19)$$

respectively, where \bar{x} is the sample mean.

3. PROPOSED METHODOLOGY

The proposed methodology is based on the main concepts addressed in the previous section about the application of the Db2 wavelet transform as well as the computation of the two statistical parameters, skewness and kurtosis.

Since the object of our study is mammogram imaging, the data considered for a cross-section, will be presented in a matrix Q , based on the two-dimensional discrete Db2 wavelet transform. The implementation in this case is a straight forward extension of the one-dimensional discrete Db2 wavelet transform, given by the Cascade Algorithm [8]. In the Algorithm 1 the direct 1D transform for L levels of decomposition is presented, where the initial vector is denoted by c_0 , and the output vector by $WT1D(c_0)$. The 2D transform is obtained by the application of the Algorithm 1 to each row of the input matrix and after-wards to each column of the resulting matrix.

Since the matrix Q has finite dimension (in our applications $2^9 \times 2^9$) and the Db2 filters h_l, g_l contain 4 values each, the computation of the coefficients $c_{j+1,k}$ and $d_{j+1,k}$ at positions near the right boundary $k = n_j - 1, 1 \leq j \leq L$ will require elements from the input vector c_j at positions that will be outside the range of variation of c_j . To overcome this issue, a periodic extension is considered. Values $c_{j,k}$, whose positions k exceed $n_j - 1$, will then be chosen from values at the beginning of the vector c_j just by considering the modulus operation $\tilde{k} = \text{mod}(k, n_j)$. The main advantage of the choice of Db2 is that this wavelet will cause the smallest interference on the boundaries because of the periodic extension. For our application this is a desirable feature, once the microcalcifications can be localized at any position of the mammography.

For our application the number of levels of decomposition L in Algorithm 1 is considered to be 1, where the wavelet coefficients will be used as a sensor to detect the regions of interest. We do not need necessarily to reconstruct the image after the analysis. Therefore after one level of decomposition, each row of Q (the input vector c_0 in Algorithm 1) will be initially

Algorithm 1: WT1D: 1D Direct wavelet transform for J decomposition levels

input : c_0 : Initial data vector; $n_0 = M = 2^{m_0}$: vector size;
 $N = 2$: for Db2, h_l and g_l : scaling and wavelet filters
output: c_J, d_J, \dots, d_1 : Set of scaling and wavelet coefficients for J levels
for $j \leftarrow 0$ **to** $J - 1$ **do**
 for $k \leftarrow 0$ **to** $n_j - 1$ **do**
 $c_{j+1,k} = \sum_{i=0}^{2N-1} h_i c_{j,2k+i}$
 $d_{j+1,k} = \sum_{i=0}^{2N-1} g_i c_{j,2k+i}$
 $n_{j+1} \leftarrow n_j / 2$
return *Wavelet decomposition vectors*: $WT1D(c_0) = (c_J, d_J, \dots, d_1)$

decomposed into two components (c_1, d_1) , producing an intermediate matrix \tilde{Q} . We transpose \tilde{Q} just as an auxiliary operation, and then once again apply Algorithm 1, this time only for one level decomposition. As the result of the two-dimensional wavelet transform (TW2D) the initial matrix Q is decomposed into four blocks, $\tilde{\tilde{Q}} = TW2D(Q) = (cc_1, dc_1, cd_1, dd_1) = (C, H, V, D)$. Here, the block $cc_1 = C$ contains the scaling coefficients, and the blocks $dc_1 = H$, $cd_1 = V$ and $dd_1 = D$, are the wavelet coefficients, which capture data variations in the horizontal, vertical and diagonal directions, respectively.

The next step in our methodology is the computation of the statistical parameters skewness and kurtosis for each block of wavelet coefficients by rows and by columns. Algorithm 2 summarizes these operations and also indicates how the regions of interest are determined.

To specify the regions associated to microcalcification clusters one needs, essentially, to determine the rows and columns associated to the highest values of kurtosis and skewness computed for each block H , V , and D . In order to have a criteria to select those highest values, which are also called significant values, the hard threshold operation Th [7] is considered. The key ingredient in this operation is the choice of the threshold parameter λ . Since this parameter is responsible for the entire decision process.

