RESEARCH　　　　　　　　　　　　　　　　　　　　　　**Open Access**

Magnitude-phase of the dual-tree quaternionic wavelet transform for multispectral satellite image denoising

Mohammed Kadiri[1,2*], Mohamed Djebbouri[3] and Philippe Carré[4]

Abstract

In this paper, we study the potential of the quaternionic wavelet transform for the analysis and processing of multispectral images with strong structural information. This new representation gives a very good division of the coefficients in terms of magnitude and three-phase angles and generalizes better the concept of analytic signal to image. Furthermore, it retains the property of shift invariant and directivity. We show an application of this transform in satellite image denoising. The proposed approach relies on the adaptation of thresholding procedures based on the dependency between magnitude quaternionic coefficients in local neighborhoods and phase regularization. In addition a non-marginal aspect of multispectral representation is introduced. Thanks to coherent analysis provided by the quaternionic wavelet transformation, the results obtained indicate the potential of this multispectral representation with magnitude thresholding and phase smoothing in noise reduction and edge preservation compared with classical wavelet thresholding methods that do not use phase or multiband information.

Keywords: Multispectral satellite image; Quaternionic wavelet analysis; Magnitude thresholding; Phase regularization; Structural similarity measure

1 Introduction

Wavelet transform have shown great success in diverse fields such as pattern recognition, image denoising, image compression, and computer graphics. The wavelet methods tend to give a good compromise for images containing such a mixture of discontinuities and texture. Previously, most researchers used the discrete wavelet transform (DWT) for image processing [1,2]. However, in many applications, it reaches its limitations, such as oscillations of coefficients at a singularity, lack of directional selectivity in higher dimensions, aliasing, and consequent shift variance. To overcome these problems, Bamberger and Smith [3] had proposed an effective filter bank for the directional decomposition of images. This filter has the important property that it can be critically sampled while achieving perfect reconstruction. Later, the undecimated

wavelet transform [4] was used in noise reduction and provides a shift invariant transformation, but at the cost of high redundancy.

More recently, the complex discrete wavelet transform (CDWT)[a] and the new quaternionic wavelet transform (QWT) employ analytic filters and propose magnitude-phase representations, shift invariance, and no aliasing. Several authors have studied the CDWT and its application to image denoising. Kingsbury [5-7] introduced a very elegant computational structure, the dual-tree complex wavelet transform (DT-CDWT), and incorporates it into the image restoration and enhancement. The DT-CDWT overcomes two drawbacks of the DWT. First is that the real and imaginary parts of CDWT associated with the pair of the Hilbert transforms are in quadrature; their magnitudes are almost shift invariant and redundancy is limited (factor 2 to compare with the undecimated wavelet transform ratio). Second, the complex phase encodes the signal location. However, in 2D, the complex representation by dual tree is not a satisfactory generalization of the analytic wavelet [8]. It has poor directional selectivity: its single phase can lead to ambiguity

* Correspondence: kadiri_univ_b@yahoo.fr
[1]Laboratory of Telecommunications and Digital Signal Processing, Department of Electronics, Faculty of Technology, Djillali Liabes University, Sidi Bel Abbes 22000, Algeria
[2]Department of Material Science, Faculty of Sciences and Technology, University of Mascara, Mascara 29000, Algeria
Full list of author information is available at the end of the article

when translating the image in two directions. Recently, the concept of generalizing complex wavelets to quaternion algebra has gained a lot of popularity [8-11]. The quaternionic wavelet transform has solved the problem of 2D localization. The phase of the QWT is represented by three angles: the first two encode horizontal and vertical orientations, while the third encodes texture information and edge. For the first application, QWT is used for multiscale image flow estimation [11]. Recently, Soulard studied the QWT [12] and its application in texture classification [13]. Gai et al. [14] used the dual-tree QWT (DT-QWT) in mono-spectral image denoising.

For denoising by classical DWT, Donoho and Johnstone have introduced the point-wise thresholding method [1,2]. In this scheme, all the wavelet coefficients below a certain value are set to zero, while the remaining ones were kept either unchanged (hardshrink) or reduced by the threshold value (softshrink). This approach offers the advantages of smoothness and adaptation. After that, several approaches which consider the influence of other wavelet coefficients on the current coefficient to be thresholded have been successively introduced. Cai and Silverman [15] proposed a thresholding algorithm by taking into account the neighboring coefficients. Their experimental results showed apparent advantages over the traditional term-by-term wavelet denoising. Chen and Bui [16] extended this idea to the multiwavelet case. They claimed that multiwavelet denoising outperforms the neighbor single-wavelet denoising for some standard test signals. Hailiang et al. [17] proved the efficacy of the multiwavelet coefficient dependency in the fault diagnosis of rolling bearings. Chinna Rao and Madhavi Latha [18] and Chen et al. [19] considered the relationship between the selective wavelet coefficients in a neighboring square window localized on the same scale. Experimental results show that these two methods produce better results in extended image denoising.

In addition to considering neighbor dependency in the same wavelet sub-band, Sendur and Selesnick [20] initiated the approach which takes into account the parent-child dependency. This idea was taken by Gai et al. [14]. For thresholding, they applied the bivariate shrinkage function to model the dependencies between current QWT coefficients and their corresponding parents. This method is based on a probabilistic estimator that seeks the relationship between the coefficients of two successive scales. They use a marginal approach applied on the real and imaginary parts of the wavelet coefficients, but the structural information (magnitude and phase) is not taken into account. In addition, only the real part is used in noise estimation.

In another work, the Bayes least squares-Gaussian scale mixture (BLS-GSM) method [21] is used for distributing visual artifacts in images during denoising. The intuition

of this method is the following: the neighborhoods of coefficients at the adjacent positions and scales are modeled as the Gaussian scale mixture. The wavelet coefficients are updated by the Bayesian least squares estimation. The contributions of this method are twofold: the full optimal BLS solution is computed for estimating coefficients, and the covariance between signal and noise is defined by the vectorial form of the linear least square (LLS). The pyramidal representation in the local model for spatial neighbors makes this algorithm efficient. However, the BLS-GSM approach requires an accurate estimation of the original signal spectrum density which makes this algorithm not adaptive. Later, new denoising algorithms based on the transforms are introduced. Dabov et al. [22] proposed a block matching and 3D filtering (BM3D) method inspired by the BLS-GSM and the non-local filters. 2D noisy image patches are separated in 3D data groups. In each group, patches have similar local structures. The 3D transform includes the 2D transform (discrete cosine transform, discrete Fourier transform, or periodized wavelet) within a group, and the 1D Haar transform in spatial dimensions which is applied to the matched 2D transformed groups. Shrinkage is done in two separate steps. In the first, hard thresholding is employed, and in the second, Wiener filter. BM3D exploits similarity between overlapping patches and the correlation of wavelet coefficients and have had optimal performances. But, when there are a few similar patches in the image, the method produces suboptimal results.

The local pixel grouping-principal component analysis (LPG-PCA) denoising procedure [23] has a similar structure to the BM3D. The difference is in the basis transform. Each pixel and his neighborhood are grouped into vector variables (LPG). This vector is PCA transformed, and the noise is removed by two shrinkage stages. The input of the second stage is filtered coefficients of the first. LPG-PCA is based on the local adaptive basis function and preserves the fine edges, whereas the previous BM3D method uses the fixed basis function which is less adapted to the local geometry of the image.

Satellite imaging has an important role in gathering information about the earth's surface. However, thermal effects, sensor saturation, quantization errors, and transmission errors generate a noise that deteriorates the quality and creates a bad effect on image analysis [24]. In [24,25], the parameters of noise in remote sensing imagery are estimated. The characteristics of the noise depend on the type of the image to be processed and on the system of acquisition. The radar remote sensing systems, such as a Synthetic Aperture Radar (SAR), are affected by multiplicative noise in addition to additive noise. In optical remote sensing multispectral imagery (the images used in our work), the noise is typically independent of the data and it is generally additive in nature. This type of noise

can be represented as a normal distribution (Gaussian), zero-mean random process. Ultimately, noise reduces the performances of important techniques of image processing such as detection, segmentation, and classification. These processes are performed by assuming that the noise is an integral part of the process. We can find some works where image denoising is made as a pretreatment. However, these approaches are not specified for satellite imaging. They are an extension of color image denoising. Luisier and Blu [26] proposed a new SURE-LET approach to image denoising. In [27], the authors extend this method to multichannel images. They used the parent-child coefficient relationship for thresholding. The efficiency of SURE-LET algorithm was demonstrated for color and satellite image processing. In [28], Saeedi et al. use the inter-channel relationship and dual-tree discrete wavelet shrinkage algorithm based on fuzzy logic. The authors have focused their work on the thresholding strategy, but they use a discrete wavelet transform which has a lack of shift variant. Chaux et al. [29] proposes a multichannel image denoising algorithm based on Stein's unbiased risk estimator [30] and on the discrete wavelet transform. A non-linear spatial estimator is proposed where this multivariate procedure operates by cleaning all components (spatial correlations are taken into account), but an inter-scale relationship is not considered. To conclude, it is interesting to note that for these three methods, the phase information is absent as in the case of the classical denoising approach. In our work, as we will see later, we propose to introduce this structural information into the denoising process.

The goal of this paper is not the comparison of different denoising method categories. More precisely, the comparison of new methods such as BM3D or LPG-PCA, which are based on bloc matching, distances us from the context of this work. We aim to show the contribution of analytic dimension and denoising based on regularization of coefficients depending on the local neighborhood and phase. At the same time, we introduce the concept of non-marginal processing in multiband case: due to the presence of potentially strong common information between the various bands, we developed a denoising method based on dual-tree quaternionic transform that supports all spectral bands simultaneously. Most of the existing algorithms apply the linear non-optimal processing separately or marginally in each band.

