

# Nonstationary Image Noise Removal (NINR)

Fatih Porikli<sup>1</sup> Akshay Soni<sup>2</sup> Kei Suwa<sup>3</sup>

<sup>1</sup> Mitsubishi Electric Research Labs, USA, fatih@merl.com

<sup>2</sup> University of Minnesota, USA, sonix022@umn.edu

<sup>3</sup> Mitsubishi Electric Corporation, Japan, suwa.kei@dx.mitsubishielectric.co.jp

**Abstract**—This paper presents a spatially-varying speckle noise removal method that does not require any noise parameters or image reconstruction factors to be predetermined by the user. It automatically estimates local noise variances by solving an optimization problem that is based on the scale invariant property of 'kurtosis' for radar imagery. Then, it aggregates multiple estimations at each pixel using a collaborative filtering approach. Experimental results demonstrate that our method outperforms the conventional procedures.

## I. INTRODUCTION

Synthetic aperture radar (SAR) images are corrupted by spatially-varying noise due to several factors, such as antenna gain and phase pattern discrepancies, spatially-varying background clutter, physical properties of the reflectors such as angular scintillation and fluctuations caused by sub-cell sized objects.

As described in [1], electromagnetic waves emitted by a SAR are no longer in phase after interaction with the reflecting surface since the waves travel different distances back from different surfaces and multiple bounce scattering occur as a result of the variance in surface roughness. The signals may also go out of phase when the synthesized antenna moves. Besides, moving surfaces cause phase differences reproduced as displacement and defocusing artifacts.

The out-of-phase waves interfere constructively or destructively to produce stronger or weaker signals. When a SAR image is formed by processing the backscatter returns from successive pulses, this effect generates a pixel-to-pixel variation in intensity. This manifests itself as a bright and dark pixel pattern called speckle, which is a multiplicative process and has a standard deviation linearly related to the mean intensity. In other words, the higher the signal strength the higher the noise.

To separate the underlying signal from the noise, existing state-of-the-art image denoising algorithms assume that the spectral properties of the original signal and the noise are known. For instance, the Wiener filter [2] seeks a linear time-invariant filter whose output would come as close to the original signal as possible. Other popular algorithms, such as the non-local means [3] and the BM3D [4] consider that the image is corrupted by additive Gaussian noise with constant variance

$$I^n = I + \eta \quad (1)$$

where  $I$  is the original image and  $\eta$  is noise. For SAR imagery, constant noise variance is not a valid assumption. Recently, a

few studies investigated nonstationary noise models [5]–[8] for different application domains.

Here, we propose the following nonstationary and multiplicative noise model

$$I_p^n = I_p \cdot \eta_p, \quad (2)$$

where  $I_p$  is the desired speckle-free original image intensity at pixel (or patch)  $p$  and  $\eta_p$  is a pixel-intensity dependent noise process, which can be approximated in sufficiently small local neighborhoods as a Gaussian noise with variance  $\sigma_p^2$ . The noise variance is spatially varying and is not constant throughout. We apply log operation to convert the above problem into an additive setting

$$I_p^n = I_p \cdot \eta_p \rightarrow \log I_p^n = \log I_p + \log \eta_p \quad (3)$$

as in [9]. To marginalize noise, we estimate noise variance within a patch at each pixel in the log transfer domain and then use the estimated noise variance when we collaboratively process multiple patches to impose intensity consistency through sparsity. Then transform the reconstructed image from log to intensity domain. We call this simple yet effective algorithm as Nonstationary Image Noise Removal (NINR).

Noise variance can be inferred by taking advantage of a statistical regularity of natural images. Since natural images tend to have spherically symmetric distributions [10], the kurtosis values in general band-pass filtered domains remain close to a positive constant [11]. Therefore, the global noise variance of the whole image can be determined by imposing kurtosis across different scales, i.e. different band-pass filtered channels of DCT or wavelets, to be a positive constant.

The local variances at each pixel location can be computed the statistics collected from surrounding pixels. Even though the noise distribution function might vary significantly within large regions, for sufficiently small local image neighborhoods, it can be modeled by a Gaussian function.

