Mammogram Diagnostics via 2-D Complex Wavelet-based Self-similarity Measures

Seonghye Jeon

Depertment of Industrial and Systems Engineeroing, Georgia Institute of Technology, US

Orietta Nicolis

Instituto de Estadistica, Universidad de Valparaiso, Chile

Brani Vidakovic

Depertment of Industrial and Systems Engineeroing, Georgia Institute of Technology, US

Abstract. Breast cancer is the second leading cause of death in women in the United States. Mammography is currently the most effective method for detecting breast cancer early; however, radiological interpretation of mammogram images is a challenging task. Many medical images demonstrate a certain degree of self-similarity over a range of scales. This scaling can help us to describe and classify mammograms.

In this work, we generalize the scale-mixing wavelet spectra to the complex wavelet domain. In this domain, we estimate Hurst parameter and image phase and use them as discriminatory descriptors to classify mammographic images to benign and malignant. The proposed methodology is tested on a set of images from the University of South Florida Digital Database for Screening Mammography (DDSM).

Keywords: Scaling; Complex Wavelets; Self-similarity; 2-D Wavelet Scale-Mixing Spectra

1. Introduction

The National Cancer Institute estimates that 1 in 8 women born today will be diagnosed with breast cancer during her lifetime (Altekruse *et al.*, 2010). Breast cancer is one of the most common forms of cancer among women in the United States, second only to non-melanoma skin cancer. A national objective has been set by the U.S Department of Health and Human Services to reduce the female breast cancer death rate from 22.9

per 100,000 females in 2007 down to 20.6 by the year 2020 – a 10% improvement (Healthy People 2020, U.S. Department of Health and Human Services). One of the most important efforts toward that goal is advancing the precision of screening technologies. Early detection is the best method for improving prognosis and also leads to less invasive options for both specific diagnosis and treatment.

Mammography is currently the most efficient and prevalent method for detecting a breast cancer early, before the disease becomes symptomatic. However, the radiological interpretation of mammogram images is a difficult task since the appearance of even normal tissue is highly variable and complex, and signs of early disease are often minute or indistinct. Reading a mammogram image is a skill that physicians develop over time, and confidently stating whether findings are cancerous or not is quite difficult and often subjective. Suspicious findings are commonly clarified by follow-up images, ultrasound, or MRI. On the other hand, it has been estimated that $10 - 30\%$ of cancers which could have been detected are missed (Oestreicher *et al.*, 2005). Thus, improving both the specificity and the sensitivity of mammographic diagnoses is an important goal in improving prognoses while also reducing the number of unnecessary procedures or surgical operations.

In high frequency and irregular data collected in real-life settings (both naturally occurring and human-made), a commonly occurring phenomenon is that of regular scaling. Examples of this have been found in a variety of systems and processes including economics (stock market, exchange rate fluctuations), telecommunications (internet data), physics (hydrology, turbulence), geosciences (wind and rainfall patterns), and several applications in biology and medicine (DNA sequences, heart rate variability, auditory nerve-spike trains). The irregular behavior of these complex structures is difficult or impossible to quantify by standard modeling techniques; but when observations are inspected at different scales, there is in fact a regular relationship between the behavior at each scale. This phenomenon has been demonstrated in many medical images, leading to its diagnostic use as a tool capable of quantifying statistical similarity of data patterns at various scales.

The standard measure of regular scaling is the Hurst exponent. This measure can also be connected to the presence of long memory and fractality in signals and images and is viewed as an informative summary. Many techniques for estimating the Hurst exponent exist, and assessing the accuracy of these estimations can be complicated. Wavelet transforms are powerful tool in estimating the Hurst exponent and modeling statistical similarity at different scales. For example, Nicolis *et al.* (2011) proposed a

method based on the wavelet spectra for extracting the self-similarity measures in an isotropic and anisotropic spaces. Ramírez-Cobo *et al.* (2011) demonstrated a wavelet-based spectra method for estimating Hurst exponent in time-varying two-dimensional rainfall maps.

For an efficient representation of an image or signal, the wavelet basis is desirable to be orthogonal, symmetric and to have compact support. (Gao and Yan, 2010). An orthogonal basis has a variety of theoretical and practical advantages: it leads to more efficient algorithms, and establishing properties of a representation is often easier with orthogonal bases. Most importantly, orthogonal bases preserve the variances, a property critical for the coherent definition of wavelet spectra. Symmetry guarantees an orientation-free representation of features, preventing distortion in the transformed domains. Moreover, the computational cost of performing wavelet transforms depends heavily on the support size of a basis. Apart from the Haar wavelet, complex wavelets with an odd number of vanishing moments are only compactly supported wavelets which are symmetric (Lawton, 1993). Due to this advantages, complex wavelet has been used in various areas including motion estimation (Magarey and Kingsbury, 1998), texture image modeling (Portilla and Simoncelli, 2000), image denoising (Achim and Kuruoglu, 2005) and NMR spectra classification (Kim *et al.*, 2008). However, to the best of our knowledge, there is no literature that investigates self-similarity measures based on the complex wavelet spectra.

The novelty of this paper is to use the scale-mixing wavelet spectra based on *complex wavelet transforms* for estimating the Hurst exponent. We focus on the estimated Hurst exponent and show its ability to differentiate cancerous from normal tissue visible in the backgrounds of mammogram images, and compare this performance with its counterpart obtained from real-valued wavelet transform. Moreover, complex wavelet transforms produce an additional measure, the phase information. We also demonstrate the classification power of the phase information and use it as an additional modality in the discriminatory analysis.