Another important point considered in our work is phase information. In most analytical wavelet denoising methods, only the magnitude of the wavelet is thresholded because the energy from the image is directed into a limited number of magnitude coefficients which 'stand out' from the noise. However, one quantity that appears to be very important in the human perception of images is phase as illustrated in [31]. The authors took the Fourier transforms of two images and used the magnitude information from one image and phase information from the other to construct a new synthetic Fourier transform which was then back-transformed to produce a new image. The features seen in such an image, while somewhat scrambled, clearly correspond to those in the image from which the phase data was obtained (see Figure 1). This idea is preserved in wavelet domain mainly for quaternionic wavelets where the phase is encoded in three angles. Regularization of this phase can greatly increase denoising results.

In this paper, we combine non-marginal DT-QWT, spatial and multiband neighboring thresholding, and phase regularization, adapted to satellite images, hence its originality.

The remainder of this article is organized as followed. The next section summarizes the theory of analytic signal and of the quaternionic wavelet transform. Section 3 explains how we can incorporate neighboring wavelet coefficients and phase regularization into image denoising. In Section 4, we propose a new algorithm by DT-QWT and neighborhood shrinkage/phase regularization function adapted to multiband or multichannel images. In Section 5, experimental results are provided, illustrating the potential of our approach for the class of real images. Finally, Section 6 is devoted to conclusions.

2 Summary of the quaternionic wavelet theory

In this section, we give the theoretical properties of the quaternionic wavelet transform which is based on the generalization of the analytic signal to image. Bulow [8] provided a strong 2D description of the analytic signal. He showed that complex algebra is only adapted to 1D signals, and 2D signal-like images are best described by quaternion algebra H.

The quaternion algebra is an extension of complex numbers to four-dimension (4D) algebra. Every element of H is a linear combination of a real scalar and three imaginary units i, j, k with real coefficients, as shown in [8]:

$$H = \{q = q_0 + i\,q_1 + j\,q_2 + k\,q_3 | q_0\,, q_1\,, q_2\,, q_3 \in R\},\tag{1}$$

with $i^2 = j^2 = k^2 = ijk = -1$, $ij = -ji = k$, $jk = -kj = i$, and $ki = -ik = j$.

In a polar form, a quaternion is defined by module and three angles which encode the phase, such as

$$q = |q|e^{i\theta}e^{j\psi}e^{k\varnothing}\tag{2}$$

$(\theta, \psi, \varnothing)$ are computed by the following formulas (for q normalized, i.e., $|q| = 1$) [10]:

$$\theta = \frac{1}{2}\arctan\left(\frac{2(q_0 q_2 + q_1 q_3)}{q_0^2 + q_1^2 - q_2^2 - q_3^2}\right)\tag{3}$$

Figure 1 Importance of phase in image processing. Sample images **(a, b)**. Second row: the phase of **(a)** appears in the left image and the phase of **(b)** appears in the right image.

$$\psi = \frac{1}{2} \arctan\left(\frac{2(q_0 q_1 + q_2 q_3)}{q_0^2 - q_1^2 + q_2^2 - q_3^2}\right) \quad (4)$$

$$\varnothing = \frac{1}{2} \arcsin(2(q_0 q_3 - q_1 q_2)) \quad (5)$$

Each quaternion phase angle is uniquely defined within the range $(\theta, \psi, \varnothing) \in [-\pi, \pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \times \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$.

For the complex case, the analytic signal $f_A(t)$ is constructed by adding to its associate 1D real signal $f(t)$ its Hilbert transform $Hf(t)$ in imaginary part. f_A and its spectrum are given by

$$f_A(t) = f(t) + i\,Hf(t) \Leftrightarrow F_A(\omega) = \begin{cases} 0 & \text{if } \omega < 0 \\ F(\omega) & \text{if } \omega = 0 \\ 2F(\omega) & \text{if } \omega > 0. \end{cases} \quad (6)$$

The modulus and the argument of f_A can be interpreted as the instantaneous magnitude and phase. Strong oscillation around one point of interest is a high magnitude, and phase indicates the relative location of this point. For generalization to 2D, Bulow introduced a definition of the quaternionic bidimensional analytic signal based on the quaternionic Fourier transform (QFT). The 2D quaternionic analytic signal for real signal f is defined as [8]

$$f_A^q(X) = f(X) + if_{Hi_1}(X) + jf_{Hi_2}(X) + kf_{Hi}(X), \quad (7)$$

where $X = (x,y)$.

The functions $\left(f_{Hi}, f_{Hi_1}, f_{Hi_2}\right)$ are, respectfully, the total Hilbert transformation and the partial Hilbert transformations, such as

$$f_{Hi}(X) = f(X) ** \frac{1}{\pi^2 xy}, f_{Hi_1}(X) = f(X) ** \frac{\delta(y)}{\pi x},$$
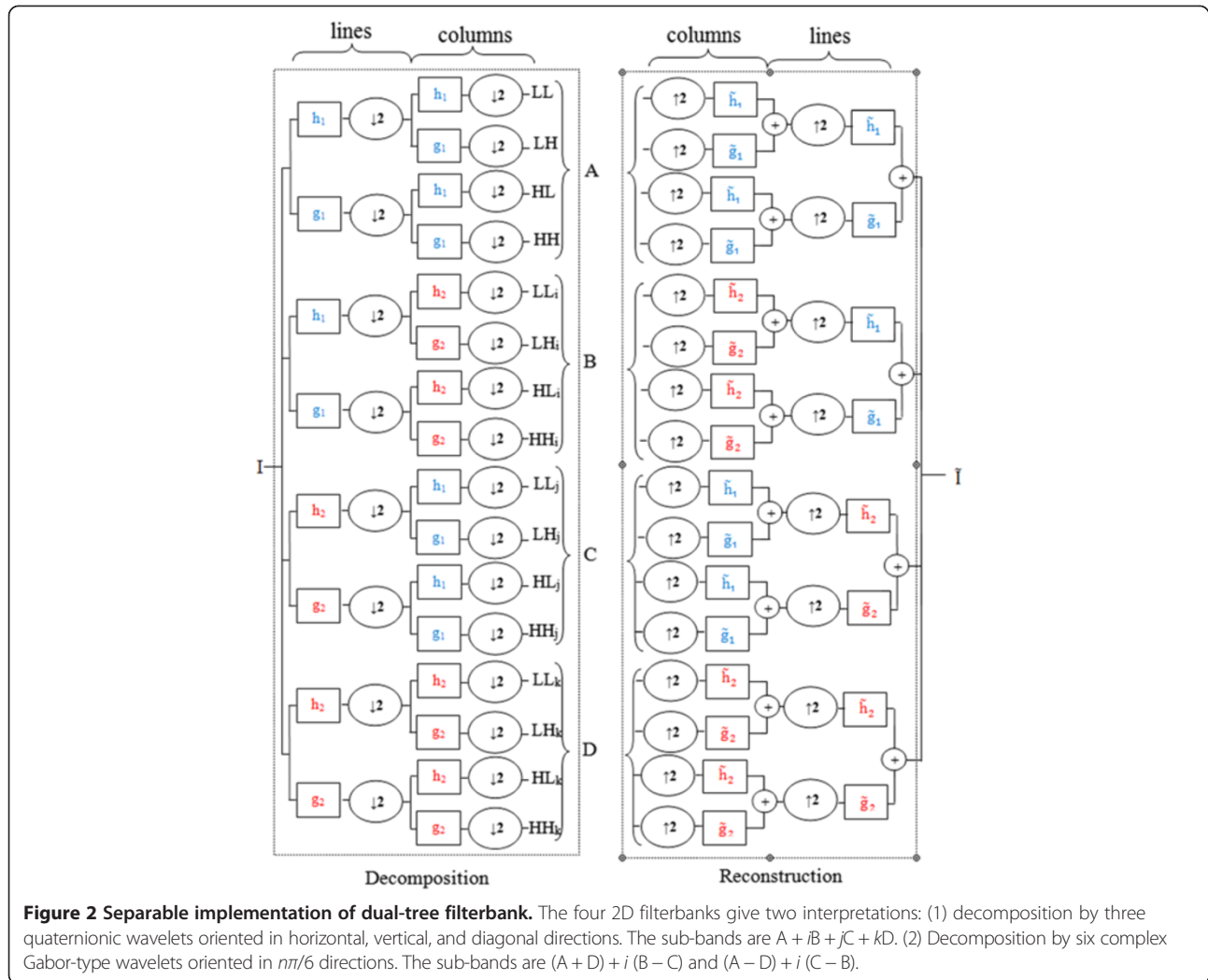$$f_{Hi_2}(X) = f(X) ** \frac{\delta(x)}{\pi y}$$

$\delta(x)$ and $\delta(y)$ are 2D Dirac distributions along the y-axis and x-axis, respectively; and ** denotes 2D convolution.

For each spatial position of the 2D analytical signal, the polar form of Equation 7 provides 2D local magnitude and phase that can be used to analyze 2D signals.

In order to obtain 2D analytical multiresolution representation, the construction of the quaternionic wavelet transform is based on the generalization of the DT scheme proposed by Kingsbury [5]. We obtain a 2D analytic wavelet and its associated quaternionic wavelet transform by organizing the four quadrature components of a 2D wavelet (real wavelet and its three Hilbert transformations: one total and two partial) as a quaternion [11].

