Multispectral Image Denoising With Optimized Vector Bilateral Filter
Honghong Peng, Raghuveer Rao, and Sohail A. Dianat

Abstract—Vector bilateral filtering has been shown to provide good tradeoff between noise removal and edge degradation when applied to multispectral/hyperspectral image denoising. It has also been demonstrated to provide dynamic range enhancement of bands that have impaired signal to noise ratios (SNRs). Typical vector bilateral filtering described in the literature does not use parameters satisfying optimality criteria. We introduce an approach for selection of the parameters of a vector bilateral filter through an optimization procedure rather than by ad hoc means. The approach is based on posing the filtering problem as one of nonlinear estimation and minimization of the Stein’s unbiased risk estimate of this nonlinear estimator. Along the way, we provide a plausibility argument through an analytical example as to why vector bilateral filtering outperforms band-wise 2D bilateral filtering in enhancing SNR. Experimental results show that the optimized vector bilateral filter provides improved denoising performance on multispectral images when compared with several other approaches.

Index Terms—Vector bilateral filtering, Stein’s unbiased risk estimator, parameter optimization.

I. INTRODUCTION

OISE is an inevitable part of most real world multispectral and hyperspectral images and it is well known that good denoising leads to performance improvement in problems such as classification, segmentation and object identification [1]–[4]. Various approaches [5]–[16] have been proposed for noise removal in such images. Recently, a vector formulation of a bilateral filter [19] was provided in [17]. It was shown to possess several advantages, the most important being a weighting mechanism that tends to preserve edges, thus contributing to improved denoising vs. edge-preservation tradeoffs compared to other approaches such as, for example, those based on the wavelet transform.

A problem with the vector bilateral filter, however, is that the performance depends on the choice of the filter parameters. In [17] the parameters were chosen as functions of the estimated noise variance of the various PCA components of the noisy hyperspectral/multispectral input image. Although there is some basis for this approach, it is largely ad hoc. In this paper, we formulate an optimization problem by adopting the perspective that the bilateral filtered output is essentially a nonlinear estimate of the underlying image. In the specific case of multivariate Gaussian noise — it does not have to be white — the risk (to be defined later) associated with a nonlinear estimator of the mean can itself be estimated using Stein’s unbiased risk estimation (SURE) [18]. We make this the basis of an approach that estimates the parameters using the estimated SURE as an optimality criterion. It is to be noted that SURE-based optimization has been used in other denoising contexts including those founded on the wavelet transform [13]–[16]. In [13], a closed-form solution of Stein’s estimation of mean squared error (MSE) is deduced for a Maximum-a-Posteriori (MAP) multivariate denoising estimator with a Bernoulli–Gaussian prior, and parameters that correspond to minimum MSE are selected as optimal parameters. A more general framework that does not assume any prior model is proposed in [14]–[15] for the MAP multivariate denoising estimator. A multi-channel SURE with linear expansion of thresholds (SURE-LET) approach is proposed in [16] to simplify parameter optimization based on SURE. The denoising function is constructed as a linear expansion of thresholds and optimized linearly according to SURE without any prior model assumption. The proposed wavelet thresholding function in [16] is “point-wise” and depends on the coefficient vector that contains coefficients of every channel in the same location and their parent coefficients in the coarser wavelet sub-band. This approach is a state of the art denoising technique for multispectral images.

The authors have proposed 2D bilateral filter parameters optimization in [30], which only optimizes a pair of fixed parameters, the geographical distance parameter and pixel similarity differences parameter, for a scalar image. In the present paper, we extend the framework in [30] to the vector bilateral image filtering case to optimize an arbitrary number of parameters. The contribution of the present paper is in showing that (a) for multi-spectral images, the vector bilateral filter demonstrates superior performance over the component-wise bilateral approach, (b) the vector bilateral filter parameter optimization can be cast into a SURE-based framework and (c) the optimized filter provides improved noise removal vs. edge preservation tradeoff compared to techniques in [16] and [30].

The paper is organized as follows. In Section II, the bilateral filter is extended to a vector form and its parameter optimization procedure is proposed based on Stein’s unbiased risk
estimator. Section III shows the experimental results. Finally, the paper is concluded in Section IV.

II. OPTIMIZED VECTOR BILATERAL FILTER

A. Vector Bilateral Filter

Let \( \mathbf{I}_{in} \) be an acquired noisy \( k \)-band multispectral/hyperspectral image (\( k = 3 \) for a color image) with spatial dimension of \( H \) rows and \( L \) columns. Let

\[
\mathbf{I}_{in}(s) = \mathbf{I}_{real}(s) + \mathbf{n}(s)
\]

where \( s \) is a 2-D vector denoting the pixel’s coordinates, \( \mathbf{I}_{in}(s) \) represents the \( k \)-dimensional vector often referred to as the spectral vector at position \( s \), \( \mathbf{I}_{real}(s) \) denotes the actual image, and \( \mathbf{n}(s) \) is a (generally non-white) Gaussian additive noise vector whose covariance matrix is \( \Gamma_n \).