The threshold operation is stated in expression (20) below for a single value s , considering λ as the threshold value. In Algorithm (2), the threshold operation given by (20) is applied for each one of the elements of the skewness vectors S_r^V , S_r^H and S_r^D computed based on considering the row values of the blocks V , H , D . Analogously for the vectors S_c^V , S_c^H and S_c^D computed by considering the columns of each block. The threshold value λ is computed according to the rule: $\lambda = 0.80 \max_{b \in \{V, H, D\}} \{ \max_j \{ MSj_b \} \}$, taking into account separately the quantities obtained by rows $j = r$ and those obtained by columns $j = c$, as indicated in Algorithm (2).

$$Th(s) = \begin{cases} s & \text{if } |s| \geq \lambda \\ 0 & \text{if } |s| < \lambda. \end{cases} \quad (20)$$

After the significant rows and columns were selected by the threshold operation, for each of the three blocks of wavelet coefficients, for $b \in V, H, D$, the row and column sets \mathcal{R}^b and \mathcal{C}^b are created. To produce the final selection of significant positions associated to microcalcification clusters on the original image, the intersections of these sets are considered, i.e., a row or a column is associated to a microcalcification region if it is significant with respect to the three blocks of wavelet coefficients, generating the index sets \mathcal{R} and \mathcal{C} . This criteria is important to avoid a considerable number of false positive results, as discussed in the next section.

The procedure executed by Algorithm (2) is now repeated substituting the skewness computations by that of the kurtosis computations of the wavelet coefficients by blocks. For the analysis based on kurtosis, the same strategy is applied, values are computed by rows and columns of each block. Again the threshold operation with threshold parameter λ , defined analogously,

Algorithm 2: Regions of interest determination via skewness

input : Q_{n_0, n_0} Initial data matrix; $n_0 = M = 2^{m_0}$

output: \mathcal{R}, \mathcal{C} : set of significant rows and columns

Selected Regions are neighborhoods of the crossing of elements in \mathcal{R} and \mathcal{C}

for b in $\{V, H, D\}$ **do**

for each row r **do**

S_r^b : skewness for row r , block b

$MSr_b \leftarrow \max_r \{S_r^b\}$: Max skewness per block (row computation)

for each column c **do**

S_c^b : skewness for column c , block b

$MSC_b \leftarrow \max_c \{S_c^b\}$: Max skewness per block (column computation)

Threshold Operation

$T_r \leftarrow \max_b \{MSr_b\}$

$T_c \leftarrow \max_b \{MSC_b\}$

$\tilde{S}_r^b \leftarrow Th(S_r^b), \lambda = 0.80T_r$

$\tilde{S}_c^b \leftarrow Th(S_c^b), \lambda = 0.80T_c$

Significant rows and columns per block

$\mathcal{R}^b = \{r : \tilde{S}_r^b \neq 0\}$

$\mathcal{C}^b = \{c : \tilde{S}_c^b \neq 0\}$

$\mathcal{R} = \bigcap_b \mathcal{R}^b$

$\mathcal{C} = \bigcap_b \mathcal{C}^b$

return Significant rows and columns \mathcal{R}, \mathcal{C}

Regions of interest are the crossing sites of significant rows and columns and their immediate neighboring surroundings

selects the most significant kurtosis values by block. As performed for the skewness parameters, the rows and columns associated to the significant kurtosis values, finally, would provide us the index sets \mathcal{R} and \mathcal{C} , determining the regions of interest. In the next section, numerical experiments considering both versions of the proposed Algorithm (2) are presented. To highlight the important role played by the association of wavelet coefficients and the skewness and kurtosis computations, other experiments with some variation in the Algorithm (2) are also designed and their results are presented and commented as well.

4. NUMERICAL RESULTS AND DISCUSSION

The mammographies presented in this section, scanned as raw format with 8-bit gray scale and 512×512 pixels size, were obtained from the University of South Florida Digital Mammography Home Page [14]. This data base also provides annotations from specialists for each image, indicating the clusters associated to malignant tissues. In order to validate our methodology, 24 images were analyzed, 6 without microcalcification clusters, and 18 containing clusters associated to malignant tissues.