To compute the QWT coefficients [12], we can use a separable 2D implementation of the dual-tree filter bank shown in Figure 2. During each stage of filtering, we

Figure 2 Separable implementation of dual-tree filterbank. The four 2D filterbanks give two interpretations: (1) decomposition by three quaternionic wavelets oriented in horizontal, vertical, and diagonal directions. The sub-bands are $A + iB + jC + kD$. (2) Decomposition by six complex Gabor-type wavelets oriented in $n\pi/6$ directions. The sub-bands are $(A + D) + i (B - C)$ and $(A - D) + i (C - B)$.

independently apply the two sets of h and g filters, two Hilbert pairs, to each dimension (x and y) of a 2D image. Therefore, the resulting 2D dual-tree implementation comprises four independent filter banks (hh, hg, gh, and gg) applied to each dimension and operating on the same 2D image. We combine the wavelet coefficients of the same sub-band from the output of each filter bank using quaternion algebra to obtain the QWT coefficients. These coefficients allow us to have a multiscale representation of analytic signal with module and phase information (see Equations 1 and 2).

3 Incorporating selective neighboring wavelet coefficients and phase regularization in image denoising

3.1 Thresholding by selective neighboring magnitude coefficients

From the 2D quaternionic wavelet transform, at every decomposition level, we get magnitude (module) of four frequency sub-bands, corresponding to an approximation

part and three detail parts. The principle is the same as that of the classical wavelets. Thresholding is applied to the coefficients of successive scales and the low-pass approximation is unchanged. Due to the linearity of the wavelet transform, the additive noise model in the image domain remains additive in the wavelet domain [18] as well as

$$w_{k,l}(x,y) = y_{k,l}(x,y) + n_{k,l}(x,y), \tag{8}$$

where $w_{k,l}(x,y)$, $y_{k,l}(x,y)$, and $n_{k,l}(x,y)$ denote noisy, noise-free wavelet coefficients, and noise components of scale k and orientation l, respectively.

As explained in Section 1, the noise is assumed Gaussian and additive. The probabilistic model adapted to the magnitude of noisy quaternionic wavelet coefficients is the *Rayleigh* distribution. The Rayleigh model is a function of the Gaussian estimation of the squared real part added to the Gaussian estimation of the squared imaginary part of noisy coefficients.

To define the denoising method, it is necessary to introduce a thresholding strategy adapted to the QWT. The basic motivation of neighbor thresholding is that if the current coefficient contains information, it is likely that the neighbor coefficients also do. (Wavelet coefficients are correlated in a small neighborhood.) We choose local windows around every coefficient of our interest, and we threshold it by using the coefficients in this neighborhood. The size of the window is predefined as a function of the image size. We shrink the magnitude of the noisy wavelet coefficients according to the following formula [18]:

$$\left| w_{k,l}(x,y) \right| = \left| w_{k,l}(x,y) \right| \times T(x,y) \qquad (9)$$

$T(x,y)$ is the shrinkage factor defined as

$$T(x,y) = 1 - \frac{\lambda^2}{S_j^2(x,y)} \qquad (10)$$

$$T(x,y) = \begin{cases} T(x,y) & \text{if } \lambda^2 < S_j^2 \\ 0 & \text{if } \lambda^2 > S_j^2, \end{cases}$$

λ is the universal threshold, with $\lambda^2 = 2\sigma^2 \log b^2$; σ is the standard deviation of corrupted coefficients; and b^2 is the size of local neighborhood window.

In Equation 10, S_j^2 is the summation of squared coefficients in the local window defined as in [19]:

$$S_j^2(x,y) = \sum_{(p,q) \in b^2} \left| w_j(p,q) \right|^2, \quad b = b_0 - j, \qquad (11)$$

where j is the level of decomposition and b_0 is a constant defined according to the size of noisy image and the support of the wavelet filter. (p,q) varies in the neighboring window centered on the coefficient $w (x,y)$. The window size b^2 varies depending on the level of decomposition because the correlation between coefficients varies in successive scales. Figure 3 illustrates a variable size neighborhood window centered at the wavelet coefficient to be thresholded. The choice of a larger size decreases the correlation between neighboring coefficients, while a smaller size brings us back to the term-by-term case (the neighbor dependency is neglected).

The shrinkage factor T of Equation 10 is a function of the adaptive sum S_j and universal threshold λ. S_j depends on the neighboring window size b^2. For each wavelet coefficient candidate to thresholding, T is calculated by comparing the sum of neighboring coefficients to λ. Then, the wavelet coefficient is either reduced or set to zero. Neighboring shrinkage is a generalization of the term-by-term thresholding.

A recent method proposed by Luisier and Blu [26], which is based on Stein's unbiased risk estimator [30], can be used to perform denoising in wavelet domain. Authors parameterize the denoising process as a sum of elementary non-linear processes with unknown weight. They minimize an estimate of the mean-squared error between the clean image and the denoised one based on the noisy data alone. However, the neighboring strategy adopted in our work is based on the direct thresholding of the coefficients. We want to place our approach among those using the same concept, but they differ in the adopted wavelet transform. We can see later that this strategy is more adapted to combination with the following phase regularization.

In some applications of image denoising, the value of the input noise variance σ^2 is known or can be measured based on the information other than the corrupted data. If this information is not available, one has to estimate it from the input data, eliminating the input of the actual signal. All frequency sub-bands of the decomposition are used in the noise estimation [1,2]. For estimating the *Gaussian* noise variance in real and in imaginary parts of noisy wavelet coefficients, we use the mean absolute deviation relation proposed by Donoho and Johnstone [2] that is denoted as

$$\sigma_g^2 = \left(\frac{\text{median}|W|}{0.6745} \right)^2, \qquad (12)$$

where median $|W|$ is the median of neighboring coefficients in the local window centered on the coefficient $w(x,y)$.
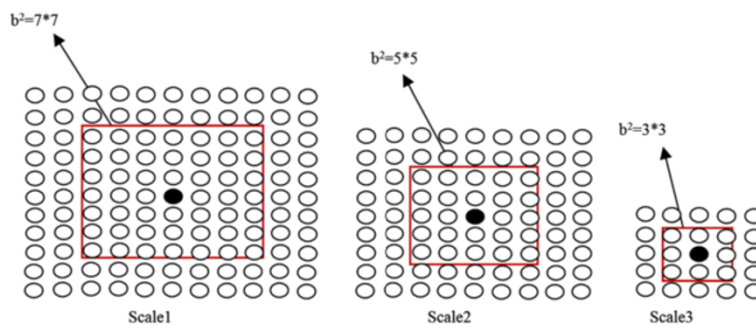
Figure 3 Neighborhood window size selection.

Then, the noise variance according to the *Rayleigh* distribution [24] is given by

$$\sigma^2 = \frac{4-\pi}{2}\sigma_g^2 \qquad (13)$$

To conclude, the algorithm described in this section is the adaptation of the method called *NeighShrink* based on the squared sum of all the processed magnitude wavelet coefficients with variable neighborhood window sizes. These sizes are in function of decomposition levels. The adaptive threshold value selected according to neighborhood provides a powerful thresholding procedure greater than the term-by-term shrinkage approach (experimental proofs for real wavelet are proposed in [18,19]).

3.2 Phase regularization

In addition to the image denoising by thresholding the magnitude of the quaternionic wavelet transform, it is important that the phase of this transform is not excluded from the process. The three quaternion phase angles (θ, ψ, $\varnothing$) for Equations 3, 4, and 5 are separable. The first two encode the shift and the third encodes the textures. More precisely, Bulow [8] defined a shift theorem for the quaternionic Fourier transform such as a shift of the image is an equivalent of an offsite of the two first terms θ and ψ of the phase.

The shift theorem for the QFT [8] approximately holds for the QWT that conducts a local QFT analysis. When a shift of image $f(X)$ to $f(X - \boldsymbol{d})$ occurs, the QFT phase undergoes the following changes:

$$(\theta(\mathbf{u}), \psi(\mathbf{u}), \varnothing(\mathbf{u})) \rightarrow (\theta(\mathbf{u}) - 2\pi u d_1, \psi(\mathbf{u}) - 2\pi v d_2, \varnothing(\mathbf{u})),$$

where $\mathbf{u} = (u,v)$ are the axes of the 2D QFT domain. $\boldsymbol{d} = (d_1, d_2)$.

Note that the 1D shift is equivalent to the structural information, but the 2D structure (e.g., corners, *T*-junctions) may be more complex than lines or edges and cannot be described by the shift of the first two angles. The author observed that when the third angle $\varnothing$ is around $\pm\frac{\pi}{4}$, the codec structure is a line or an edge oriented along a diagonal. The angle $\varnothing$ can be interpreted as the relative amplitude of signal energy along the 1D which manifolds in two orthogonal directions.

Chan et al. [11] demonstrated the importance of the quaternionic wavelet transform phase in image processing and analysis. Chan and his co-authors also developed a multiscale flow/motion estimation algorithm that computes a disparity flow map between two images with respect to local object motion [32]. Soulard and Carré have developed an efficient method for texture classification, thanks to coherent multiscale analysis brought by the magnitude and phase of the quaternionic wavelet

transform [13]. In their approach, the authors used a global measure of energy from the magnitude, and they combine it with the weighted standard deviation of the third-angle quaternionic phase. They observed that this last measure phase contains structural information that contributed to improving the classification.

From those analyses, we observed that the combination of QWT magnitude and phase is effective in several image processing tools. In our algorithm, by adjusting only $\varnothing$ of the quaternionic sub-band coefficients, a potential interesting change can be observed in image quality, and therefore, we can improve denoising performance. In our knowledge, there are very few methods that use phase in the process of denoising. With analytical decomposition, the only proposition is the Miller and Kingsbury approach [33]. They have modeled discontinuities in image by using wavelet coefficients derotated by twice the phase in local scale and the next coarser scale at the same spatial location. In our work, we propose a regularization of the phase information. This approach is sufficient if we want to reduce complexity. We use a typical first-order regularizer $R(\varnothing) = C\varnothing$ to enforce spatial smoothness [34]. From this concept, quaternionic wavelet coefficients become

$$w_{k,l}(x, y) = \left(\left| w_{k,l}(x,y) \right|_T \right) e^{i\theta} e^{j\psi} e^{kC\varnothing} , \qquad (14)$$

where $\left| w_{k,l}(x, y) \right|_T$ is the thresholded magnitude coefficient from the NeighShrink method.

