Using BayesShrink, BiShrink, Weighted BayesShrink, and Weighted BiShrink in NSST and SWT for Despeckling SAR Images

Nikou Farhangi and Sedigheh Ghofrani[*]

Abstract

Synthetic aperture radar (SAR) images are inherently degraded by multiplicative speckle noise where thresholding-based methods in the transform domain are appropriate. Being sparse, the coefficients in the transformed domain play a key role in the performance of any thresholding methods. It has been shown that the coefficients of nonsubsampled shearlet transform (NSST) are sparser than those of stationary wavelet transform (SWT) for either clean or noisy images. Therefore, it is expected that thresholding-based methods in NSST outperform those in the SWT domain. In this paper, BayesShrink, BiShrink, weighted BayesShrink, and weighted BiShrink in NSST and SWT domains are compared in terms of subjective and objective image assessment. As BayesShrink try to find the optimum threshold for every subband, BiShrink uses coefficients, name "parent," to clean up coefficients called "child," and the weighted methods consider the coefficients' noise efficiency, which implies that subbands in the transform domain may be affected by noise differently. Two models for considering the parent in the NSST domain are proposed. In addition, for both BayesShrink and BiShrink, considering the weighting factor (coefficients noise efficiency) would improve the performance of the corresponding methods as well. Experimental results show that the weighted-BiShrink despeckling approach in the NSST domain gives an outstanding performance when tested with both artificially speckled images and real SAR images.

Keywords: BayesShrink, BiShrink, Weighted BayesShrink, Weighted BiShrink, Nonsubsampled shearlet transform, Stationary wavelet transform, SAR images despeckling

1 Introduction

Synthetic aperture radar (SAR) can be used in a wide variety of applications in the military, geology, scientific discoveries, mapping, and surveillance of Earth. The main advantages of SAR is its ability to operate under diverse weather conditions such as darkness, rain, snow, fog, and dust, where SAR exhibits speckle noise. It must be stressed that speckle is noise-like, but it is not noise; it is a real electromagnetic measurement. Therefore, when the radar scans a uniform surface, the SAR images emerge as dramatic changes in gray, with some resolution cells shown as a dark spot, and others shown as a bright spot, depicting granular ups and downs. The

spots rooted in a coherent superposition of the radar echo are called speckle noise.

For any coherent imaging like SAR, such as sonar and ultrasound, despeckling is an important process for image enhancement. Removing speckles and preserving edges are the main goals of enhancement approaches. In general, the despeckling of SAR images is carried out in either the spatial or transformed domain [1]. Despite their low computational complexity, the performance of spatial domain filters is often not as well as the transformed domain algorithms [2]. Wavelet [3] is a well-known multiscale transform that can effectively mitigate point singularities for one-dimensional signals. For linear singularities in images, two-dimensional separable wavelets were used. However, the lack of directionality motivated researchers to propose the curvelet [4] and contourlet [5, 6] methods, which use the multiscale transform followed

* Correspondence: s_ghofrani@azad.ac.ir
Electrical and Electronic Engineering Department, Islamic Azad University, South Tehran Branch, No. 9, 14-th Street, Shahid Beheshti Street, Tehran 15317-64611, Iran

by the directional filter bank. Their basis functions with wedge-shaped or rectangular support regions provide good sparse representations for high dimensional singularities. Recently, shearlet transform (ST) based on an affine system [7], which can sparsely represent an image and has flexible orientation, has been proposed [8]. This new representation is based on a simple and rigorous mathematical framework that not only provides a more flexible theoretical tool for the geometric representation of multidimensional data, but is also easy to implement. In addition, shearlet exhibits highly directional sensitivity and is spatially localized [8–10]. ST has been applied in various practical problems such as total variation for denoising [11], deconvolution [12], SAR despeckling [2, 13], and Bayesian shearlet shrinkage for SAR despeckling via sparse representation [14]. Further, Markarian and Ghofrani [15] proposed a new method based on compressive sensing for speckle reduction of SAR images. However, image processing and video coding have made remarkable progress in recent years [16, 17].

Thresholding is a common method for denoising in the transform domain [18], where finding the optimum threshold value is the main problem. Regarding the methods, VISUShrink [19] obtains the universal threshold value, whereas SUREShrink [20], BayesShrink [21, 22], and bivariate shrinkage (BiShrink) [23–26] obtain the threshold values adaptively for every subband. Among these approaches, BayesShrink is a well-known method used in the nonsubsampled shearlet transform (NSST) domain [27], and BiShrink functions using the Bayesian estimation theory applied in the wavelet [20, 23, 28], contourlet [24, 29], and shearlet [25] domains are used as well.

In this paper, we compare the performances of BayesShrink, BiShrink, weighted BayesShrink, and weighted BiShrink in NSST and stationary wavelet transform (SWT) domains in terms of subjective and objective image assessment. As BayesShrink tries to find the optimum threshold for every subband, BiShrink uses coefficients named parent to clean up coefficients called child, and the weighted methods consider the coefficients' noise efficiency, which imply that the subbands in the transform domain may be affected by noise differently. Two models for considering the parent in the NSST domain are proposed. In addition, for both BayesShrink and BiShrink, considering the weighting factor (coefficients noise efficiency) would improve the performance of the corresponding methods as well. The novel Bishrink despeckling method named BI-NSST is developed, and the weighted Bishrink are used in NSST and SWT domains for the first time, where the approaches are named WBI-NSST and WBI-SWT, respectively. Considering the coefficients' noise efficiency in SWT and NSST to obtain the weighting factor and the optimum threshold value is the main

contribution of this paper. However, the performance of three proposed methods in SWT and four new approaches in NSST are compared with five state-of-the-art papers ([2, 13, 15, 27, 30]) in terms of subjective and objective image evaluations when artificially speckled and real SAR images are denoised.

The paper is organized as follows: Section 2 presents the preliminary of the speckle noise model, the BayesShrink and BiShrink methods, as well as noise estimations and signal variances. Further, ST and the following NSST are explained in Section 3. The proposed methods based on BiShrink in the NSST domain are explained in Section 4. Section 5 shows the experimental results and finally, Section 6 concludes the paper.

2 Threshold-based SAR image despeckling

Determining the optimum threshold value is the main problem in any thresholding-based method. VISUShrink [19] obtains the universal threshold value whereas SUREShrink [20], BayesShrink [21, 22, 28], and BiShrink [23–26, 31] determine the adaptive threshold values for every subband. In the following, the speckle noise model is introduced, the BayesShrink and BiShrink methods are explained, and finally, the median for power noise estimation in the transformed domain is expressed.

2.1 Speckle noise model

In general, for any coherent imaging systems such as SAR, multiplicative speckle noise degrades the image. The speckle noise is modeled as,

$$I_y = I_x + I_x N_s \tag{1}$$

where I_x and I_y refer to the noise free signal and the observed signal, and N_s is speckle noise in spatial domain. Writing Eq. (1) as, $I_y = I_x(1 + N_s)$. in order to convert the multiplicative noise into additive noise, the homomorphic framework is used, which means using the logarithm transform before processing and the exponential transform at the end (see Fig. 1),

$$y = x + n \tag{2}$$

where $y = \log(I_y)$, $x = \log(I_x)$, and $n = \log(1 + N_s)$ are in order noisy signal, noise free signal, and additive noise.

2.2 BayesShrink denoising in transformed domain

As said above, the multiplicative noise is converted into the additive noise by using logarithm. So, applying a linear transform to Eq. (2), we have,

$$Y_k = X_k + N_k \tag{3}$$

where k refers to the decomposition subband, Y_k, X_k, and N_k are noisy, clean, and noise coefficients in

Fig. 1 Block diagram of BiShrink despeckling in transformed domain
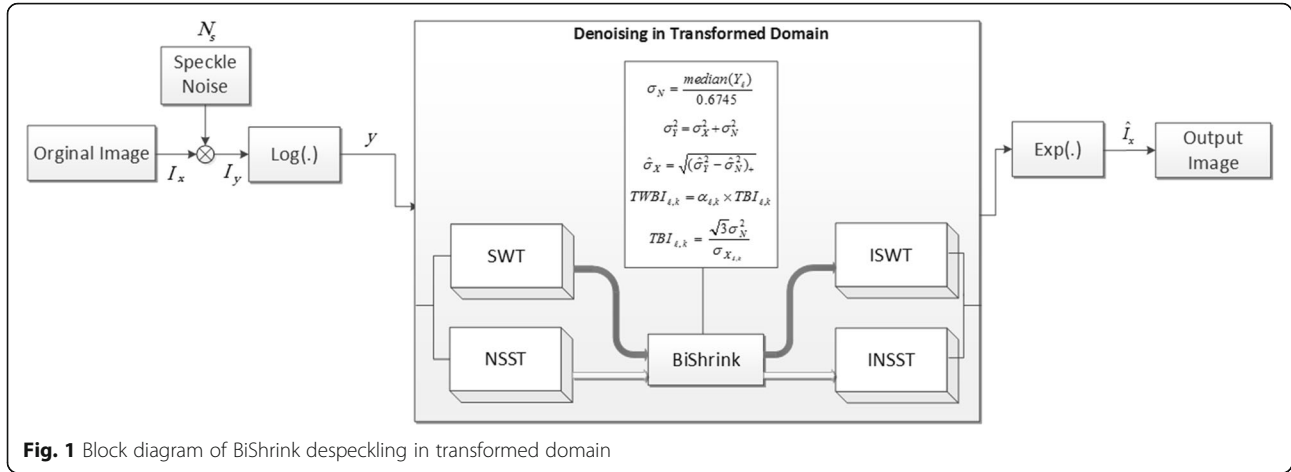
order. The goal of Bayesian denoising method [21] is estimating $\hat{X}_k$, from the observed data, Y_k. In order to simplify the notations, the superscript k, which indicates the subband, is dropped off (i.e., $Y = X + N$). For Eq. (3), Bayesian maximum a posteriori (MAP) estimator is [5, 23, 32],

$$\hat{X}(Y) = \arg \max_X [p_{X|Y}(X|Y)] \qquad (4)$$

According to the Bayes rule, the conditional probability density function (PDF) is $p(X|Y) = \frac{p(Y|X)p(X)}{p(Y)}$; therefore, ignoring $p(Y)$ because of being constant, $\hat{X}(Y)$ is,

$$\hat{X}(Y) = \arg \max_X \left[p_{Y|X}(Y|X)p_X(X) \right]$$
$$= \arg \max_X [p_N(Y-X)p_X(X)] \qquad (5)$$

where p_X is the prior distribution of the noise free coefficients and p_N is the noise PDF assumed zero-mean Gaussian with variance σ_N^2 [13, 27], $p_N(N) = \frac{1}{\sigma_N \sqrt{2\pi}} \exp\left(-\frac{N^2}{2\sigma_N^2}\right)$.

By applying the logarithm function, Eq. (5) is written as,

$$\hat{X}(Y) = \arg \max_X \left[-\frac{(Y-X)^2}{2\sigma_N^2} + f(X) \right] \qquad (6)$$

where $f(X) = \log(p_X(X))$ [33]. Finding $\hat{X}$ is equivalent to solve $\frac{Y-\hat{X}}{\sigma_N^2} + f'(\hat{X}) = 0$.