A further novelty of our work is the use of the information contained in the background tissue of images. Most of the references found in literature dealing with breast cancer detection methods are based on microcalcifications (Wang and Karayiannis, 1998; Netsch and Peitgen, 1999; Kestener *et al.*, 2001; El-Naqa *et al.*, 2002). Only recently the information contained in the background is taken into consideration (Nicolis *et al.*, 2011; Hamilton *et al.*, 2011). This classifying measure based on the irregularity measure of background tissue promises to be a new tool to be used in combination with existing clinical diagnostic tools, thus improving discriminatory power of non-invasive diagnostic techniques.

The paper is organized as follows. In Section 2, we briefly describe the data set used in the analysis. In Section 3, the complex wavelet based scale-mixing wavelet spectra is proposed as new tools for estimating the self-similarity indices. Also the importance of the phase information is discussed. Section 4 deals with the classification of mammogram images using the wavelet descriptors obtained from Section 3 as classifiers. Finally Section 5 contains discussion and some concluding remarks.

2. The Dataset

The collection of digitized mammograms we analyzed was obtained from the University of South Florida's Digital Database for Screening Mammography (DDSM)

        http://marathon.csee.usf.edu/Mammography/Database.html.

The DDSM is described in detail in Heath *et al.* (2000). Images from this database containing suspicious areas are accompanied by pixel-level "ground truth" information relating locations of suspicious regions to what was assessed and verified through biopsy. We selected 45 normal (benign) cases and 79 cancer (malignant) cases scanned on the HOWTEK scanner at the full 43.5 micron per pixel spatial resolution. Each case contains four mammograms (two for each breast: the craniocaudal (CC) and mediolateral oblique (MLO) projections) from a screening exam. We considered only the CC projections, using either side of the breast image. Five sub-images of size 1024×1024 was taken from each case for analysis. An example of mammogram image and the location of sub-images is provided in Fig. 1. For the classification purpose, we only used one of the five sub-images from each mammogram.

3. Discrete complex wavelets

This section discusses discrete complex wavelet transforms. Unlike the popular method used in this context (Selesnick *et al.*, 2005), the proposed method is unitary (complex orthogonal) and minimal.

3.1. **Complex wavelet basis.** The construction of the complex wavelet basis associated with multiresolution analysis follows the usual approach proposed by Mallat (1998) and Daubechies (1992). Details on the construction and properties of complex wavelets can be found in Lawton (1993), Lina and Mayrand (1995), Strang and Nguyen (1996), and Zhang *et al.*
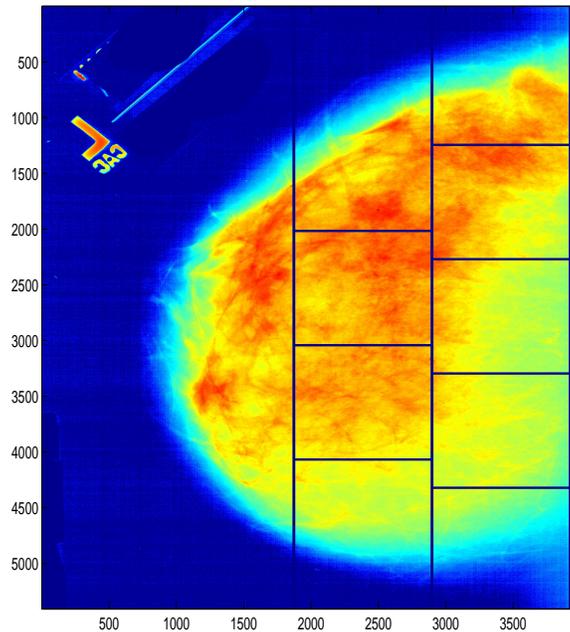
FIGURE 1. An example of mammogram image with five sub-images of size 1024×1024 considered for the analysis.

(1999). In analogy to the real case, the wavelet function $\psi(x)$ for the complex wavelet is given by

$$\psi(x) = \frac{1}{\sqrt{2}} \sum_k (-1)^k 2h_{1-k}^* \phi(2x - k) \tag{1}$$

where ϕ is the scaling function, h is the low pass filter and the * indicates the complex conjugate. The representation of wavelets in 2D can be done through the tensor product of univariate scaling functions and wavelets as follows:

$$\begin{aligned}
\phi(x,y) &= \phi(x) \cdot \phi(y) \\
\psi_h(x,y) &= \phi(x) \cdot \psi(y) \\
\psi_v(x,y) &= \psi(x) \cdot \phi(y) \\
\psi_d(x,y) &= \psi(x) \cdot \psi(y)
\end{aligned} \tag{2}$$

where symbols h, v, d in (2) stand for horizontal, vertical and diagonal directions, respectively. The atoms capture image features in the corresponding directions.

3.2. **The complex scale-mixing 2-D wavelet transform.** The discrete complex wavelet transform (DCWT) can be considered as a complex-valued extension of the standard discrete wavelet transform (DWT). It uses complex-valued filtering (analytic filter) for transforming the real/complex signals. Complex wavelet coefficients can be computed by Mallat's algorithm (Mallat, 1998; Morettin, 1997)

$$c_{j-1,l} = \sum_k h^*_{k-2l} c_{j,k} \tag{3}$$

and

$$d_{j-1,l} = \sum_k g^*_{k-2l} c_{j,k} \tag{4}$$

where h is as in (1) and g is the quadrature mirror filter. The $*$ denotes the complex conjugate.