We want to extract unique value that defines the global direction in a sub-band and has structural information at the same time. For this, the finite matrix C is chosen as a simple median filter with variant size. The size of smoothing matrix C changes according to the scale. It should be noted that the regularization of the phase by median filter is applied to the thresholded magnitude coefficients; consequently, the phase regularization is controlled by the value of the magnitude.

To conclude, we note that the denoising method does not increase the computational cost dramatically. If a real wavelet transform spends N operations, the construction of QWT would need $4N$ operations. Moreover, denoising-based real wavelet requires the estimation of the threshold and the thresholding operation for each coefficient, with the quaternionic transform. This process is applied for two informations: magnitude and phase. Finally, the new operation is the polar conversion.

The process of denoising is illustrated in Figure 4. The experimental efficiency of phase regularization is shown in Section 5.

4 Multispectral image denoising by the DT-QWT and the NeighShrink/phase regularization algorithm

In the previous sections, we defined the quaternionic wavelet transform and the thresholding/regularization

Figure 4 Principle of the dual-tree quaternionic wavelet denoising.
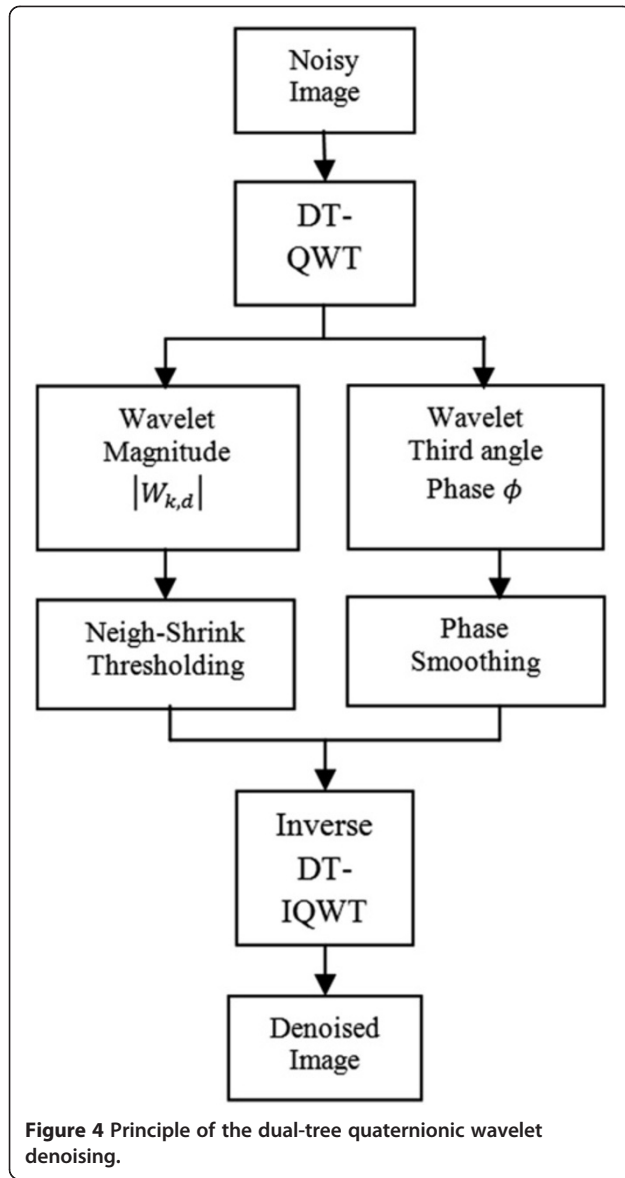
strategy for mono-spectral image. In multispectral image, different bands are correlated: an image discontinuity from one band is most likely to occur in at least some of the remaining bands. It should be noted that in order to avoid confusion between the spectral bands of the wavelet transform and the multiband image, the second is called multichannel.

For denoising, there are two main conceivable strategies: the first one consists of marginally applying a denoising process; the second is to devise specific non-separable multichannel denoising algorithms. Our interest is focused on the latter strategy. Therefore, we defined a *non-linear method* which generalizes a mono-channel approach by taking into account the relationship between channels (it is not a marginal approach).

The 'clean' multichannel wavelet coefficients contain $M \in N^*$ components y^m with $m \in [1, ..., M]$. Typically, M is equal to three in RGB images. It might be larger for satellite images. Therefore, the multichannel noisy observation in the wavelet domain is as follows:

$$W_{k,l}(x,y) = Y_{k,l}(x,y) + N_{k,l}(x,y) , \qquad (15)$$

where $Y \triangleq (y^{(1)}, ..., y^{(M)})$ is the noise-free wavelet vector, $N \triangleq (n^{(1)}, ..., n^{(M)})$ is the noise vector, and $W \triangleq (w^{(1)}, ..., w^{(M)})$ is the noisy wavelet vector. (x,y) are the coordinates of the coefficient in the corresponding sub-band, and k and l are scale and orientation, respectively.

We see that each coefficient located in position (x, y) and scale k is taken in the vector W (vectors Y, N) with the coefficients of the remaining channels according to the same position and the same scale.

In color imaging, it is important to treat pixels as color components, not as three separate RGB colors. When only the separate channels are considered, more artifacts are introduced. For thresholding in the mth channel, the wavelet coefficient $w_{k,l}^m(x, y)$ must be modified according to its spatial neighboring but also depending on the corresponding coefficients of the same scale in the remaining channels. For this, we propose to combine multichannel information and spatial information. As in single channel thresholding, the intra-scale/inter-channel shrinkage factor T is a function of squared summation S_j^2 and universal threshold λ^2 (return to Equation 10). These parameters are defined in a multichannel case according to the proposed formulas:

$$S_j^2(x,y) = \sum_{(p,q) \in b^2} \left| w_j^{(m)}(p,q) \right|^2 + \sum_{\substack{i=1 \\ i \neq m}} \left| w_j^i(x,y) \right|^2 \qquad (16)$$

In Equation 16, we sum neighboring coefficients inside a window of size b^2 in scale j and channel m. This first result is added to the sum of coefficients in the same position (x,y) but of all M channels. The threshold is defined such that $\lambda^2 = 2\sigma^2 \log b^2$ and the noise variance σ^2 is given by

$$\sigma^2 = \frac{4-\pi}{2} \left(\frac{\text{median}|W_c| + \text{median}|W_{x,y}|}{0.6745} \right)^2 , \qquad (17)$$

where median $|W_c|$ and median $|W_{x,y}|$ are the median of neighboring coefficients in the same channel m and the median of coefficients of all channels in the same spatial position, respectively.

In Equations 16 and 17, we give a new formulation of the parameters that allows us to calculate the multichannel threshold value T. The first term specified an intra-scale relationship (spatial), and the second defined an inter-channel correlation. We note that in the second term of S_j^2 and σ^2,

the sum and the median, respectively, are made on all channels. When the number of channels is very high, e.g., for hyperspectral images, we can define two approaches: First, only the adjacent correlated channels are considered. However, when the correlated channels are not adjacent, we can search correlated bands with a block matching approach [22]. The proposed algorithm can be adapted to this second case but with an increase of complexity.

Previous multichannel magnitude thresholding is combined with linear phase regularization. It is important to notice that the multichannel phases are smoothed separately following the mono-channel strategy (median filter). Indeed, inter-correlation between the phases of different channels is not known, and the formulation of this relationship is not yet established (this work is in progress).

Proposed multichannel denoising process by the dual tree quaternionic wavelet transform and the NeighShrink/phase regularization algorithm is shown in Figure 5 (we considered a multispectral image with three channels and decomposed it into three scales).

5 Results and discussion

Different tests are accomplished to rate the effectiveness of the proposed algorithm in reducing noise and compare it with known techniques. In the following section, we present the denoising results in both single-channel and multichannel cases. This section is intended to illustrate the contribution of the quaternionic wavelet transform, the multiband information in spectral and spatial thresholding, and the phase smoothing compared to the methods based on classical real neighboring coefficient regularization.

5.1 Single-channel denoising

We compare neighborhood thresholding and the phase regularization method (proposed algorithm called *NeighShrink/phase-smooth*) with different thresholding techniques (soft shrinkage [2], neighboring shrinkage without phase regularization [18], and bivariate shrinkage [14], called *VisuShrink*, *NeighShrink*, and *BiShrink*, respectively). For implementation software of the bivariate thresholding method, we refer to the homepage [35], thanks to Shihua Cai and Keyong Li. We note that in single-channel denoising, analysis and synthesis of images over all denoising processes are made by the same dual-tree quaternionic wavelet transform with five levels of decomposition. We change only the thresholding methods listed above.

