

SUPPLEMENT PAPER OF HYPERSPECTRAL IMAGE DENOISING USING NON-CONVEX FRACTION FUNCTION

Tianyu Liu¹, Zhi Wang^{1,}, Jianping Gou¹, Wu Chen²*

¹College of Computer and Information Science, Southwest University, China

²College of Software, Southwest University, China

1. SUPPLEMENT

In this section, we describe proofs, and results that are covered in our paper but cannot be expanded in the main text due to space constraints.

1.1. Proofs of Theorem 1

In this section, we provide a complete proof of Theorem 1 proposed in the original paper. Theorem 1 guarantees the convergence of our proposed algorithm.

Lemma 1. The sequence $\{\Lambda_{k+1}\}$ is bounded

Proof. The optimal S_{k+1} needs to satisfy the first-order local optimality condition,

$$\begin{aligned} 0 &\in \partial_S \mathcal{L}(L_{k+1}, S, \Lambda_k; \mu_k) |_{S_{k+1}} \\ &= \partial_S (\lambda \|S\|_1) |_{S_{k+1}} - \Lambda_k - \mu_k (D - L_{k+1} - S_{k+1}) \\ &= \partial_S (\lambda \|S\|_1) |_{S_{k+1}} - \Lambda_{k+1} \end{aligned}$$

The subgradient of $\|S\|_1$ can be written as

$$\partial_S (\|S\|_1) |_{S_{k+1}} = \begin{cases} 0, & \text{if } S_{k+1}(i, j) = 0 \\ \pm 1 & \text{otherwise.} \end{cases}$$

It is clearly that $\partial_S (\|S\|_1) |_{S_{k+1}}$ is bounded. Thus, $\{\Lambda_{k+1}\}$ is bounded, which completes the proof. \square

Lemma 2. Sequence $\{D_k\}$ and $\{S_k\}$ are bounded if $\sum_{i=1}^{\infty} \left(\frac{\mu_i + \mu_{i+1}}{\mu_i} \right)^2 < \infty$.

Proof. After some simple deduction, we can get

$$\begin{aligned} &\mathcal{L}(L_k, S_k, \Lambda_k; \mu_k) \\ &= \mathcal{L}(L_k, S_k, \Lambda_{k-1}; \mu_{k-1}) + \frac{\mu_k - \mu_{k-1}}{2} \|D - L_k - S_k\|_F^2 \\ &\quad + \langle \Lambda_k - \Lambda_{k-1}, D - L_k - S_k \rangle \\ &= \mathcal{L}(L_k, S_k, \Lambda_{k-1}; \mu_{k-1}) + \frac{\mu_k + \mu_{k-1}}{2(\mu_{k-1})^2} \|\Lambda_k - \Lambda_{k-1}\|_F^2 \end{aligned} \tag{1}$$

*zhiw@swu.edu.cn.

Then, by using (6) in the original paper, we have

$$\begin{aligned}
& \mathcal{L}(L_{k+1}, S_{k+1}, \Lambda_k; \mu_k) \\
& \leq \mathcal{L}(L_{k+1}, S_k, \Lambda_k; \mu_k) \\
& \leq \mathcal{L}(L_k, S_k, \Lambda_k; \mu_k) \\
& = \mathcal{L}(L_k, S_k, \Lambda_{k-1}; \mu_{k-1}) + \frac{\mu_k + \mu_{k-1}}{2(\mu_{k-1})^2} \|\Lambda_k - \Lambda_{k-1}\|_F^2.
\end{aligned} \tag{2}$$

Iterating the inequality chain (2) for k times, we obtain

$$\begin{aligned}
& \mathcal{L}(L_{k+1}, S_{k+1}, \Lambda_k; \mu_k) \\
& \leq \mathcal{L}(L_1, S_1, \Lambda_0; \mu_0) + \sum_{i=1}^k \frac{\mu_i + \mu_{i-1}}{2(\mu_{i-1})^2} \|\Lambda_i - \Lambda_{i-1}\|_F^2
\end{aligned} \tag{3}$$