For a pre-defined domain kernel \( \Omega \), a \( n \times n \) neighborhood centered at the target pixel at position \( s \), the vector bilateral filter takes the form [18]:

\[
\mathbf{I}_{out}(s) = f_{vec-bilateral}(\mathbf{I}_{in}(s), (\sigma_d, \Sigma_s)) = \sum_{p \in \Omega} g_d(|p - s|, \sigma_d) g_s(D(p, s), \Sigma_s) \mathbf{I}_{in}(p)
\]

\[
= \sum_{p \in \Omega} g_d(|p - s|, \sigma_d) g_s(D(p, s), \Sigma_s)
\]

where \( \mathbf{I}_{in}(s) \) and \( \mathbf{I}_{out}(s) \) are the input and output images of the vector bilateral filter, \( \mathbf{p} \) is a 2-D vector representing the pixel coordinates, \( D(p, s) = \mathbf{I}_{in}(p) - \mathbf{I}_{in}(s) \) is the pixel value difference, \( g_d(x, \sigma_d) \) and \( g_s(x, \Sigma_s) \) are weight functions for geometric distance and pixel value difference, respectively. \( g_d(x, \sigma_d) \) is defined as a Gaussian function:

\[
g_d(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right)
\]

It is very similar in form to the geometric distance weight function of a scalar bilateral filter [20] for gray scale images. Except for the vectorized input and output, the major difference between the vector and scalar cases lies in the weight function for the pixel value distance metric, for which the authors proposed using [17]

\[
g_s(x, \Sigma) = \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} \exp\left(-\frac{(x^T\Sigma^{-1}x)}{2}\right)
\]

where \( \Sigma \) is a positive definite matrix that we will refer to as the correlation matrix or the range kernel matrix. From Eqs. (2) and (4) it is seen that the exponent of \( g_s(x, \Sigma) \) has the squared Mahalanobis distance between \( \mathbf{I}_{in}(p) \) and \( \mathbf{I}_{in}(s) \).

Compared to a component-wise 2D bilateral filter, one important advantage of the proposed vector bilateral filter is that it can exploit correlation between bands more efficiently, and as shown below, this advantage increases as the number of bands increases. Considering that edges are the salient structures in an image, we demonstrate this point here using an image that has a single sharp edge. Assume we have a \( k \)-band multi-spectral image with similar perfect edge images in each band. The \( ith \) band image, for any \( i \) between 1 and \( k \) is assumed to have the form shown in Fig. 1 with signal edge value \( I_{i,j} \) and constant background value \( I_{b,j} \). In this way, the correlation between any two band images is 1.

Note that the edge transition is assumed to occur at the same locations in all the bands. However the edge height, determined by the difference between the top or signal value \( I_{i,j} \) and the bottom or background value \( I_{b,j} \), varies with the bands.

Also assume that the signal image is contaminated with zero mean additive uniform Gaussian white noise to form the actual observed multispectral image. The noise covariance matrix is \( \Gamma_n = \sigma^2 \mathbf{E}_k \), where \( \sigma \) is the identical noise standard deviation for each channel and \( \mathbf{E}_k \) is a \( k \times k \) identity matrix. Then for a center pixel with value \( \mathbf{I}_s(s) = I_{i,j} + n(s) \), right on the edge position \( s \) in the observed \( ith \) band image, the corresponding denoising estimation result \( \hat{I}(s) \) can be expressed as:

\[
\hat{I}(s) = \frac{\sum_{j \in \Omega^-} W_j^- (I_{b,j} + n_{j,i}) + \sum_{l \in \Omega^+} W_l^+ (I_{i,j} + n_{l,i})}{\sum_{j \in \Omega^-} W_j^- + \sum_{l \in \Omega^+} W_l^+}
\]

where \( W_j^- \) is the positive weighting factor for pixels that belong to background kernel area \( \Omega^- \) with value of noise \( n_{j,i} \) plus background \( I_{b,j} \), and \( W_l^+ \) is the positive weighting factor for pixels that belong to signal kernel area \( \Omega^+ \) with the value of signal \( I_{i,j} \) plus noise \( n_{l,i} \). As assumed before, the entire signal has the same value, i.e. \( I(s_j) = I_{i,j} \). Accordingly, we can derive signal to error (SER) measure for this estimation as:

\[
\text{SER}_{edge} = \left| \frac{I_{i,j} - \hat{I}(s)}{I_{i,j} - I(s)} \right|
\]

\[
= \frac{\left| \sum_{j \in \Omega^-} W_j^- (I_{b,j} + n_{j,i}) + \sum_{l \in \Omega^+} W_l^+ (I_{i,j} + n_{l,i}) \right|}{\sum_{j \in \Omega^-} W_j^- + \sum_{l \in \Omega^+} W_l^+}
\]

\[
= \frac{\left| \sum_{j \in \Omega^-} W_j^- (I_{b,j} - I_{i,j}) \right|}{\sum_{j \in \Omega^-} W_j^-}
\]