4.1. Results for the proposed methodology. A mammography with normal tissues is presented in Figure (2a). Figure (2b) shows a mammography presenting microcalcification clusters and also the annotation according to [14] (given by the red curve) for the clusters associated to malignant tissues.

Figure (2c) presents the same mammography, as in Figure (2b), this time however, including annotations in yellow. The yellow lines were obtained as the result of the Algorithm (2) applied

for both skewness and kurtosis computations. The intersections of the lines indicate the neighborhoods where microcalcification clusters are more alike. In our simulation two regions were detected, a cluster associated to the malignant tissue, as well as a second cluster, also associated to microcalcifications according to [14], but the latter one without malignant tissues.

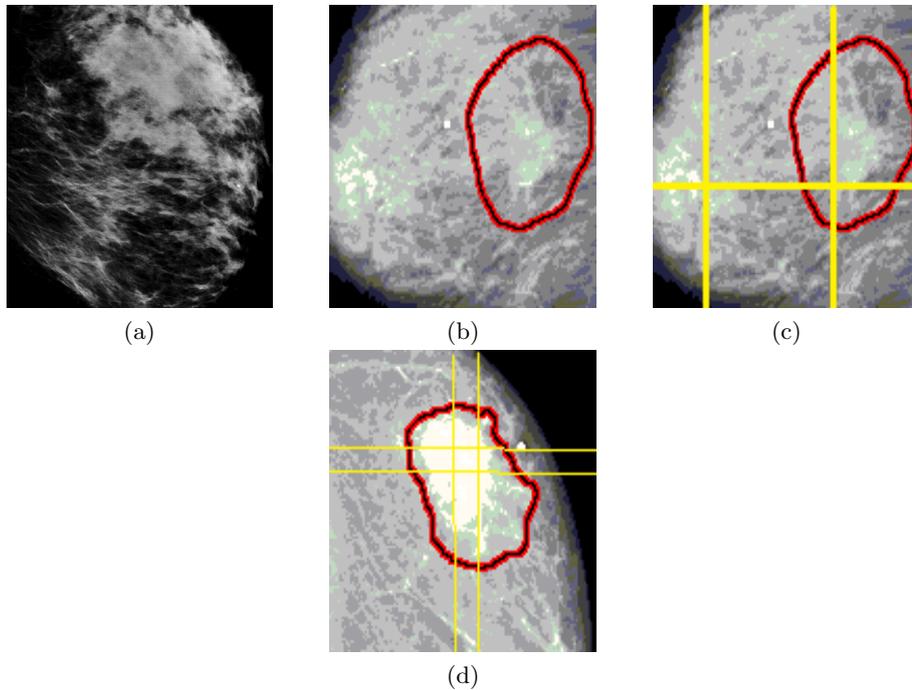


Figure 2: (a)Left breast mammography with normal tissues. (b),(c), (d)Examples with microcalcification: Regions of interest are neighborhoods of the significant rows and columns crossings (Yellow lines). Red curves given by [14]

It is also important to point out that the main purpose of our study is not to determine whether a cluster is associated to malignant tissues or not. The main goal of our methodology instead is to provide a support for specialists to pay attention to regions that might not be observed, but could be of interest. Since in some cases, such regions are not easily detectable, as can be seen in the examples provided in [14].

Figure (2d) is yet another example of mammography containing microcalcifications cluster. Again the red curve is the annotation provided by the data bank [14] and the yellow lines are the results for the analysis obtained by the proposed methodology. In this case, the intersections of the yellow lines determine the same cluster. In our statistical approach such results are not considered as false positive results, even-though over determined, the localization of the microcalcification cluster were done correctly.

In order to collect statistics about the number of correct detections, 24 mammographies from [14] were analyzed, 6 without any cluster, 18 with annotations for regions with malignant tissues. For the test group considered normal, the algorithm did not detect any region of interest, obtaining therefore 100% of correct detection.