Because $\|\Lambda_i - \Lambda_{i-1}\|_F^2$ is bounded, all terms on the right side of the inequality are bounded. So, $\mathcal{L}(L_{k+1}, S_{k+1}, \Lambda_k; \mu_k)$ is the upper bound. By adding a $\frac{1}{2}\mu_k \|\Lambda_k\|_F^2$ term to (3) in the original paper, we have

$$\begin{aligned}
& \mathcal{L}(L_{k+1}, S_{k+1}, \Lambda_k; \mu_k) + \frac{1}{2\mu_k} \|\Lambda_k\|_F^2 \\
& = \mathcal{P}_a(L_{k+1}) + \lambda \|S_{k+1}\|_1 \\
& \quad + \frac{\mu_k}{2} \left\| Y - L_{k+1} - S_{k+1} + \frac{\Lambda_k}{\mu_k} \right\|_F^2
\end{aligned} \tag{4}$$

According to lemma 1, the left side is bounded. Since every term on right side of above inequality are bounded, $\{L_k\}$ and $\{S_k\}$ are bounded. \square

Theorem 2. Suppose a function $F(X)$ can be represent as $F(X) = f \circ \sigma(X)$, where $X \in \mathbb{R}^{m \times n}$, and f is differentiable. The gradient of $F(X)$ at X is

$$\frac{\partial F(X)}{\partial X} = U \text{diag}(\theta) V^\top \tag{5}$$

where $\theta = \left. \frac{\partial f(y)}{\partial y} \right|_{y=\sigma(X)}$.

Proof of Theorem 1: Let $\{L_k, S_k, \Lambda_k\}$ is the sequence generated by Algorithm 1, it is proved to be bounded by lemma 1 and lemma 2. According to Bolzano-Weierstrass theorem, the sequence has at least one accumulation point, and define as $\{L^*, S^*, \Lambda^*\}$. We assume that $\{L_k, S_k, \Lambda_k\}$ converges to $\{L^*, S^*, \Lambda^*\}$.

Since $\{\Lambda\}$ is bounded, we have

$$D - L_{k+1} - S_{k+1} = \lim_{k \rightarrow \infty} (\Lambda_{k+1} - \Lambda_{k+1}/\mu_k) = 0$$

Then, $Y = L^* + S^*$ holds. According to first-order optimality condition, for L^{k+1}

$$\begin{aligned}
& \partial_L \mathcal{L}(L, S_k, \Lambda_k; \mu_k)|_{L_{k+1}} \\
& = \partial_L (\mathcal{P}_a(L))|_{L_{k+1}} - \Lambda_k - \mu_k (D - L_{k+1} - S_k) \\
& = \partial_L (\mathcal{P}_a(L))|_{L_{k+1}} - \Lambda_{k+1} + \mu_k (S_k - S_{k+1}) \\
& = 0
\end{aligned}$$

If the SVD of L is $U[\text{diag}(\sigma_i)]V^\top$, according to Theorem 2 above, we have

$$\partial_L \mathcal{P}_a(L)|_{L^{k+1}} = U \text{diag}(\theta) V^\top \tag{6}$$

where

$$\theta_i = \begin{cases} \frac{a^2}{(a\sigma_i + 1)^2}, & \sigma_i \neq 0 \\ a^2, & \text{otherwise,} \end{cases}$$

Here $\theta_i \in (0, a^2]$ is obviously infite, $\partial_L \mathcal{P}_a$ is bounded. According to lemma 4, $\{\Lambda\}$ is bounded, $\lim_{k \rightarrow \infty} \mu_k (S^{k+1} - S^k)$. Under the assumption that $\lim_{k \rightarrow \infty} \mu_k (S^{k+1} - S^k) \rightarrow 0$, we have

$$\partial_L \mathcal{P}_a(L^*) + \Lambda^* = 0$$

Consequently, $\{L^*, S^*, \Lambda^*\}$ satisfies Karush—Kuhn—Tucker (KKT) conditions of $\mathcal{L}(L, S, \Lambda; \mu)$. Thus, $\{L^*, S^*\}$ is a stationary point of original problem (3), which completes the proof.