(6)
Assuming the edge signal, $I_{1,i} - I_{b,i}$, is significantly larger than the noise, we can further approximate the SER$_{edge}$ as:

$$\text{SER}_{edge} \approx \frac{I_{1,i} - I_{b,i}}{I_{b,i} - I_{1,i}} \left(1 + \frac{\sum_{j \in \Omega^+} W_j^+}{\sum_{j \in \Omega^-} W_j^-}\right)$$

$$= \frac{I_{1,i} - I_{b,i}}{I_{b,i} - I_{1,i}} \left(1 + \sum_{j \in \Omega^+} W^j R_l\right)$$

where

$$W^j R_l = \frac{W_j^+}{\sum_{j \in \Omega^-} W_j} \quad (7)$$

Thus, the SER$_{edge}$ measure is positively related to each individual $W^j R_l$. The larger each $W^j R_l$ is, the better the noise filtering result will be. The ideal denoising filter weighting kernel for the surrounding pixels would intuitively be such that the pixels with similar values to the center pixel are assigned higher weight while pixels with significant different values are assigned lower weight. In the perfect edge case, the ratio of weights between the signal and background areas is very similar for observed signal region and background noise region. Thus, we can only consider one ratio to predict the estimation kernel performance. The higher the value of this ratio, the better SER and filtering results will be.

If we only consider the pixel value weighting kernel, the fixed geometric weighting will not affect the final result. For the vector bilateral filter case, if we set the range covariance matrix $\Sigma$ to be a diagonal matrix with parameter $\Sigma_{s,d}$ for the $d$th band, the filtering weight ratio would be given by:

$$W_{l,vec} = \exp \left( \sum_{d=1}^{k} \frac{-(n(s)d-n_{s,d})^2}{2\sigma^2_{s,d}} \right)$$

$$\sum_{j \in \Omega^-} \exp \left( \sum_{d=1}^{k} \frac{-(l_{1,d} - l_{b,d} + n(s)d - n_{s,d})^2}{2\sigma^2_{s,d}} \right)$$

$$\approx \frac{1}{\sum_{j \in \Omega^-} \prod_{d=1}^{k} \exp \left( \frac{(l_{1,d} - l_{b,d})^2}{2\sigma^2_{s,d}} \right)}$$

$$= J \prod_{d=1}^{k} \exp \left( \frac{(l_{1,d} - l_{b,d})^2}{2\sigma^2_{s,d}} \right) \quad (8)$$

where $n(s)d, n_{l,d}, n_{j,d}$ are additive noise at the center pixel, the signal side pixel and background side pixel at band $d$ image respectively. $J$ is the number of pixels belong to background kernel area $\Omega^-$. The approximation is made with high signal to noise ratio assumption. With previous hypothesis, if 2D bilateral filter is applied separately to $i$th band image, it will be a special case of vector bilateral filter, the corresponding $WR_{l,2D}$ is simplified to $J \exp((l_{1,i} - l_{b,i})^2/2\sigma^2_{s,i})$. As for the assumed idea scenario, $\exp((l_{1,i} - l_{b,i})^2/2\sigma^2_{s,i})$ is always larger than 1 for each band, the vector bilateral filter will always have a larger $WR_l$ and hence outperform component wise 2D bilateral filter by utilizing more band images, this performance advantage will increase with band numbers. Since for multispectral images there is usually a strong degree of correlation of spatial edge features between the bands, the proposed vector bilateral filter should commonly benefit from this fact.

### B. Mean-Squared Error (MSE) and Stein’s Unbiased Risk Estimator (SURE)

We now address the problem of finding $\sigma_d$ and $\Sigma$ such that the vector bilateral filter offers optimum performance in terms of tradeoff between noise removal and edge-smearing across all spectral band images.

For a (noisy) observed $k$-dimensional data vector in Eq. (1), suppose $I_{out,\phi}(s)$ is a denoised image from $I_{in}(s)$ obtained as:

$$I_{out,\phi}(s) = f(I_{in}(s), \theta) = f(I_{real}(s) + n(s), \theta), \quad s \in T \quad (9)$$

where $f$ is a nonlinear estimator of $I_{real}(s)$, $\theta$ is a parameter vector associated with this estimator, and $T$ is the set of spatial indices of the whole image ($T = \{T_1, T_2 \ldots T_{HL}\}$). The total number of pixels in the image is $HL$. The quality of this denoising estimator $f$ is often evaluated using sample mean square error (MSE) measure expressed in L2 norm as:

$$MSE = \frac{1}{HL} \sum_{s \in T} \left[ I_{out,\phi}(s) - I_{real}(s) \right]^2 \quad (10)$$