If $p_X(X)$ is assumed Gaussian with zero mean and variance σ^2, then $f(X) = -\log\left(\sqrt{2\pi}\sigma\right) - X^2/2\sigma^2$ and the estimated $\hat{X}(Y)$ is [23, 34],

$$\hat{X}(Y) = \frac{\sigma^2}{\sigma^2 + \sigma_N^2} Y \qquad (7)$$

If $p_X(X)$ is assumed Laplace with zero mean and variance σ^2, $p_X(X) = \frac{1}{\sqrt{2}\sigma} \exp\left(\frac{-\sqrt{2}|X|}{\sigma}\right)$, then $f(X) = -\log\left(\sigma\sqrt{2}\right) - \sqrt{2}|X|/\sigma$ and the estimated $\hat{X}(Y)$ is [23],

$$\hat{X}(Y) = \text{sign}(Y)\left(|Y| - \frac{\sqrt{2}\sigma_N^2}{\sigma} \right)_+ \qquad (8)$$

Eq. (8) is the classical soft shrinkage function [33] defined as,

$$\text{soft}(g, \tau) = \text{sign}(g)(|g| - \tau)_+ \qquad (9)$$

$$(|g| - \tau)_+ = \begin{cases} |g| - \tau & \text{if } |g| \geq \tau \\ 0 & \text{if } |g| < \tau \end{cases} \qquad (10)$$

Based on the classical soft shrinkage function, Eq. (8) is rewritten $\hat{X}(Y) = \text{soft}\left(Y, \frac{\sqrt{2}\sigma_N^2}{\sigma}\right)$.

2.3 BiShrink denoising in transformed domain

In 2002, a novel denoising method named BiShrink was proposed [23, 35]. In fact, BiShrink is a simple nonlinear shrinkage function in the transformed domain, with subbands known as child and parent. To obtain the threshold value for denoising a child subband, BiShrink also uses the coefficients of the parent subband.

Suppose X_1 and X_2 as child and parent coefficients in transformed domain, then the vector form of Eq. (3) is,

$$Y = X + N \qquad (11)$$

where $X = (X_1, X_2)$, $Y = (Y_1, Y_2)$, and $N = (N_1, N_2)$. Similar to the Bayesian MAP estimator explained in Section 2.2, the PDFs of the clean coefficients $p_X(X)$ and the noise coefficients $p_N(N)$ are to be known for estimating the noise free coefficients $\hat{X}(Y)$ from the observed data. In literatures [23, 24, 36], noise is often assumed Gaussian,

$$p_N(N) = \frac{1}{2\pi\sigma_N^2} \exp\left(-\frac{N_1^2 + N_2^2}{2\sigma_N^2}\right) \qquad (12)$$

and in [23], the noise-free coefficients $p_X(X)$ were given by,

$$p_X(X) = \frac{3}{2\pi\sigma^2} \exp\left(-\frac{\sqrt{3}}{\sigma}\sqrt{X_1^2 + X_2^2}\right) \quad (13)$$

According to Eq. (6), the bivariate MAP estimator is,

$$\hat{X}(Y) = \arg \max_{X_1, X_2} \left[-\frac{(Y_1 - X_1)^2}{2\sigma_N^2} - \frac{(Y_2 - X_2)^2}{2\sigma_N^2} + f(X)\right] \quad (14)$$

Estimating $\hat{X}_1$ and $\hat{X}_2$ needs solving equations, $\frac{Y_1 - \hat{X}_1}{\sigma_N^2} + f_1(\hat{X}) = 0$ and $\frac{Y_2 - \hat{X}_2}{\sigma_N^2} + f_2(\hat{X}) = 0$ where f_1 and f_2 represent the derivative of $f(X)$ with regard to X_1 and X_2, respectively. According to Eq. (13), we have $f(X) = \log\left(\frac{3}{2\pi\sigma^2}\right) - \frac{\sqrt{3}}{\sigma}\sqrt{X_1^2 + X_2^2}$, $f_1(X) = \frac{df(X)}{dX_1} = -\frac{\sqrt{3}X_1}{\sigma\sqrt{X_1^2 + X_2^2}}$, and $f_2(X) = \frac{df(X)}{dX_2} = -\frac{\sqrt{3}X_2}{\sigma\sqrt{X_1^2 + X_2^2}}$. Under assuming that the noise power of different subbands is invariant [36], i.e., $\sigma_{N_1} = \sigma_{N_2} = \sigma_N$, and defining $r = \sqrt{X_1^2 + X_2^2}$, we have:

$$Y_1 = \hat{X}_1\left(1 + \frac{\sqrt{3}\sigma_N^2}{\sigma r}\right) \quad (15)$$

$$Y_2 = \hat{X}_2\left(1 + \frac{\sqrt{3}\sigma_N^2}{\sigma r}\right) \quad (16)$$

$$r = \left(\sqrt{Y_1^2 + Y_2^2} - \frac{\sqrt{3}\sigma_N^2}{\sigma}\right)_+ \quad (17)$$

Substituting Eq. (17) into Eq. (15), the joint shrinkage function is [36],

$$\hat{X}_1(Y) = \frac{\left(\sqrt{Y_1^2 + Y_2^2} - \frac{\sqrt{3}\sigma_N^2}{\sigma}\right)_+}{\sqrt{Y_1^2 + Y_2^2}} Y_1 \quad (18)$$

In BiShrink denoising method based on Eq. (18), dead zone region where $\hat{X}_1(Y) = 0$ is,

$$\text{Dead zone} = (Y_1,, Y_2) : (\sqrt{Y_1^2 + Y_2^2} \leq \frac{\sqrt{3}\sigma_N^2}{\sigma}\} \quad (19)$$

To compare the NSST and SWT BiShrink functions and to observe the circular shape dead zone, we have chosen the "Barbara" image of size 512×512 pixels and 256 Gy levels. The test image is corrupted by the additive Gaussian noise with two noise levels ($\sigma_N = 10$, 30), and the corresponding dead zone region based on Eq. (19) are shown in Fig. 2. As shown in Eq. (19) and observed in the image, a direct relationship exists between the radius of the circular-like dead zone and the noise power. Since NSST is sparser than SWT, it yields the smaller radius of dead zone.

2.4 Estimating noise and signal variance

The noise and signal variance $\hat{\sigma}_N^2$ and $\hat{\sigma}^2$, respectively, are required for the implementation both the BayesShrink [21] and BiShrink [23] algorithms. Using the median estimator [34] in the transformed domain for noise variance estimation is common:

$$\hat{\sigma}_N = \frac{\textbf{median}(|Y_\ell|)}{0.6745} \quad (20)$$

where Y_ℓ refers to the coefficients of ℓth decomposition level. For any arbitrary subband in the transformed domain, $\sigma_Y^2 = \sigma^2 + \sigma_N^2$. Since the observed signal in the transformed domain is modeled as zero mean [27], the variance of each subband is obtained by

$$\hat{\sigma}_Y^2 = \frac{1}{M}\sum_{i,j=1}^{M}(Y_\ell^2)_{i,j} \quad (21)$$

where M is the size of square-shaped window. The variances of every window are obtained, and the average value is computed. Using $\hat{\sigma}_Y^2$ and $\hat{\sigma}_N^2$, the signal standard deviation $\hat{\sigma}$ is [33]

$$\hat{\sigma} = \sqrt{(\hat{\sigma}_Y^2 - \hat{\sigma}_N^2)_+} \quad (22)$$

3 Nonsubsampled shearlet transform (NSST)

A novel multi-scale directional representation system called shearlets was proposed in 2005 [7]. Two properties, multi-resolution and sparsity, render the ST attractive in science and engineering [10, 11, 37]. In the following, the continuous and discrete STs are explained briefly.

The continuous shearlet transform for an arbitrary signal f is:

$$S_f(a, s, t) = \langle f, \psi_{ast}\rangle \quad (23)$$

where $a \in R^+$, $s \in R$, and $t \in R^2$ refer to the scaling, shearing, and translation parameters, respectively, and the shearlet function is given by $\psi_{ast}(x) = a^{-3/4}\psi(M_{as}^{-1}(x-t))$ and $M_{as} = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix} = \begin{pmatrix} a & \sqrt{a}s \\ 0 & \sqrt{a} \end{pmatrix}$.

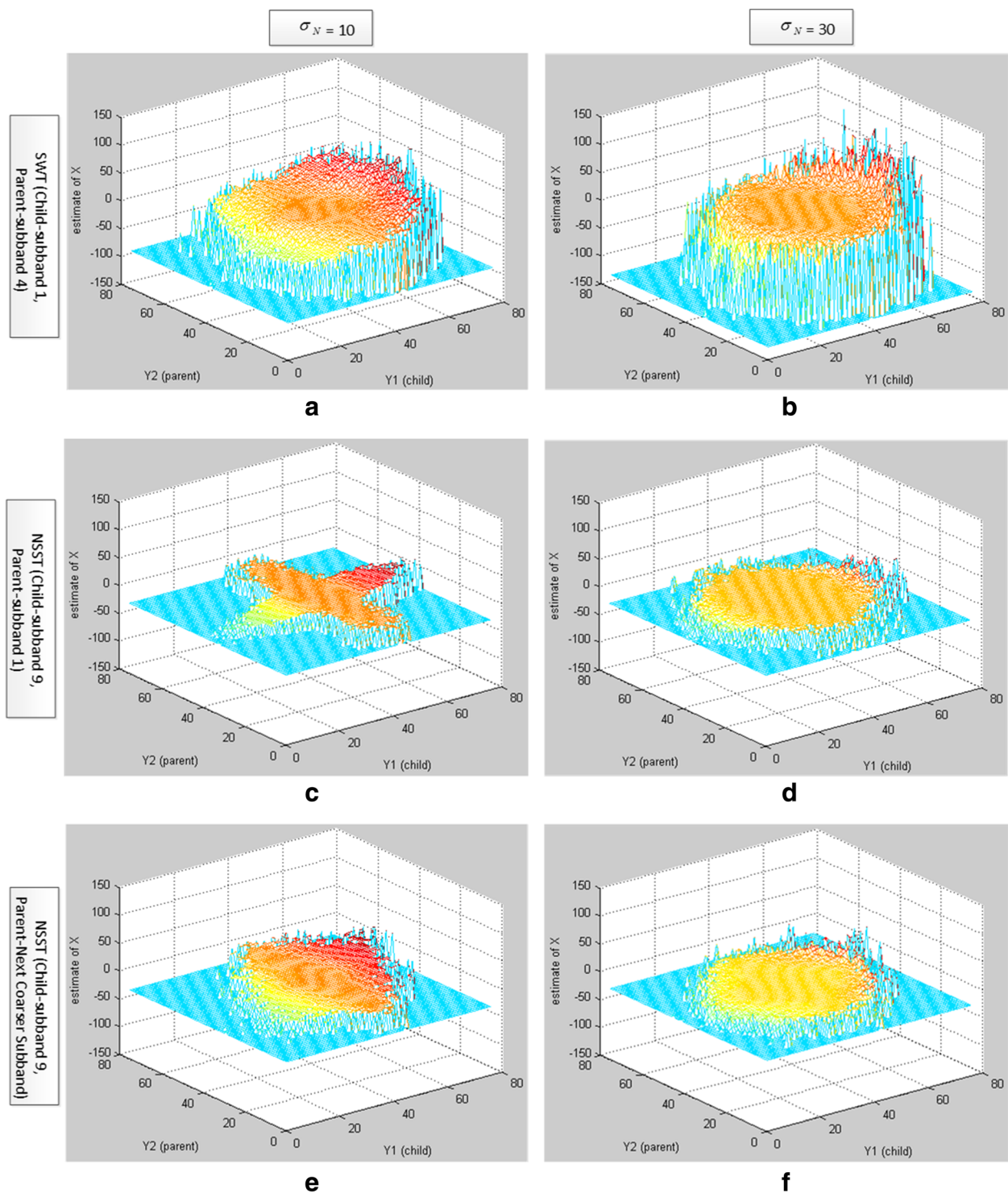The discrete version of ST [10] used in digital signal processing is

$$S_f(j, \ell, k) = \langle f, \psi_{j,\ell,k}\rangle \quad (24)$$

where $j, \ell \in Z$, $k \in Z^2$, $\psi_{j,\ell,k} = |\det A_0|^{j/2}\psi(B_0^j A_0^\ell x - k)$, and

$$A_0 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \text{ and } B_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

In general, using subsampling operations causes variant shifts in a transform. Therefore, by omitting the up- and down-sampling blocks, SWT [30] in 2003 and NSST [10] in 2008 were proposed. In the

Fig. 2 a–f Dead zone of BiShrink for Barbara in SWT and NSST domains and for two noise variances

nonsubsampled transform version, as the coefficients do not decimate between the decomposition levels, all subband sizes are the same as the original input image. Therefore, SWT and NSST require more computation and storage room in comparison with the conventional WT and ST.