1.2. Enlarged Experimental Results

The size of the images of the experimental results in our paper may be too small due to the length limit of the article. We have enlarged them to better show the experimental results.

Fig. 1: Denoising results of Washington DC Mall. The false color image is composed by bands (R: 69, G: 106, and B: 151).

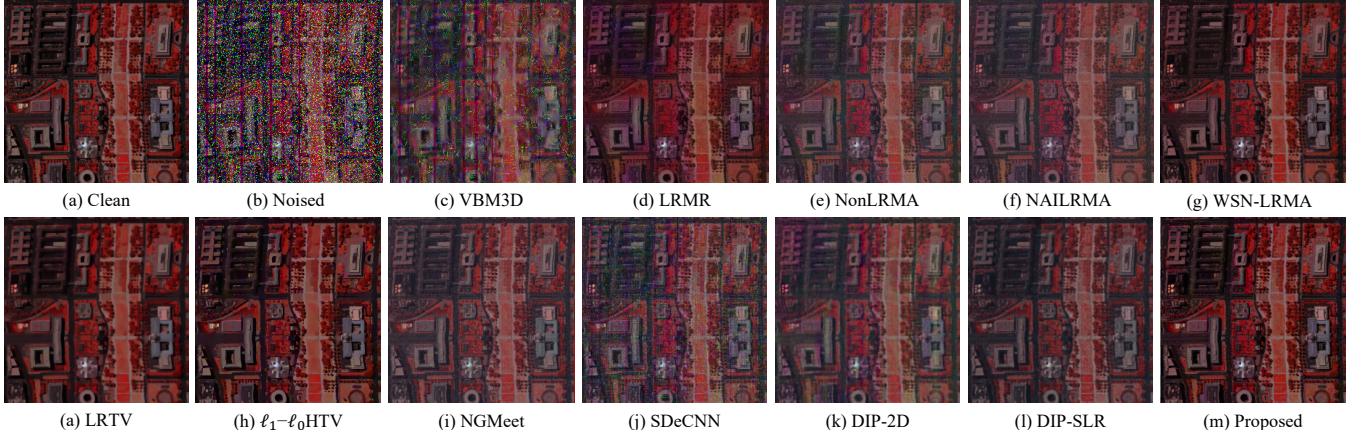


Fig. 2: Denoising results of Urban dataset. The false color image is composed by bands (R: 71, G: 107, and B: 207).