The images used in our experimentation are the second green band (left image), the first red band (middle image), and the fourth infrared band (right

Figure 5 Global scheme multispectral image denoising by DT-QWT and NeighShrink/phase regularization method. The phase smoothed is the third angle. The two other angles remain unchangeable.
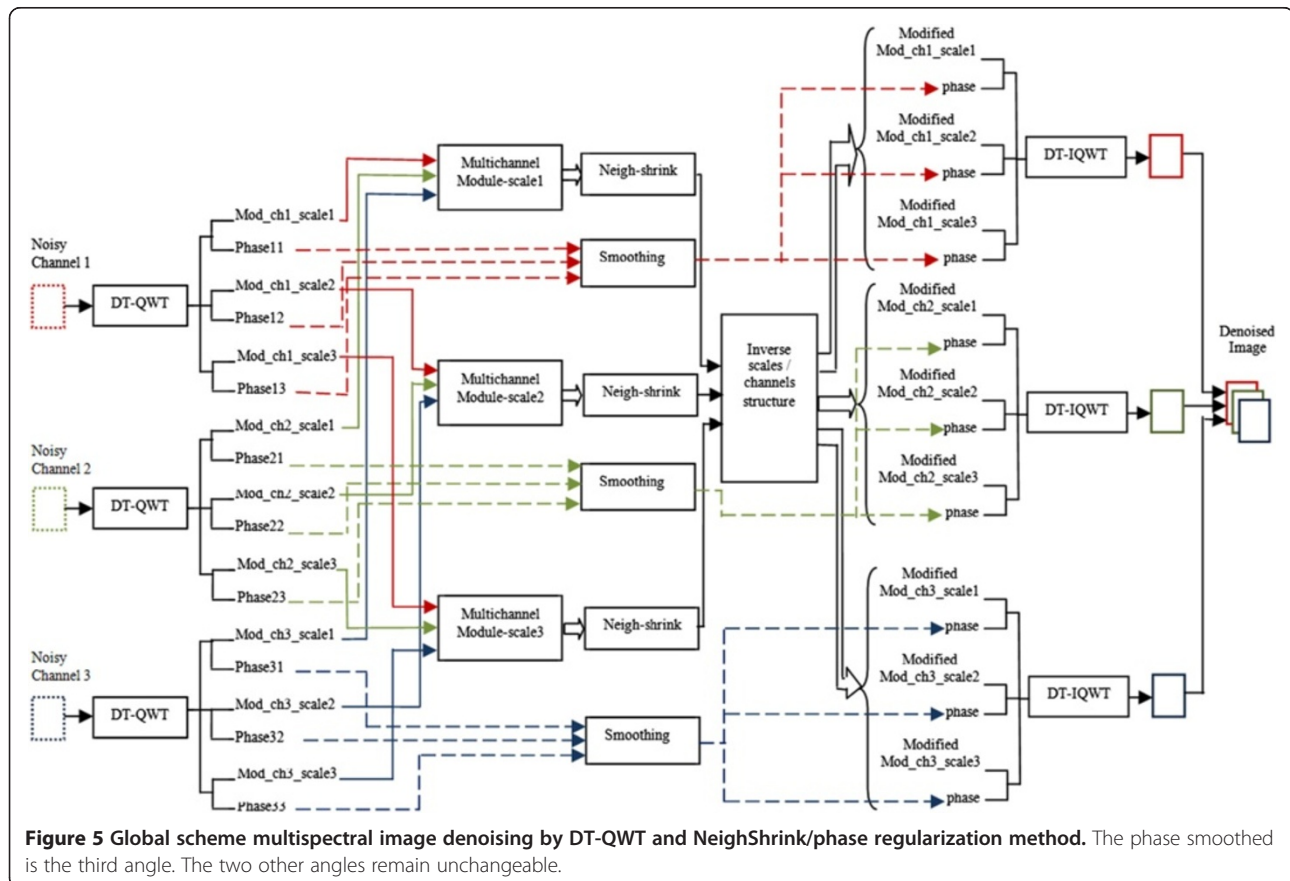
image) of satellite images [36]. The first one covers the area called *Sebkha*, part of Oran City in western Algeria, the second is one band of satellite image that covers part of Mouhammadia City in Algeria, and the last represents another area of Oran City (Figure 6). In mono-channel experiments, we have chosen three independent bands. There is no correlation between those data, and they are derived from three different areas. This choice will allow us to see the potential of our method to denoise various structures in the images. These single-channel images have the same size of 400×400. Following the model of Equation 8, normally distributed, uncorrelated, and zero-mean additive noise was generated for six levels. Then, each band is contaminated with computer-generated additive Gaussian noise $(0, \sigma_n^2)$ to simulate a noisy image. Inherent low-level noise in the original image was considered as a part of a data. More details on multispectral images are given in the next section.

The proposed approach has been evaluated using visual analysis and objective peak signal-to-noise ratio (PSNR), a criterion which is commonly used as a measure of noise suppression:

$$\text{PSNR} = 10 \log_{10} \frac{255^2 NM}{\sum_{x=1}^{N} \sum_{y=1}^{M} \left(I(x,y) - \hat{I}(x,y) \right)^2}, \tag{18}$$

where I and $\hat{I}$ are noisy and denoised images, respectively. $N \times M$ is the size of the images.

The PSNR is simple to calculate, and it is mathematically convenient in the context of optimization. However, this objective metric is not very well matched to perceived visual quality. The structural similarity index (SSIM) is a very powerful tool which is based on structural information of distorted images and converges in the same results as the visual perception. This measure was highly adapted in our algorithm. It takes into account the structural dependencies between neighboring pixels when the PSNR based on the MSE is calculated pixel by pixel.

The *SSIM* [37] between reference image I and processed image $\hat{I}$ is given by

$$\text{SSIM}(x,y) = l(x,y)c(x,y)s(x,y) \tag{19}$$

The term $l(x,y)$ stands for the luminance comparison function, $c(x,y)$ for the contrast comparison function, and $s(x,y)$ for the structure comparison.

These functions are given by the following formulas:

$$l(x,y) = \frac{2\bar{\mu}_I.\bar{\mu}_{\hat{I}} + C_1}{(\bar{\mu}_I)^2 + (\bar{\mu}_{\hat{I}})^2 + C_1} \tag{20}$$

$$c(x,y) = \frac{2\,\sigma_I.\sigma_{\hat{I}} + C_2}{\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2} \tag{21}$$

$$s(x,y) = \frac{\sigma_{I\hat{I}} + C_3}{\sigma_I.\sigma_{\hat{I}} + C_3} \tag{22}$$

with $C_1 = (LK_1)^2$, $C_2 = (LK_2)^2$, and

$$C_2 = (LK_2)^2 \tag{23}$$

where $\bar{\mu}_I$ and $\bar{\mu}_{\hat{I}}$ are the mean intensities of I and $\hat{I}$, respectively. σ_I and $\sigma_{\hat{I}}$ are the standard deviations used in the estimation of image contrast, and $\sigma_{I\hat{I}}$ corresponds to the covariance between the two images. L is the dynamic range of luminance (usually the maximum gray level). K_1 and K_2 are two constant parameters to adjust the metric variation (the Matlab implementation by the authors in [37] used the values of 0.01 and 0.03, respectively).

Tables 1 and 2 summarize the obtained results in PSNR (dB) and SSIM (%). In Table 3, we give the average gain of the two various metric comparisons.

Several conclusions can be drawn from these experiments:

1. VisuShrink does not have any denoising power or very low performance when the noise level is low (noise variance: 15, 20).
2. The effect of using only the magnitude neighboring thresholding (*NeighShrink*) for the three images is generally a considerable PSNR and SSIM gain compared to classical *VisuShrink* thresholding.
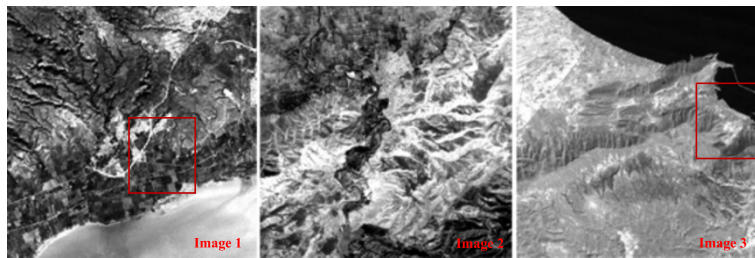
Figure 6 Test images used in the single-channel denoising experiments. Images referred to as image 1 to image 3 (left to right). Red squares represent the parts displayed in Figures 7 and 8. The images corresponding to different spectral bands are captured in gray. Additional processing gives these images false color (green, red…).

Table 1 Comparison of various denoising methods by PSNR (dB)

Methods	Noise variance					
	15	20	25	30	40	50
Image 1						
Noisy image	24.87	22.40	20.78	18.19	16.69	14.91
VisuShrink	22.93	21.51	21.49	19.71	19.90	17.90
NeighShrink	25.23	23.43	22.89	21.93	20.63	19.58
BiShrink	25.32	24.48	23.19	23.58	22.43	21.93
Proposed algorithm	26.24	25.65	25.02	24.36	23.12	22.99
Image 2						
Noisy image	24.90	22.42	20.76	18.18	16.64	14.82
VisuShrink	24.28	21.72	21.68	20.82	20.01	19.65
NeighShrink	27.37	24.95	24.02	22.31	19.80	23.12
BiShrink	26.74	25.47	25.01	20.48	23.40	23.53
Proposed algorithm	27.74	25.34	24.90	23.95	23.36	22.63
Image 3						
Noisy image	25.07	22.58	20.90	18.23	16.63	14.81
VisuShrink	24.25	22.29	25.04	22.92	22.65	21.87
NeighShrink	27.32	26.26	25.45	22.59	22.46	19.13
BiShrink	28.22	27.37	25.15	23.29	23.26	23.21
Proposed algorithm	28.02	28.42	26.55	25.02	25.03	24.92

Wavelet transform: dual-tree quaternionic wavelet transform with five levels.