NSST, as a multi-scale directional representation, is able to give a good sparse representation of an image. To show the sparsity of NSST in contrast with SWT, the Barbara of size 512×512 pixels and 256 Gy levels is used. The original test image and the one corrupted by Gaussian noise are decomposed by SWT and NSST into three levels. As shown in Fig. 3, for SWT in each decomposition level, there are three subbands, whereas for NSST there are 16, eight, and four subbands at the first, second, and third levels, respectively. Figure 3 shows the histogram of SWT (1st and 4th subbands) and NSST (9th and 1st subbands, all subbands to the next coarser) coefficients. The histograms in Fig. 3 conclude that the NSST is

sparser than the SWT for both noisy and noise-free images. Sparsity, which means that most coefficients are approximately zero, plays a key role in any thresholding algorithm. The sparsity is a typical characteristic of the transform domain where noise is uniformly spread throughout all coefficients, and the data is represented by a small subset of big coefficients [38]. Therefore, coefficients with small magnitudes can be considered as noise and set to zero. The approach in which each coefficient is compared with a threshold in order to decide whether it constitutes a desired part of the original data is called the thresholding approach. Obviously, using a thresholding approach in a sparser transform is outstanding; since the

Fig. 3 Histogram for noise-free and noisy ($\sigma_N^2 = 0.05$) coefficients in SWT and NSST domains
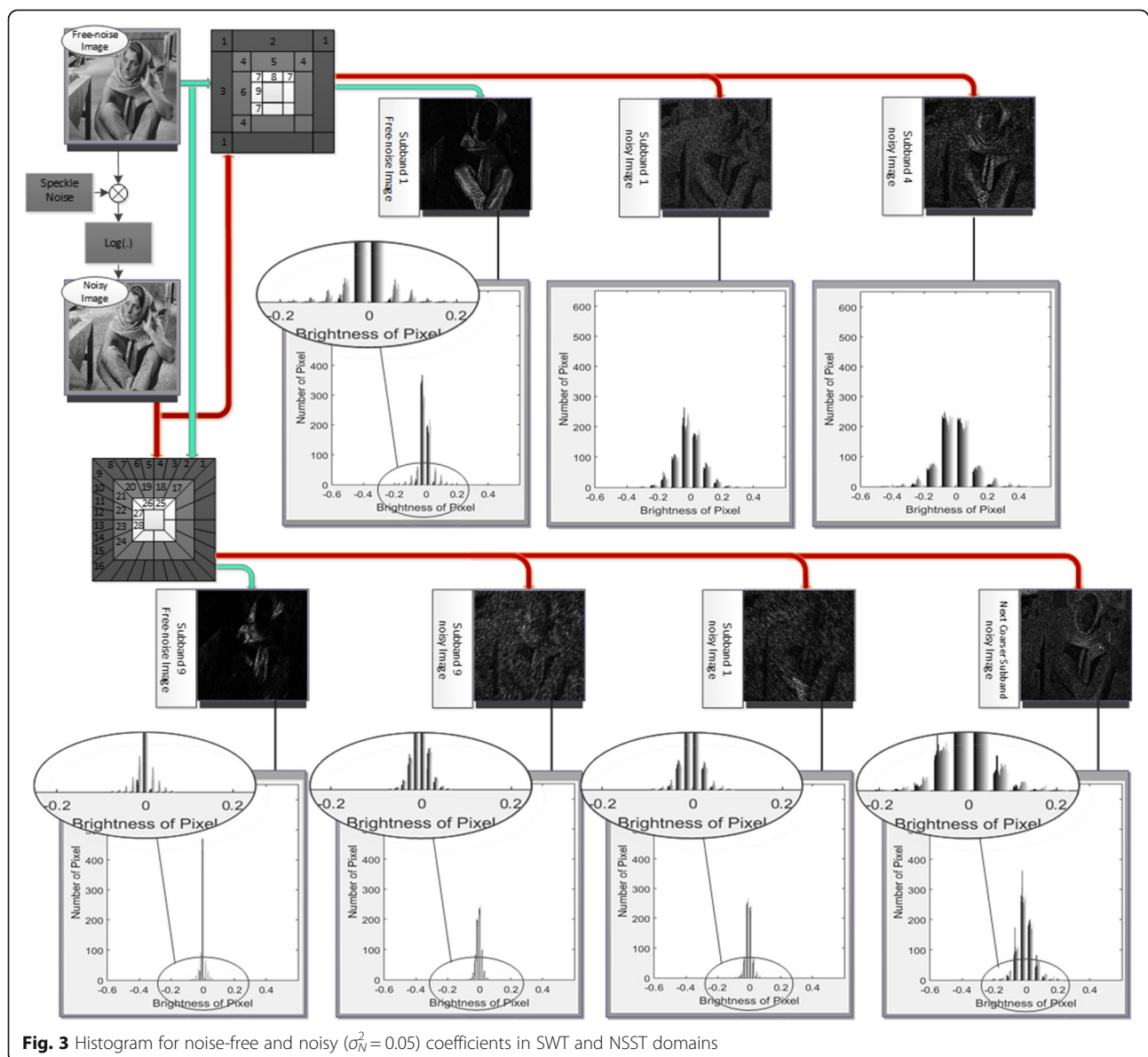
Table 1 The Sd parameter for noise-free image and three different noise power

|	Sd	Free noise	$\sigma_N^2 = 0.05$	$\sigma_N^2 = 0.1$	$\sigma_N^2 = 0.15$
Barbara	SWT	0.1864	0.1936	0.1969	0.1986
	NSST	0.1563	0.1760	0.1812	0.1843
Lena	SWT	0.1731	0.1913	0.1950	0.1969
	NSST	0.1366	0.1759	0.1813	0.1841

basic functions of the NSST have multi-directional wedge-shaped support regions [2], it provides better sparse representation than the SWT.

In this paper, in addition to showing the histograms, we also compute the average standard deviation (Sd) for an objective comparison between NSST and SWT in terms of sparsity. In this regard, the Sd

parameter of all normalized subbands are obtained and the average Sd for noise-free and noise-variant "Lena" and Barbara images with sizes 512 × 512 are presented in Table 1. Increasing the power noise corresponds to a bigger Sd value. However, for both test images, and under different noise powers, the NSST is sparser than the SWT. Therefore, as anticipated, the performance of the threshold-based denoising filter is better in the NSST domain than the SWT domain.
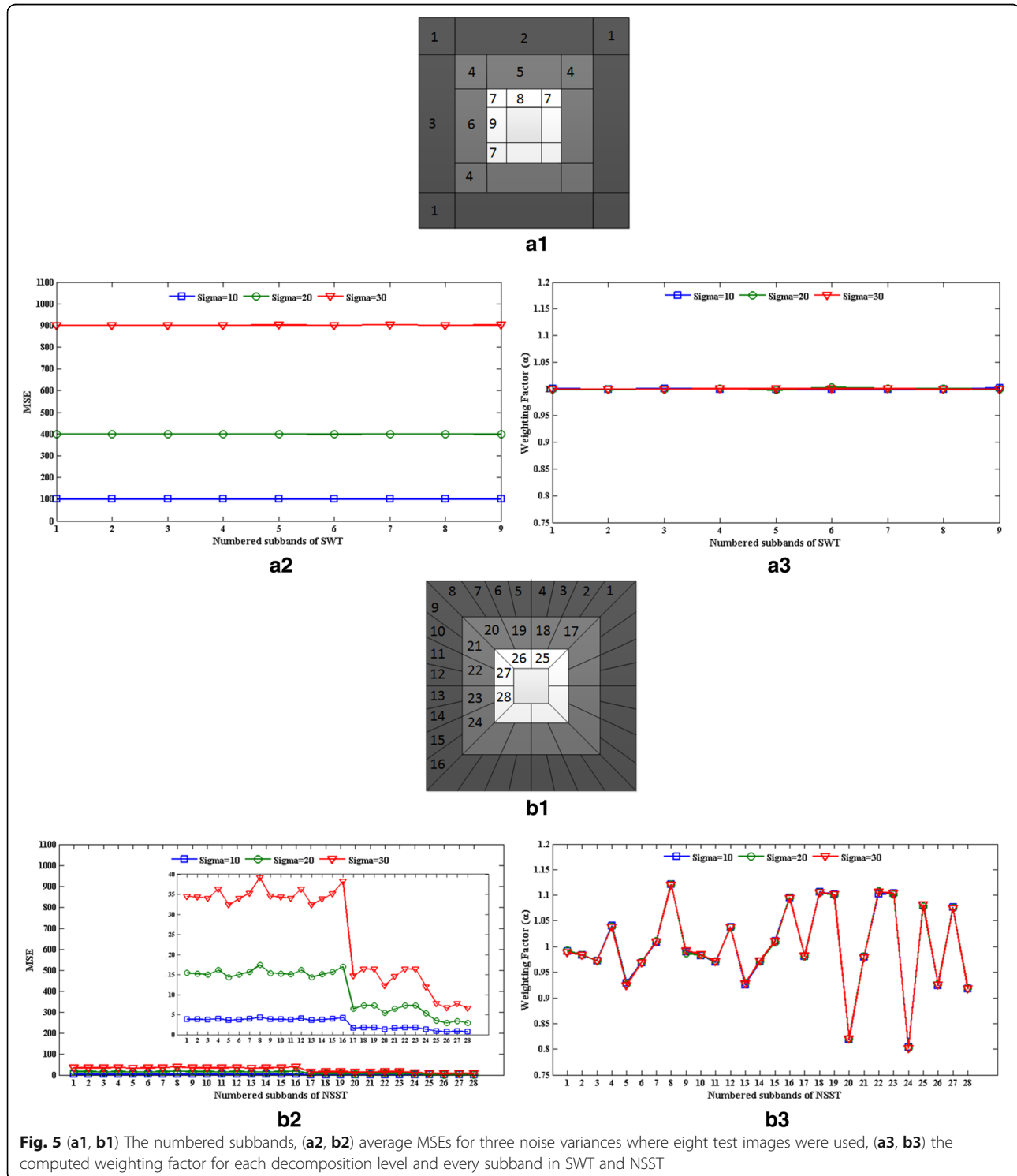
4 Proposed methods
In the first part of this section, the image assessment parameters to evaluate denoising methods are explained, and the mutual information (MI) to
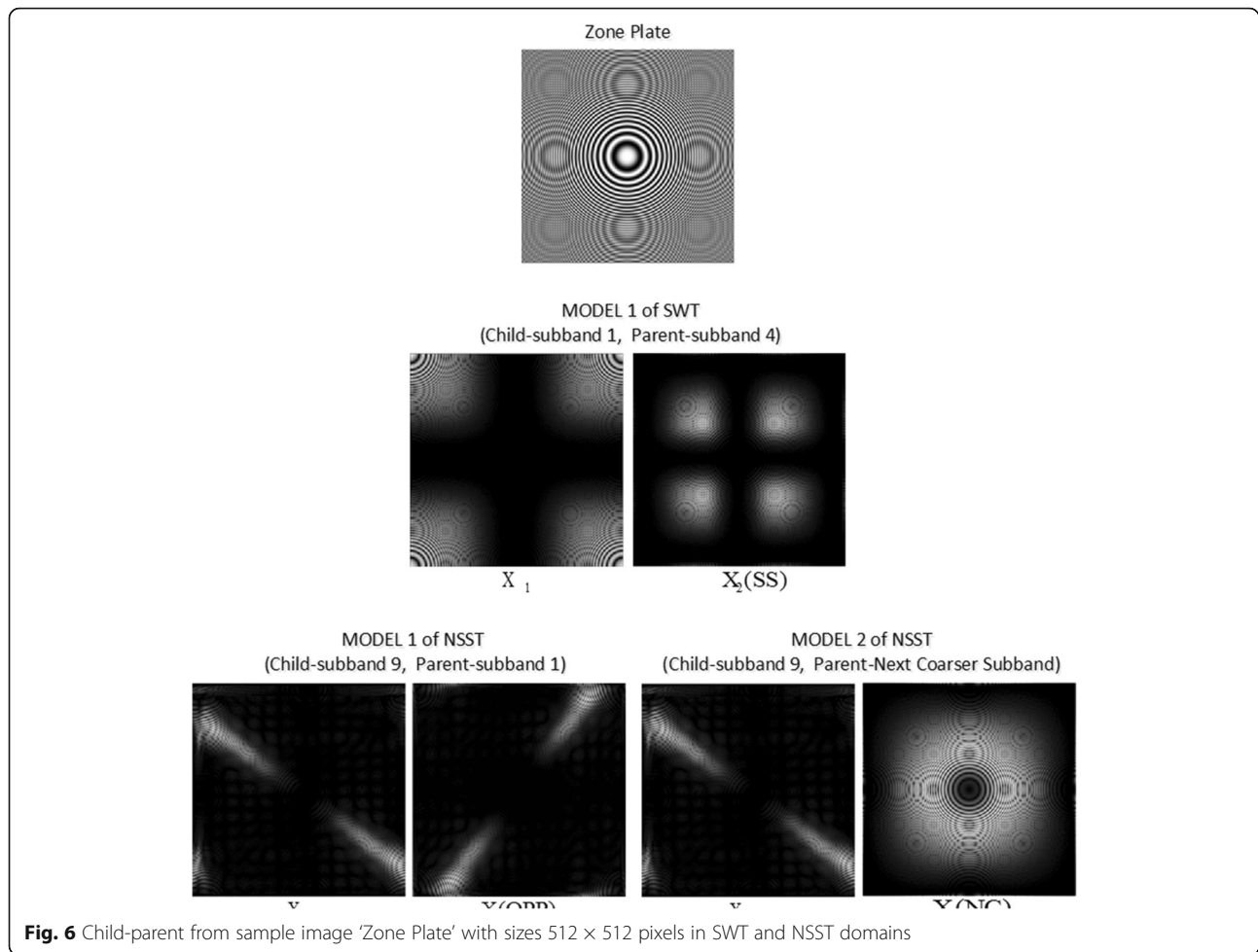

