# "Multispectral Images Denoising by Intrinsic Tensor Sparsity Regularization": Supplementary Material

Qi Xie[1], Qian Zhao[1], Deyu Meng[1,*], Zongben Xu[1], Shuhang Gu[2], Wangmeng Zuo[3], Lei Zhang[2]
[1]Xi'an Jiaotong University; [2]The Hong Kong Polytechnic University; [3]Harbin Institute of Technology
xq.liwu@stu.xjtu.edu.cn timmy.zhaoqian@gmail.com {dymeng zbxu}@mail.xjtu.edu.cn
shuhanggu@gmail.com wmzuo@hit.edu.cn cslzhang@comp.polyu.edu.hk

## Abstract

*In this supplementary material, we provide the proofs to Theorems 1 and 2 presented in the maintext. We also present more clarifications on the parameter settings in our experiment. Furthermore, we show more experimental results to further substantiate the effectiveness of our method.*

## 1. Proofs to Theorems 1 and 2

We first give the following Lemma [4].

**Lemma 1.** (*von Neumann's trace inequality*) *For any $m \times n$ matrices $A$ and $B$, with $\sigma(A) = [\sigma_1(A), \sigma_2(A), ...\sigma_r(A)]^T$ and $\sigma(B) = [\sigma_1(B), \sigma_2(B), ...\sigma_r(B)]^T$, where $r = \min(m, n)$, being the singular values of $A$ and $B$, Then*

$$tr(A^T B) \leq \sigma(A)^T \sigma(B).$$

*The equality is achieved when there exists unitarians $U$ and $V$ that $A = U\Sigma_A V^T$ and $B = U\Sigma_B V^T$ are SVDs of $A$ and $B$, respectively.*

Based on the result of Lemma 1, we can deduce the following theorem.

**Theorem 1.** *$\forall\, A \in \mathbb{R}^{m\times n}$, the following problem:*

$$\max_{U^T U} \langle A, U\rangle, \tag{1}$$

*has the following closed-from solution $\hat{U} = BC^T$, where $A = BDC^T$ denotes the condensed SVD of $A$.*

*Proof.* Without loss of generality, we assume $m > r$. For any $U \in \mathbb{R}^{m\times n}$ satisfying $U^T U = I$, it is easy to see that all of its singular values are 1. Thus, the condensed SVD of $U$ can be written as

$$U = \hat{B} I_{r\times r} \hat{C}^T,$$

where $B \in \mathbb{R}^{m\times r}, C \in \mathbb{R}^{r\times r}$ and $I_{r\times r} \in \mathbb{R}^{r\times r}$ is the unitary matrix. By *von Neumann's trace inequality*, we can easily deduce that $\langle A, U\rangle = \text{trace}(A^T U)$ attains its upper bound when $\hat{B} = B$ and $\hat{C} = C$. Then we can obtain $\hat{U} = BC^T$.

When $m \leq r$ we can obtain the result in the similar way.■

Before proving Theorem 2, we first give two lemmas. It should be mentioned that, Gong et al. [2] has proved similar result as Lemma 2, but less rigorous.

**Lemma 2.** *Let $0 < \lambda$ and $0 < \varepsilon < \min\left(\sqrt{\lambda}, \frac{\lambda}{y}\right)$, the following problem:*

$$\min_x f(x) = \lambda \log\left(|x| + \varepsilon\right) + \frac{1}{2}(x - y)^2 \tag{2}$$

*has a local minimal*

$$\mathrm{D}_{\lambda,\varepsilon}(y) = \begin{cases} 0 & \text{if} \quad c_2 \leq 0 \\ \text{sign}(y)\left(\frac{c_1 + \sqrt{c_2}}{2}\right) & \text{if} \quad c_2 > 0 \end{cases}, \tag{3}$$

*where $c_1 = |y| - \varepsilon$ and $c_2 = (c_1)^2 - 4(\lambda - \varepsilon|y|)$.*

*Proof.* We first consider the situation that $y \geq 0$. It is easy to see that when $x < 0$, $f(x) \geq f(0)$.