Table 2 Comparison of various denoising methods by SSIM (%)

Methods	Noise variance					
	15	20	25	30	40	50
Image 1						
Noisy image	83.96	76.70	71.03	60.24	53.26	44.57
VisuShrink	84.63	65.96	66.45	53.19	55.87	36.55
NeighShrink	73.40	76.50	74.66	71.38	62.47	58.13
BiShrink	83.09	81.57	76.71	88.85	72.04	74.51
Proposed algorithm	86.53	85.57	83.61	92.80	74.93	75.78
Image 2						
Noisy image	89.28	82.84	77.25	66.08	58.24	48.22
VisuShrink	83.87	70.32	70.47	64.97	58.80	57.56
NeighShrink	92.52	87.06	84.67	79.73	71.56	82.86
BiShrink	91.53	88.23	88.27	75.59	83.88	82.15
Proposed algorithm	93.81	90.02	88.41	87.83	86.60	86.26
Image 3						
Noisy image	70.12	59.94	52.42	39.91	32.68	24.97
VisuShrink	61.30	53.84	68.96	52.23	51.75	48.07
NeighShrink	80.80	75.50	71.96	56.63	54.79	52.36
BiShrink	83.92	80.93	73.37	64.72	64.55	59.46
Proposed algorithm	84.36	84.62	77.41	69.67	74.01	68.02

Wavelet transform: dual-tree quaternionic wavelet transform with five levels.

3. *NeighShrink* is not efficient as opposed to the *bivariate* denoising method in all cases.

4. The addition of phase smoothing to the magnitude neighboring shrinkage mostly outperforms other approaches with fixed wavelet. In Table 3, the comparison for image 1 shows that the average PSNR and SSIM improvement gained by the proposed method over *NeighShrink* (without phase smoothing) are 2.45 dB and 13.78%, respectively. When our method is compared to *BiShrink*, we gain 1.07 dB and 3.73%.

5. For high levels of noise (40, 50), PSNR comparisons for image 2 and image 3 are not adequate with these conclusions. However, for the same noise variance values, the SSIM gives a better result which corresponds to visual observations. In image 1 (Figure 6), the edges of the squared vegetation are naturally very disenable over other structures. PSNR and SSIM are perfectly adapted with this image and allow very good comparisons. But, image 2 and image 3 (see Figure 6) have mixed structures and in some noise levels, only the SSIM, which is a structural metric, gives results that correspond to visual analysis.

For visual evaluation, there are two important criteria: the visibility of processing artifacts and preserving image edges. Figures 7 and 8 illustrate denoising results of single-channel image 1 and image 3, respectively, from different methods. For a better visualization of the details and differences between denoising results, only partial parts of the images are displayed (see red square in Figure 6). The *NeighShrink* approach surpasses classical *VisuShrink* thresholding for the two images, but the noise is still present (Figures 7c,d and 8c,d).

The bivariate shrinkage reduces the noise more effectively than the NeighShrink but details are very smooth

Table 3 Average PSNR and SSIM differences between denoising algorithms for the three images

		PSNR (dB)	SSIM (%)
NeighShrink/VisuShrink	Image 1	1.33	8.98
	Image 2	2.23	15.4
	Image 3	0.69	9.31
BiShrink/NeighShrink	Image 1	1.20	10.03
	Image 2	0.51	1.87
	Image 3	1.22	5.81
Proposed algorithm/NeighShrink	Image 1	2.45	13.78
	Image 2	1.05	5.75
	Image 3	2.45	11.00
Proposed algorithm/BiShrink	Image 1	1.07	3.73
	Image 2	0.39	3.88
	Image 3	1.24	5.19
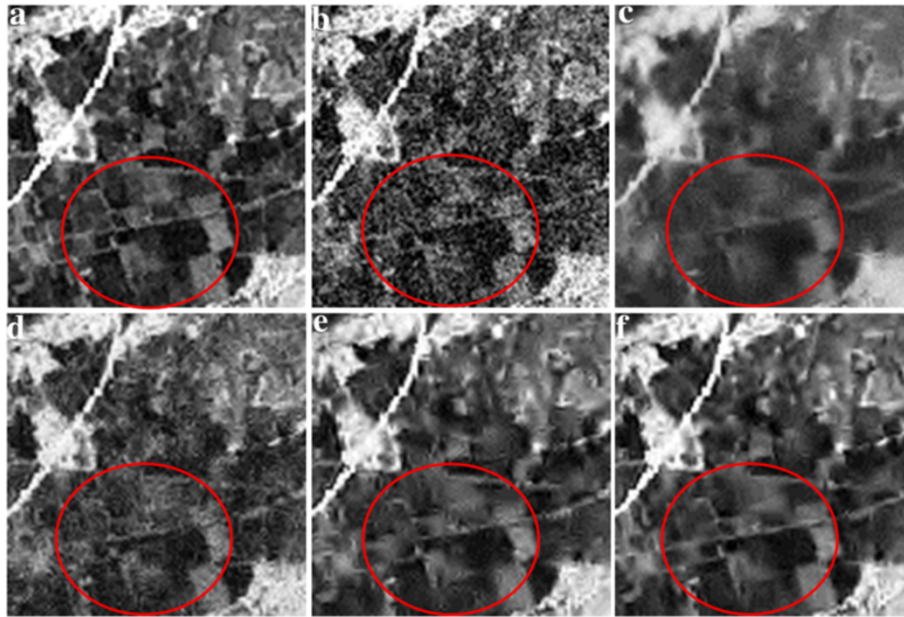
Figure 7 Single-channel denoising comparison results in PSNR/SSIM of image 1. (a) Partial noise-free image, **(b)** noisy image 18.19 dB/60.24%, **(c)** VisuShrink 19.71 dB/53.19%, **(d)** NeighShrink 21.93 dB/71.38%, **(e)** BiShrink 23.58 dB/88.85%, and **(f)** proposed algorithm 24.36 dB/92.80%.

(see Figures 7e and 8e). Better results are obtained with the NeighShrink/PhaseSmooth algorithm which can effectively distinguish the regions of interest from noise (square vegetation edges in Figure 7f and bottom structures in Figure 8f enclosed in red circles are more contrasted), meaning the correction of the third quaternionic angle is key to realizing the full potential of the algorithm (Figure 7f).

5.2 Multichannel satellite image denoising

We propose in this section to study the adaptation of the single-channel algorithm to multispectral satellite images. There are several sources of noise in optical satellite images (photonic, electronic, quantization error, etc.), and the additive zero-mean Gaussian noise model is a realistic approximation as shown in [24,25].

We perform the multichannel algorithm based on the DT-QWT and the NeighShrink/PhaseSmooth denoising strategy where a non-marginal aspect is highlighted. In order to compare different possible wavelet choices, the experimental results are derived from the DWT and the DT-CWT (for these representations, the thresholding approach is NeighShrink). Phase smoothing cannot be applied to the DWT (no phase) and the DT-CWT (the unique phase of this transform is a location information and phase smoothing adds nothing to denoising). In addition, the proposed algorithm is compared to the DT-QWT-Neighboring shrinkage and the DT-QWT-Bivariate shrinkage. We specify that the non-separable denoising is only done by our method. In all other approaches, the analysis/thresholding/synthesis scheme is marginally (linearly) performed channel by channel.

The experiments in this section have been carried out on two seven-band satellite images shown in Figures 9 and 10, which represent two regions called *Sebkha* and *Sea*, parts of Oran City in western Algeria [36]. The first *Thematic Mapper* image contains a lake and vegetation with several roads. The second includes sea and mountains. The coverage areas are 30 × 30 km with resolution of 30 m and size of 400 × 400 × 7. We note that for our comparison, denoising methods are applied to the seven bands of the satellite images. However, only three channels are used in the display of visual results (red, blue, and infrared for the Sebkha image and red, blue, and green for the Sea image; these bands allow differentiation between soil, vegetation species, coastal areas, sea, and biomass). This choice is justified by the fact that these denoising results are subsequently used in the following processes such as segmentation or compression or simply a visual interpretation of information contained in the images. In this case, we will need only three bands which correspond to areas of our interest (vegetation, lakes, sea, mountains, roads, etc.).

We measured the experimental results by the PSNR, objectively, which is an extension of the definition given by Equation 18 as

$$\mathrm{PSNR} = 10\log_{10}\frac{255^2 NMC}{\sum_{k=1}^{C}\left(\sum_{x=1}^{N}\sum_{y=1}^{M}\left(I(x,y)-\hat{I}(x,y)\right)^2\right)},$$

(24)

where C is the number of channels.
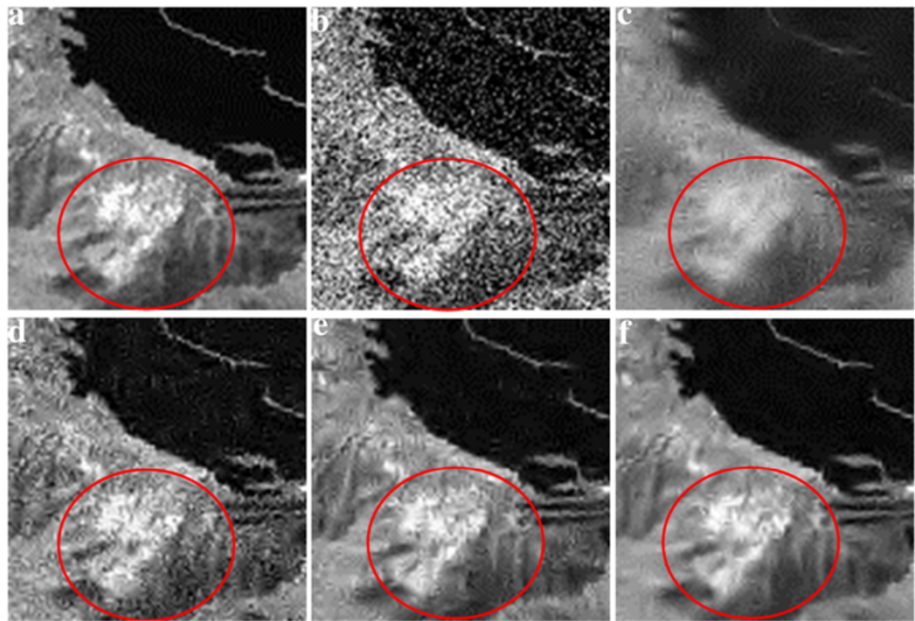
Figure 8 Single-channel denoising comparison results in PSNR/SSIM of image 3. (a) Partial noise-free image, **(b)** noisy image 14.81 dB/24.94%, **(c)** VisuShrink 21.87 dB/48.07%, **(d)** NeighShrink 19.13 dB/52.36%, **(e)** BiShrink 23.21 dB/59.46%, and **(f)** proposed algorithm 24.92 dB/68.02%.