Conversely, the reconstruction is given by

$$c_{j,k} = \sum_l c_{j-1,l} h_{k-2l} + \sum_l d_{j-1,l} g_{k-2l}. \tag{5}$$

Moreover, the real and imaginary coefficients are used to compute the modulus and phase information. The wavelet coefficients can be written as

$$d_{j,k} = Re(d_{j,k}) + i \cdot Im(d_{j,k})$$

with magnitude

$$|d_{j,k}| = \sqrt{Re(d_{j,k})^2 + Im(d_{j,k})^2}$$

and phase

$$\angle d_{j,k} = \arctan\left(\frac{Im(d_{j,k})}{Re(d_{j,k})}\right)$$

when $|Re(d_{j,k})| > 0$.

There are many versions of the 2-D wavelet transforms which lead to different tessellations, or tilings (Ramírez-Cobo *et al.*, 2011). Here we define the complex wavelet atoms as follows

$$\phi_{(j_1,j_2),\mathbf{k}}(\mathbf{x}) = 2^{(j_1+j_2)/2} \phi(2^{j_1}x - k_1, 2^{j_2}y - k_2) \tag{6}$$

$$\psi_{\delta,(j_1,j_2),\mathbf{k}}(\mathbf{x}) = 2^{(j_1+j_2)/2} \psi_\delta(2^{j_1}x - k_1, 2^{j_2}y - k_2), \tag{7}$$

where δ is one of directions h, v, or d, and $(j_1, j_2) \in \mathbb{Z}^2$. Then, any function $f \in \mathcal{L}_2(\mathbb{R}^2)$ can be represented as

$$
\begin{aligned}
f(\mathbf{x}) \;=\; & \sum_{\mathbf{k}} c_{(J_0,J_0),\mathbf{k}} \; \phi_{(J_0,J_0),\mathbf{k}}(\mathbf{x}) \\
& + \sum_{j>J_0} \sum_{\mathbf{k}} d_{(J_0,j),\mathbf{k}} \; \psi_{h,(J_0,j),\mathbf{k}}(\mathbf{x}) \\
& + \sum_{j>J_0} \sum_{\mathbf{k}} d_{(j,J_0),\mathbf{k}} \; \psi_{v,(j,J_0),\mathbf{k}}(\mathbf{x}) \\
& + \sum_{j_1,j_2>J_0} \sum_{\mathbf{k}} d_{(j_1,j_2),\mathbf{k}} \; \psi_{d,(j_1,j_2),\mathbf{k}}(\mathbf{x}),
\end{aligned}
$$

and a 2-D wavelet transform, which we call the *scale-mixing wavelet transform* is obtained. The motivation for the name *scale-mixing* is obvious. Unlike the atoms in the traditional 2-D wavelet transform, the scale-mixing atoms mix the scale indices thus capturing the "energy flux" between the scales.

The scale-mixing detail coefficients are defined as

$$
\begin{aligned}
d_{(J_0,j),\mathbf{k}} \;&=\; 2^{(J_0+j)/2} \int f(\mathbf{x}) \, \psi_h^*(2^{J_0}x - k_1, 2^j y - k_2) \; dx \, dy, \\
d_{(j,J_0),\mathbf{k}} \;&=\; 2^{(j+J_0)/2} \int f(\mathbf{x}) \, \psi_v^*(2^j x - k_1, 2^{J_0} y - k_2) \; dx \, dy, \\
d_{(j_1,j_2),\mathbf{k}} \;&=\; 2^{(j_1+j_2)/2} \int f(\mathbf{x}) \, \psi_d^*(2^{j_1}x - k_1, 2^{j_2} y - k_2) \; dx \, dy, \quad (8)
\end{aligned}
$$

where ψ^* is a complex conjugate of ψ. Note that (j_1, j_2) in (6) and (7) can be indexed as $(j_1, j_1 + s)$ with $s \in \mathbb{Z}$.

Similar to the traditional one- and two-dimensional cases, the complex scale-mixing detail coefficients are linked to the original image (2-D signal) through a matrix equation. Suppose that a $2^n \times 2^n$ image (matrix) A is to be transformed into the wavelet domain. The complex wavelet matrix W is first composed by the complex scaling and wavelet filter coefficients h_k and g_k as in Vidakovic (1999). Note that the wavelet filter is given by $g_k = (-1)^k h_{1+N-k}^*$ and N is a shift parameter which affects the location of the wavelet. Then the rows of A are transformed by a one-dimensional transform given by the wavelet matrix W, resulting in $WA^\dagger$. The same is repeated on the rows of $WA^\dagger$. The result is

$$
B = W(WA^\dagger)^\dagger = WAW^\dagger, \tag{9}
$$

the scale-mixing wavelet transform of matrix A, which will be the basis for defining the scale-mixing spectra. Here $C^\dagger$ denotes Hermitian transpose

(a)   (b)
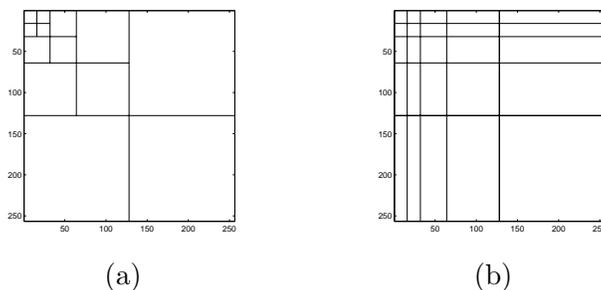
FIGURE 2. Tessellations for 2-D wavelet transforms. (a) Traditional 2-D transform of depth 4; (b) Scale-mixing wavelet transform of depth 4.

of a complex matrix C. The representation in (9) is a finite-dimensional implementation of (8) for signal $f(\mathbf{x})$ sampled in a matrix form.