Fig. 4 MI for **a** SWT and **b** NSST
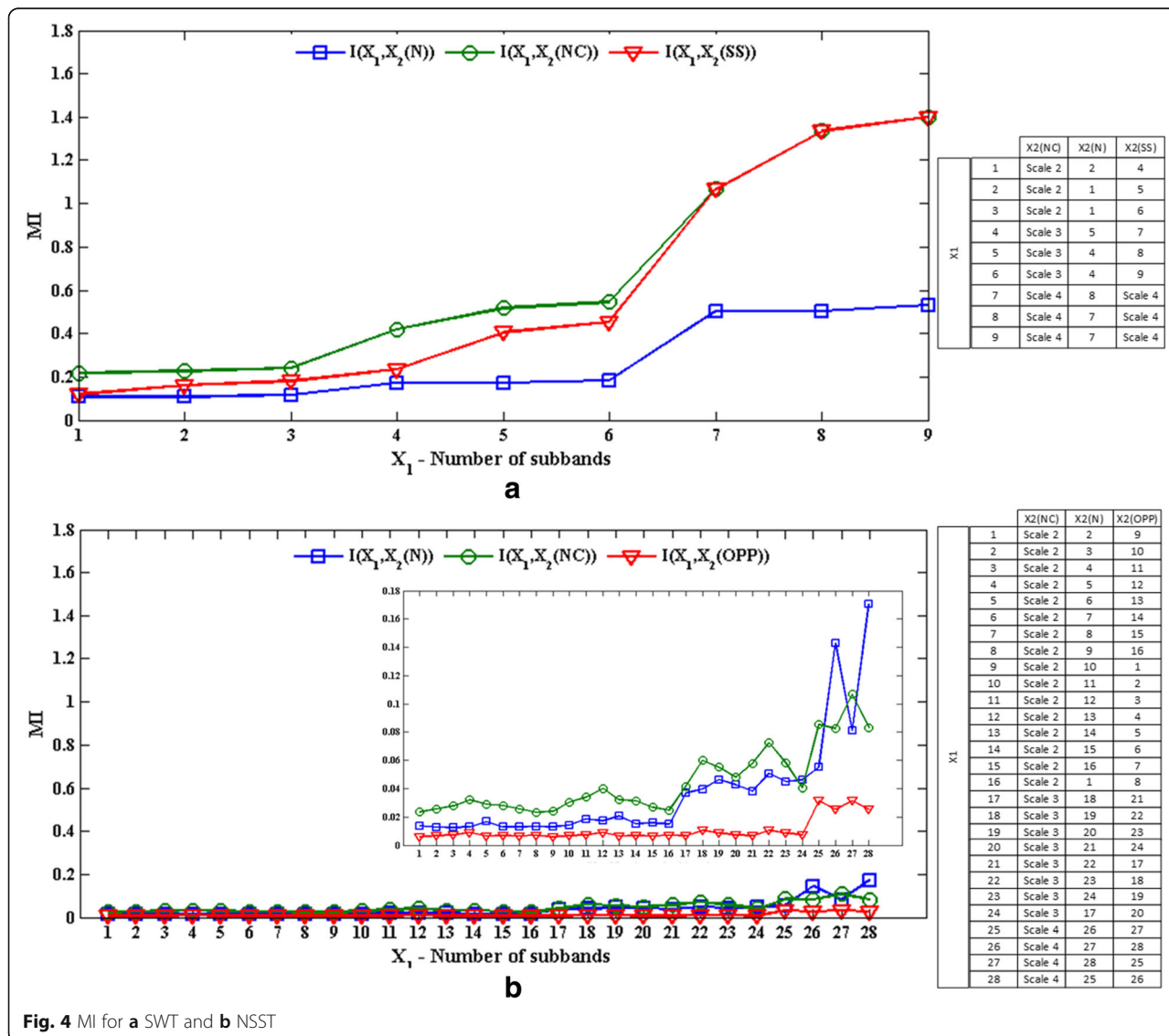
measure the statistical dependency between a child and its corresponding parent coefficients is expressed. Subsequently, the models for BiShrink in the transformed domain are proposed and the weighted BiShrink in the NSST and SWT domains are applied for the first time.

For Figs. 4 and 5, we have used eight test images, named Barbara, Lena, House, Boat, Goldhill, Fingerprint, Cameraman, and Peppers, of size 512×512 pixels and 256 Gy levels. To achieve measurement reliability in a noisy environment, the algorithm was run 10 times for
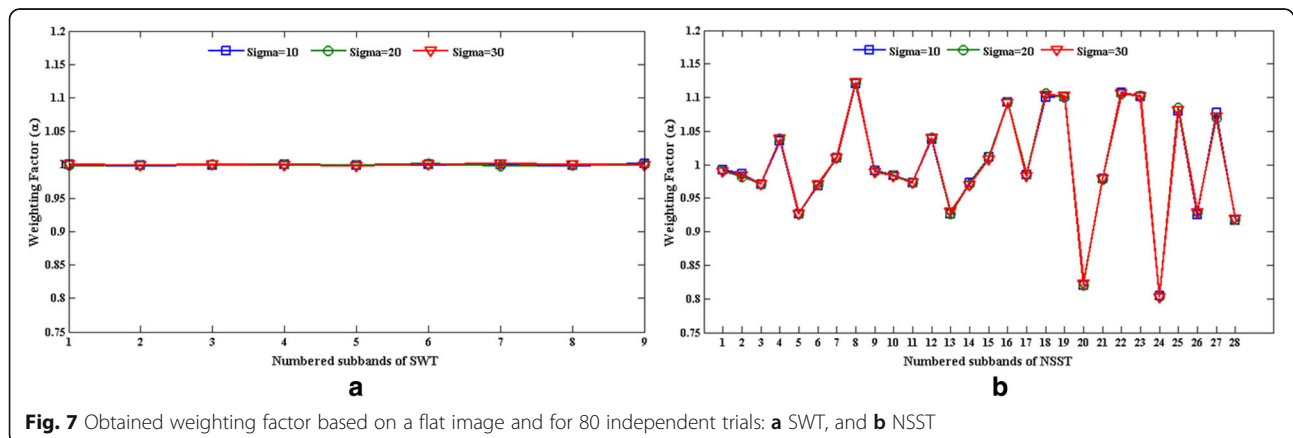

Fig. 5 (**a1**, **b1**) The numbered subbands, (**a2**, **b2**) average MSEs for three noise variances where eight test images were used, (**a3**, **b3**) the computed weighting factor for each decomposition level and every subband in SWT and NSST