The difficulty in applying this measurement metric to the observed noisy image is that the underlying image, $I_{real}(s)$, is unknown. The MSE is a random variable depending on noise, and the expected value of $MSE$ in Eq.(10) is referred to as the Risk $R_0$:

$$R_0 = E[MSE] \quad (11)$$

The problem of estimating Risk without access to ground truth image is circumvented to some extent with Stein’s Unbiased Risk Estimator (SURE) [15]–[16], [18]. With additive multivariate Gaussian noise hypothesis, SURE provides an analytical means for unbiased estimation of MSE. It is given by:

$$\hat{R}_0 = \frac{1}{HL} \sum_{s \in T} \left[ I_{in}(s) - I_{out,\phi}(s) \right]^2$$

$$- Tr(\Gamma_n) + 2 \frac{1}{HL} \sum_{s \in T} Tr \left( \Gamma_n J_f(I_{in}(s)) \right) \quad (12)$$

where $\Gamma_n$ is the noise covariance matrix and $J_f(I_{in}(s))$ is the Jacobian matrix with respect to $I_{in}(s)$. The matrix element in the $i$-th row and $j$-th column of $J_f(I_{in}(s))$ is given by:

$$J_f(I_{in}(s))_{i,j} = \frac{\partial f_i(I_{in}(s), \theta)}{\partial I_{in,j}(s)} \quad (13)$$

It is an unbiased estimator for the expectation of MSE in(10):

$$R_0 = E[MSE] = E \left[ \hat{R}_0 \right] \quad (14)$$

For image denoising purpose, we regard $\hat{R}_0$ as a reliable estimate of the MSE for optimization, as the total number of pixels, $HL$, in an image is usually a very large number.
C. Vector Bilateral Filtering Parameter Optimization

By substituting the $f(I_{in}(s), \theta)$ in eq. (12) with the proposed vector bilateral function $f_{vec-bilateral}(I_{in}(s), (\sigma_d, \Sigma_s))$ in eq. (2), we obtain an expression of SURE for the proposed vector bilateral filter [20]:

$$
\hat{R}_{(\sigma_d, \Sigma_s)} = \frac{1}{HL} \sum_{s \in T} \left[ \left\| I_{in}(s) - I_{out,(\sigma_d, \Sigma_s)}(s) \right\|^2 \right] - Tr(\Gamma_n) + 2 \frac{1}{HL} \sum_{s \in T} \left[ Tr \left( \Gamma_n^T J_{f_{vec-bilateral}} (I_{in}(s)) \right) \right]
$$

(15)

The parameter set $(\sigma_d, \Sigma_s)$ corresponds to $\theta$ in Eq. (9). A closed form expression in terms of observed signal $I_{in}(s)$ and parameters $(\sigma_d, \Sigma_s)$ for the gradient vector $J_{f_{vec-bilateral}}(I_{in}(s))$ is derived in Appendix as:

$$
J_{f_{vec-bilateral}}(I_{in}(s))_{i,j} = \left( \sum_{p \in \Omega} A(p) \left( (I_{in}(p)-I_{in}(s))^T \left( \Sigma_s^{-1}+\Sigma_i^{-1}T \right) \right) I_{in,j}(s) \right) + \delta_{i,j}
$$

$$
J_{f_{vec-bilateral}}(I_{in}(s))_{i,j} = \left( \sum_{p \in \Omega} A(p) \left( (I_{in}(p)-I_{in}(s))^T \left( \Sigma_s^{-1}+\Sigma_i^{-1}T \right) \right) I_{in,j}(s) \right) + \delta_{i,j}
$$

(16)

where $\delta_{i,j}$ is the delta function and $A(p)$ is defined as:

$$
A(p) = g_d(\|p-s\|, \sigma_d)
$$

$$
\times \exp \left( -\frac{1}{2} (I_{in}(p)-I_{in}(s))^T \Sigma_s^{-1} (I_{in}(s)-I_{in}(p)) \right)
$$

(17)

The noise covariance matrix $\Gamma_n$ can either be measured from sensor calibration data or be estimated efficiently with the median absolute deviation method [21]–[23]. The diagonal terms of the noise covariance matrix $\Gamma_n(j, j)$ are estimated with [23]:

$$
\hat{\Gamma}_n(j, j) = \left( \frac{1.4826 \text{median}(\|I_{in,j} - \text{median}(I_{in,j})\|)}{j \in \{1, \ldots, k\}} \right)^2
$$

(18)