For the 18 abnormal cases results are shown in Table (1). The detection results based on the Algorithm (2), considering skewness, matches to the results obtained when the analysis were done based on the kurtosis computation, i.e. the same regions were detected when kurtosis was taken into account. This information is also provided in Table (1), and specified by the wording: “*same region*”.

The multiple (more than one) regions identified related to the same microcalcification clusters that do not characterize false positive results in our statistics, are referred to as *over determined*

regions. In our experiment the number of false detections were relatively small, three cases in 18 analyzed mammography with malignant tissues. The column labeled as Annotations in Table (1) provides the information given by [14], indicating the number of malignant clusters identified by specialists. We include a remark concerning our over determined detections in the Annotations columns.

Abnormal Mammographies - Algorithm 2				
Image	Detection		skewness and kurtosis	Annotations [14]
	skewness	kurtosis		
1	1	1	same region	1
2	4	4	same region	4
3	1	1	same region	1
4	2	2	same region	1 (over determined)
5	2	2	same region	2
6	1	1	same region	1
7	2	2	same region	2
8	2	2	same region	1 (over determined)
9 (Fig. 2d)	4	4	same region	1 (over determined)
10	4	4	same region	2 (1 false positive)
11	4	4	same region	2 (1 false positive)
12	2	2	same region	2
13	1	1	same region	1
14	6	6	same region	6
15	3	3	same region	1 (1 false positive)
16 (Fig. 2c)	2	2	same region	1 (over determined)
17	2	2	same region	1 (over determined)
18	4	4	same region	1 (over determined)

Table 1: Performance of the proposed detection methodology for abnormal mammography

4.2. Additional simulations. To provide a more precise justification of our methodology, some additional simulations were performed in order to numerically analyze the efficiency of the combination of the wavelet transform with the computation of the statistical parameters, skewness and kurtosis.

4.2.1. Test 1: skewness and kurtosis computed directly from the original input matrix. Now we apply Algorithm (2) with one major modification, without performing the two-dimensional wavelet transform. It means that, the values of skewness and kurtosis for rows and columns are computed directly for the input image data. The same set of 24 mammogram images (6 normal and 18 with abnormal regions) is considered for the analysis. Table (2) presents the results obtained for this Test (4.2.1), where now false detections were performed for the set of normal images. Then the first part of Table (2) shows the results for the three examples with wrong detections. The detection results for the remaining three normal cases were omitted, since the detection of the zero cluster was correct.

For the set of 18 images with abnormal clusters the results are presented in the second part of the Table (2). The column labeled by *Identified Malignant regions* indicates the number of matching identified regions according to the number of correct microcalcification regions specified on the data bank.

The intersections of rows and columns associated to the significant values for skewness and kurtosis indicate once again the detected regions. Nevertheless, for this experiment the number

Normal mammography				
Image number	Detection		skewness and kurtosis detections	Identified Malignant regions
	skewness	kurtosis		
19	5	5	same	5
20	30	30	same	30
21	12	12	same	12
Abnormal Mammograms				
Image number	Detection		skewness and kurtosis detections	Identified Malignant regions
	skewness	kurtosis		
1	2	2	same	0
2	1	1	same	0
3	4	4	same	0
4	6	9	not the same	0
5	12	12	same	0
6	9	9	same	9
7	12	12	same	0
8	16	8	not the same	0
9	12	9	not the same	0
10	15	15	same	0
11	16	16	same	0
12	24	24	same	12 (over determined)
13	5	5	same	5
14	15	15	same	15
15	8	8	same	0
16	30	30	same	0
17	28	28	same	0
18	24	24	same	0

Table 2: Performance using Test 1 on 20 images where 17 of them have abnormal tissues.

of false positives results were much higher than the previous one and the regions eliminated by the performed analysis, based on skewness and kurtosis, do not match for all analyzed images, as provided by the Algorithm (2). Another drawback of this test is that: now in some cases, regions in mammography were assigned to calcification clusters, where in fact no cluster was present, see for example Figure (3c). The original image is presented in Figure (4a) and the detection results for the proposed Algorithm (2) are presented in Figure (3b). This experiment indicates the relevance of the computation of the wavelet coefficients in combination with the statistical parameters for the correct analysis.