After obtaining pixel-level local noise variance, NINR determine multiple clusters of similar image patches. It then filters each cluster of patches, which are arranged into 3D data structure, in a way similar to the process described in [4]. However, unlike [4], NINR estimates the local noise variances.

Our experiments demonstrate that NINR significantly outperforms the BM3D and the NL both in terms of PSNR and SSIM scores. In comparison to the BM3D<sup>e</sup> (BM3D with a global noise variance estimator), NINR provides significant gain while preventing over- and under-smoothing.

## II. PIXEL-LEVEL VARIANCE ESTIMATION

Noise variance at each pixel is estimated using the noisy image without any additional prior. We use a kurtosis constancy based method, which has been proposed for image splicing and forgery detection [10].

Kurtosis  $\kappa$  is a descriptor of the shape of a probability distribution. For example, a Gaussian distribution has kurtosis value of zero. It is defined as

$$\kappa = \bar{\mu}_4(\sigma^2)^{-2} - 3, \quad (4)$$

for a random variable  $X$ , where

$$\sigma^2 = E[(X - E[X])^2], \quad (5)$$

and

$$\bar{\mu}_4 = E[(X - E[X])^4]. \quad (6)$$

For a natural image, the kurtosis is nearly constant over its bandpass filtered domains such as DCT or wavelet decompositions.

Let the noisy image  $\log I^n$  to be transformed into the frequency domain  $\log I^n \rightarrow f^n$ . We apply  $K$  band-pass filters to obtain  $K$  bands. The kurtosis of the original and the noisy image in the  $k^{\text{th}}$  band are denoted as  $\kappa_k$  and  $\bar{\kappa}_k$  respectively. These statistics are related as

$$\bar{\kappa}_k = \kappa_k \left( \frac{\bar{\sigma}_k^2 - \sigma^2}{\bar{\sigma}_k^2} \right)^2 \quad (7)$$

for  $k = 1, \dots, K$  where  $\bar{\sigma}_k^2$  is the variance of the  $k^{\text{th}}$  channel. The statistical regularity of natural images in the bandpass filtered domains imposes positive kurtosis values. Considering near constant kurtosis values over different scales we have

$$\kappa_k \approx \kappa. \quad (8)$$

To find  $\kappa$  and  $\sigma^2$  that minimize the difference between the two sides of (7) over all scales we solve this minimization problem

$$\sigma_p^2 = \arg \min_{\sqrt{\kappa}, \sigma^2} \sum_{k=1}^K \left[ \sqrt{\bar{\kappa}_k} - \sqrt{\kappa} \left( \frac{\bar{\sigma}_k^2 - \sigma^2}{\bar{\sigma}_k^2} \right) \right]^2, \quad (9)$$

whose minimizer provides the solution for noise variance. (9) is convex and it has a closed form solution.

Pixel-wise noise variance  $\sigma_p^2$  within a patch around each pixel can be estimated by computing variance and kurtosis of each overlapping patch for each band

$$\sigma^2 = \mu_2 - \mu_1^2, \quad (10)$$

and

$$\kappa = \frac{\mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4}{\mu_2^2 - 2\mu_2\mu_1^2 + \mu_1^4}. \quad (11)$$

where uncentered moments  $\mu_i = E[X^i]$  are obtained using spatial averaging. Then, the closed form solution of (9) is applied to find the pixel-wise noise variance. The noise variance of a patch can be taken as the maximum of the estimated noise variances.



Fig. 1: Thresholding by pixel-wise noise variances.

To illustrate the power of the pixel-wise noise estimation, in Fig. 1-left we show an image corrupted by a different variance noise. We used a simple k-means to clustering on the estimated variance values to determine the segments shown on the right.

## III. COLLABORATIVE FILTER

After, the pixel-level variances are determined, the next task is to seek a linear filter whose output is as close to the original signal as possible. For this, we employ Wiener filter with the estimated noise variance. However, a direct application of such a local linear filter would disregard the natural correlation between similar image patches that can elevate additional regularization constraints.