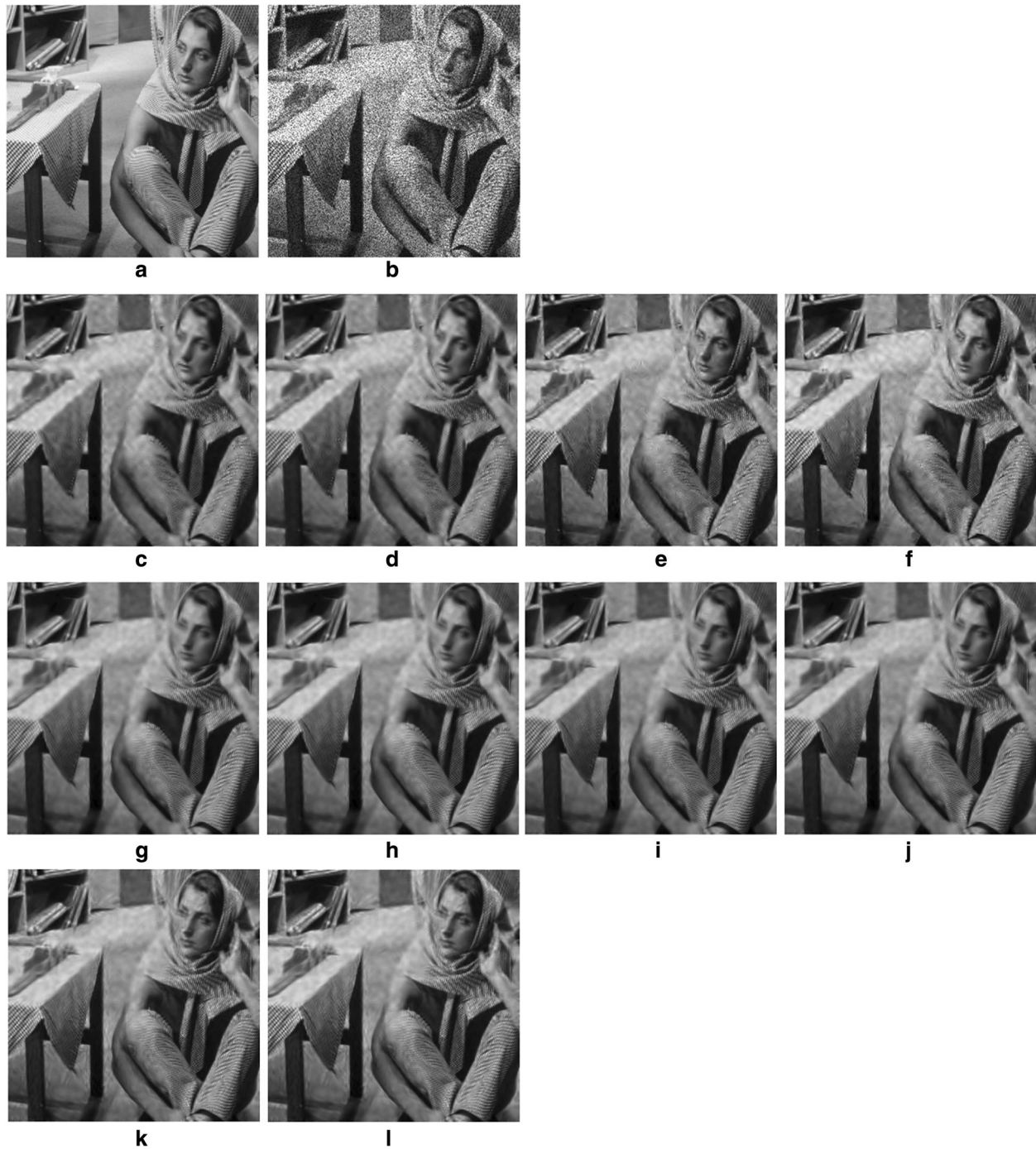**Fig. 6** Child-parent from sample image 'Zone Plate' with sizes 512 × 512 pixels in SWT and NSST domains

the processing of every image; therefore, the average values for 80 (i.e., 8×10) completely independent trials are obtained.

Figure 1 shows the block diagram of the BiShrink despeckling method in the NSST and SWT domains.

In general, the BiShrink denoising algorithm consists of a three-step process: estimating $\hat{\sigma}_N$ for every subband according to Eq. (20), estimating $\hat{\sigma}_Y^2$ and $\hat{\sigma}^2$ based on Eqs. (21) and (22), and obtaining the noise-free coefficients using Eq. (18).

Fig. 7 Obtained weighting factor based on a flat image and for 80 independent trials: **a** SWT, and **b** NSST

Fig. 8 a Barbara as the original test image, **b** noisy image ($\sigma_N^2 = 0.1$), **c–f** despeckled images by B-SWT [30], WB-SWT, BI-SWT, WBI-SWT, and **g–l** despeckled images by B-NSST [2], WB-NSST [27], BI-NSST(1), WBI-NSST(1), BI-NSST(2), and WBI-NSST(2)

4.1 Image assessment parameters

Among the image assessment parameters used to evaluate the performance of a despeckling algorithm, in this paper, we have chosen the peak signal-to-noise ratio (PSNR) [39] and structural similarity (SSIM) [40] as full references and equivalent number of looks (ENL) [13], respectively; and mean square difference (MSD) [41] and edge save index (ESI) [13] as no references.

PSNR measures an image quality:

$$\text{PSNR} = 20 \log_{10}\left(\frac{256}{\sqrt{MSE}} \right) \tag{25}$$

Table 2 Two objective assessment parameters for Barbara to compare with methods in SWT and NSST domains

		$\sigma_N^2 = 0.05$		$\sigma_N^2 = 0.1$		$\sigma_N^2 = 0.15$	
		PSNR (dB) ↑	SSIM ↑	PSNR (dB) ↑	SSIM ↑	PSNR (dB) ↑	SSIM ↑
SWT	B-SWT [30]	26.2485	0.999962	24.5411	0.999883	23.5183	0.999734
	WB-SWT	26.3588	0.999963	24.5631	0.999886	23.6646	0.999738
	BI-SWT	27.6081	0.999963	25.4481	0.999894	24.0742	0.999740
	WBI-SWT	27.6893	0.999963	25.4621	0.999894	24.1642	0.999741
NSST	B-NSST [2]	28.2209	0.999969	26.1606	0.999894	24.7003	0.999747
	WB-NSST [27]	28.2413	0.999969	26.1921	0.999894	24.7371	0.999746
	BI-NSST(1)	28.2823	0.999969	26.2552	0.999894	24.8769	0.999742
	WBI-NSST(1)	28.2971	0.999969	26.2861	0.999894	24.9112	0.999744
	BI-NSST(2)	28.6433	0.999974	26.5448	0.999904	25.0982	0.999764
	WBI-NSST(2)	28.6819	0.99974	26.5694	0.999905	25.1537	0.999763

where $\text{MSE} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \left[\hat{I}_x(i,j) - I_x(i,j) \right]^2$ is the mean square error (MSE), mn is the image size, and I_x and $\hat{I}_x$ are the input and retrieved images, respectively (see Fig. 1 for the mentioned notations).

The SSIM index measures the similarity between the original and the despeckled image through a local statistical analysis (i.e., mean, variance, and covariance between the unfiltered and despeckled pixel values from the sliding window). SSIM ∈ (-1, 1) and a bad similarity between the original and the despeckled image corresponds to SSIM → - 1, whereas a good similarity will be indicated by values SSIM → 1.

ENL and MSD both measure the speckle suppression:

$$\text{ENL} = \frac{\overline{\hat{I}_x}^2}{\text{NV}} \quad (26)$$

$$\text{MSD} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \left[\hat{I}_x(i,j) - I_y(i,j) \right]^2 \quad (27)$$

where $\text{NV} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \left[\hat{I}_x(i,j) - \overline{\hat{I}_x} \right]^2$ and $\overline{\hat{I}_x} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \hat{I}_x(i,j)$. As ENL carries no information about the image resolution degradation, it is often used jointly with

other parameters such as MSD. Large values for ENL and MSD indicate significant filtering. As ENL is to be computed over a uniform region, the image is divided into cells of size 16 × 16 and 25 × 25 pixels, where the ENL is computed for every block and finally, averaged to obtain the ENL value.

The ESI [13] reflects the edge preservation capability of a despeckling technique and is measured in both the horizontal and vertical directions:

$$\text{ESI}^h = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n-1} | \hat{I}_x(i, j+1) - \hat{I}_x(i,j) |}{\sum_{i=1}^{m} \sum_{j=1}^{n-1} | I_y(i, j+1) - I_y(i,j) |} \quad (28)$$

$$\text{ESI}^v = \frac{\sum_{j=1}^{n} \sum_{i=1}^{m-1} | \hat{I}_x(i+1, j) - \hat{I}_x(i,j) |}{\sum_{j=1}^{n} \sum_{i=1}^{m-1} | I_y(i+1, j) - I_y(i,j) |} \quad (29)$$

MI [25, 42–44], a parameter for measuring the dependency between X_1 as the child and X_2 as the parent, is expressed as

Table 3 Obtained PSNRs for denoising the eight test images in NSST domain when $\sigma_n^2 = 0.1$. The algorithm was run 30 times for every image, and the average PSNR is reported

		PSNR(dB)							
		Barbara	Lena	House	Boat	Goldhill	Fingerprint	Cameraman	Peppers
NSST	B-NSST	26.16	28.27	28.41	25.89	26.13	20.66	28.32	27.75
	WB-NSST	26.19	28.28	28.41	25.86	26.10	20.68	28.32	27.76
	BI-NSST(1)	26.25	28.42	28.52	26.06	26.28	20.81	28.39	27.77
	WBI-NSST	26.28	28.42	28.52	26.05	26.25	20.84	28.40	27.78
	BI-NSST(2)	26.54	28.47	28.46	26.33	26.74	22.06	28.01	27.50
	WBI-NSST(2)	26.56	28.50	28.52	26.34	26.72	22.08	28.03	27.52
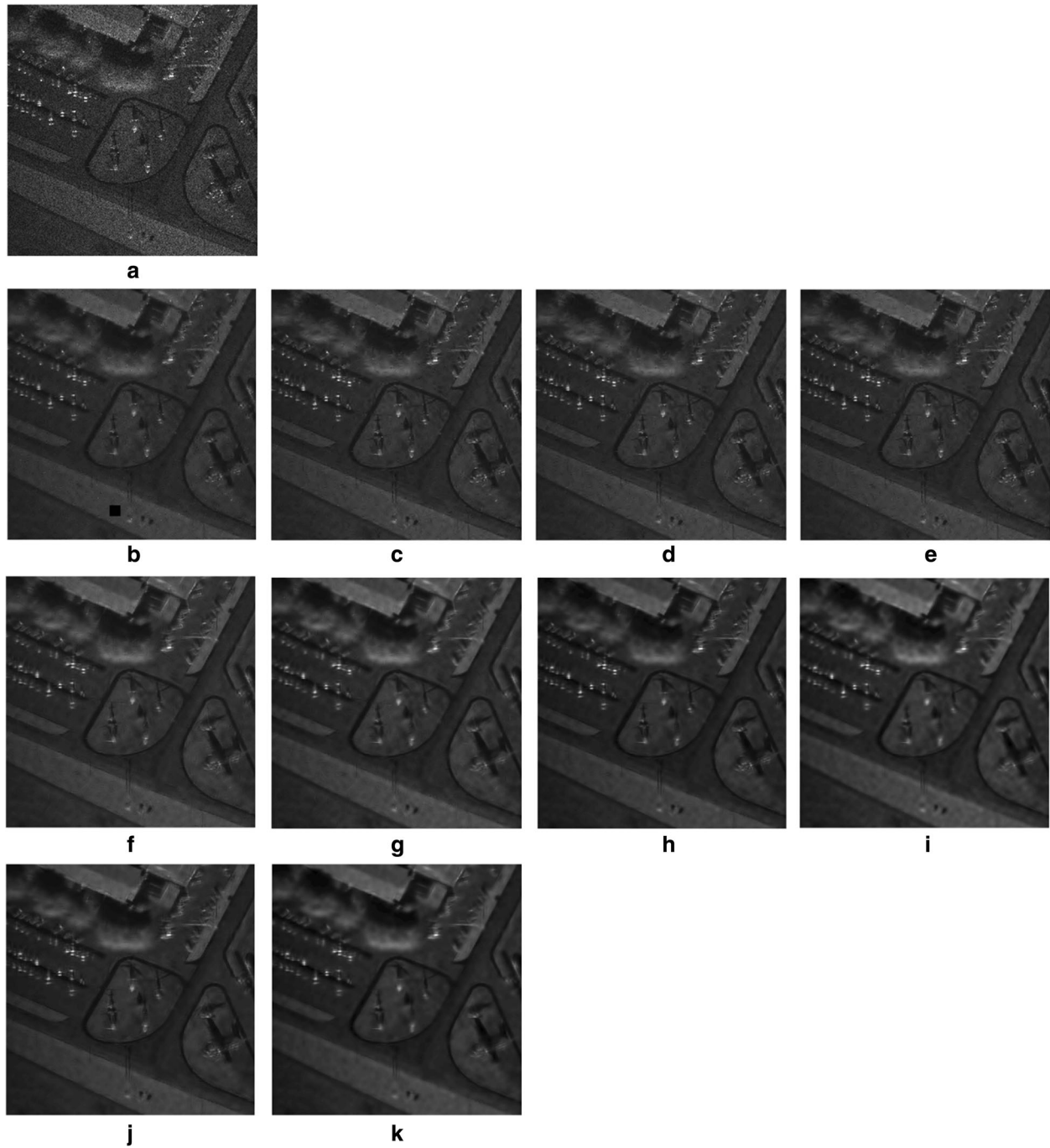
Fig. 9 a original SAR image Air Craft with size 512 × 512, **b–e** despeckled images by B-SWT [30], WB-SWT, BI-SWT, WBI-SWT, and **f–k** despeckled images by B-NSST [2], WB-NSST [27], BI-NSST(1), WBI-NSST(1), BI-NSST(2), and WBI-NSST(2)