4.2.2. *Test 2.* This test is basically to investigate whether, or not, the wavelet transform applied after the computation of skewness and kurtosis, according to the Test (4.2.1), could improve the analysis or perhaps even avoid the misdirections that were previously obtained in Test 1. Therefore, for the rows and columns, skewness (S^r , S^c) and kurtosis (K^r , K^c) were obtained as described in (4.2.1), and then, the one dimensional wavelet transform is computed by Algorithm (1), for one decomposition level, generating vectors $WT1D(S^r)$, $WT1D(S^c)$, $WT1D(K^r)$ and $WT1D(K^c)$. Then to each one of these vectors the threshold operation (20) is applied. Once again the threshold value λ is given as in Algorithm (2). We obtain the lines and columns associated to the most significant values after thresholding (the analysis are preserved heuristically). Again the obtained information is transferred in a way that they are related to the original positions of input image.

In Table (3) results for the numerical experiments with 4 abnormal mammographies and 1 normal image are presented, which illustrates the failure of the Test 2 approach.

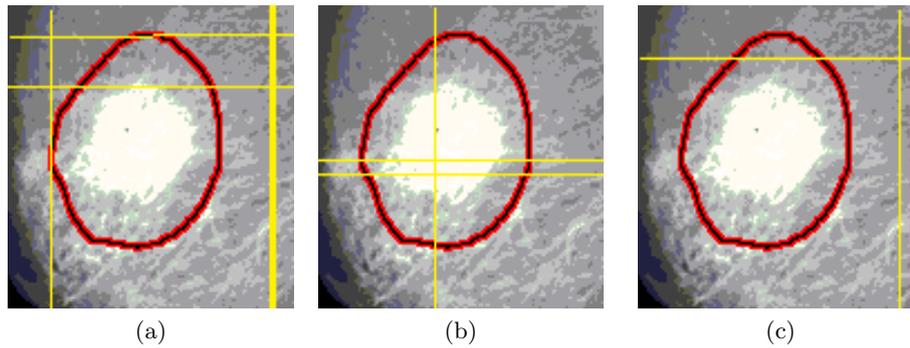


Figure 3: (a) Mammography with malignant tissue marked by physicians. (b) Two regions of interest selected as being neighborhood of the crossing of significant row and columns obtained by calculating skewness and kurtosis values of wavelet coefficients. (c) A region selected as being a neighborhood of the crossing of significant row and columns obtained by the analysis of skewness and kurtosis, computed directly from the image data indicating no interesting tissue.

Normal mammographies				
Image number	Detection		Skewness and kurtosis detections	Identified regions
	skewness	kurtosis		
1	6	6	same	6
Abnormal Mammograms				
Image number	Detection		Skewness and kurtosis detections	Identified regions
	skewness	kurtosis		
2	6	6	same	0
3	12	12	same	0
4	10	10	same	0
5	16	8	not the same	0

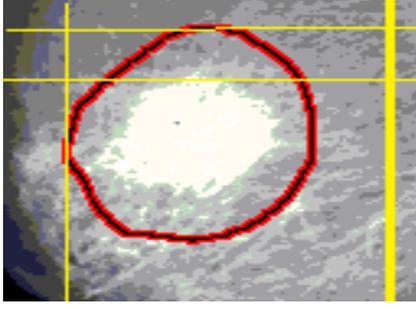
Table 3: Performance using Test 2 on 5 images where 4 of them have abnormal tissues.

The application of the wavelet transform associated to the threshold operation after the computation of skewness and kurtosis did not improve any results from the analysis obtained for Test (4.2.1). Figure (6) also illustrated the detected regions, indicating the failure in the computation of skewness and kurtosis directly from the original input data.