When $x \geq 0$, $f(x)$ is differentiable, and by re-arranging $f'(x) = 0$ we can get:

$$\lambda \frac{1}{x + \varepsilon} + x - y = 0,$$

which is equivalent to

$$(x + \varepsilon) f'(x) = x^2 + (\varepsilon - y)x + \lambda - \varepsilon y = 0. \tag{4}$$

Thus we have:

1) if $c_2 < 0$, (4) has no real solution, and it's easy to see that $f'(x) > 0, x \in (0, +\infty)$. Since $f(x)$ is a continuous function, its global minimum is $x = 0$;

2) if $c_2 = 0$, (4) has the solutions $x = \frac{c_1}{2}$, and it's easy to see that $f'(x) \geq 0, x \in (0, +\infty)$. Since $f(x)$ is a continuous function, its global minimum is $x = 0$;

3) if $c_2 > 0$, (4) has two solutions, $x_1 = \frac{c_1 - \sqrt{c_2}}{2}$ and $x_2 = \frac{c_1 + \sqrt{c_2}}{2}$. Besides, since $0 < \varepsilon < \min\left(\sqrt{\lambda}, \frac{\lambda}{y}\right)$, we

*Corresponding author.

can also obtain that:

$$\begin{aligned} c_2 = (y-\varepsilon)^2 - 4(\lambda - \varepsilon y) &> 0 \\ y^2 - 2\varepsilon y + \varepsilon^2 &\geq 4C - 4\varepsilon y \\ (y+\varepsilon)^2 &> 4C \\ y - \varepsilon &> 2\sqrt{C} - 2\varepsilon \\ c_1 &> 0. \end{aligned} \tag{5}$$

Thus $x_2 > x_1 > c_1 > 0$, and it's then easy to obtain that $f'(x) > 0$ when $0 < x < x_1$ or $x > x_2$ and $f'(x) < 0$ when $x1 < x < x_2$. Therefore, $x = x_2$ is the only local minimum of $f(x)$ over $(0, +\infty)$.

For $y < 0$, we can prove in a similar way. Therefore, Theorem 1 holds.■

**Lemma 3.** *Given $Y \in \mathbb{R}^{m\times n}$, $m \geq n$, let $Y = U\text{diag}(\sigma_1, \sigma_2, ..., \sigma_n)V^T$ be the SVD of $Y$. Define $d_i$ as $i$-th singular value of $X$, the optimum to the following problem:*

$$\min_{X\in\mathbb{R}^{m\times n}} g\left(d_1, d_2, ..., d_n\right) + \frac{1}{2}\|X - Y\|_F^2 \tag{6}$$

*can be expressed as $\hat{X} = U\text{diag}(\hat{d}_1, \hat{d}_2, ..., \hat{d}_n)V^T$, where $(\hat{d}_1, \hat{d}_2, ..., \hat{d}_n)$ is the optimum to the following optimization problem:*

$$\begin{aligned} \min_{d_1,d_2,...,d_n} \sum_{i=1}^{n} (\sigma_i - d_i)^2 + g(d_1, d_2, ..., d_n) \\ \textit{s.t. } d_1 \geq d_2 \geq .... \geq d_n \geq 0. \end{aligned} \tag{7}$$

*Proof.* For any $X \in \mathbb{R}^{m\times n}$, denote $\bar{U}D\bar{V}^T$ as the SVD of $X$, where $D = \text{diag}(d_1, d_2, ..., d_n)$ is the singular value matrix, with $d_1 \geq d_2 \geq .... \geq d_n \geq 0$. Based on the property of *Frobenius* norm, we have

$$\begin{aligned} & g\left(d_1, d_2, ...d_n\right) + \frac{1}{2}\|X - Y\|_F^2 \\ =& \frac{1}{2}\|Y\|_F^2 - Tr\left(Y^T X\right) + \frac{1}{2}\|X\|_F^2 + g\left(d_1, d_2, ...d_n\right) \\ =& - Tr\left(Y^T X\right) + \frac{1}{2}\sum_{i=1}^{n}(\sigma_i^2 + d_i^2) + g\left(d_1, d_2, ...d_n\right). \end{aligned}$$

Therefore, (6) is equivalent to

$$\begin{aligned} & \min_{\bar{U},D,\bar{V}} -Tr\left(Y^T X\right) + \frac{1}{2}\sum_{i=1}^{n}(\sigma_i^2 + d_i^2) + g\left(d_1, d_2, ...d_n\right) \Leftrightarrow \\ & \min_{D}\left\{-\max_{\bar{U},\bar{V}} Tr\left(Y^T X\right) + \frac{1}{2}\sum_{i=1}^{n}(\sigma_i^2 + d_i^2) + g\left(d_1, d_2, ...d_n\right)\right\} \end{aligned}$$

Based on *von Neumann's trace* inequality, we know that $Tr\left(Y^T X\right)$ achieves its upper bound $\sum_{i=1}^{n}(\sigma_i d_i)$ if $\bar{U} = U$ and $\bar{V} = V$. Thus (6) is equivalent to

$$\begin{aligned} & \min_{D} \sum_{i=1}^{n}(\sigma_i^2 + 2\sigma_i d_i + d_i^2) + g\left(d_1, d_2, ...d_n\right) \\ \Leftrightarrow & \min_{d_1,d_2,...,d_n} \sum_{i=1}^{n}(\sigma_i - d_i)^2 + g(d_1, d_2, ..., d_n) \\ & \text{s.t. } d_1 \geq d_2 \geq .... \geq d_n \geq 0. \end{aligned}$$

From the above derivation, we can see that the optimal solution to (6) is

$$\hat{X} = U\text{diag}(\hat{d}_1, \hat{d}_2, ..., \hat{d}_n)V^T$$

where $(\hat{d}_1, \hat{d}_2, ..., \hat{d}_n)$ is the solution to (7).

