A Non-Local Low-Rank Framework for Ultrasound Speckle Reduction

Anonymous CVPR submission

Paper ID 2395

Abstract

‘Speckle’ noise refers to the granular patterns that occur in ultrasound images due to wave interference. Speckle removal can greatly improve the visibility of the underlying structures in an ultrasound image and enhance subsequent post-processing. We present a novel framework for speckle removal based on low-rank non-local filtering. Our approach works by first computing a guidance image that assists in the selection of candidate patches for non-local filtering in the face of significant speckle noise. The candidate patches are further refined using a low-rank minimization estimated using a truncated weighted nuclear norm (TWNN) and structured sparsity. We show that the proposed filtering framework produces results that outperform state-of-the-art methods both qualitatively and quantitatively. This framework also provides better segmentation results when used for pre-processing ultrasound images.

1. Motivation and Related Work

Medical ultrasound is a widely used noninvasive imaging modality that can reveal internal anatomic structures. Ultrasound makes use of a transducer to emit ultra-high-frequency sound waves, which change direction when a reflective surface is encountered. Careful timing of the emitted sound signal and its observed echo is used to determine the anatomical structures. One drawback of ultrasound imaging is the noise that results from wave interference when the scattered waves constructively and destructively combine to produce the black and white ‘speckle’ pattern characteristic of ultrasound images [3, 14]. Figure 1 shows a typical ultrasound image and the granular pattern appearance of the speckle noise.

The presence of speckle noise lowers the overall image quality and makes the interpretation of ultrasound images challenging for nonspecialists [22, 30]. Speckle noise can also adversely affect the identification of normal and pathological tissues by trained specialists [8, 19]. Furthermore, it lowers the accuracy of computer-aided diagnosis [5] and adversely affects subsequent image processing tasks such as segmentation [2, 5].

Over the last two decades there have been a number of methods proposed to reduce speckle noise. A number of wavelet-based methods have been proposed to decompose the ultrasound image into frequency subbands and then use various strategies to filter wavelet coefficients associated with speckle noise (see [7] for an overview of wavelet-based methods). However, these frequency domain approaches tend to oversmooth the image details by filtering excessive frequencies, or produce ringing artifacts due to removal of incorrect bands [32].

Another popular strategy for speckle removal are local image filtering methods. Among these methods, the most successful ones are those based on anisotropic diffusion (e.g., [19, 8, 31]) and the bilateral filter (e.g., [2]). While local filters are successful for speckle reduction, their performance suffers in the presence of strong noise, which corrupts the correlations between neighboring pixels [10]. In addition to local filtering, non-local filtering methods have also been proposed. Methods such as non-local means (NLM) [5, 32, 29] leverage the entire image by finding similar patches in a larger neighborhood of a target pixel. The collection of patches is then used to filter the target pixel.
These non-local self-similarity (NSS) approaches are sensitive to the quality of the selected patches and can produce blurry results if poorly similar patches are selected.

Recently, a number of NLM filters have been developed for various image processing tasks by combining NSS and low-rank priors—e.g., image denoising [9], video denoising [12], multispectral image denoising [26], and image deblurring [6]. These methods, however, target natural images and often have problems in finding candidate patches due to the severity of the speckle noise patterns present in ultrasound images. The success of these non-local methods serves as the starting point of our filtering framework.

**Contributions.** We propose a novel non-local filtering framework for speckle noise reduction. Due to the noisy nature of ultrasound images, non-local filtering methods could perform poorly when selecting candidate patches. To overcome this problem, our approach first pre-filters the input image to produce a guidance image to improve the patch selection quality. We further formulate a low-rank optimization model to process the selected patches, where the noise is considered to be sparse with the clear patch being low-rank. We describe how to modify existing low-rank optimization methods to accommodate the noisy nature of speckle noise. To verify the effectiveness of the proposed method, we test it on a number of synthetic and clinical ultrasound images, and compare our results against several state-of-the-art methods. We also evaluate our method in terms of a segmentation accuracy. Our approach shows notable improvement on a range of image quality metrics.