$$I(X_1, X_2) = \sum_{x_1 \in X} \sum_{x_2 \in X_2} p(x_1, x_2) \log \frac{p(x_1, x_2)}{p(x_1)p(x_2)} \qquad (30)$$

where (x_1, x_2) is a pair of random variables with joint $p(x_1, x_2)$ and marginal $p(x_1)$ and $p(x_2)$ PDFs. The MI or

$I(X_1, X_2)$ is zero if the child and parent are totally independent. However, increasing the MI means more dependency between the child and parent coefficients in the transformed domain. Therefore, the best child-parent coefficients are those that are totally dependent (i.e., the most positive MI).

Table 4 Four no reference parameters for Peninsula SAR image to compare with methods in SWT and NSST domains

		Peninsula (real SAR image)			
		ENL ↑	MSD ↑	ESI^v ↑	ESI^h ↑
SWT	NOISY	2.9993	0.0000	1.0000	1.0000
	B-SWT [30]	21.2362	0.0005	0.3020	0.0829
	WB-SWT	25.4732	0.0005	0.2524	0.1022
	BI-SWT	32.8736	0.0005	0.3580	0.2509
	WBI-SWT	60.3220	0.0006	0.2920	0.3107
NSST	B-NSST [2]	58.4842	0.0006	0.8174	0.6288
	WB-NSST [27]	97.9584	0.0006	0.8125	0.6401
	BI-NSST(1)	106.9630	0.0006	0.8164	0.6299
	WBI-NSST(1)	119.1139	0.0006	0.8071	0.6407
	BI-NSST(2)	112.0674	0.0006	0.8096	0.6558
	WBI-NSST(2)	152.9135	0.0006	0.7936	0.6706

4.2 Parent and child coefficient models

Although a fair amount of research on image denoising in the transformed domain has been carried out [31, 35], thresholding due to simplicity is still attractive [19, 20]. However, thresholding in a bivariate MAP that exploits the dependency between coefficients [25, 45] gives appropriate results.

The implementation of BiShrink for denoising, based on Eq. (18), requires a child coefficient X_1 and its parent coefficient X_2. In the SWT transformed domain, for an arbitrary coefficient considered as child, X_1, three parents named $X_2(N)$, $X_2(SS)$, and $X_2(NC)$ can be considered where $X_2(N)$ refers to the neighbor subband at the same level, $X_2(SS)$ is a subband at the same orientation but at the next coarser level, and $X_2(NC)$ denotes all the subbands that belong to the next coarser level. For example, by noticing the numbered SWT subbands shown in Fig. 3, if the 1st subband is X_1, the 2nd or 3rd subband can be considered as $X_2(N)$, the 4th subband is $X_2(SS)$, and all 4th–6th subbands are $X_2(NC)$. Similarly, in the NSST transformed domain, for an arbitrary child X_1, three parents called $X_2(N)$, $X_2(NC)$, and $X_2(OPP)$ are considered, in which $X_2(OPP)$ refers to a subband at the same level as X_1 but in the opposite orientation. For example, by noticing the numbered subbands for the NSST shown in Fig. 3, if the 9th subband is X_1, the 8th or 10th subband can be considered as $X_2(N)$, and all 17th–24th subbands are $X_2(NC)$, and the 1st subband is $X_2(OPP)$. For the sample image "Zone Plate," the child-parent in the SWT and NSST domains are shown in Fig. 6.

As mentioned in Section 4.1, the best child-parent coefficients are those that are totally dependent (the most value of MI). We now use the eight test images mentioned and add Gaussian noise with zero mean and standard deviation $\sigma_N = 30$. Noisy images are decomposed by SWT and NSST into three levels, and the MI for different subbands as a child considering the three introduced models are obtained. As expected and shown in Fig. 4, the MI for both transformed domains is not zero because of the dependency between a child and the inter- or intra-subbands. Figure 4 shows that $I(X_1, X_2(NC)) > I(X_1, X_2(SS)) > I(X_1, X_2(N))$ for the SWT, and $I(X_1, X_2(NC)) > I(X_1, X_2(N)) > I(X_1, X_2(OPP))$ for the NSST. In this paper, we use $X_2(SS)$ as the classical parent in the SWT domain, based on previous studies [23, 26, 28, 35]. For the NSST domain, we propose the following two following:

Model 1: considering $X_2(OPP)$ as the parent, the method is named BI-NSST (1).
Model 2: considering $X_2(NC)$ as the parent, the method is named BI-NSST (2).

Although model 1 was proposed for the ST domain [25], we used it in the NSST domain as well. In addition, proposing model 2 in the NSST domain according to the MI shown in Fig. 4 is the contribution herein, where model 2 is expected to outperform model 1 in the NSST domain.

4.3 Weighted shrinkage method

Most shrinkage denoising techniques, including BiShrink [23, 25, 31, 35], assume that the noise power for different subbands are the same. Although this assumption for WT [23] is true, it was shown to be not entirely correct for the nonsubsampled contourlet transform [42] and NSST [13, 27] (i.e., noise power of different subbands are not the same). Although the weighted BayesShrink in the NSST domain [27] was previously used, the weighted BiShrink in the NSST domain, called the WBI-NSST method, is proposed in this paper. To show the validity of the assumption above, the original noise-free test images and the noisy images are decomposed to three levels by SWT and NSST, and the MSE of noise-free and noisy image coefficients are obtained for each decomposition level ℓ and subband k,

$$\text{MSE}_{\ell,k} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \left[Y_{\ell,k}(i,j) - X_{\ell,k}(i,j) \right]^2 \quad (31)$$

As the nonsubsampled version is used, all subband sizes are the same as the input image (i.e., mn). The MSE of all subbands for the eight mentioned test images in the SWT and NSST domains is obtained, and the average values are shown in Fig. 5. As expected, all subbands of the SWT (Fig. 5a2) are affected by noise approximately equal to the subbands in NSST where some subbands are more robust against noise than others (Fig. 5b2).
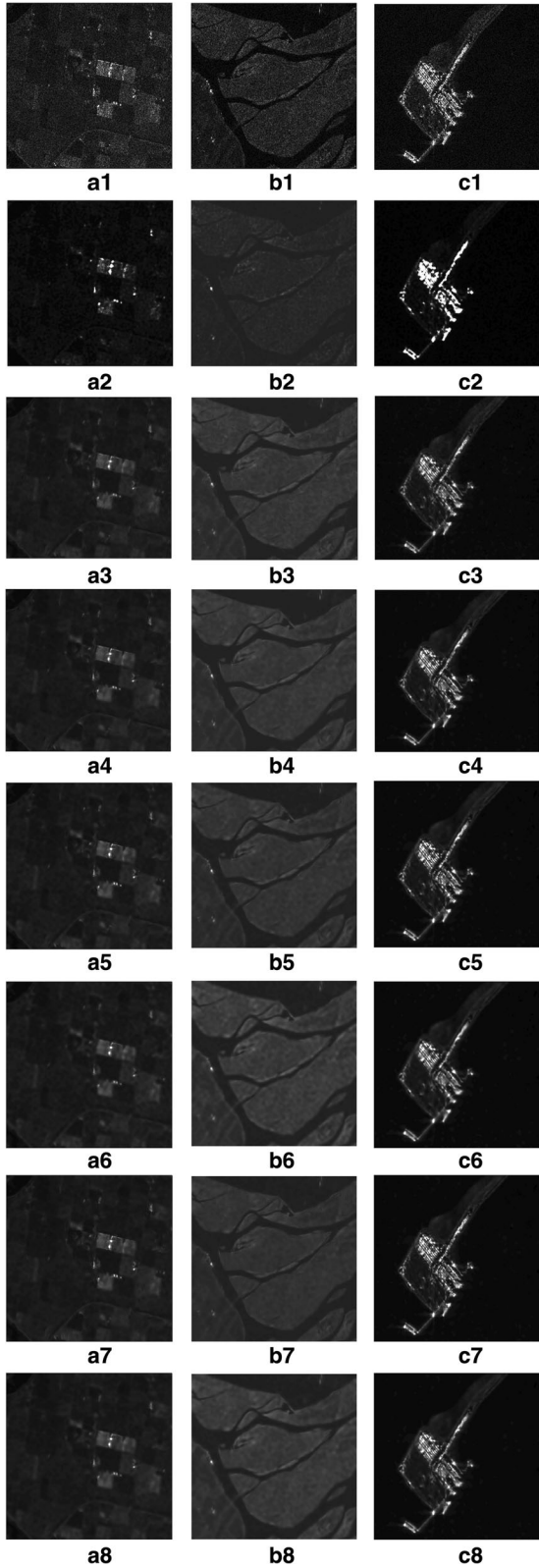
Fig. 10 Real SAR images called Farmland, Peninsula, and Shipping Terminal are shown on the first row, the results for High-TV [15], GГD - NSST [13], NIG - NSST [13], our proposed methods, BI-NSST(1), WBI-NSST(1), BI-NSST(2), and WBI-NSST(2) are shown in order from rows 2 to 7

According to Eq. (18), the BiShrink threshold values [36] named TBI for every decomposition level ℓ and subband k is

$$\mathrm{TBI}_{\ell,k} = \frac{\sqrt{3}\sigma_N^2}{\sigma_{X_{\ell,k}}} \tag{32}$$

where $\sigma_{X_{\ell,k}}^2$ (or σ^2) is the power of the noise-free signal in the transformed domain. In this paper, the noise variance is approximated by the robust median estimator of Eq. (20), and the power of the noise-free signal is estimated using Eq. (22); thus, the Bishrink weighted threshold named TWBI is

$$TWBI_{\ell,k} = \alpha_{\ell,k} TBI_{\ell,k} \tag{33}$$

where α is the weighting factor that depends on the decomposition level ℓ and subband k, expressed as

$$\alpha_{\ell,k} = \frac{\mathrm{MSE}_{\ell,k}}{\overline{\mathrm{MSE}_\ell}} \tag{34}$$

where $\overline{\mathrm{MSE}_\ell}$ is the average MSE of all subbands that belong to the same level, $\overline{\mathrm{MSE}_\ell} = \frac{1}{K_\ell}\sum_{k=1}^{K_\ell}\mathrm{MSE}_{\ell,k}$ and K_ℓ is the number of subbands for the ℓ – th level decomposition.