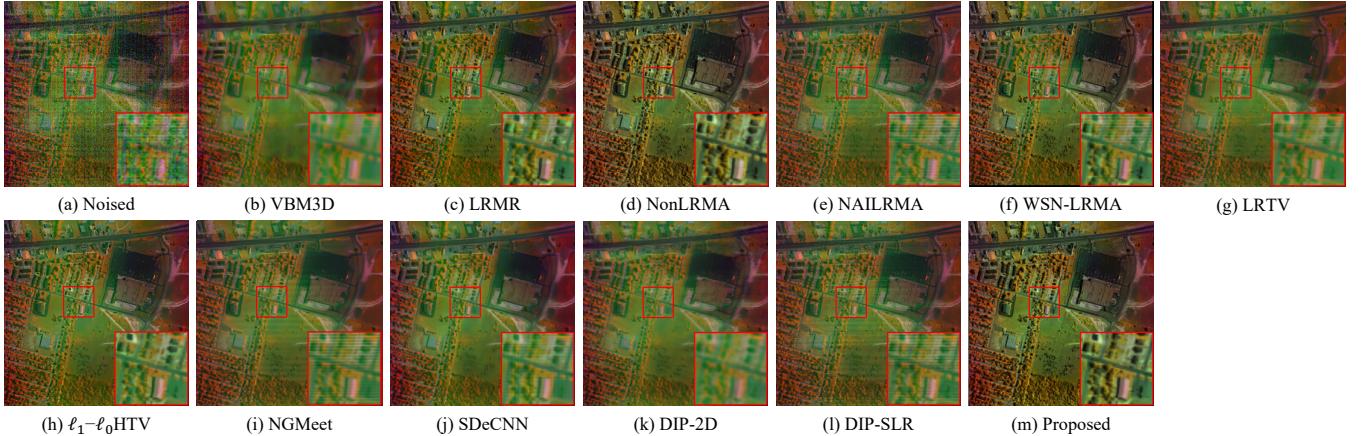


Fig. 3: Denoising results of Urban dataset. The Horizontal mean profiles on band 207.

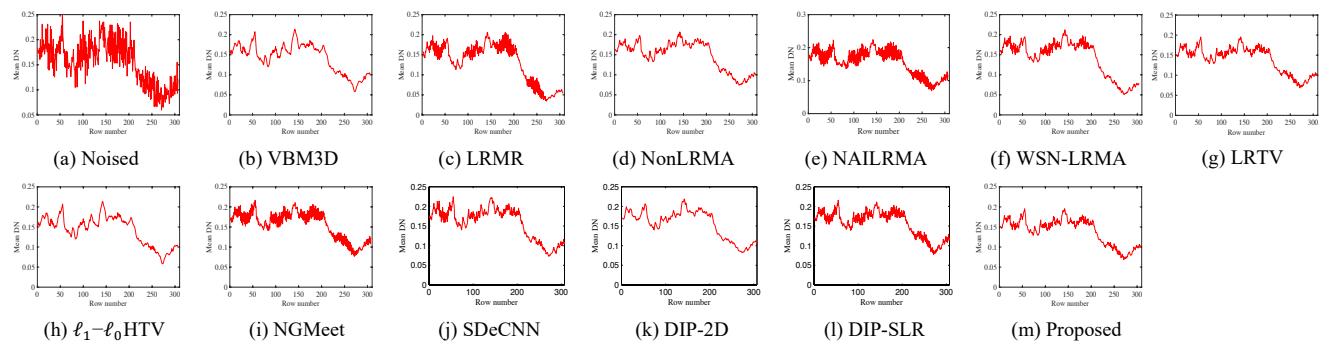
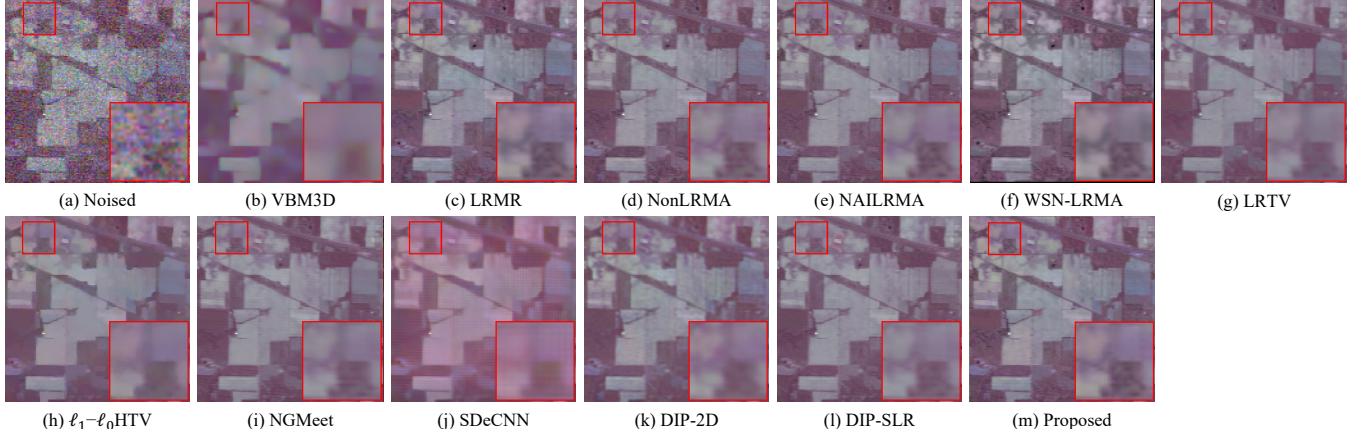