To take advantage of this correlation, we exploit the concepts of *group matching* and *collaborative filtering*. Note that, an image is sparse in transform domain if most of the transform domain coefficients are either zero or are very small in magnitude, so that they can be neglected. In that case, the image can be well approximated as a linear combination of a few basis elements that correspond to pixel-wise consistent patterns. Denoised image can be obtained by keeping only the large transform coefficients, which are mainly due to original signal, and discarding small coefficients, which are mainly due to noise. Collaborative filtering takes advantages of the sparsity by group matching.

For collaborative filtering, we incorporate the approach described in [4]. As in that approach, we apply filtering twice. After the first filtering, we select similar patches and apply Wiener filter with the estimated pixel-wise variances. For a patch  $\log I_p^n$ , we find the most similar patches  $\log I_q^n$  in its search neighborhood and determine clusters  $C_p^\phi$ . Since this cluster is determined from the noisy image patches, it might contain incorrect patches. To remove the noise in these patches, we apply transform domain filtering, that is, 2D DCT  $f_{2D}$  followed by hard-thresholding  $\theta$  (setting zero) of higher-frequency coefficients

$$\phi(f_{2D}(\log I_p^n - \mu_p)). \quad (12)$$

Using the above filtered patches, we evaluate patch similarity. We arrange similar these patches into 3D data  $D_p$  on which a 1D transform and hard-thresholding is applied a second time to the values of the pixels at the same patch locations

$$\phi(f_{1D}(D_p)). \quad (13)$$

This second transform domain hard-thresholding along each pixel incorporates collaborative information from multiple patches. By mapping back the inverse 1D transformed patches

onto the image coordinates and combining the pixel-wise responses, we obtain an intermediate image  $\log I^m$ .

Next, we determine the clusters of patches  $C_p^m$ , this time from the intermediate image  $\log I^m$ . We arrange  $\log I_p^m$  and  $\log I_p^n$  into  $D_p^m$  and  $D_p^n$  to use  $D_p^m$  for a more accurate computation of the Wiener filter coefficients. We apply these coefficients to clusters formed from the unfiltered noisy patches  $D_p^n$ . The Wiener deconvolution coefficients in Fourier transform domain are defined from the energy of the transform domain coefficients as

$$W_p = \frac{|f_{3D}(D_p^m)|^2}{|f_{3D}(D_p^m)|^2 + \sigma_p^2} \quad (14)$$

where  $f_{3D}$  is DFT. When the noise variance at certain frequencies increases the Wiener filter attenuates the response. Note that, we use the pixel-wise noise variances here. Element-by-element multiplication of  $W$  with the transform domain coefficients  $f_{3D}(D_p^n)$  gives the Wiener filtered response in the transform domain, which is then mapped back to spatial domain. These filtered patches are aggregated at each pixel location with weights  $\omega_p$  inversely proportional to the Wiener coefficients  $W_p$  and pixel-wise variance values

$$\omega_p = \frac{1}{\sigma_p^2 \|W_p\|_2^2} \quad (15)$$

to make pixels with higher uncertainty to contribute less. We have used the parameter values and transforms similar to what is proposed in original BM3D implementation and we direct user to [4] for details.

#### IV. EXPERIMENTAL ANALYSIS

We performed two sets of experiments for quantitative and qualitative evaluations.

In the quantitative evaluations, we used a corpus of  $256 \times 256$  noise-free, gray-level aerial images, some of which are shown in Fig. 2 first column and reported in Tables I and II. We corrupted each input image with a multiplicative noise that has a variance proportional to the image intensity value. Samples can be seen in Fig. 2 second column.

We tested the proposed NINR and compared its PSNR and SSIM scores to BM3D [4], BM3D<sup>e</sup> [5], and NLM [3]. BM3D requires the input variance to be given and the BM3D<sup>e</sup> computes it using a global variance estimator. We used the variance value that gives the most pleasing results for BM3D. NLM requires both the global noise variance estimator and the number of patches, which was set to 5 in all experiments. We applied log transform to convert the multiplicative problem to additive for BM3D, BM3D<sup>e</sup> and NLM for most objective evaluations.