Using the weighting factor, $\alpha_{\ell,\ k}$ results in the optimum threshold value in Eq. (33), which is then applied to the coefficients as the soft thresholding in Eq. (18). The corresponding methods are named WBI-SWT, WBI-NSST(1), and WBI-NSST(2).

Obtaining the values of $\alpha_{\ell,\ k}$ is the main challenge for implementing the weighted Shrinkage method. In this paper, two approaches for determining the weighting factor, $\alpha_{\ell,\ k}$, for test images and real SAR images are proposed. For the test images, the MSE between the noiseless and noisy coefficients gives the optimum weighting factor, $\alpha_{\ell,\ k}$ at each decomposition level and for every subband as well, see Fig. 5a3, b3. In real applications, including SAR images, clean or noise-free signals are nonexistent. Therefore, a white or a flat image (whose pixels have the same gray scale) with the same size as the input image (512 × 512 pixels herein) is used as the noiseless signal. As mentioned above and with regard to Eq. (34), the MSE between the noiseless and noisy coefficients gives the optimum weighting factor, $\alpha_{\ell,\ k}$ at each decomposition level and for every subband as well, see Fig. 7. According to Fig. 5a3, b3, we conclude that the values of the weighting factor are irrespective of either image type or noise variance, but depend on the transform

or the coefficients' noise efficiency. In the next section, the obtained weighting factors, $\alpha_{\ell,\ k}$ shown in Fig. 7, are used for any weighted versions of shrinkage methods in the NSST or SWT domains when the purpose is to despeckle either the test images or the true SAR images.

5 Experimental results and discussion

In this paper, we used the images Barbara, Lena, House, Boat, Goldhill, Fingerprint, Cameraman, and Peppers of size 512×512 pixels and 256 Gy levels as the test images: images Farmland, Peninsula, and Shipping Terminal [46] of sizes 500×500 pixels and Aircraft [47] of size 512×512 pixels as the real SAR images. The three images (Farmland, Peninsula, and Shipping Terminal) [46] are part of the SAR images collected by RADARSAT-1 in the Fine Beam 2 mode on June 16, 2002. Most of the illuminated scenes was in Delta, British Columbia, Canada. The radar was operating in the C-band with HH polarization. These three parts have good image characteristics such as having grains, as well as many high- and low-frequency parts. The mini SAR image "Aircraft" [47] was collected from the Kirtland AFB region on August 27, 2007 in the Ka-band and Ku-band.

An input signal is decomposed into three levels using SWT and NSST. According to the numbered subbands shown in Fig. 5a1, b1, the SWT has three subbands in each level whereas the NSST has 16, eight, and four subbands for the 1st, 2nd, and 3rd decomposition levels, respectively. The block diagram in Fig. 1 indicates that the homomorphic framework (using the logarithm function at the first and exponential at the end) is used for both test images and real SAR images. The method to obtain the threshold value for the shrinkage methods and the weighted versions are explained in detail in Section 4.3.

Here, in the SWT domain, we evaluate the performance of BayesShrink (B-SWT) [30], weighted BayesShrink (WB-SWT), BiShrink (BI-SWT), and weighted BiShrink (WBI-

Table 5 Four no reference parameters to compare the performance of our proposed methods and references [13] and [15] for despeckling a real SAR image called Farmland

	Farmland (real SAR image)			
	ENL ↑	MSD ↑	ESI^v ↑	ESI^h ↑
NOISY	3.0833	0.000000	1.0000	1.0000
High-TV [15]	43.9698	0.001603	0.0200	0.0476
GΓD-NSST [13]	70.3518	0.001639	0.0994	0.2848
NIG-NSST [13]	88.3271	0.001693	0.0626	0.1908
BI-NSST(1)	81.4881	0.001693	0.0705	0.1877
WBI-NSST(1)	91.3149	0.001638	0.1085	0.3195
BI-NSST(2)	84.9580	0.001703	0.0651	0.1632
WBI-NSST(2)	110.4275	0.001639	0.1093	0.3162

Table 6 ENL^h to compare the performance of our proposed methods and references [13] and [15] for despeckling a real SAR image called Farmland, Peninsula, and Terminal Shipping

ENL^h	Farmland	Peninsula	Terminal Shipping
High-TV [15]	68.8133	51.8599	90.1469
GΓD-NSST [13]	116.3365	82.1719	110.0757
NIG-NSST [13]	145.8000	120.8965	218.7877
BI-NSST(1)	128.2550	111.9838	185.3405
WBI-NSST(1)	133.9873	120.9091	224.8171
BI-NSST(2)	145.1024	130.5634	196.1714
WBI-NSST(2)	167.4749	154.0545	306.1236

SWT); in the NSST domain, we present the results achieved from BayesShrink (B-NSST) [2], weighted BayesShrink (WB-NSST) [27], BiShrink based on model 1 and model 2 ((BI-NSST(1), BI-NSST(2)), and weighted BiShrink based on model 1 and model 2 ((WBI-NSST(1), WBI-NSST(2)) in terms of subjective and objective criteria. Furthermore, two methods in the NSST domain, GΓD‑NSST [13], NIG‑NSST [13], and a high-order total variation method based on compressive sensing called High-TV [15] are also applied to compare the achieved performance among the proposed methods and state-of-the-art papers under the same circumstances. For this purpose, we had to implement all the methods ([2, 13, 15, 27, 30]).
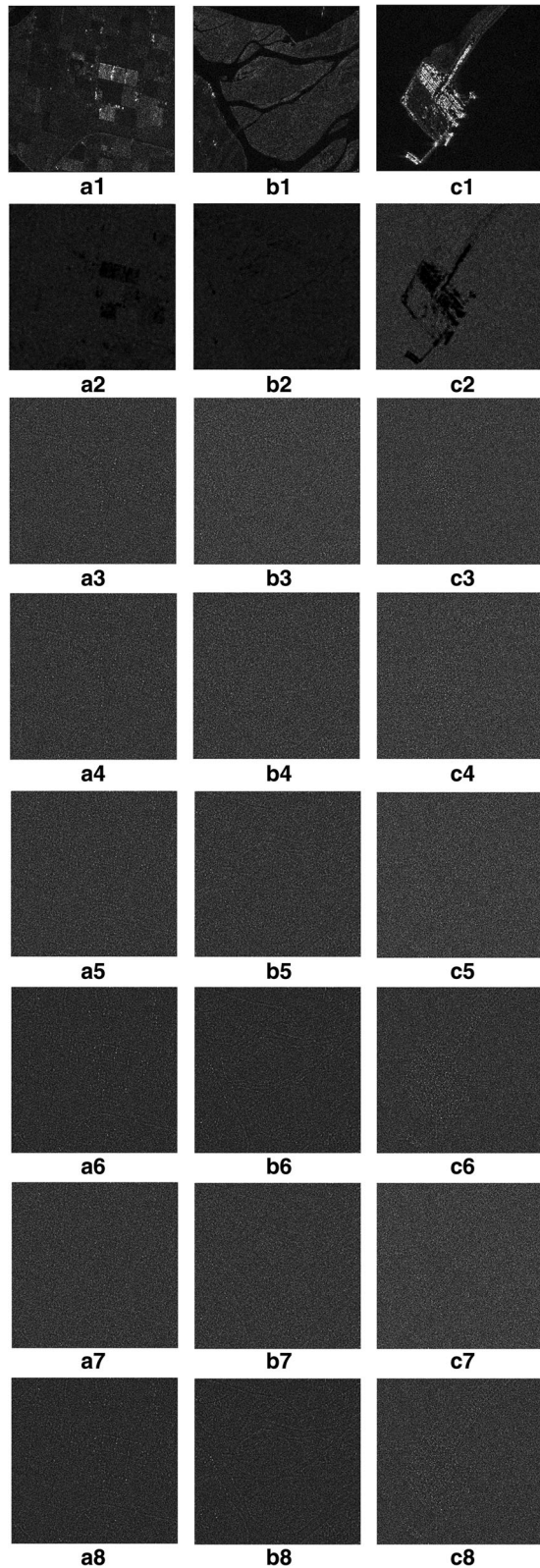
For Barbara as a sample test image, which is corrupted by speckle noise with variance $\sigma_N^2 = 0.1$, the image results in the SWT and NSST domains are shown in Fig. 8. The objective full reference parameters [39, 40] (PSNR, SSIM) are obtained (for average values, the algorithm was run 30 times) for the test images and presented in Tables 2 and 3. Although all methods in the NSST domain outperform those in the SWT domain, no considerable improvement of weighted versions over the direct ones (for example WB-NSST versus B-NSST) was observed for artificial speckle noise.

Using Air Craft of size 512×512 pixels as a real SAR image (see Fig. 9) shows that the proposed methods in the NSST domain is better than the approaches in the SWT domain based on visual qualification, i.e., noise reduction and edge preservation. Table 4, according to four no reference parameters, also proves that for "Peninsula" of size 512×512 pixels as a real SAR image, not only do the methods in the NSST domain outperform those in the SWT domain but also the performance of the weighted versions are significantly better than the directed ones; for example, see WBI-NSST(2) and BI-NSST(2).

According to the results above for artificial speckle noise, and the shrinkage methods in the SWT domain, in the following, we only considered the real SAR images, using four proposed methods in the

Fig. 11 Real SAR images called Farmland, Peninsula, and Shipping Terminal are shown on the first row, ratio images of the results for High-TV [15], GГD‑NSST [13], NIG‑NSST [13], our proposed methods, BI-NSST(1), WBI-NSST(1), BI-NSST(2), and WBI-NSST(2) are shown in order from rows 2 to 7

NSST domain, and compared them with [13] and [15]. While the visual results shown in Fig. 10 do not distinguish the methods precisely, the no reference parameters presented in Table 5 indicate that WBI-NSST(2) is the best approach. Since the ENL is one of the very important parameters that indicate speckle suppression in real SAR images, in this paper, we compute this parameter in two ways: (1) splitting an image into blocks with sizes 16×16 and 25×25 pixels for the real SAR images with size 500×500 and 512×512, respectively, obtaining the ENL for every block, and then writing the mean value ENL in Tables 4 and 5; (2) considering a homogeneous region of size 50×50 pixels and obtaining the ENL value named as $\mathrm{ENL}^{\mathrm{h}}$, see Table 6.

At the end, Fig. 11 shows the ratio images [2, 48] for Farmland, Peninsula, and Shipping Terminal using four proposed methods in the NSST domain and compared with [13] and [15] by considering $I_{\mathrm{ratio}} = I_y / \widehat{I}_x$, where I_y is a real SAR image and $\widehat{I}_x$ is the despeckled image. In general, the ratio image provides significant information on speckle suppression and edge preservation. Any geometric structures or details correlated with the original image in the ratio image indicates that some possible relevant information (e.g., edges or bright scatterers) have been removed or modified using the despeckling method. Note that any content within the ratio image apart from the pure speckle indicates that some modification on the nonhomogeneous areas (edges, mainly) has been performed by the filter. An ideal filter would not alter such edges or bright scatterers, and therefore, ratio images would show a pure speckle pattern. Therefore, if the ratio image does not have any structure (completely noisy shape) or does not show any edges or details from the inside, the algorithm is appropriate. In this case, see Fig. 11 that shows the image ratios of our proposed methods in the NSST domain in comparison with references [13] and [15].