Fig. 4: Denoising results of Indian Pines. The false color image is composed by bands (R: 220, G: 149, and B: 109).



1.3. Additional Experimental Results

In this section, we put some additional qualitative comparison experimental results to demonstrate the superior performance of our proposed algorithm.

1.3.1. Synthesis Experiments

In order to measure the effect of different algorithms' recovery more intuitively, we compare the digital number values of vertical and horizontal mean profiles with the original clean image. As shown in Fig. 5 and 6, the red curve represents the mean digital number value of the original image, while the blue one stands for the value after denoising by different algorithms. From the results, we find that the curve shape of VBM3D has changed, which means that the image may be altered. NonLRMA, NAILRMA, and NGMeet are shifted in value, meaning that the impulse noise is not removed. LRMR does not fit the original curve well. WSN-LRMA, ℓ_1 - ℓ_0 HTV, and LRTV results are similar, with poor recovery at some peaks. Our algorithm fits the original image the best and has excellent performance in recovery.

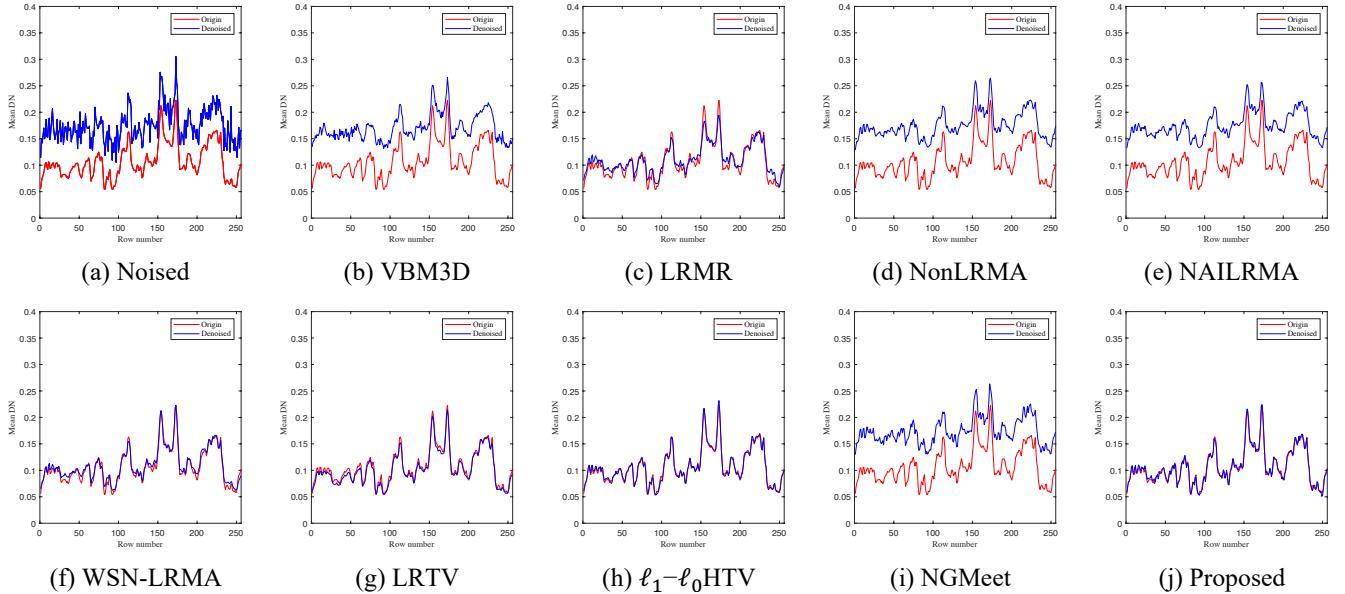


Fig. 5: Vertical mean profiles of band 160 in Washington DC dataset.