## 2. Proposed Filtering Framework

Figure 2 provides an overview of our proposed filtering framework. The framework begins by computing a guidance image to improve the search for candidate patches as described in Sec. 2.1. A ‘clear patch’ is estimated from the candidate patches by estimating a low-rank and sparse representation of the patch collection as described in Sec. 2.2. Lastly, the final despeckled image is produced by aggregating the restored patches as described in Sec. 2.3.

### 2.1. Non-local Patch Selection

Given a reference patch in the input image, the non-local patch selection process aims to find a group of patches similar to the reference patch based on some distance metric. Due to the large intensity variations caused by the speckle noise, direct application of the Euclidean distance is not effective. In other medical imaging modalities, such as MRI, it has been demonstrated that a pre-filtered version of the input can serve as a guidance image to assist in non-local patch selection [17]. The key is to find an appropriate method to generate a guidance image for the image modality at hand. Since speckle noise has a granular texture-like pattern, we employ the windowed inherent variation (WIV) measure [28] to generate the guidance image, $G$, from the input image $I$ as follows:

$$G(p) = \sqrt{\left| \sum_q g_{p,q} \cdot (\partial_x I)_q \right|^2 + \left| \sum_q g_{p,q} \cdot (\partial_y I)_q \right|^2},$$  \hspace{1cm} (1)

where $p$ is a pixel in $I$, $q$ is a pixel in the rectangular neighborhood centered at $p$, and $g_{p,q}$ is a weighting function based on spatial affinity, which is defined as:

$$g_{p,q} \propto \exp\left(-\frac{\text{dist}_{p,q}}{2\sigma_w^2}\right),$$  \hspace{1cm} (2)

where $\sigma_w$ controls the spatial scale of the neighboring rectangle.

In general, a patch dominated by speckle noise has a small $G$ value compared to patches with structure and features. The reason is that speckle noise is observed as a texture-like pattern with dark and bright intensities. Such a pattern leads to a large amount of positive and negative partial derivatives in all directions, while structured edges in a patch contribute to gradients in more similar directions. With the WIV-guidance image, we compute the distance between two non-local patches centered at pixels $p$ and $q$ as:

$$\text{dist}(p,q) = \|P_I(p) - P_I(q)\| + \|P_G(p) - P_G(q)\|,$$  \hspace{1cm} (3)

where $\|\cdot\|$ represents the $L_2$ norm, and $P_I(p), P_I(q), P_G(p)$ and $P_G(q)$ are the vectorized patches centered at pixels $p$ and $q$ in image $I$ and guidance image $G$, respectively.

Figure 2: An overview of our non-local low-rank filtering framework. First, we compute a guidance image to help locate the candidate patches (Sec. 2.1). Then, we refine the patches to recover the low-rank structure (Sec. 2.2) and aggregate the low-rank patches to construct the final filtered result (Sec. 2.3).
The purpose of the guidance image is to improve patch selection in noisy ultrasound images. As demonstrated in Fig. 3, the speckle noise is heavily suppressed in the WIV map, since features have larger filter response than speckle noise. While this filtered image is not suitable as a despeckled result, it serves as a good guidance image. In particular, we can see that the despeckled results are improved when using the WIV in Fig. 3. Only using the input image to measure patch similarity is inefficient to separate patches centered at features from speckle noise dominated patches, leading to feature blurring. With the patch distance defined in Eq. 3, we select the $K$ most similar patches for each patch in the input image. In our implementation, the window for searching similar patches is $(2 \times S_w + 1) \times (2 \times S_w + 1)$ with $S_w=20$ to reduce the computation time. In all the experiments, we set $K=30$ and patch size as $7 \times 7$.