As shown in Tables I and II, NINR outperforms other methods both in terms of PSNR and SSIM. The best results are marked in bold, and the second best in blue color. On average, NINR has +2dB to +5dB improvement over BM3D<sup>e</sup> and others. It prevents over-smoothing that BM3D generates. Note that, in all images, the stationary noise assumption of BM3D,

	noisy	NINR	BM3D <sup>e</sup>	BM3D	NLM
sat-1	18.14	<b>26.55</b>	22.58	<b>24.59</b>	23.18
sat-2	24.80	<b>29.91</b>	25.56	<b>26.87</b>	26.75
sat-3	26.78	<b>31.80</b>	28.34	28.66	<b>29.17</b>
sat-4	19.95	<b>24.29</b>	21.38	20.90	<b>21.85</b>
sat-5	19.03	<b>27.42</b>	23.56	<b>24.78</b>	24.17
sat-6	20.50	<b>25.53</b>	23.05	23.19	<b>23.59</b>
sat-7	24.07	<b>29.04</b>	24.90	25.12	<b>25.95</b>
sat-8	21.64	<b>28.08</b>	26.59	25.90	<b>27.10</b>
sat-9	20.47	<b>25.20</b>	<b>24.95</b>	22.53	24.52
sat-10	20.67	<b>26.58</b>	25.50	23.89	<b>25.68</b>

TABLE I: PSNR (dB) Results. Best: **bold**, second best: **blue**.

	noisy	NINR	BM3D <sup>e</sup>	BM3D	NLM
sat-1	0.490	<b>0.731</b>	0.572	<b>0.637</b>	0.578
sat-2	0.656	<b>0.894</b>	0.799	0.788	<b>0.821</b>
sat-3	0.661	<b>0.867</b>	0.825	0.773	<b>0.830</b>
sat-4	0.726	<b>0.832</b>	<b>0.763</b>	0.657	0.760
sat-5	0.455	<b>0.759</b>	0.606	<b>0.618</b>	0.5982
sat-6	0.633	<b>0.784</b>	<b>0.749</b>	0.599	0.742
sat-7	0.686	<b>0.885</b>	0.777	0.760	<b>0.796</b>
sat-8	0.492	<b>0.799</b>	<b>0.775</b>	0.689	0.759
sat-9	0.651	<b>0.800</b>	<b>0.800</b>	0.609	0.756
sat-10	0.565	<b>0.778</b>	<b>0.749</b>	0.632	0.729

TABLE II: SSIM Results. Best: **bold**, second best: **blue**.

BM3D<sup>e</sup> and NLM, fails to provide a satisfactory filtering both bright and dark regions. As visible, NINR successfully balances filtering over whole image.

For qualitative evaluations, we tested NINR on sample SAR images as shown in Fig. 3. Again, NINR provided the most pleasing results.

#### REFERENCES

- [1] F. Qiu, J. Berglund, J. Jensen, P. Thakkar, and D. Ren, "Speckle noise reduction in SAR imagery using a local adaptive median filter," *GIScience and Remote Sensing*, vol. 41, no. 3, pp. 244–266, 2004.
- [2] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.
- [3] A. Buades, B. Coll, and J. Morel, "A review of image denoising algorithms, with a new one," *Simul*, vol. 4, 2005.
- [4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, 2007.
- [5] A. Foi, V. Katkovnik, and K. Egiazarian, "Signal-dependent noise removal in pointwise shape-adaptive DCT domain with locally adaptive variance," *EUSIPCO*, 2007.
- [6] A. Mencattini, G. Rabottino, M. Salmeri, R. Lojacono, and B. Sciu-nzi, "Denoising and enhancement of mammographic images under heteroscedastic additive noise by an optimal subband thresholding," *IJWMI*, vol. 8, no. 5, 2010.
- [7] I. Tosic, B. Olshausen, and B. Culpepper, "Learning sparse representations of depth," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, 2011.
- [8] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "A nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 119–133, 2013.
- [9] R. Sundaresan and F. Porikli, "Additive noise removal by sparse reconstruction on image affinity nets," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2012.
- [10] X. Pan, X. Zhang, and S. Lyu, "Exposing image splicing with inconsistent local noise variances," in *IEEE International Conference on Computational Photography*, 2012.
- [11] D. Zoran and Y. Weiss, "Scale invariance and noise in nature image," in *International Conference on Computer Vision*, 2009.