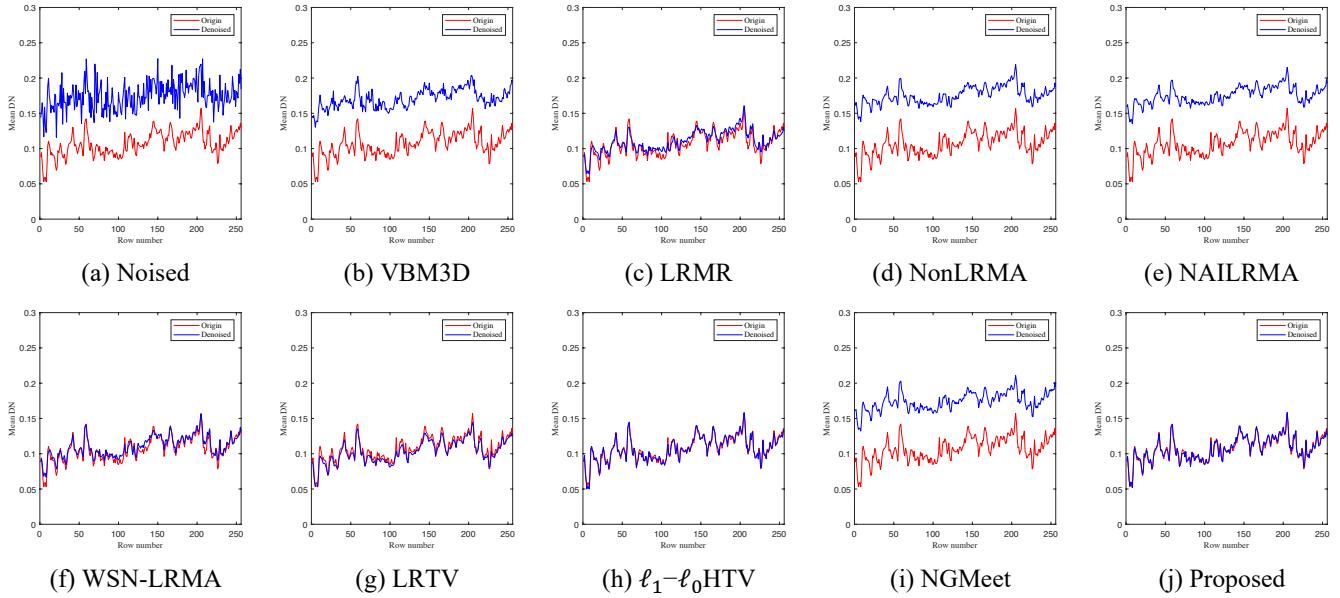


Fig. 6: Horizontal mean profiles of band 160 in Washington DC dataset.