### 2.2. Low-Rank Patch Recovery

After finding the $K$ most similar patches $\{P_i\}_{i=1}^{K}$ (in image $I$) for a given reference patch $P_{ref}$, we construct a patch group (PG) matrix $\Psi_I$:

$$
\Psi_I = [V(P_{ref}), V(P_1), V(P_2), ..., V(P_K)],
$$

where $V(\cdot)$ vectorizes a patch as a 49-element column vector. Similarly, we denote $\Psi_D$ as the PG matrix for each pixel in the final despeckled image $D$. Our observation on the ultrasound images is that $\Psi_D$ should be a low-rank matrix due to the strong correlation between patches after speckle removal. However, due to the speckle noise, the rank of $\Psi_I$ (the raw input) tends to be high. Therefore, we formulate a low-rank recovery process to estimate $\Psi_D$ from $\Psi_I$. That is, we decompose $\Psi_I$ into a low-rank component ($\Psi_D$) and a sparse component ($\Psi_{\eta}$) by solving:

$$
\min_{\Psi_D, \Psi_{\eta}} \text{rank}(\Psi_D) + \alpha \|\Psi_{\eta}\|_0 \quad \text{s.t.} \quad \Psi_I = \Psi_D + \Psi_{\eta},
$$

where $\text{rank}(\Psi_D)$ denotes the rank of $\Psi_D$, which equals to the $L_0$ norm of the singular values of $\Psi_D$; and $\alpha$ is a weight to balance the two regularization terms. The second sparse term is introduced to the low-rank recovery process to improve the robustness of the method against outliers caused by noise artifacts and patch grouping error.

**Truncated weighted nuclear norm.** The foregoing $L_0$ optimization is known to be NP-hard [1]. Robust principal component analysis (RPCA) [25] is a common way to solve it in a tractable way by approximating the rank operation $\text{rank}(\Psi_D)$ using the nuclear norm $\|\Psi_D\|_*$, which is defined as the sum of all the singular values of $\Psi_D$. Note that in our implementation, we use SVD to decompose $\Psi_D$ to obtain the singular values of $\Psi_D$.

In practice, the rank operation may not be well approximated using the nuclear norm [33, 9], since it minimizes all singular values equally. As a result, important image features, which correspond to large singular values, will become blurred because their corresponding singular values will be minimized extensively according to the nuclear norm. Therefore, we should assign smaller weights to larger singular values, so that their magnitude can be maintained after the minimization; this is also suggested in [9]. Similarly, the smallest singular values, which correspond to noise, can be simply removed, as suggested in [33].

In this regard, we formulate a truncated and weighted nuclear norm (TWNN), $\|\cdot\|_{tw}$, to better approximate the rank operator by combining the strength of the truncated

---

**Figure 3:** Comparing the final despeckled results with and without the use of guidance image on a clinical (top) and synthetic (bottom) ultrasound image as inputs; the synthetic image input is corrupted by a synthetic speckle noise model [5].
nuclear norm [33] and the weighted nuclear norm [9]:

\[ ||\Psi_D||_{tw} = \sum_{i=1}^{M} w_i \sigma_i(\Psi_D), \quad (6) \]

where \( M \) is the total number of the singular values; and \( w_i \) is the weight for the \( i \)-th singular value \( \sigma_i \) of \( \Psi_D \). Since we use SVD, \( M \) equals to the minimum of \( K+1 \) and \( d^2 \), which are the dimensions of the PG matrix \( \Psi_I \).

Since large singular values usually correspond to major components of the matrix (important image features) while small singular values usually correspond to noise, a natural way is to set \( w_i \) to be inversely proportional to the magnitude of the singular value [9] and to zeroize the \( w_i \)'s corresponding to the smallest singular values; hence, we define

\[ w_i = \begin{cases} 
0 & \text{if } i \leq \lambda \\
\frac{\theta}{\sigma_i(\Psi_D) + \varepsilon} & \text{otherwise} 
\end{cases}, \quad (7) \]

where \( \lambda \) and \( \theta \) are parameters, and \( \varepsilon \) is set to be 0.00001 to avoid division by zero. In all the experiments, we empirically set \( \lambda \) as 9 and \( \theta \) as \( 5\sqrt{2} \).