The tessellation induced by transform in (9) is shown in Figure 2 (b). A more general transform can be obtained as an iterative repetition of the transform in (9) with depth k, applied only on the "smooth part" of the previous iterative step. This general approach unifies the traditional and scale-mixing 2-D wavelet transforms.

The scale-mixing 2-D transform is operationally appealing. Constructing appropriate W is computationally fast and, since W is unitary, the inverse transform is straightforward:

$$A = W^\dagger BW.$$

By inspecting the tessellation in Figure 2, several hierarchies of detail spaces can be identified. The diagonal hierarchy interfaces coefficients with the same component scales and coincides with the diagonal hierarchy in the traditional 2-D spectra. One level above and below the diagonal hierarchy are hierarchies of detail spaces that interface the scales that differ by 1. For the hierarchy above the diagonal, the scales along x-direction are interfaced by the next coarser scale along y-direction. For the hierarchy below the diagonal, the roles of x and y are interchanged.

The unitarity of W implies

$$\text{trace}(AA^\dagger) = \text{trace}(BB^\dagger)$$

for $B = WAW^\dagger$, implying the total energy in the image A

$$E = \text{trace}(AA^\dagger)$$

is preserved.
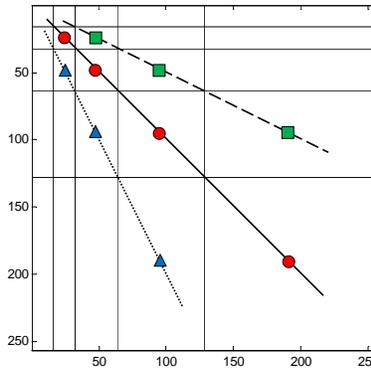
FIGURE 3. (a) Three detail-space hierarchies generating the scale-mixing 2-D transform, where (j_1, j_2) is indexed as $(j, j+s)$, $s \in \mathbb{Z}$. Circles correspond to $s = 0$, triangles to $s = 1$, and squares to $s = -1$. The scales (j_0, j), $j_0 = 7$ (squares), and (j, j_0), $j_0 = 6$ (triangles) are shown in the figure.

3.3. **The complex scale-mixing wavelet spectra.** The scale-mixing spectra is defined in terms of the complex scale-mixing coefficients in (8)

$$S(j) = \log_2 \mathbb{E} \left(|d_{(j,j+s),\mathbf{k}}|^2 \right), \tag{10}$$

where $j, s \in \mathbb{Z}$ are fixed. Note that $s = 0$ in (10) corresponds to the diagonal 2-D spectra.

To calibrate the scale-mixing spectra, consider now a 2-D fractional Brownian motion, $B_H(\mathbf{u})$. For such a process, the scale-mixing detail coefficients are given by

$$d_{(j,j+s);\mathbf{k}} = 2^{j+\frac{s}{2}} \int B_H(\mathbf{u})\psi^* \left(2^j u_1 - k_1, 2^{j+s} u_2 - k_2 \right) d\mathbf{u}. \tag{11}$$

where ψ^* denotes the complex conjugate of ψ^d, the wavelet atom in the diagonal direction defined in (7). These coefficients are random variables with zero mean and variance (Heneghan *et al.*, 1996), which leads to

$$\mathbb{E} \left[|d_{(j,j+s);\mathbf{k}}|^2 \right] = 2^{2j+s} \int \psi \left(2^j u_1 - k_1, 2^{j+s} u_2 - k_2 \right)$$

$$\times \psi^* \left(2^j v_1 - k_1, 2^{j+s} v_2 - k_2 \right) \mathbb{E} \left[B_H(\mathbf{u}) B_H(\mathbf{v}) \right] d\mathbf{u} \, d\mathbf{v}. \tag{12}$$

As in Veitch and Abry (1999), it is assumed here that the coefficients within and across the scales are uncorrelated.

From (12), it can be shown that

$$\mathbb{E}\left[|d_{(j,j+s);\mathbf{k}}|^2\right] = 2^{-j(2H+2)} \, V_{\psi,s}(H), \tag{13}$$

where $V_{\psi,s}(H)$ is an expression depending on ψ, H and s, but not on the scale j,

$$V_{\psi,s}(H) = -\frac{\sigma_H^2}{2} \int \int \psi(\mathbf{p}+\mathbf{q}) \, \psi^*(\mathbf{q}) \, |\mathbf{p}_s|^{2H} 2^{-s} \, d\mathbf{p} \, d\mathbf{q}. \tag{14}$$

A proof of (13) is provided in the Appendix. By taking logarithms in (13)

$$\log_2 \mathbb{E}\left[|d_{(j,j+s);\mathbf{k}}|^2\right] = -(2H+2)j + \log_2 V_{\psi,s}(H) \tag{15}$$

for $j \in \mathbb{Z}$, the Hurst exponent can be estimated from the slope of the linear equation (15). Finally, the empirical counterpart of (15) is a regression defined on pairs

$$\left(j, \log_2 \frac{1}{n} \sum_{j,j+s} \left| d_{(j,j+s),\mathbf{k}} \right|^2 \right), \quad j, s \in \mathbb{Z}. \tag{16}$$

The slope of the regression would estimate the Hurst exponent, i.e., $H = -(slope + 2)/2$. Instead of the sample mean in (16), different location measures could be used, such as the median.