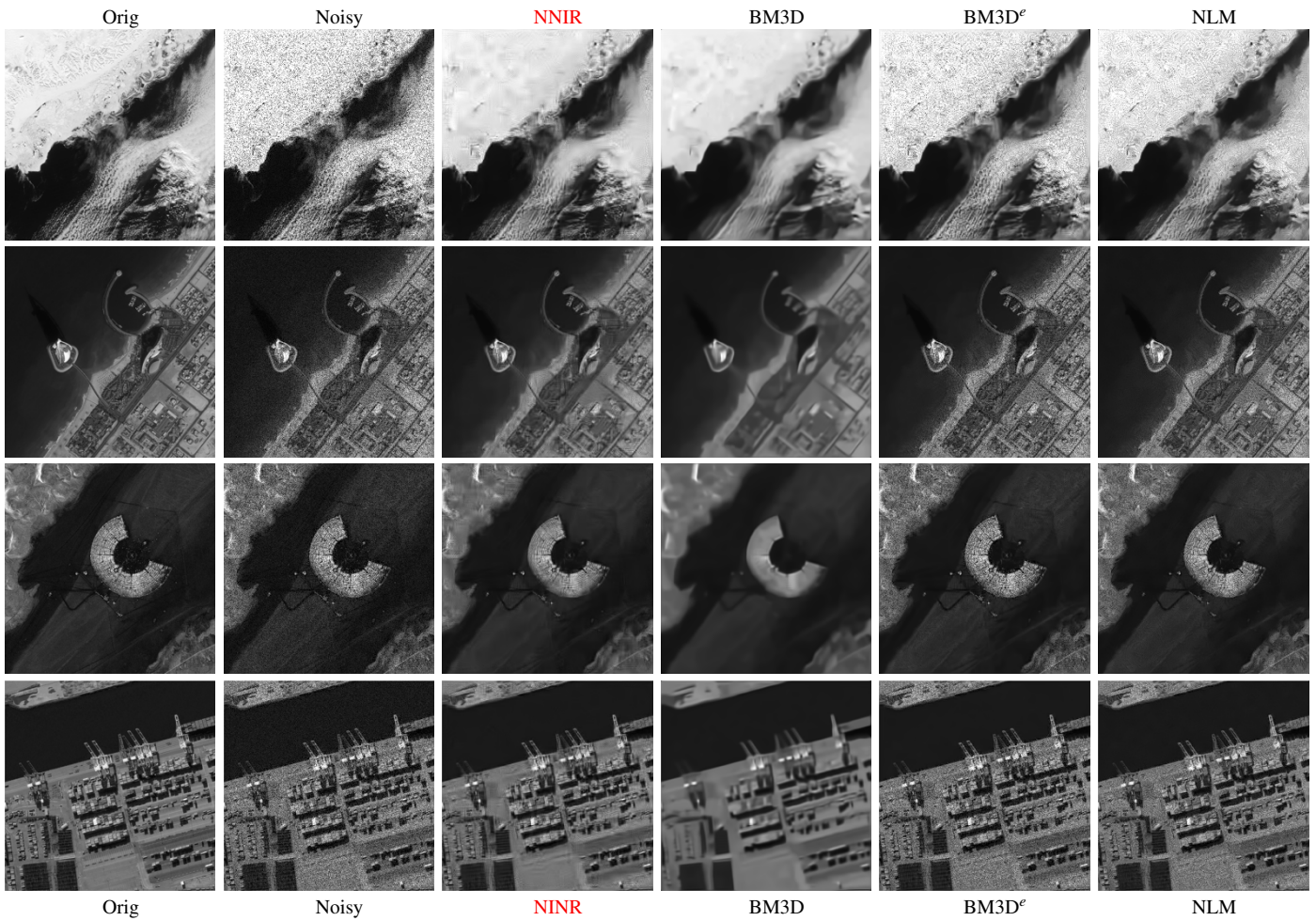


Fig. 2: Existing methods fails for nonstationary noise. BM3D (BM3D<sup>e</sup>, NLM) over- (under-) smooths whole image. Since NINR locally adapts, it provides the most pleasing results (for best assessment, please view on digital display).