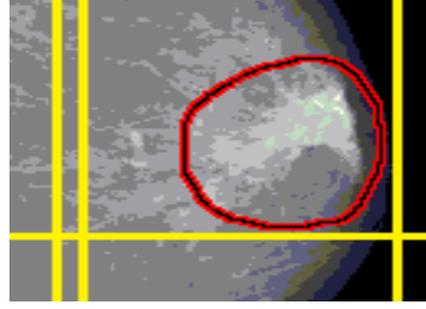
4.2.3. *Test 3.* As a final numerical experiment, we would like to confirm the importance of the application of the wavelet transform in the formulation of the proposed methodology presented by Algorithm (2). Therefore, in this test we substitute the wavelet transformation by considering to compute another quantity: *the Gradient of the original matrix*. We initially compute the gradient of the image in (x, y) -geometry (GI). The remaining analysis is kept unaffected. Hence, we follow the Algorithm (2) until the detection part is concluded.

Table (4) shows the performance of Test 3 for 10 mammographies with malignant tissues.

The results presented in Table (4) are for the computation of skewness, computed with respect to the coefficients obtained by the gradient estimation. In the case where the derivative in x -direction is computed, five of the presented images were detected as not having any malignant cluster. And for four other images the number of false positive detections was considerably high, as presented in Table (4). For the derivative in y -direction, no malignant detection was performed for three images and for the remaining six images the number of false detections was again considerably high. The results obtained for the computation of kurtosis were omitted,



(a) Image 2 in Table 3



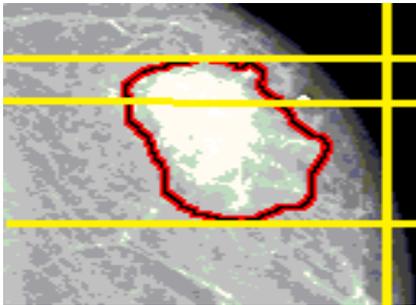
(b) Image 3 in Table 3

Figure 4: Test 2 selected regions, where detected clusters do not match with interested microcalcification clusters.

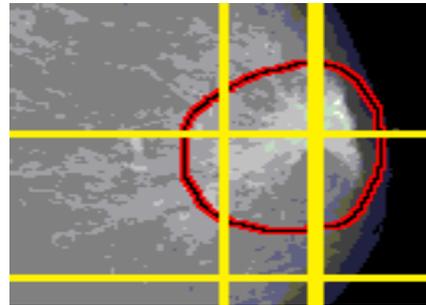
Abnormal Mammograms				
Image number	Detection		Detected region by skewness of G_x	Detected region by skewness of G_y
	skewness of G_x	skewness of G_y		
1	4	4	4	4
2	9	8	2	1
3	25	24	0	8
4	40	16	0	8
5	56	24	20	0
6	30	12	0	3
7	36	21	15	9
8	6	8	0	0
9	28	20	0	0
10	42	20	8	3

Table 4: Performance using Test 3 for 10 images with abnormal regions.

since the quality of the detection considering the modification proposed in this test was neither satisfactory.



(a) Image 10 in Table 4



(b) Image 3 in Table 4

Figure 5: Regions selected as being neighborhood of the crossing of significant rows and columns obtained by the analysis of derivatives in y -direction of skewness and kurtosis computed from the image data. (a) indicates no interesting tissue and (b) indicates the over detected of malignant tissue and plenty of false detections.

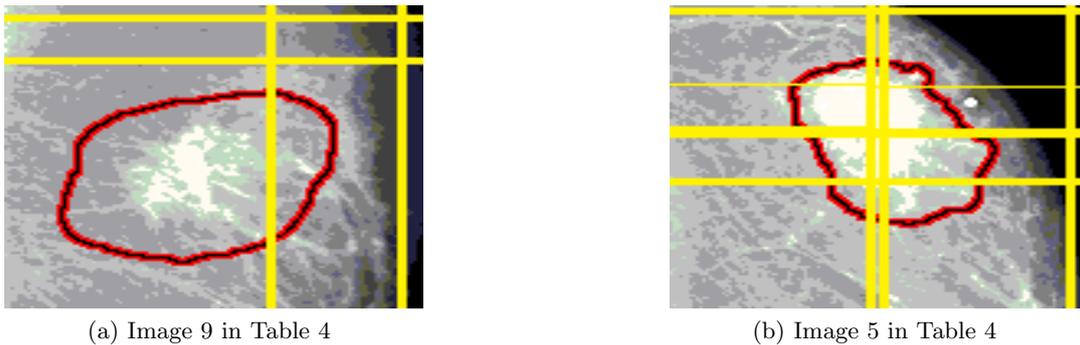


Figure 6: Regions selected as being neighborhood of the crossing of significant row and columns obtained by the analysis of derivative in x- direction of skewness and kurtosis computed from the image data, where (a) indicates no interesting tissue and (b) indicates the over detected of malignant tissue with plenty of false detections.