Although (16) is traditionally fitted by the ordinary least squares (OLS) regression, the level-wise variance of wavelet coefficients is not constant. Therefore one can improve the estimator by using more robust approach that accounts for the heteroscedasticity. Veitch and Abry (1999) used weighted linear regression to improve the estimator. This method weights each level by the inverse of the variance of that level. Hamilton *et al.* (2011) proposed estimation methods that are based on the Theil regression, that is, a weighted average of all pairwise slopes s_{ij} between levels i and j. Given a weight w_{ij}, the estimator of the overall slope in (16) is then $\sum_{i,j} w_{ij} s_{ij} / \sum_{i,j} w_{ij}$. Different types of weights are proposed, from which we obtain more robust estimation methods. For example, one can use harmonic average weighted slope which is theoretically optimal, or level-enhanced weights that favor slopes from finer levels. In this paper we adopted robust estimators proposed in the literature along with the OLS regression based estimator to perform more comprehensive comparison.

3.4. **The complex phase information.** It is known that phase and spectrum are intertwined in a nontrivial way to describe the data. While phases encode most of the coherent (in space and scale) structure of the image, the spectrum mostly encode the strength of local information that could be corrupted with noise (Clonda *et al.*, 2004). For this reason phase information have been used in edge detection and in the reconstruction of images.

A classical illustration is given in Oppenheim and Lim (1981) where the image reconstruction is more driven by the phase of the Fourier transform rather than by the magnitude.

To elucidate this point we conducted a simple experiment, illustrated in Figure 4. Noiseless images of Lenna and Barbara are transformed to the complex wavelet domain using Daubechies 6-tap complex wavelet filter. Each image resulted in magnitude and phase matrices of the same sizes as the original image. Next, we switched the phase matrices, and transformed the objects back to the original domain. This resulted in two surrogate images, the first from Lenna modulus and Barbara phase, and the second from Barbara modulus and Lenna phase. Although the modulus carries information about the "energy" of the image, the phase information dominates the reconstruction, as evident from the second row of images in Figure 4.
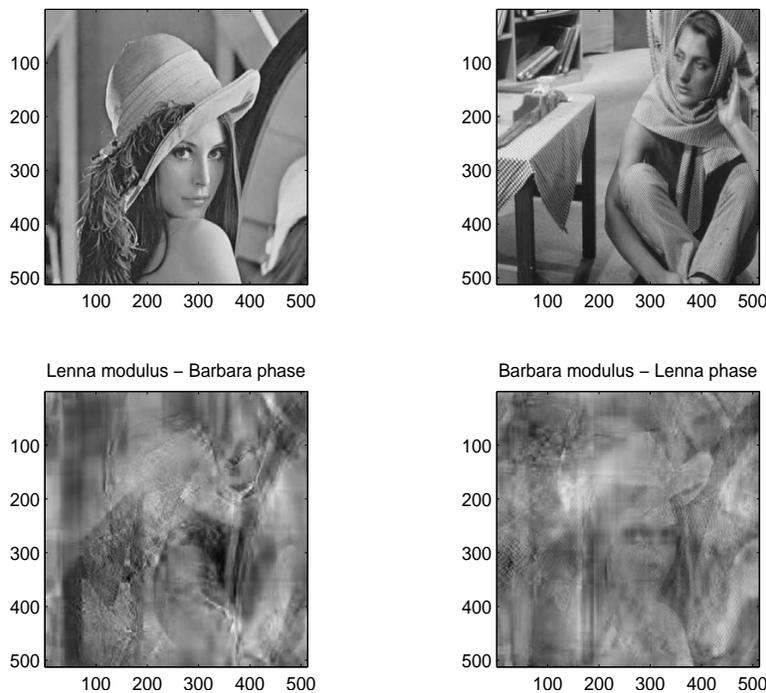
FIGURE 4. Images of Lenna and Barbara (first row) are decomposed using Daubechies complex 6 tap filter. In the wavelet domain the phase information is switched and images transformed back. Note that phase information dominates in the back-transformed images.

Recently, many research studies have focused on using phase information from the complex wavelet transforms (Anderson *et al.*, 2005; Hua and Orchard, 2008; Miller and Kingsbury, 2008; Rakvongthai and Oraintara, 2008; Remenyi *et al.*, 2014). In the wavelet domain the phase of a coefficient near an isolated feature varies linearly with its distance from the feature. Despite of the numerous literatures focused on the usage of phase in detecting edges and the feature orientations, the discriminatory power of phase in the complex wavelet domain has not yet been systematically studied.

In contrast to magnitudes of complex wavelet coefficients, their phases do not exhibit any scaling property. However, some summary statistics of the phases turn out to be discriminatory. In the following section, we demonstrate how the phase information could be used as a classification modality.

4. Mammogram Classification

In this section we illustrate how the complex wavelet-based spectra and the phase information can be used to classify digitized mammograms. We demonstrate that the spectra slope and phase, as descriptors of digitized images, have good discriminatory power. It is straightforward to implement the described analysis in various scientific areas in which 2-D data are instrumental, such as geoscience or industrial applications.

For every sub-image of size 1024×1024, we performed discrete real-valued wavelet transform (DWT) and discrete complex wavelet transform (DCWT) using Daubechies 6 tap filter. After each transform, we estimated the slope of wavelet spectra using traditional ordinary least squares regression (OLS) along with three robust estimation methods described in Section 3.3. The robust estimation approaches include Abry-Veitch weighted regression (AV), level enhanced OLS (EOLS) and harmonic average weighted slopes (HA). For more details on these robust estimators, we refer the readers to Veitch and Abry (1999); Hamilton *et al.* (2011). Note that the wavelet spectra slope is used as a predictor instead of the Hurst exponent. It is because the estimated Hurst exponent H is empirical, and the slowly decaying spectra (with slope > -2) caused by the presence of colored noise, could render H to be negative.

For each classification method, we randomly selected 67% of the data as a training set to fit the classifier and used the remaining 33% of the data to test performance. The random selection of training and testing sets was repeated 10,000 times, so the reported prediction measures are averaged over 10,000 runs. Performance was compared in terms of sensitivity, specificity, and overall correct classification rate.