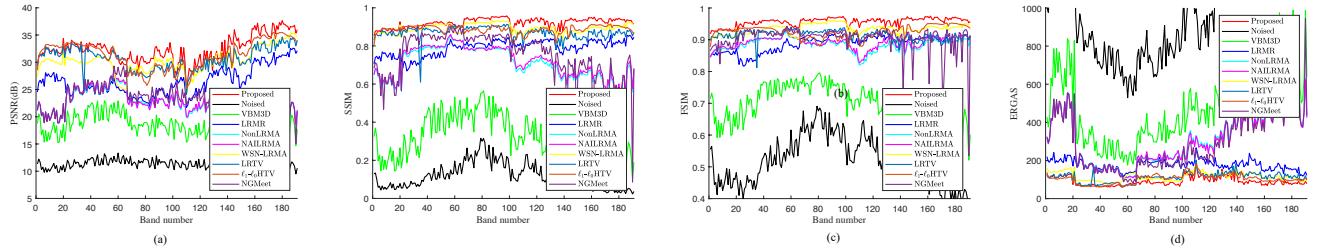


Fig. 7: PSNR, SSIM, FSIM, and ERGAS values of each band of the denoised results in Case 3. (a) PSNR values. (b) SSIM values. (c) FSIM values. (d) ERGAS values.

Fig. 7 shows the values of PSNR, SSIM, FSIM, and ERGAS for each band of the Washington DC dataset in Case 3 for all comparison algorithms. These indexes reflect the overall performance of the algorithms in the recovery process rather than just performing well in particular cases. As the images illustrate, our algorithm always gets the top performance in most cases in each criterion, which also proves its robustness.

1.3.2. Real Experiments

In the HYDICE Urban dataset, band 147 is severely corrupted by the mixing noises. Fig. 8 shows the comparison results of band 147 before and after the recovery. From the results, we see that VBM3D, NAILRMA, LRTV and NGMeet fail to recover the images. LRMR, WSN-LRMA, ℓ_1 - ℓ_0 HTV only recover the images partially with plenty noise remaining. NonLRMA has a good result, but chromatic aberration appears in the edge areas. Our algorithm has a better visual effect.

In the AVIRIS Indian Pines Dataset, band 150 is corrupted by severe Gaussian noise and impulse noise, which is tough to recover. Fig. 9 shows the denoising results of band 150. From the results, it can be seen that VBM3D does not succeed in recovering the image. LRMR and WSNM remove some of the noise, but there is still too much blur in some of the areas. LRTV, and ℓ_1 - ℓ_0 HTV over-smooth the image and details are almost missing. NonLRMA and NAILRMA are able to remove noises but some details are lacking. NGMeet adds non-existent ripples. Our algorithm achieves a more desirable result.