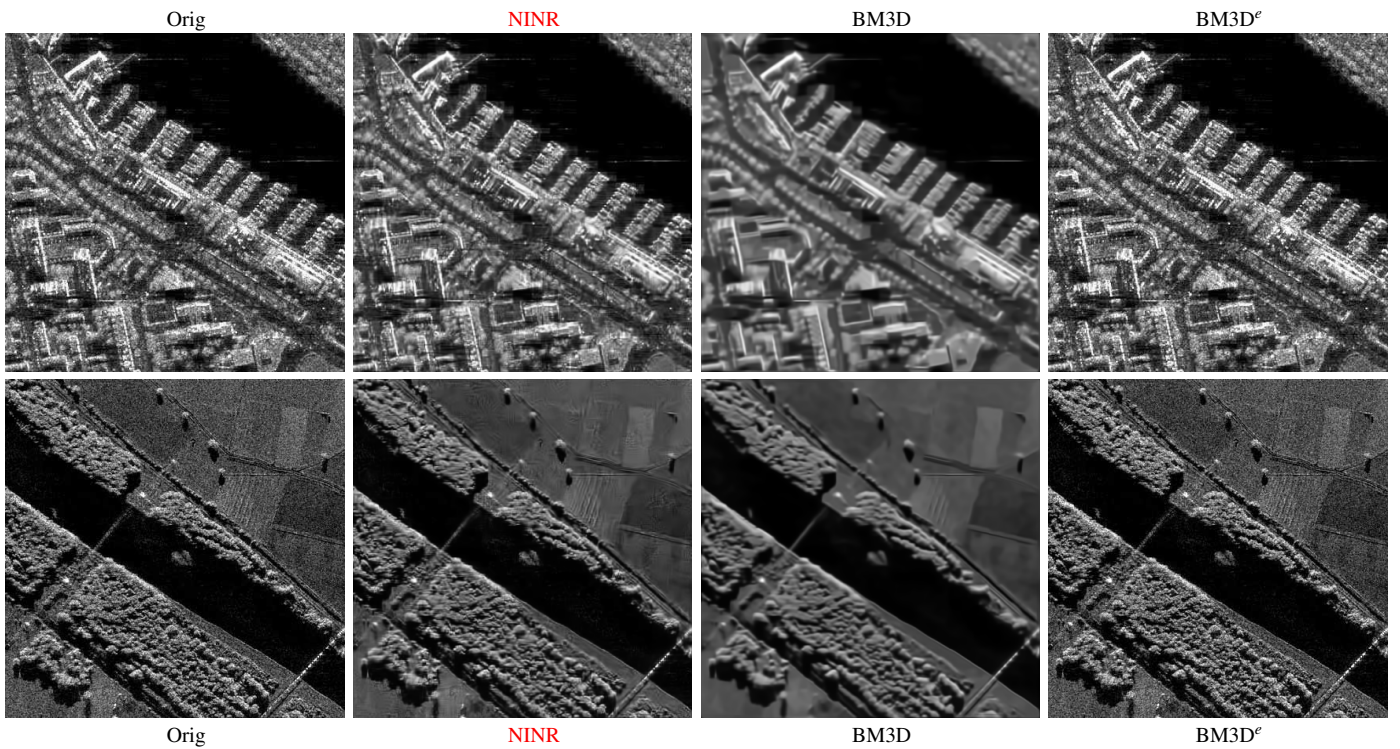


Fig. 3: BM3D removes important SAR image details. BM3D<sup>e</sup> under-estimate variance. NINR accurately removes noise.