The off-diagonal terms of the estimated noise covariance matrix $\hat{\Gamma}_n(i, j)$ are defined as:

$$
\hat{\Gamma}_n(i, j) = \frac{1.4826^2}{4ab} \times \left( \text{median} \left( \|aI_{in,i} + bI_{in,j} - \text{median}(aI_{in,i} + bI_{in,j})\| \right)^2 - \text{median} \left( \|aI_{in,i} - bI_{in,j} - \text{median}(aI_{in,i} - bI_{in,j})\| \right)^2 \right)
$$

$$
a = (\hat{\Gamma}_n(i, i))^{-1/2}, \quad b = (\hat{\Gamma}_n(j, j))^{-1/2}
$$

(19)

As $\hat{R}_0$ is analytically defined and can be computed numerically, we pose the bilateral filter parameter optimization problem as a constrained optimization problem as defined as:

$$
(\sigma_{d, opt}, \Sigma_{s, opt}) = \arg \min_{(\sigma_d, \Sigma_s)} \hat{R}_{(\sigma_d, \Sigma_s)}, \quad \text{s.t. } \sigma_d > 0, \Sigma_s > 0
$$

(20)

Algorithm 1 Optimized Vector Bilateral Filter Algorithm (OVBF)

Given the input vector image $I_{in}$

**Step 1. Initialization:**

- Set initial parameter $\sigma_{d, opt}$, $\Sigma_{s, opt}$.
- Set start $t = 0$. Set maximum iteration number $t_{max}$ and stop threshold $\epsilon$.

**Step 2. Iteration:**

1. Calculate $R_t = \hat{R}_{(\sigma_d, \Sigma_s)}$ by (15)
2. Calculate $I_{out}$ by (2)
3. $t = t+1$
4. Update $\sigma_{d, opt}$ with SQP
5. Update $\Sigma_{s, opt}$ with SQP
6. Calculate $R_{t+1} = \hat{R}_{(\sigma_{d, opt}, \Sigma_{s, opt})}$

**Step 3. Output optimal $I_{out}$ with minimal $\hat{R}_{(\sigma_d, \Sigma_s)}$**

This constrained non-linear optimization is solved numerically using the sequential quadratic programming (SQP) method. In essence, the solution to Eq. (20) maximizes the signal to noise ratio since the risk is an estimate of the noise power.

Our optimized vector bilateral filter algorithm is summarized as shown in Algorithm 1.

The SQP calculation can be implemented with MATLAB function `fmincon`. As the calculation of vector bilateral filter output $I_{out}$ and risk estimation $\hat{R}_{(\sigma_d, \Sigma_s)}$ share common elements, their computation can be incorporated together to increase efficiency. For optimized MATLAB implementation, the computation of vector bilateral filter on a 256 $\times$ 256 $\times$ 3 image takes 5 s on a Dell XPS laptop with 2.4 GHz Intel i7 processor, while the risk estimation takes additional 3 s.

### III. Experiments

**A. Experiment Design**

To compare and analyze the performance of the proposed approach, Monte Carlo simulation experiments were carried out on ground truth multispectral images which were contaminated with simulated multivariate zero mean additive Gaussian noise. The noise covariance matrices were varied to generate different test images. The denoising performances were evaluated with the metric:

$$
PSNR = 20 \log_{10} \left( \frac{\text{bitdepth} - 1}{\sqrt{MSE}} \right) dB.
$$

(21)

where bitdepth refers to the maximum ground truth image bit depth across all bands, $MSE$ is defined in Eq. (10).

**B. Experiment Ground Truth Targets**

Instead of randomly selecting specific natural scene images as ground truth images, we employed the dead leaves model to generate synthetic images that are more representative of the overall natural image statistics. The original dead leaves model
is proposed in [24], where it has been demonstrated that specific dead leaves models can reproduce most known statistics of natural images [25]–[27]. We generated the multispectral image test target with modified code for grayscale dead leaves target. Circle objects with uniformly random distributed size and gray level in each band are arbitrarily placed and occluded in a fixed size blank image with a looking up manner. By this means, the combination space of contrast and texture size change is well covered with enough samples. In our experiment, the PSNR results are averaged over one hundred fifty synthetic implementations of the color dead leaves model to reinforce complete coverage of the sampling space. For each implementation, both target image and noise realization change to simulate the averaged general result trend with large variety of nature images. An example of the synthetic color dead leaves image with size of 256 × 256 × 3 is shown in Fig. 2(a).

We have also conducted experiment on several real world multispectral images extracted from Hyperspectral Digital Imagery Collection Experiment (HYDICE) set [28]–[29]. For each target test image, the test result is recorded as average of twenty times simulation realization. For each simulation, the target image keeps the same, only noise realization changes to ensure a reliable quantitative assessment of the specific target image. The selected 306 × 306 pixel image covers a qualitatively complex suburban scene with a ground resolution of two meters. One band image extracted from the original multispectral image is shown in Fig. 2(a).