For the association of the wavelet coefficients with the computation of the third and fourth statistical moments (skewness and kurtosis), the ordering in the computations does matter for the entire algorithm performance; and the computation of these two statistical measurements based on wavelet coefficients is more efficient for the determination of malignant regions in mammographies than the computation of skewness and kurtosis directly from the image or from another transformation.

5. CONCLUSION

This work presents an approach to detect microcalcifications of mammography based on the two dimensional Daubechies wavelet transform and the computation of two statistical parameters, skewness and kurtosis. The potential of the three wavelet subbands of the mammography decomposition is explored to identify regions associated to microcalcifications clusters on the original image.

The importance of the use of the decimated wavelet transform with two vanishing moments (Db2) is argued analytically, as well as numerically. The performance of the proposed methodology, involving the analysis of the computed statistical parameters, was tested by reliable results for the detection of regions associated to microcalcifications. The set of 24 images used for the numerical tests were obtained from the American Cancer Society.

Also different associations were verified, though they were in poor quality and were not considered to be satisfactory. These variations were explained with Tests 1, 2 and 3. The first was the computation of skewness and kurtosis directly from the input matrix, the second was the application of the one-dimensional wavelet transform after the computation of the statistical parameters, and for the last test a different transformation at the initial stage of the proposed Algorithm (2) was considered. None of the modifications had produced results comparable to the performance presented by our Algorithm (2), what provides here an evidence of the relevance of the proposed algorithm to perform the analysis.

According to the numerical evidences presented by the tables and illustrated by the images, the methodology proposed and discussed in this paper is an easy to implement and effective alternative for the analysis of mammography, with the goal of detecting regions associated to microcalcification clusters.

REFERENCES

- [1] M. Asadzadeh, *Lecture Notes in Wavelets with Fourier Analysis*, Electronic version, Chalmers (2010).
- [2] Jöran Bergh, Fredrik Ekstedt, Martin Lindberg, *Wavelets*, Studentlitterature, Lund (1999).

- [3] Albert Boggess and Francis J. Narcowich, *A first Course in Wavelets with Fourier Analysis*, Prentice Hall, (2001).
- [4] M. Garcia, A. Jemal, *Global Cancer Facts and Figures 2011*, Atlanta, GA: American Cancer Society (2011).
- [5] Ingrid Daubechies, *Ten Lectures on Wavelets* (1992).
- [6] Lokenath, Debnath, *Wavelets Transforms and their Applications*, Birkhäuser, (2002).
- [7] David L. Donoho, *De-noising by Soft-Thresholding* IEEE Transactions on Information Theory, Vol. 41 No. 3, May 1995.
- [8] Mallat, S. *A Wavelet tour of signal Processing*, Academic Press, San Diego, 1998.
- [9] Moors, J. J. A. *A Quantitative Alternative for Kurtosis*, J. the Royal Stat. Soc. Ser. D (The Statistician) 37, pp. 25-32 (1988).
- [10] H. Oja, *On location, scale, skewness and kurtosis of Univariate Distributions*, Scand, J. Stat. pp 1-15 (1980).
- [11] K.Prabhu Shetty, V. R. Udupi and K. Saptalakar, *Wavelet Based Microcalcification Detection on Mammographic Images*, International Journal of Computer Science and Network Security, VOL. 9 No. 7, July 2009.
- [12] Ted C. Wang, Nicolaos B. Karayiannis, *Detection of Microcalcifications in Digital Mammograms Using Wavelets*, IEEE Transactions on medical imaging, VOL. 17, No.4 (1998).
- [13] <http://www.cancer.org/Cancer/BreastCancer/>, American Cancer Society
- [14] <http://marathon.csee.usf.edu/Mammography/Database.html>, University of South Florida Digital Mammography Home Page