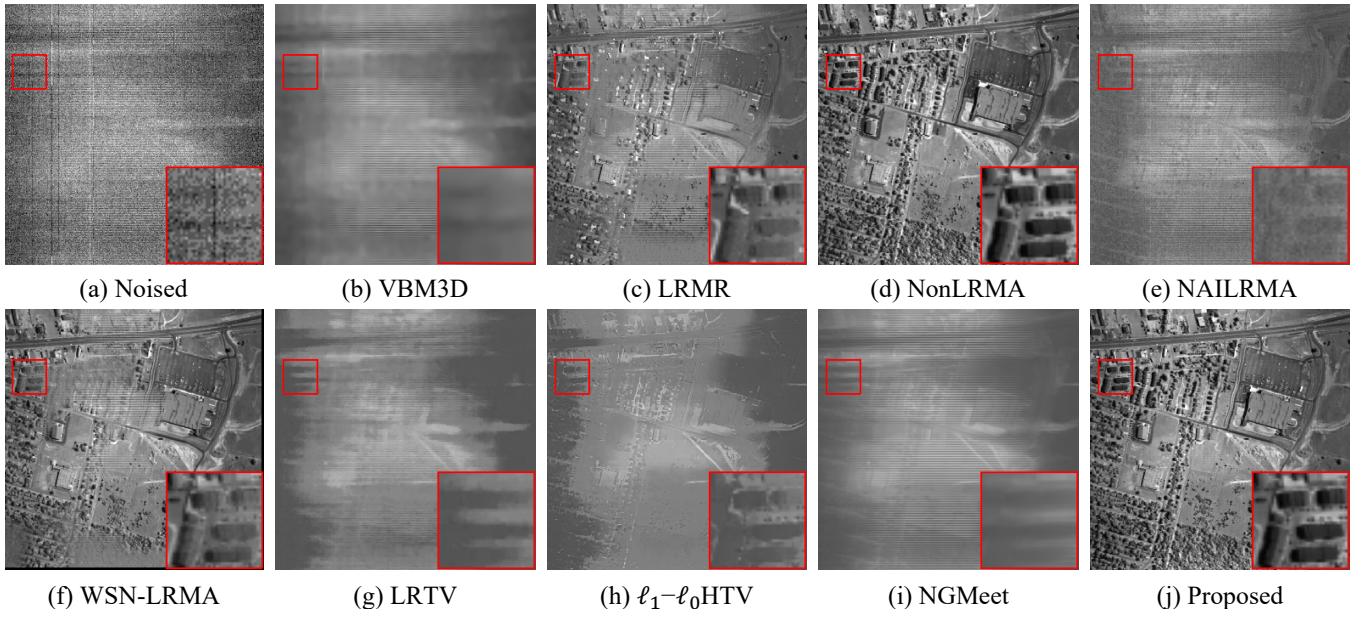


Fig. 8: Denoising results of band 147 in the HYDICE Urban dataset.

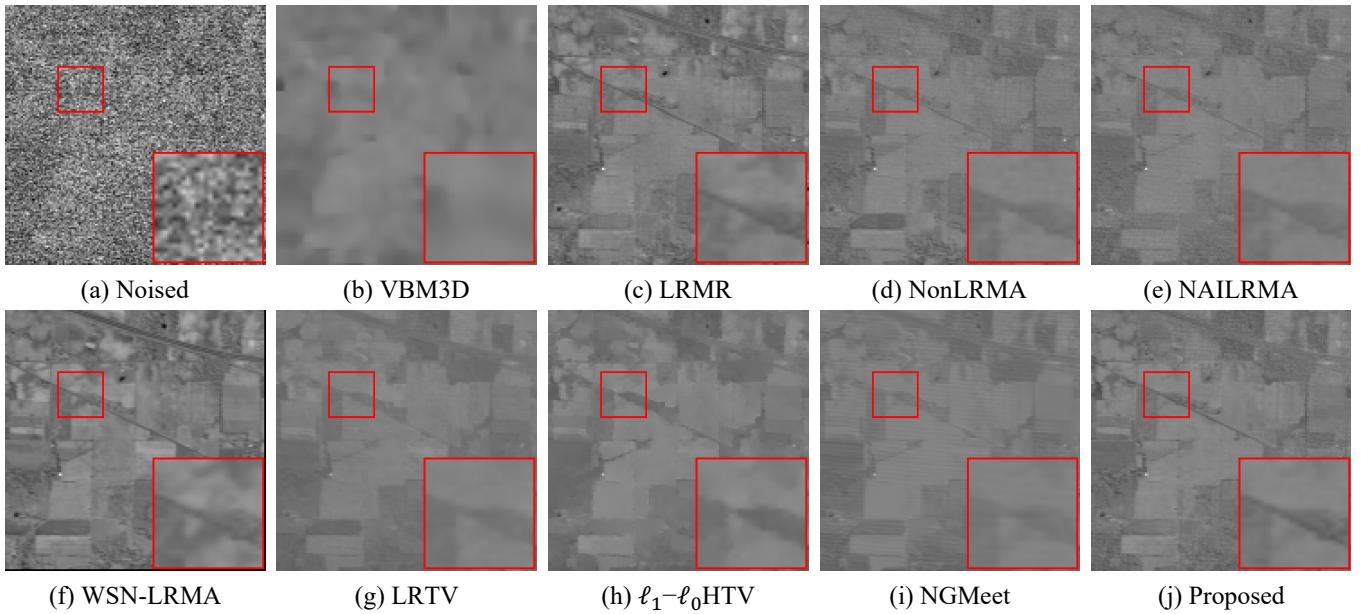


Fig. 9: Denoising results of band 150 in the AVIRIS Indian Pines dataset.