The initialization of \( \sigma_i(\Psi_D) \). To iteratively solve Eq. 5 with TNN, we need to initialize \( \sigma_i(\Psi_D) \), but we do not have \( D \) at the beginning. Hence, before proceeding to iteratively minimize Eq. 5, we initialize \( \sigma_i(\Psi_D) \) as

\[ \sigma_i(\Psi_D) = \sqrt{\max(\sigma_i^2(\Psi_I) - \beta, 0)}, \quad (8) \]

where \( \beta \) is a parameter that estimates the noise component. We empirically set \( \beta \) as a value in [5, 50], using a larger \( \beta \) for ultrasound images with stronger noise.

Modeling the \( ||\Psi_n||_0 \) term. Usually, this term is approximated by the \( L_1 \) norm, as in the RPCA method [25]. However, since the \( L_1 \) norm treats each element in \( \Psi_n \) independently, it does not take into account the spatial connections among groups of elements in \( \Psi_n \). Hence, we propose to employ the structured sparsity \( \Omega_n \) [16, 13] to approximate \( ||\Psi_n||_0 \) for ultrasound speckle reduction, since \( \Omega_n \) can encode the structure prior knowledge of \( \Psi_n \) by involving overlapping submatrices in \( \Psi_n \) (which is actually a \( d^2 \)-by-\((K+1)\) matrix):

\[ \Omega_n = \sum_{g \in \Psi_n} ||g||_\infty, \quad (9) \]

where \( g \) is each \( 3 \times 3 \) submatrix in \( \Psi_n \), and \( ||.||_\infty \) is the maximum value over all the elements in \( g \). Hence, each pair of adjacent groups (or submatrices) have six overlapping elements in \( \Psi_n \).

Our model. By putting the TNN (Eq. 6) and the structured sparsity (Eq. 9) into Eq. 5, we obtained the final objective function to recover the underlying low-rank matrix:

\[ \min_{\Psi_D, \Psi_n} \sum_{i=1}^{M} w_i \sigma_i(\Psi_D) + \alpha \sum_{g \in \Psi_n} ||g||_\infty \quad \text{s.t. } \Psi_I = \Psi_D + \Psi_n, \quad (10) \]

where \( \alpha \) is set to be 1.0 in the current implementation. In Fig. 4, we compare the despeckling performance of our method with the original RPCA [25]. Our model models the low-rank regularization term with the TNN and the sparsity term using structure sparsity, so we can better preserve the features than that with the original RPCA.

Optimization. We have developed an efficient optimization procedure using the alternating direction method of multipliers (ADMM) to minimize the objective function in Eq. 10. Due to space limit, we provide the details of our optimization strategy in the supplemental material. Note that our Matlab implementation will be made publicly available upon publication of this work.

2.3. Final Recovery

The procedure outlined in Sec. 2.2 is applied iteratively. In the beginning of each subsequent iteration, we adopt an iterative regularization method [27, 9] to generate a new result by adding part of the filtered speckle noise back to the current despeckled image as follows:

\[ I_h = D_{h-1} + \delta \cdot (I - D_{h-1}), \quad (11) \]

where \( I \) is the original ultrasound image; \( D_{h-1} \) is the despeckled result after \((h-1)\) iterations; and \( \delta \) denotes the amount of filtered component that is to be added back to the result to avoid oversmoothing (\( \delta=0.13 \) in our experiments). The \( I_h \) is the generated ultrasound image with the iterative regularization in Eq. 11.