C. Comparisons Method

Two denoising algorithms are used to compare with the proposed optimized vector bilateral filter:

1) The Vector SURE-LET Multichannel image denoising algorithm [16], for which good results have been reported and hence is a good reference for evaluation. The general MAP multivariate SURE denoising estimator proposed in [15] reported slighter better performance than the SURE-LET implementation with average SNR advantage around 0.2 and maximum SNR difference around 0.8, but the neither the code or the test data are available. Considering the small difference in performance shown in the results of [15], these two algorithms can be regarded close in performance. To be fair, we would like to mention the SURE-LET implementation has the potential of additional performance gain with undecimated wavelet transform.

2) The optimal 2D bilateral filter [30] applied separately to each channel of the multispectral image. This reference will demonstrate the advantage of proposed vector bilateral filter when the number of available bands increases.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard deviation of noise σ per channel</td>
<td></td>
<td></td>
<td>26.42</td>
</tr>
<tr>
<td>σ = 20</td>
<td>22.11</td>
<td>29.42</td>
<td>32.32</td>
</tr>
<tr>
<td>σ = 30</td>
<td>18.59</td>
<td>25.60</td>
<td>27.51</td>
</tr>
<tr>
<td>σ = 40</td>
<td>16.09</td>
<td>23.18</td>
<td>24.40</td>
</tr>
<tr>
<td>σ = 50</td>
<td>14.15</td>
<td>21.73</td>
<td>22.50</td>
</tr>
<tr>
<td>σ = 100</td>
<td>8.13</td>
<td>18.50</td>
<td>18.62</td>
</tr>
</tbody>
</table>

D. Experiment Test Case Results and Discussion

1) White Noise Image Test Case: To illustrate how the proposed method compares to other methods for different values of noise power, 256 × 256 × 3 color dead leaves targets are generated, and corrupted with zero-mean white Gaussian noise of different uniform noise covariance matrices \( \Gamma_{n,1} = \sigma^2 I_3 \), where \( \sigma \) changes from 20 to 100 and \( I_3 \) denotes the 3 × 3 identity matrix. The obtained results are provided in Table I.

Judging from quantitative PSNR measure, the proposed vector bilateral filter outperforms the comparison methods clearly throughout most of different noise strength range. In the case of medium noise strength (\( \sigma \) ranges from 20 to 50), the advantage is around 1 dB gain. With severe noise presence (\( \sigma \) up to 100), the advantage narrows to around 0.1 dB. This narrowing difference trend can be explained by the following observation: if the noise variance is really large, the contamination will strongly demolish weak part of the original signal to an unrecoverable point across all band images. For each target test image, the test result is recorded as average of twenty times simulation realization. This trend is illustrated in Fig. 2. As the results are based on the average of many different texture image instances, i.e. different combination of dead leaf patterns and noise realization, this Fig. 2 should well represent the denoising performance of different method on typical noisy texture images. And it evidently shows the recognizable advantage of proposed method in medium and low noise situations.
One real color image rendering example is provided in Fig. 3 for visual inspection. An instance of the color dead leaves model is corrupted with medium noise strength of \( \sigma \) equals 30. It is shown in Fig. 3(c) that the propose vector bilateral filter preserved small details better. More small circle objects can be discerned from the denoising results compared to other methods and the boundaries of the vector bilateral filter rendering are more naturally smoothed, while ringing effect can be inspected in the vector SURE-LET results.

Similar test is carried out on a multispectral subset of eleven band images (ranging from 450nm to 650nm with 20 nm interval in the visible wavelength), which is extracted from the clear target HYDICE image and corrupted with zero-mean additive white Gaussian noise of different noise covariance matrices \( \Gamma_{n,2} = \sigma^2 I_{11} \). \( \sigma \) varies from 20 to 100 and \( I_{11} \) denotes the \( 11 \times 11 \) identity matrix. The results are shown in Table II.

The quantitative results show same trend with synthetic color image results, the advantage of the vector bilateral filter is stronger in this HYDICE image as more band images are utilized.

2) Colored Noise Test Case: As the noise in the multispectral image is not always white, we also compare all the algorithms for color Gaussian noise adding, zero mean additive colored Gaussian noise with covariance matrix:

\[
\Gamma_{n,4} = 30^2 \begin{bmatrix}
1 & 0.5 & 0.5 \\
0.5 & 1 & 0.5 \\
0.5 & 0.5 & 1
\end{bmatrix}
\]

to targeting color dead leaves images. The comparison is shown in Table III. Same conclusion as in the white Gaussian noise case can be drawn.
Fig. 4. $\hat{R}_0$ vs. MSE plot.