The most parsimonious classification approach would be logistic regression involving only the wavelet spectra slope as a predictor. The result is summarized in Table 1. Regardless of estimation methods, the performance is comparable to the status quo with rate ranging from 0.62 to 0.68. And this diagnostic classification does not require any skills or training and it only uses background regularity of the image summarized by wavelet spectra slope.

TABLE 1. Logistic classification based on the wavelet spectra slope; four different estimation methods were compared, each with real-valued (DWT) and complex (DCWT) wavelet transform.

	Method	Sensitivity	Specificity	Correct Classification
OLS	DWT	0.93	0.10	0.63
	DCWT	0.94	0.07	0.62
AV	DWT	0.86	0.37	0.68
	DCWT	0.85	0.35	0.66
EOLS	DWT	0.93	0.13	0.63
	DCWT	0.93	0.10	0.63
HA	DWT	0.87	0.30	0.66
	DCWT	0.87	0.29	0.65

One of the interesting findings is that the phase contains information to classify normal images from malignant. Since the features and directions of background tissue is best preserved in the level of finest detail, we focus on the phase information of finest detail. Figure 4 shows how the estimated density of the phase average and coefficient of variation (CV) at the finest level from normal images differ from cancer images. In this case, we hypothesize control mammograms have no clusters of consistent features and edges in the detailed wavelet space.

To validate the discriminatory power of the phase information, we assessed the logistic models by using each of the three summary statistics (average, variance, coefficient of variation) of the phases at the finest level. As Table 2 shows, any single predictor classifies malignant and normal images with correct classification rate of 61-65%.

We then conducted classification analysis based on the complex wavelet spectra slope and the phase average. Figure 6 shows a scatter plot of cases by complex spectra slope versus phase average, illustrating the differentiation between benign and malignant cases.
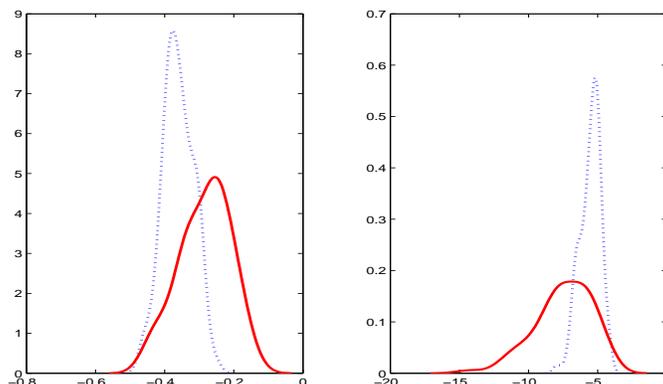
FIGURE 5. *Left panel*: Estimated density of phase average at the finest level. *Right panel*: Estimated density of phase coefficient of variation at the finest level. The solid line corresponds to malignant cases and the dotted line to normal cases.

TABLE 2. Logistic classification based on summary statistics of phase information at the finest level.

	Method	Sensitivity	Specificity	Correct Classification
Phase	Average	0.82	0.35	0.65
	Variance	0.87	0.15	0.61
	CV	0.78	0.43	0.64

We combined complex wavelet spectra obtained from four different estimation methods with the phase average. For each of the four pairs, we performed logistic, linear and quadratic classification. By combining wavelet spectra slope with phase average, we obtain classification rate of up to 0.68 with stable sensitivity and specificity rate. Several wavelet bases and level combinations have been compared but the results and the conclusion remain the same.
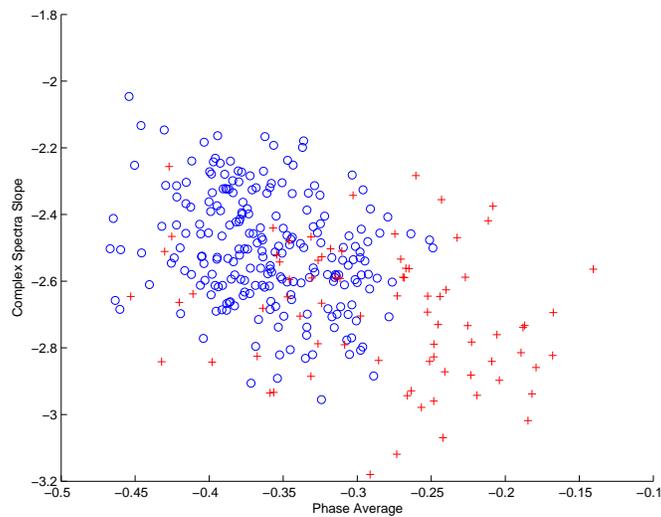
FIGURE 6. Scatter plot of Complex spectra slope (obtained by AV estimation method) versus Phase average. The symbols denote: *circles* for normal mammographies, *crosses* for malignant mammographies.

5. **Conclusions**

In this paper we proposed a complex scale-mixing 2-D wavelet transform in the context of assessing regularities of 2-D objects. The proposed transform is implemented by matrix operations, and it guarantees unitarity, low computational cost, directional insights, interplay between the scales, and a straightforward inverse transform. We then explored the spectra and self-similarity measures based on the proposed complex wavelet transform and demonstrated its utility in the context of mammogram image classification. The procedure is based on assessing the regularity of background tissues of images rather than mammogram features such as microcalcification and tumor masses. This approach is an unused diagnostic modality in the field.