Fig. 5 shows intermediate despeckled results of our method on a synthetic ultrasound image. As the iteration progresses, speckle noise is gradually suppressed while image features are revealed.
3. Experiments

We evaluate the performance of our method on a number of synthetic and clinical ultrasound images by comparing with the following state-of-the-art despeckling filters: (1) speckle reducing anisotropic diffusion (SRAD) [31], (2) squeeze box filter (SBF) [21], (3) optimized Bayesian non-local means (OBNLM) [5], (4) anisotropic diffusion guided by Log-Gabor filters (ADLG) [8], and (5) non-local mean filter combined with local statistics (NLMLS) [29].

We evaluate our approach on a total of 60 clinical images: 20 liver images, 20 breast images, and 20 gall bladder images. See supplemental material for all the results. In our implementation, all but two parameters are fixed, so only $\beta$ in Eq. 8 and the number of iterations ($H$) in final recovery (see Sec. 2.3) need to be tuned. In detail, $H$ is empirically set as $[5, 10]$, depending on the noise level. The value of $\beta$ also depends on the noise level, and we use a larger $\beta$ for ultrasound images with high speckle noise level. For all the other methods, we also tune their associated parameters until we can produce the best result. We obtain code of SRAD, OBNLM, and ADLG from their project webpages. For SBF, we obtain its code from the author, while for NLMLS, we implement the method based on the paper.

3.1. Synthetic Images

We first start with synthetic results, since it is possible to have quantitative measurements and comparisons.

**Quantitative Metrics.** We use five metrics to compare the performance of our method against others: peak signal-to-noise ratio (PSNR), Pratt’s figure of merit (FOM) [31], universal quality index (UQI) [23], structural similarity (SSIM) [24], and visual information fidelity (VIF) [20].

**Results.** For the purpose of quantitative comparisons, we generate noise over ground truth images by employing the synthetic speckle noise model in [5], which is a multiplicative Gaussian $\mathcal{N}(0, \sigma^2_s)$, where $\sigma_s$ controls the noise level. We set $\sigma^2_s$ as 0.2, 0.2, and 0.1 for the cases shown in Fig. 3 (bottom), Fig. 5, and Fig. 6, respectively.
3.2. Clinical Images

We also visually compared our method with others on a number of clinical ultrasound images. Fig. 7 shows an example on an ultrasound image with polycystic liver. We present the despeckled image and its removed speckle noise component in the first and second rows, respectively. Our method produces better despeckling results by preserving image features, while others tend to oversmooth those features. In addition, our noise layer is much more consistent and does not contain excessive structure details as compared with the others.

Table 1: Quantitative comparison for results in Fig. 6.

<table>
<thead>
<tr>
<th>Method</th>
<th>PSNR</th>
<th>FOM</th>
<th>UQI</th>
<th>SSIM</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRAD</td>
<td>32.75</td>
<td>0.9242</td>
<td>0.6933</td>
<td>0.9812</td>
<td>0.6964</td>
</tr>
<tr>
<td>SBF</td>
<td>30.77</td>
<td>0.2960</td>
<td>0.2842</td>
<td>0.2761</td>
<td>0.2730</td>
</tr>
<tr>
<td>OBNLM</td>
<td>29.14</td>
<td>0.7794</td>
<td>0.3951</td>
<td>0.9548</td>
<td>0.5403</td>
</tr>
<tr>
<td>ADLG</td>
<td>28.97</td>
<td>0.7423</td>
<td>0.1318</td>
<td>0.9611</td>
<td>0.4138</td>
</tr>
<tr>
<td>NLMLS</td>
<td>28.68</td>
<td>0.5207</td>
<td>0.1246</td>
<td>0.9484</td>
<td>0.3564</td>
</tr>
<tr>
<td>Ours</td>
<td>30.32</td>
<td>0.5238</td>
<td>0.1970</td>
<td>0.9455</td>
<td>0.3618</td>
</tr>
</tbody>
</table>

Table 2: Comparison of PSNR values for despeckled results using synthetic noise on Fig. 6(a) at different noise levels.