Table 4 summarizes the obtained results. We observe the following:

1. Quaternionic and complex wavelet transforms outperform the discrete wavelet transform when the thresholding strategy (NeighShrink) is the same

(average gain 1.14 and 1.06 dB for image Sebkha and Sea, respectively).

2. DT-QWT and DT-CWT have very close results.

3. As in the single-channel experiment, the bivariate shrinkage is more efficient than the neighboring shrinkage without phase smoothing.
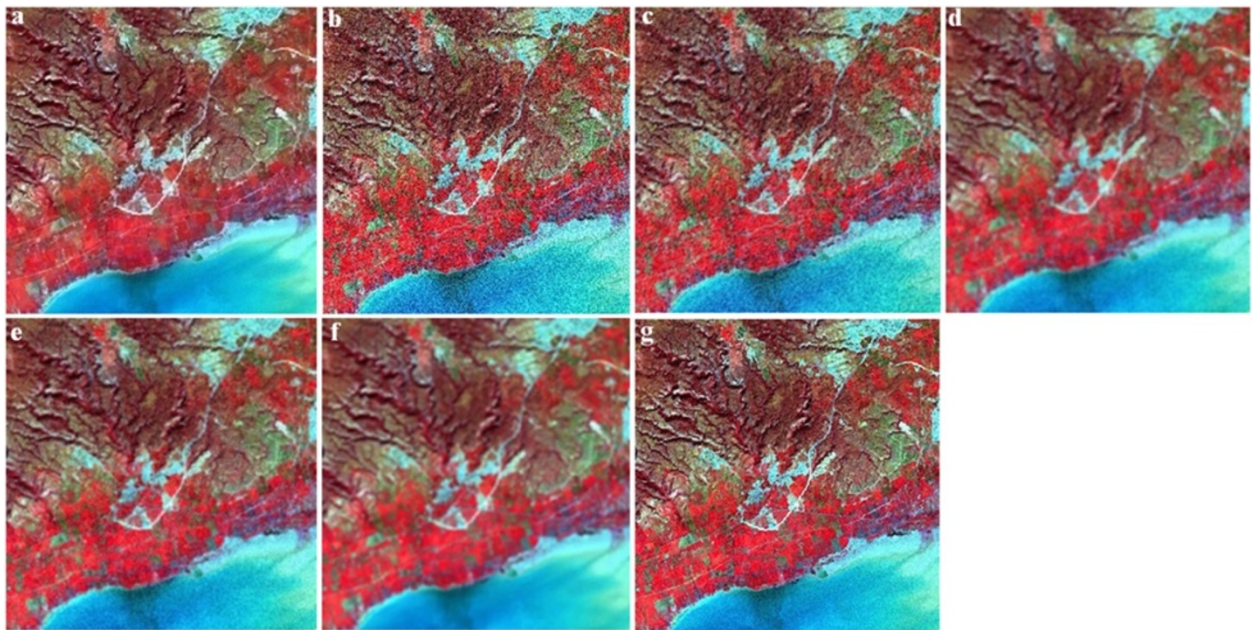
Figure 9 Multichannel Sebkha satellite image denoising. (a) Noise-free image. **(b)** Noisy image 21.26 dB. **(c)** Results of DWT and NeighShrink 22.18 dB. **(d)** Results of DT-CWT and NeighShrink 23.32 dB. **(e)** Results of DT-QWT and NeighShrink 23.36 dB. **(f)** Results of DT-QWT and BiShrink 24.57 dB. **(g)** Results of DT-QWT and NeighShrink/PhaseSmooth 26.14 dB.
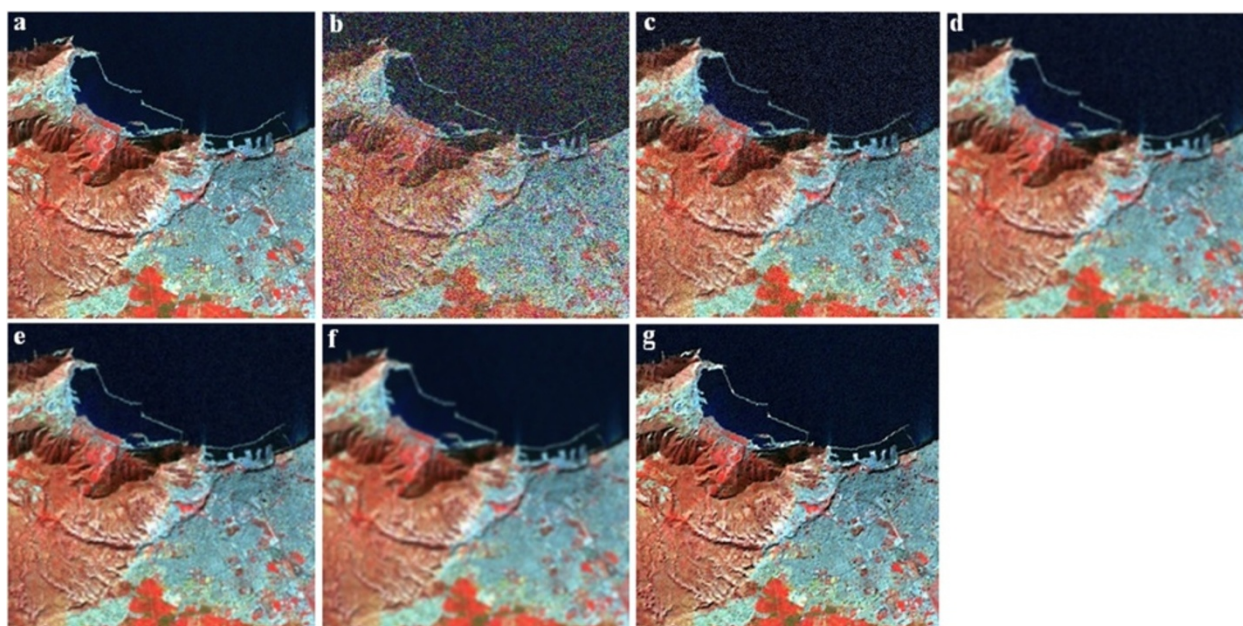
Figure 10 Multichannel Sea satellite image denoising. (a) Noise-free image. **(b)** Noisy image 15.28 dB. **(c)** Results of DWT and NeighShrink 16.12 dB. **(d)** Results of DT-CWT and NeighShrink 17.17 dB. **(e)** Results of DT-QWT and NeighShrink 17.17 dB. **(f)** Results of DT-QWT and BiShrink 19.21 dB. **(g)** Results of DT-QWT and NeighShrink/PhaseSmooth 20.87 dB.

4. Compared to the DT-QWT with neighboring channel-by-channel shrinkage, the proposed inter-channels DT-QWT NeighShrink/PhaseSmooth achieves an improved performance and yields a larger total PSNR gain (average 2.79 and 3.03 dB for the two images, respectively). The PSNR gain values are greater than the results obtained in Section 5.1 (2.45 dB).

5. When we compare our method to the DT-QWT with bivariate shrinkage, we gain 1.06 and 1.66 dB for the Sebkha and Sea images, respectively. Again, the multichannel algorithm gives better results than the single channel (1.24 dB) for the second image.

Figures 9 and 10 illustrate the comparative results among different multichannel denoising methods and proposed algorithm applied to the two images. In Figures 9c and 10c (discrete wavelet transform), the noise is very present. Noise is reduced in Figures 9d,e,f and 10d,e,f), but these three methods have the tendency to smooth discontinuities. We note that the quaternionic wavelet transform is greater than the DWT and very close to the CWT, while the thresholding strategy is only the neighboring shrinkage. The proposed methods preserve the edges of each structure near the discontinuities. This is demonstrated in Figures 9g and 10g where the algorithm incorporating inter-channel thresholding and linear phase smoothing produces a sharper image than the DT-QWT with bivariate shrinkage and neighboring shrinkage for both the Sebkha and Sea images. Vegetation and squares are identified in Figure 9g, and lines of mountains are shown in Figure 10g. In most cases, noise is not entirely removed by our method, but it is significantly reduced and the edges are sharper.

To conclude, we can say that the new formulation of the threshold factor of the quaternionic magnitude coefficients,

Table 4 PSNR (dB) comparison of multichannel image denoising algorithms

Methods	Noise variance					
	15	20	25	30	40	50
Sebkha						
Noisy image	25.14	22.89	21.26	18.73	17.06	15.56
DWT (NeighShrink)	26.09	23.79	22.18	19.63	17.94	16.45
DT-CWT (NeighShrink)	27.25	24.95	23.32	20.77	19.06	17.57
DT-QWT (NeighShrink)	27.29	24.98	23.36	20.79	19.06	17.57
DT-QWT (BiShrink)	28.50	26.19	24.57	21.99	20.25	18.75
DT-QWT (NeighShrink/ PhaseSmooth)	30.09	27.77	26.14	23.55	21.80	20.29
Sea						
Noisy image	25.87	23.16	21.18	18.91	17.14	15.28
DWT (NeighShrink)	26.74	24.03	22.04	19.76	17.97	16.12
DT-CWT (NeighShrink)	27.81	25.10	23.11	20.82	19.02	17.17
DT-QWT (NeighShrink)	27.86	25.15	23.17	20.83	19.02	17.17
DT-QWT (BiShrink)	29.09	26.38	24.39	22.07	20.26	19.21
DT-QWT (NeighShrink/ PhaseSmooth)	30.77	28.06	26.04	23.74	21.92	20.87

the estimation of noise variance based both on spatial and multichannel dependencies and the multiplication of the third-angle quaternionic phase by smoothing matrix have a great impact on satellite image denoising compared to the classical methods and advanced methods which do not use the information contained in the phase. All these experimentations demonstrate that a coherent analysis is associated with the quaternionic wavelet transformation and the potential of this multispectral representation with magnitude thresholding and phase smoothing for noise reduction and features preservation.