To obtain the estimates of Hurst exponent, we used three robust estimation methods (AV, EOLS and HA) along with the ordinary least squares estimator. We also studied phase statistics at the finest level as a classifier. From the logistic classification model, we found that robust Hurst exponent estimates and the phase average have enough power to differentiate benign from malignant cases, with correct classification rate of 62-68%. We also showed that by combining spectra slope estimates with phase average,

TABLE 3. Logistic, Linear and Quadratic classification based on the pair of complex wavelet spectra and the phase average.

	Method	Sensitivity	Specificity	Correct Classification
(OLS; phase average)	Logistic	0.83	0.35	0.65
	Linear	0.59	0.58	0.58
	Quadratic	0.58	0.73	0.62
(AV; phase average)	Logistic	0.82	0.43	0.67
	Linear	0.59	0.68	0.62
	Quadratic	0.65	0.76	0.68
(EOLS; phase average)	Logistic	0.83	0.36	0.65
	Linear	0.59	0.58	0.59
	Quadratic	0.59	0.73	0.63
(HA; phase average)	Logistic	0.83	0.40	0.67
	Linear	0.60	0.63	0.61
	Quadratic	0.64	0.73	0.66

we obtained high classification rate with stable sensitivity and specificity. Therefore, complex wavelet transform greatly appeals to investigators seeking more robust and reliable classification method for various application area.

It is well known that for the real wavelets there is no symmetric and compactly supported scaling function defining an orthogonal MRA. Complex wavelets assure symmetry, compact support and orthogonality (that is unitarity) of decomposing atoms which is desirable in image representations. Since complex wavelets provide the only compactly supported and symmetric basis which are unitary, it better represents the image than real-valued wavelet transform. By using proposed method, complex wavelet transforms can be readily performed and applied to various domains to summarize, classify or interpret highly complex data.

Appendix

Derivation of expression (13). The scale-mixing detail coefficients of a 2-D fBm (11) are

$$d_{(j,j+s);\mathbf{k}} = 2^{j+\frac{s}{2}} \int B_H(\mathbf{u})\psi\left(2^j u_1 - k_1, 2^{j+s} u_2 - k_2\right) d\mathbf{u}.$$

These coefficients are random variables with zero mean and variance (Heneghan *et al.*, 1996)

$$\mathbb{E}\left[|d_{(j,j+s);\mathbf{k}}|^2\right] = 2^{2j+s} \int \psi\left(2^j u_1 - k_1, 2^{j+s} u_2 - k_2\right)$$

$$\times \psi^*\left(2^j v_1 - k_1, 2^{j+s} v_2 - k_2\right) \mathbb{E}\left[B_H(\mathbf{u})B_H(\mathbf{v})\right] d\mathbf{u}\, d\mathbf{v}. \qquad (17)$$

Since

$$\mathbb{E}\left[B_H(\mathbf{u})B_H(\mathbf{v})\right] = \frac{\sigma_H^2}{2}\left(|\mathbf{u}|^{2H} + |\mathbf{v}|^{2H} - |\mathbf{u}-\mathbf{v}|^{2H}\right),$$

and

$$\int \psi\left(2^j u_1 - k_1, 2^{j+s} u_2 - k_2\right) d\mathbf{u} = \int \psi\left(2^j v_1 - k_1, 2^{j+s} v_2 - k_2\right) d\mathbf{v} = 0,$$

it can be easily seen that (17) becomes

$$\mathbb{E}\left[|d_{(j,j+s);\mathbf{k}}|^2\right] = -\frac{\sigma_H^2}{2}\, 2^{2j+s} \int\int \psi\left(2^j u_1 - k_1, 2^{j+s} u_2 - k_2\right)$$

$$\times \psi^*\left(2^j v_1 - k_1, 2^{j+s} v_2 - k_2\right) |\mathbf{u}-\mathbf{v}|^{2H} d\mathbf{u}\, d\mathbf{v}.$$

Next, define substitutions

$$\mathbf{p} = (p_1, p_2) = \left(2^j(u_1 - v_1), 2^{j+s}(u_2 - v_2)\right),$$

Then, if $\mathbf{p}_s \equiv (p_1, 2^{-s}p_2)$,

$$\mathbb{E}\left[|d_{(j,j+s);\mathbf{k}}|^2\right] = -\frac{\sigma_H^2}{2}\, 2^{2j+s} \int\int \psi(\mathbf{p}+\mathbf{q})$$

$$\psi^*(\mathbf{q})\, 2^{-2jH}\, |\mathbf{p}_s|^{2H}\, 2^{-4j-2s}\, d\mathbf{p}\, d\mathbf{q}$$

$$= -\frac{\sigma_H^2}{2}\, 2^{-j(2H+2)} \int\int \psi(\mathbf{p}+\mathbf{q})\, \psi^*(\mathbf{q})\, |\mathbf{p}_s|^{2H}2^{-s}\, d\mathbf{p}\, d\mathbf{q}$$

$$= 2^{-j(2H+2)}\, V_{\psi,s}(H),$$

where

$$V_{\psi,s}(H) = -\frac{\sigma_H^2}{2}\, \int\int \psi(\mathbf{p}+\mathbf{q})\, \psi^*(\mathbf{q})\, |\mathbf{p}_s|^{2H}2^{-s}\, d\mathbf{p}\, d\mathbf{q}, \qquad (18)$$

is an integral depending on ψ, H, and s, but not on the scale j.

References

Achim, A. and Kuruoglu, E. (2005). Image denoising using bivariate α-stable distributions in the complex wavelet domain. *IEEE Signal Processing Letters*, **12**, 17–20.

Altekruse, S., Kosary, C., Krapcho, M., Neyman, N., Aminou, R., and Waldron, W. (2010). Seer cancer statistics review: 1975-2007.