TABLE V

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>18.59</td>
<td>25.60</td>
<td>27.19</td>
<td>28.42</td>
</tr>
<tr>
<td>5</td>
<td>18.59</td>
<td>25.60</td>
<td>28.70</td>
<td>29.67</td>
</tr>
<tr>
<td>7</td>
<td>18.59</td>
<td>25.60</td>
<td>29.62</td>
<td>30.46</td>
</tr>
<tr>
<td>9</td>
<td>18.59</td>
<td>25.60</td>
<td>30.24</td>
<td>30.94</td>
</tr>
<tr>
<td>11</td>
<td>18.59</td>
<td>25.60</td>
<td>30.83</td>
<td>31.57</td>
</tr>
</tbody>
</table>

TABLE VI

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>18.59</td>
<td>27.34</td>
<td>24.39</td>
<td>30.22</td>
</tr>
<tr>
<td>5</td>
<td>18.59</td>
<td>27.34</td>
<td>24.40</td>
<td>30.43</td>
</tr>
<tr>
<td>7</td>
<td>18.59</td>
<td>27.33</td>
<td>24.40</td>
<td>30.52</td>
</tr>
<tr>
<td>9</td>
<td>18.59</td>
<td>27.33</td>
<td>24.41</td>
<td>30.65</td>
</tr>
<tr>
<td>11</td>
<td>18.59</td>
<td>27.33</td>
<td>24.41</td>
<td>30.73</td>
</tr>
</tbody>
</table>

Repeating the test on target HYDICE image with noise covariance matrix:

$$\Gamma_{\text{n},4} = 30^2 \begin{bmatrix} 1 & 0.5 & \ldots & 0.5 \\ 0.5 & 1 & \ldots & 0.5 \\ \vdots & \vdots & \ddots & \vdots \\ 0.5 & 0.5 & \ldots & 1 \end{bmatrix}$$

The result is shown in Table IV.

3) Increased Number of Bands Case: To show how the number of bands will impact the performance of different algorithms on complex image, two experiments are conducted.

In the first test, one grayscale dead leaves image is created and duplicated with random scale to multiple bands to be used as target signal image. Gaussian white noise with same noise standard deviation of 30 for each band is added to generate the final test image. In this way, the signal is totally correlated. The results are shown in Table V.
Fig. 6. Experiment results of HYDICE waste band image: (a) Waste band image at 1946.23 nm (band 151) (b) Result of proposed optimal vector bilateral filter (c) Result of component wise optimal 2D bilateral filtered (d) Result of vector SURE-LET.

As predicted by theory, the vector bilateral filter shows significant performance improvement (almost 3 dB) as the number of bands increases (from 3 to 11), while the performance of 2D bilateral remains unchanged. It is noticeable that the vector SURE-LET approach also benefits considerably from the increasing band number in this ideal case.

In the second test, a real world situation is examined. For the target HYDICE image, the first 3, 5, 7, 9 and 11 band images are chosen as the target signal images. The correlation between band images is measured by the average correlation coefficient defined as:

\[
C_{\text{mean}} = \text{mean} \left( \sum_{i=1}^{k} \sum_{j=1, j \neq i}^{k} C(i, j) \right)
\]

where \(C(i, j)\) is correlation coefficient between band \(i\) and \(j\). Gaussian white noise with same noise standard deviation of 30 for each band is added for test. In our test signal image, no matter how many band images are selected, there is always considerable amount of correlation between the signal bands images, the average correlation coefficients are all above 99%.

From the results provided in Table VI, while other methods show little or no improvement, the proposed vector bilateral filter shows again the best performance and noticeable performance improvement (about 0.5 dB) with the increase of the number of bands (from 3 to 11 bands). One band image example (at wavelength 490 nm) is also shown in Fig. 5 for visual comparison. It clearly shows that as the total number of bands rises from 3 to 11, for the same individual band, the denoising result is sharper in edges and more small details (Fig. 5(d)) are recovered. As very often, real world multi-spectral images show strong correlation between bands and majority of multi-spectral image bands have high SNR, these two experiments should reflect the nature of real world multi-spectral image denoising well and demonstrate the strong advantage of the proposed vector bilateral filter in denoising multispectral images, especially for hyperspectral images with large number of bands.

4) Waste Band Image Recovery Case: We also demonstrate the performance of proposed vector bilateral filter with the whole HYDICE urban hyper spectral image of 210 bands, in which many band images are considered as waste band images with little signal and strong noise. The results for waste band image at wavelength 1946.23 nm (band 151) are shown in Fig. 6. It is clearly shown in visual comparison that the optimized vector bilateral filter successfully recovers important signal features such as edge and prominent small features.

IV. CONCLUSION

In this paper, we have proposed an optimized vector bilateral filter approach based on Stein’s principle for denoising multispectral images corrupted by additive Gaussian noise. The basis for the approach is the viewpoint that the output of the bilateral filter is a nonlinear estimate of the underlying noiseless image. Our experiments show that this method leads to improved performance in both quantitative and visual quality measures compared with the wavelet based multispectral image denoising approach as well as the component wise optimized 2D bilateral filter technique. Furthermore, the proposed method demonstrates amplified performance advantage over component wise approach as the number of bands increases.