6 Conclusions

In this article, we introduce the 2D multiscale quaternionic wavelet transform for satellite image denoising application. We reintroduce the fact that this new representation is particularly efficient for the description of image features and more efficient for the detail representation than the discrete wavelet transform or the complex wavelet transform. As we have reviewed, quaternionic transformation generalizes 1D complex wavelet to higher dimensions and offers more information: a phase feature associated with 'texture' characteristics. Redundancy brought by the QWT phase adds complete structural information about local features of images contrary to the undecimated wavelet transform that is only associated with the translation invariance property.

The QWT is not straightforward to interpret, but here, we gave an application study crossing the gap between that framework and the way to use this tool by showing its superiority over standard wavelets in a denoising context. For this, a denoising method based on the DT-QWT with single-channel and multichannel selective neighboring coefficient thresholding and linear phase smoothing is presented. The proposed algorithm applied both in separate bands and multispectral satellite images reduces noise and keeps the edges sharp.

The obtained results confirm the efficacy of intra-channel and inter-channel dependency in thresholding and the phase regularization in comparison to the term-by-term classical shrinkage algorithm and the bivariate approach. A non-marginal strategy developed in our work outperforms existing methods, both from computational and from a quality point of view. This improvement is due to the shift invariance of the QWT magnitude together with the use of the QWT phase that contains useful structural information for image analysis. The proposed multichannel model has the potential to be extended to hyperspectral images and to introduce more information about phase.

Another question that should be investigated in a future work is the ability of the proposed method to exploit the parent-child relationship or inter-scale dependencies in addition to neighboring intra-scale and inter-channel correlations. Also, it may be possible to use a non-linear dependency of phase and study the relationship between successive phases on different scales.

Endnotes

[a]In this article, we only analyze the invertible discrete representation in order to build a denoising method. For this, the complex continuous wavelet representation (for example, complex Morlet) is not described.

Author details
[1]Laboratory of Telecommunications and Digital Signal Processing, Department of Electronics, Faculty of Technology, Djillali Liabes University, Sidi Bel Abbes 22000, Algeria. [2]Department of Material Science, Faculty of Sciences and Technology, University of Mascara, Mascara 29000, Algeria. [3]Laboratory of Telecommunications and Digital Signal Processing, Department of Electronics, Faculty of Technology, Djillali Liabes University, Sidi Bel Abbes 22000, Algeria. [4]XLIM-SIC Laboratory, Department of Signal, Image, and Communications, XLIM Institute, CNRS UMR 6172, UFR Sciences-SP2MI, University of Poitiers, Futuroscope Chasseneuil 86073 Poitiers CEDEX9, France.

References
1. DL Donoho, IM Johnstone, Ideal spatial adaptation by wavelet shrinkage. Biometrica **81**(3), 425–455 (1994)
2. DL Donoho, IM Johnstone, Adapting to unknown smoothness via wavelet shrinkage. J. Roy. Statist. Soc. **92**(44), 1413–1421 (1997)
3. RH Bamberger, MJT Smith, A filter bank for the directional decomposition of image: theory and design. IEEE Trans. Image Processing **40**(4), 882–893 (1992)
4. M Lang, H Guo, J Odegard, C Burrus, R Wells, Noise reduction using an undecimated discrete wavelet transform. IEEE Signal Processing Lett. **3**(1), 10–12 (1996)
5. NG Kingsbury, *The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters* (IEEE Digital Signal Proc (Workshop on DSP, Bryce Canyon, USA, 1998). pp. 2543–2560
6. NG Kingsbury, *The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement, in the 9th European Signal Processing Conference (EUSIPCO)* (Sept, Rhodes, 1998). pp. 319–322
7. NG Kingsbury, A dual-tree complex wavelet transform with improved orthogonality and symmetry properties, in *Proceedings of IEEE ICIP, Vancouver, 10–13 Sept 2000, vol. 2*, pp. 375–378
8. T Bulow, *Hypercomplex Spectral Signal Representations for the Processing and Analysis of Images* (Christian Albrechts University of Kiel, Dissertation, 1999)
9. EB Corrochano, Multi-resolution image analysis using the quaternion wavelet transform. J. Num. Algo. **39**(1), 35–55 (2005)
10. EB Corrochano, The theory and use of quaternion wavelet transform. J. Math. Imaging Vis. **24**(1), 19–35 (2006)
11. WL Chan, H Choi, R Baraniuk, Quaternion wavelets for image analysis and processing, in *The International Conference on Image Processing, Singapore, 11 Oct 2004, vol. 5*, pp. 3057–3060
12. R Soulard, *Quaternions et algèbres géométriques pour le traitement d'images* (University of Poitiers, France, Dissertation, 2009)
13. R Soulard, P Carré, Quaternionic wavelets for texture classification. Pattern Recog. Lett. **32**, 1669–1678 (2011)

14. S Gai, P Liu, J Liu, X Lang, A new image denoising algorithm via bivariate shrinkage based on quaternion wavelet transform. J. Comput. Inf. Sys. **6**(11), 3751–3760 (2010)
15. TT Cai, BW Silverman, Incorporating information on neighboring coefficients into wavelet estimation. Sankhya Series **63**(2), 127–148 (2001)
16. GY Chen, TD Bui, Multiwavelets denoising using neighboring coefficients. IEEE Signal Processing Lett. **10**(7), 211–214 (2003)
17. S Hailiang, ZI Yanyang, HE Zhengjia, W Xiaodong, Y Jing, Translation-invariant multiwavelet denoising using improved neighbouring coefficients and its application on rolling bearing fault diagnosis, in the 9th International Conference on Damage Assessment of Structures(DAMAS). J. Phys. Conf. Ser. **305**, 012012 (2011)
18. B Chinna Rao, M Madhavi Latha, Selective neighbouring wavelet coefficients approach for image denoising. Int. J. Computer Science Com. **2**(1), 73–77 (2011)
19. GY Chen, TD Bui, A Krzyak, Image denoising with neighbour dependency and customized wavelet and threshold. Pattern Recognition **38**, 115–124 (2005)
20. L Sendur, IW Selesnick, Bivariate shrinkage with local variance estimation. IEEE Signal Processing Lett. **9**(12), 438–441 (2002)
21. J Portilla, V Stela, MJ Wainwright, EP Simoncelli, Image denoising using scale mixture of Gaussians in the wavelet domain. IEEE Trans. Image Processing **12**(11), 1338–1351 (2003)
22. K Dabov, A Foi, V Katkovnik, K Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Processing **16**(8), 2080–2095 (2007)
23. L Zhang, W Dong, D Zhang, G Shi, Two-stage image denoising by principal component analysis with local pixel grouping. Pattern Recognition **43**(4), 151–1549 (2010)
24. BR Corner, M Narayanan, SE Reichenbach, Noise estimation in remote sensing imagery using data masking. Int J Remote Sensing **24**(N4), 689–702 (2003)
25. A Jalobeanu, LB Féraud, J Zerubia, Estimation of blur and noise parameters in remote sensing, in ICASSP. Orlando, FL, USA **13–17**, 3580–3583 (May 2002)
26. F Luisier, T Blu, A new SURE approach to image denoising: interscale orthonormal wavelet thresholding. IEEE Trans. Image Processing **16**(3), 593–606 (2007)
27. F Luisier, T Blu, SURE-LET multichannel image denoising: interscale orthonormal wavelet thresholding. IEEE Trans. Image Processing **17**(4), 482–492 (2008)
28. J Saeedi, MH Moradi, K Faez, A new wavelet-based fuzzy single and multi-channel image denoising. Image Vis. Comput. **28**, 1611–1623 (2010)
29. C Chaux, AB Benyahia, JC Pesquet, Use of Stein's principle for multichannel image processing, in *IEEE-EURASIP International Symposium on Control, Communication. and Signal Processing, Marrakech, Morocco, 13–15 March 2006.*
30. C Stein, Estimation of the mean of a multivariate normal distribution. Ann. Stat. **9**(N6), 1135–1151 (1981)
31. AV Oppenheim, JS Lim, The importance of phase in signals. Proc. IEEE **69**, 529–541 (1981)
32. WL Chan, H Choi, R Baraniuk, Coherent multiscale image processing using dual-tree quaternion wavelets. IEEE Trans. Image Processing **17**(7), 1069–1082 (2008)
33. M Miller, K Kingsbury, Image denoising using derotated complex wavelet coefficients. IEEE Trans. Image Processing **17**(9), 1500–1511 (2009)
34. JA Fessler, DC Noll, Iterative image reconstruction in MRI with separate magnitude and phase regularization. Proc. IEEE Int. Symp. Biomed. Imaging **1**, 209–212 (2004)
35. S Cai, K Li, Bivariate shrinkage functions for wavelet based denoising. http://eeweb.poly.edu/iselesni/WaveletSoftware/denoise2.html
36. Algerian Space Agency. http://www.asal.dz
37. Z Wang, A Bovik, H Sheikh, E Simoncelli, Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Processing **13**(4), 600–612 (2004)