<table>
<thead>
<tr>
<th>σ²</th>
<th>PSNR</th>
<th>FOM</th>
<th>UQI</th>
<th>SSIM</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15</td>
<td>26.64</td>
<td>25.56</td>
<td>25.08</td>
<td>23.75</td>
<td>23.89</td>
</tr>
<tr>
<td>0.2</td>
<td>27.75</td>
<td>26.04</td>
<td>25.65</td>
<td>24.11</td>
<td>24.59</td>
</tr>
<tr>
<td>0.25</td>
<td>28.97</td>
<td>27.08</td>
<td>26.56</td>
<td>24.89</td>
<td>25.98</td>
</tr>
<tr>
<td>0.3</td>
<td>29.14</td>
<td>27.98</td>
<td>27.13</td>
<td>25.48</td>
<td>26.48</td>
</tr>
<tr>
<td>Ours</td>
<td>30.77</td>
<td>29.60</td>
<td>28.42</td>
<td>27.61</td>
<td>27.61</td>
</tr>
</tbody>
</table>

Fig. 6, we start with a clean image and add speckle noise to it; then we quantitatively compare the despeckled results of our method with SRAD, SBF, OBNLM, ADLG, and NLMLS. Visual inspection shows our method better preserves boundaries compared to the other methods, while Table 1 reports the corresponding metric values for the despeckled results. Clearly, our method outperforms others for all the five metrics. High PSNR shows that our result is more consistent with the noise-free clean image. The high FOM shows that our method has better performance in terms of edge preservation. Our method also achieves the highest UQI, SSIM and VIF values, implying that our result has the best visual quality compared to others. Moreover, we test another four noise levels σ² = {0.15; 0.2; 0.25; 0.3} on the same clean image over all the methods. Table 2 lists the resulting PSNR values, showing that our method achieves consistently high performance.
Figure 9: Comparison of speckle reduction on an ultrasound image with multiple liver cysts at various sizes. (a) Original image; results by (b) SRAD [31], (c) SBF [21], (d) OBNLM [5], (e) ADLG [8], (f) NLMLS [29], and (g) Our method.

Figure 10: Despeckled results of our method on ultrasound images of four different tissue regions. First row: original images; second row: despeckled images; and third row: removed speckle noise component.

3.3. Pre-processing for Segmentation

Our method is also effective as a pre-processing step for the breast ultrasound (BUS) image segmentation. BUS is commonly used to distinguish between benign and malignant tumors that can be characterized by the shape or contour features of segmented breast lesions [18][4][8].

Quantitative Metrics. Four metrics are used to evaluate the segmentation accuracy: a combined accuracy of true and false positive rate (AC) [3], Hausdorff distance (HD) [3], Hausdorff mean (HM) [3], and root mean square symmetric distance (RMSD) [11]. A good segmentation result should have large AC, and small HD, HM, and RMSD.

We compare the segmentation accuracy of breast lesion on the speckled image and corresponding despeckled results using a widely adopted level-set segmentation method by Li et al. [15]. The segmentation performances are evaluated on ten BUS images with different lesions. Two ex-
One reason why the segmentation performances of the other methods degrade is that their results are more blurry, and therefore, the level-set function is not able to more accurately stop at the lesion boundaries.

### 4. Conclusion

We propose a non-local low-rank filtering framework for speckle noise reduction. To overcome problems with non-local patch selection, we use a guidance image based on windowed inherent variation (WIV) filtering. To remove speckle noise within a group of similar patches, we decompose it into a low-rank component with a proposed truncated and weighted nuclear norm (TWNN) and a sparse component with the structured sparsity regularization. We also devise an efficient optimization based on the ADMM framework to solve the minimization. Both quantitative and qualitative evaluations on various synthetic and clinical images demonstrate that our method is able to effectively remove speckle noise and better preserve features compared with the state-of-the-art speckle reduction techniques. In addition, segmentation comparisons on a number of breast ultrasound images reveal that the despeckled results of our method can better facilitate breast lesion segmentation than results produced from the compared despeckling methods.
References