6 Conclusions

In this paper, three methods in SWT and four approaches in NSST are developed according to BayesShrink, BiShrink, weighted BayesShrink, and weighted BiShrink. For BiShrink implementation in the NSST domain, two models to choose the child-parent are proposed with regard to the MI parameter. Although the model recommended by the MI value outperforms for synthesized image with highly

detailed content, it is not appropriate for true SAR images and synthesized images with many smooth regions. Because, in this study, we showed that any thresholding-based methods in the NSST domain outperform the SWT domain, finding new parameters to choose the suitable child-parent in the NSST is future research work.

Availability of data and materials
The three images (Farmland, Peninsula, and Shipping Terminal) [46] are part of the SAR images collected by RADARSAT-1 in the Fine Beam 2 mode on June 16, 2002. Most of the illuminated scenes were in Delta, British Columbia, Canada. The radar was operating in the C-band with HH polarization. The mini SAR image Aircraft provided by Sandia National Laboratories [47] was collected from the Kirtland AFB region on August 27, 2007 in the Ka-band and Ku-band.

Authors' contributions
We confirm that this work is original and has not been published nor is it currently under consideration for publication elsewhere. Both authors read and approved the final manuscript.

Authors' information
Nikou Farhangi was born in Macoo, Iran, in 1987. She received the B.Sc. degree in telecommunication engineering from Islamic Azad University, Urmiya branch in 2010 and the M.Sc. degree in communication systems from Islamic Azad University, South Tehran Branch, Iran, in 2017, respectively. Her research interests include digital signal and image processing. Sedigheh Ghofrani was born in 1968 in Ghochan. She received the B.S. degree in electronic engineering from Tehran University, Iran, in 1991, the M.S. degree in communication from Islamic Azad University, South Tehran Branch, Iran, in 1997, and the Ph.D. degree in electronic from Iran University of Science and Technology, Tehran, Iran, in 2004. She has been an Assistant Professor in the Department of Electronic and Electrical Engineering, Islamic Azad University, South Tehran Branch from 2004 to 2011 and an Associate Professor since 2012. In 2003, she spent 8 months at the School of Electronic and Electrical Engineering, the University of Leeds, UK, supported by the British Council foundation. In 2012, she spent 8 months at the Center for Advanced Communications (CAC) at Villanova University, PA, USA, as Visiting Research Professor. Her area of research includes image processing and signal processing.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

References
1. H Chen, Y Zhang, H Wang, C Ding, Stationary-wavelet-based despeckling of SAR images using two-sided generalized gamma models. IEEE Geosci. Remote Sens. Lett. **9**(6), 1061–1065 (2012)
2. B Hou, X Zhang, X Bu, H Feng, SAR image despeckling based on nonsubsampled shearlet transform. IEEE J. Selected Topics Appl. Earth Observations Remote Sensing **5**(3), 809–823 (2012)
3. DL Donoho, De-noising by soft-thresholding. IEEE Trans. Inf. Theory **41**(3), 613–627 (1995)
4. EJ Candes, DL Donoho, *Curvelets: A Surprisingly Effective Nonadaptive Representation for Objects with Edges* (DTIC Document, 1999)
5. X Zhang, X Jing, Image denoising in contourlet domain based on a normal inverse Gaussian prior. Digit. Signal Process. **20**(5), 1439–1446 (2010)
6. N Minh, *Directional Multiresolution Image Representations* (University of Canberra Bachelor of Engineering in Computer Engineering Citeseer, 2002)
7. D. Labate, W.-Q. Lim, G. Kutyniok and G. Weiss, Sparse multidimensional representation using shearlets, International Society for Optics and Photonics, pp. 59140U-59140U-9, 2005.
8. K Guo, D Labate, Optimally sparse multidimensional representation using shearlets. SIAM J. Math. Anal. **39**(1), 298–318 (2007)
9. W-Q Lim, The discrete shearlet transform: a new directional transform and compactly supported shearlet frames. IEEE Trans. Image Process. **19**(5), 1166–1180 (2010)
10. G Easley, D Labate, W-Q Lim, Sparse directional image representations using the discrete shearlet transform. Appl. Comput. Harmon. Anal. **25**(1), 25–46 (2008)
11. GR Easley, D Labate, F Colonna, Shearlet-based total variation diffusion for denoising. IEEE Trans. Image Process. **18**(2), 260–268 (2009)
12. VM Patel, GR Easley, DM Healy Jr, Shearlet-based deconvolution. IEEE Trans. Image Process. **18**(12), 2673–2685 (2009)
13. S Jafari, S Ghofrani, Using two coefficients modeling of nonsubsampled Shearlet transform for despeckling. J. Appl. Remote. Sens. **10**(1), 015002–015002 (2016)
14. SQ Liu, SH Hu, Y Xiao, YL An, Bayesian shearlet shrinkage for SAR image de-noising via sparse representation. Multidim. Syst. Sign. Process. **25**(4), 683–701 (2014)
15. H Markarian, S Ghofrani, High-TV based CS framework using MAP estimator for SAR image enhancement. IEEE J. Selected Topics Appl. Earth Observations Remote Sensing **10**(9), 4059–4073 (2017)
16. C Yan, Y Zhang, J Xu, F Dai, J Zhang, Q Dai, F Wu, Efficient parallel framework for HEVC motion estimation on many-core processors. IEEE Trans. Circuits Syst. Video Technol. **24**, 2077–2089 (2014)
17. C Yan, Y Zhang, J Xu, F Dai, L Li, Q Dai, F Wu, A highly parallel framework for HEVC coding unit partitioning tree decision on many-core processors. IEEE Signal Process. Lett. **21**, 573–576 (2014)
18. M. C. Motwani, M. C. Gadiya, R. C. Motwani and F. C. Harris, Survey of image denoising techniques, Proceedings of GSPX, pp. 27-30, 2004.
19. DL Donoho, JM Johnstone, Ideal spatial adaptation by wavelet shrinkage. Biometrika **81**(3), 425–455 (1994)
20. DL Donoho, IM Johnstone, Adapting to unknown smoothness via wavelet shrinkage. J. Am. Stat. Assoc. **90**(432), 1200–1224 (1995)
21. HA Chipman, ED Kolaczyk, RE McCulloch, Adaptive Bayesian wavelet shrinkage. J. Am. Stat. Assoc. **92**, 1413–1421 (1997)
22. Q. Liu and L. Ni, Image denoising using Bayesian shrink threshold based on weighted adaptive directional lifting wavelet transform, International Society for Optics and Photonics in International Conference on Graphic and Image Processing, pp. 828570-828570-6, 2011.
23. IW Selesnick, Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. IEEE Trans. Signal Process. **50**(11), 2744–2756 (2002)
24. Z De-xiang, W Xiao-pei, G Qing-wei, G Xiao-jing, SAR image despeckling via bivariate shrinkage based on contourlet transform. IEEE Int. Symp. Comput. Intell. Design **2**, 12–15 (2008)
25. Q Guo, S Yu, X Chen, C Liu, W Wei, Shearlet-based image denoising using bivariate shrinkage with intra-band and opposite orientation dependencies. IEEE Int. Joint Conf. Comput. Sci. Optimization **1**, 863–866 (2009)
26. S. Chitchian, M. Fiddy and N. M. Fried, Denoising during optical coherence tomography of the prostate nerves via bivariate shrinkage using dual-tree complex wavelet transform, International Society for Optics and Photonics in SPIE BiOS, Biomedical Optics, pp. 716112-716112-4, 2009.
27. S Jafari, S Ghofrani, M Sheikhan, Comparing undecimated wavelet, nonsubsampled contourlet and shearlet for SAR images despeckling. Majlesi J. Electr. Eng. **9**(3) (2015)
28. X. Xu, Y. Zhao, W. Zhou and Y. Peng, SAR image denoising based on alpha-stable distribution and Bayesian wavelet shrinkage, International Society for Optics and Photonics in Sixth International Symposium on Multispectral Image Processing and Pattern Recognition, pp. 74951U-74951U-8, 2009.
29. W Hongzhi, H Cai, Locally adaptive bivariate shrinkage algorithm for image denoising based on nonsubsampled contourlet transform. IEEE International Conf. Comput. Mechatronics Control Electron. Eng. **10**(6), 33–36 (2010)

30. X Wang, RS Istepanian, YH Song, Microarray image enhancement by denoising using stationary wavelet transform. IEEE Trans. NanoBioscience **2**(4), 184–189 (2003)
31. D Min, Z Jiuwen, M Yide, Image denoising via bivariate shrinkage function based on a new structure of dual contourlet transform. Signal Process. **109**, 25–37 (2015)
32. F. Lenzen, "Statistical regularization and denoising," Dissertation in Mathematics submitted for the degree of Doctor rerum naturalium, Chap. 1 (2006)
33. A Hyvärinen, Sparse code shrinkage: denoising of nonGaussian data by maximum likelihood estimation. Neural Comput. **11**(7), 1739–1768 (1999)
34. D-X Zhang, Q-W Gao, X-P Wu, Bayesian based speckle suppression for SAR image using contourlet transform. J. Electron Sci. Technol. China **6**(1), 79–82 (2008)
35. L Şendur, IW Selesnick, Bivariate shrinkage with local variance estimation. IEEE Signal Process. Lett. **9**(12), 438–441 (2002)
36. S Xing, Q Xu, D Ma, Speckle denoising based on bivariate shrinkage functions and dual-tree complex wavelet transform. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. **38**, 1–57 (2008)
37. S Dahlke, G Kutyniok, P Maass, C Sagiv, H-G Stark, G Teschke, The uncertainty principle associated with the continuous shearlet transform. Int. J. Wavelets Multiresolution Inf. Process. **6**(02), 157–181 (2008)
38. G. Andria, F. Attivissimo, A. M. L. Lanzolla, and M. Savino, A suitable threshold for speckle reduction in ultrasound images, IEEE Transaction on Instrumentation and Measurement, vol. 62, no. 8, Agust 2013.
39. Z Wang, AC Bovik, HR Sheikh, EP Simoncelli, Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
40. D Brunet, ER Vrscay, Z Wang, On the mathematical properties of the structural similarity index. IEEE Trans. Image Process. **21**, 1488–1499 (2012)
41. J Zhang, TM Le, S Ong, TQ Nguyen, No-reference image quality assessment using structural activity. Signal Process. **91**(11), 2575–2588 (2011)
42. M Kiani, S Ghofrani, Two new methods based on contourlet transform for despeckling synthetic aperture radar images. J. Appl. Remote. Sens. **8**(1), 083604–083604 (2014)
43. X-Y Wang, Y-C Liu, H-Y Yang, Image denoising in extended shearlet domain using hidden Markov tree models. Digit. Signal Process. **30**, 101–113 (2014)
44. TM Cover, JA Thomas, *Elements of Information Theory* (John Wiley & Sons, 2012)
45. GS Shin, MG Kang, Wavelet-based denoising considering interscale and intrascale dependences. Opt. Eng. **44**, 067002–067009 (2005)
46. IG Cumming, FH Wong, Digital processing of synthetic aperture radar data. Artech house **1**, 3 (2005)
47. Armin W. Doerry Automatic compensation of antenna beam roll-off in SAR images, Department of Energy in United States, 2006.
48. L Gomez, ME Buemi, JC Jacobo-Berlles, ME Mejail, A new image quality index for objectively evaluating despeckling filtering in SAR images. IEEE J. Selected Topics Appl. Earth Observations Remote Sensing **9**, 1297–1307 (2016)