This work can potentially be extended to other bilateral filter types [31] and the non-Gaussian noise case [32], [33] in future research.

APPENDIX

Detailed derivation for Eq.(16):

\[
J_{\text{vec-bilateral}} (I_{in}(s))_{i,j} = \frac{\partial J_{\text{vec-bilateral},i}(I_{in}(s), \theta)}{\partial I_{in,j}(s)}
\]

\[
= \frac{\sum_{p \in \Omega} g_d(p-s) \exp \left( -\frac{1}{2} (I_{in}(s) - I_{in}(p))^T \frac{1}{\Sigma_1} (I_{in}(s) - I_{in}(p)) \right) I_{in,j}(s)}{\sum_{p \in \Omega} g_d(p-s) \exp \left( -\frac{1}{2} (I_{in}(s) - I_{in}(p))^T \frac{1}{\Sigma_1} (I_{in}(s) - I_{in}(p)) \right)}
\]

\[
= \frac{\sum_{p \in \Omega} A(p) I_{in,j}(s) + \delta_{i,j}}{\sum_{p \in \Omega} A(p)}
\]

\[
= \left( \sum_{p \in \Omega} \frac{\partial A(p)}{\partial I_{in,j}(s)} I_{in,j}(s) + \delta_{i,j} \right)
\]

\[
- \left( \sum_{p \in \Omega} \frac{\partial A(p)}{\partial I_{in,j}(s)} \right) \left( \sum_{p \in \Omega} A(p) I_{in,j}(s) \right) \frac{\sum_{p \in \Omega} A(p)}{\sum_{p \in \Omega} A(p)}
\]
where
\[ A(p) = g_A(p - s, \sigma_I) \]
\[ \times \exp \left( \frac{-1}{2} (I_{in}(s) - I_{in}(p))^T \Sigma_{s}^{-1} (I_{in}(s) - I_{in}(p)) \right) \]
and
\[ \delta_{i,j} = \begin{cases} 
1, & \text{if } i = j \\
0, & \text{if } i \neq j 
\end{cases} \]
\[ \frac{\partial A(p)}{\partial I_{in,j}(s)} = A(p) \left( (I_{in}(p) - I_{in}(s))^T \frac{1}{2} \left( \Sigma_{s}^{-1} + \Sigma_{s}^{-1} \right) \right)_{j} \]

ACKNOWLEDGMENT

H. Peng thanks Dr. J. Chen for many fruitful discussions.

REFERENCES


Honghong Peng is currently a Ph.D. Candidate with the Center of Image Science, Rochester Institute of Technology. He received the M.S. degree in optical engineering from the Beijing Institute of Technology, Beijing, China, in 2001. His current research interests include image/video processing, computer vision, and pattern recognition.

Raghuveer Rao received the B.E. degree in electronics and communication from Mysore University, Mysore, India, the M.E. degree in electrical communication engineering from the Indian Institute of Science, and the Ph.D. degree in electrical engineering from the University of Connecticut. From 1987 to 2008, he was with the Rochester Institute of Technology, where he was a Professor of electrical engineering and imaging science. His research contributions are in the areas of signal/image processing and digital communications. Since 2008, he has been with the Sensors and Electron Devices Directorate, Army Research Laboratory, Adelphi, MD, USA, where he is the Chief of the Image Processing Branch. He has held visiting appointments with the Indian Institute of Science, Princeton University, the Naval Surface Warfare Center, and the Air Force Research Laboratory. He is fellow of SPIE.
Sohail A. Dianat is currently a Professor and the Department Head with the Department of Electrical and Microelectronic Engineering and a Professor with the Imaging Science Center, Rochester Institute of Technology. He received the B.S. degree in electrical engineering from the Arya-Mehr University of Technology, Tehran, Iran, in 1973, and the M.S. and D.Sc. degrees in electrical engineering from George Washington University, in 1977 and 1981, respectively. In 1981, he joined the Rochester Institute of Technology, where he is currently a Professor of electrical engineering and imaging science. His current research interests include digital signal/image processing and wireless communication. He is the author or co-author of three books and holds 15 patents in the field of control for digital printing. He is a member of the IEEE Signal Processing and Communication Society. He is a fellow of the SPIE: The International Society for Optical Engineering. He has served as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING. He was the Organizer and Co-Chairman of the Fifth IEEE Signal Processing Society’s Workshop on Spectrum Estimation and Modeling, Rochester, NY, USA, in 1990. He has served as a Technical Program Committee Member for ICIP 2003. He has served as a Chair for the track on Wireless Sensing, Localization, and Processing at the SPIE Defense and Security Symposium Conferences, Orlando, FL, USA, from 2000 to 2013. He has served as a Finance Chair of ICIP2013. He has numerous publications in the fields of signal/image processing and digital communications.