Anderson, R., Kingsbury, N., and Fauqueur, J. (2005). Determining multiscale image feature angles from complex wavelet phases. *International conference on Image processing (ICIP), Toronto, Canada, September 28-30.*

Clonda, D., Lina, J.-M., and Goulard, B. (2004). Complex daubechies wavelets: properties and statistical image modelling. *Signal Processing*, **84**, 1–23.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. Number 61 in CBMS-NSF Series in Applied Mathematics. SIAM, Philadelphia.

El-Naqa, I., Yang, Y., Wernick, M., Galatsanos, N., and Nishikawa, R. (2002). A support vector machine approach for detection of microcalcifications. *IEEE Transactions on medical imaging*, **21**(12), 1552–1563.

Gao, R. and Yan, R. (2010). *Wavelets: Theory and Applications for Manufacturing*. Springer.

Hamilton, E. K., Jeon, S., Cobo, P. R., Lee, K. S., and Vidakovic, B. (2011). Diagnostic classification of digital mammograms by wavelet-based spectral tools: A comparative study. *The Proceedings of the 2011 IEEE International Conference on Bioinformatics and Biomedicine*, pages 384–389.

Heath, M., Bowyer, K., Kopans, D., Moore, R., and Kegelmeyer, P. (2000). The digital database for screening mammography. *The Proceedings of the 5th International Workshop on Digital Mammography, Toronto, Canada, Medical Physics Publishing (Madison, WI).*

Heneghan, C., Lown, S., and Teich, M. (1996). Two dimensional fractional Brownian motion: Wavelet analysis and synthesis. *Image analysis and interpretation, proceedings of the IEEE Southwest Symposium*, pages 213–217.

Hua, G. and Orchard, M. T. (2008). Image reconstruction from the phase or magnitude of its complex wavelet transform. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3261–3264.

Kestener, P., Lina, J., Saint-Jean, P., and Arneodo, A. (2001). Wavelet-based multifractal formalism to assist in diagnosis in digitized mammograms. *Image analysis and stereology*, **20**(3), 169–175.

Kim, S., Wang, Z., Oraintarac, S., Temiyasathita, C., and Wongsawatc, Y. (2008). Feature selection and classification of high-resolution nmr spectra in the complex wavelet transform domain. *Chemometrics and Intelligent Laboratory Systems*, **90**, 161–168.

Lawton, W. (1993). Applications of complex valued wavelet transforms to subband decomposition. *IEEE Transactions on Signal Processing*, **41**, 3566–3568.

Lina, J.-M. and Mayrand, M. (1995). Complex daubechies wavelets. *Applied and Computational Harmonic Analysis*, **2**, 219–229.

Magarey, J. and Kingsbury, N. (1998). Motion estimation using a complex-valued wavelet transform. *IEEE Transactions on Signal Processing*, **46**, 1069–1084.

Mallat, S. G. (1998). *A Wavelet Tour of Signal Processing*. Academic Press, San Diego.

Miller, M. and Kingsbury, N. (2008). Image modeling using interscale phase properties of complex wavelet coefficients. *IEEE Transactions on Image Processing*, **17**(9), 1491–1499.

Morettin, P. A. (1997). *Wavelets in Statistics*. 3rd International Conference on Statistical Data Analysis Based on L_1 Norm and Related Methods. Neuchâtel, Switzerland.

Netsch, T. and Peitgen, H. (1999). Scale-space signatures for the detection of clustered microcalcifications in digital mammograms. *IEEE Transactions on medical imaging*, **18**(9), 774–786.

Nicolis, O., Ramírez-Cobo, P., and Vidakovic, B. (2011). 2-D wavelet-based spectra with applications. *Computational Statistics and Data Analysis*, **55**(1), 738–751.

Oestreicher, N., Lehman, C., Seger, D., Buist, D., and White, E. (2005). The incremental contribution of clinical breast examination to invasive cancer detection in a mammography screening program. *American Journal of Roentgenology*, **184**(2), 428–432.

Oppenheim, A. and Lim, J. (1981). The importance of phase in signals. *The Proceedings of the IEEE*, **69**, 529–541.

Portilla, J. and Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International journal of computer vision*, **40**, 49–70.

Rakvongthai, Y. and Oraintara, S. (2008). On phase statistics of complex wavelet coefficients at edges. *European Conference on Signal Processing*, pages 2–6.

Ramírez-Cobo, P., Lee, K. S., Molini, A., Porporato, A., Katul, G., and Vidakovic, B. (2011). A wavelet-based spectral method for extracting self-similarity measures in time-varying two-dimensional rainfall maps. *Journal of Time Series Analysis*, **32**(4), 351–363.

Remenyi, N., Nicolis, O., Nason, G., and Vidakovic, B. (2014). Image denoising with 2-d scale-mixing complex wavelet transforms. *To Appear in IEEE Transactions on Signal Processing*.

Selesnick, I. W., Baraniuk, R. G., and Kingsbury, N. G. (2005). The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine*,

22(6).

Strang, G. and Nguyen, T. (1996). *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Wellesley, MA.

Veitch, D. and Abry, P. (1999). A wavelet-based joint estimator of the parameters of long-range dependence. *IEEE Transactions on Information Theory*, **45**, 878–897.

Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. John Wiley & Sons, Inc.

Wang, T. and Karayiannis, N. (1998). Detection of microcalcifications in digital mammograms using wavelets. *IEEE Transactions on medical imaging*, **17**(4), 498–509.

Zhang, X.-P., Desai, M., and Peng, Y. (1999). Orthogonal complex filter banks and wavelets: Some properties and design. *IEEE Transactions on Signal Processing*, **47**(4), 1039–1048.