The proof is then completed. ■

Based on the above two lemmas, we can now present the following theorem.

**Theorem 2.** *Given $Y \in \mathbb{R}^{m\times n}$, $m \geq n$, let $Y = U\text{diag}(\sigma_1, \sigma_2, ..., \sigma_n)V^T$ be the SVD of $Y$. Let $0 < \lambda$, $0 < \varepsilon < \min\left(\sqrt{\lambda}, \frac{\lambda}{\sigma_1}\right)$, and define $d_i$ as $i$-th singular value of $X$. The solution of the following problem:*

$$\min_{X\in\mathbb{R}^{m\times n}} \lambda\sum_{i=1}^{n}\log(d_i + \varepsilon) + \frac{1}{2}\|X - Y\|_F^2 \tag{8}$$

*can then be expressed as $\hat{X} = U\text{diag}(\hat{d}_1, \hat{d}_2, ..., \hat{d}_n)V^T$, where $\hat{d}_i = \text{D}_{\lambda,\varepsilon}(\sigma_i), i = 1, 2, ..., n$.*

*Proof.* According to Lemma 3, the optimum to (8) can expressed as $\hat{X} = U\textit{diag}(\hat{d}_1, \hat{d}_2, ..., \hat{d}_n)V^T$, where $(\hat{d}_1, \hat{d}_2, ..., \hat{d}_n)$ is the optimum to the following optimization problem:

$$\begin{aligned} \min_{d_1,d_2,...,d_n} \sum_{i=1}^{n}\left\{(\sigma_i - d_i)^2 + \lambda\log(d_i + \varepsilon)\right\} \\ \text{s.t. } d_1 \geq d_2 \geq .... \geq d_n \geq 0. \end{aligned} \tag{9}$$

We first consider the following unconstrained problem:

$$\min_{d_1,d_2,...,d_n} \sum_{i=1}^{n}\left\{(\sigma_i - d_i)^2 + \lambda\log(d_i + \varepsilon)\right\}. \tag{10}$$

According to Lemma 2, a local optimum to (10) is $\bar{d}_i = \text{D}_{\lambda,\varepsilon}(\sigma_i)$, and it's easy to see that $\bar{d}_1 \geq \bar{d}_2 \geq .... \geq \bar{d}_n \geq 0$ holds. $\hat{d}_i = \bar{d}_i, i = 1, 2, ..., n$ then corresponds to a local optimum to (9).

The proof is completed. ■

## 2. Parameter setting details

In our experiments we set $\beta = cv^{-1}$, and in this section, we will provide clarifications on this parameter setting strategy. Let's first consider the following matrix-based proximal problem:

$$\hat{X} = \arg\min_{X\in\mathbb{R}^{m\times n}} \frac{1}{\beta}\log\left(|d_i| + \varepsilon\right) + \frac{1}{2}\|X - Y\|_F^2, \tag{11}$$

where $d_i$ is the $i$-th singular value of $X$. Based Theorem 2, we can define $\hat{X} = U\hat{D}V^T$ as the SVD of $\hat{X}$, and $Y = U\Sigma V^T$ as the SVD of $Y$, where $\hat{D} = \text{diag}(\hat{d}_1, \hat{d}_2, ..., \hat{d}_n)$ and $\Sigma = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_n)$. Approximately let $\varepsilon = 0$, and then we have

$$\begin{aligned}
&\langle \hat{X}, Y - \hat{X} \rangle \\
=&Tr(\hat{X}^T(Y - \hat{X})) \\
=&Tr\left(\left(U\hat{D}V^T\right)^T U\left(\Sigma - \hat{D}\right)V^T\right) \\
=&\sum_{i=1}^{r} \hat{d}_i(\sigma_i - \hat{d}_i) \\
=&\sum_{i=1}^{r} \left(\frac{\sigma_i + \sqrt{\sigma^2 - 4\beta^{-1}}}{2}\right)\left(\frac{\sigma_i - \sqrt{\sigma^2 - 4\beta^{-1}}}{2}\right) \\
=&\sum_{i=1}^{r} \frac{\sigma_i^2 - (\sigma_i^2 - 4\beta^{-1})}{4} \\
=&r\beta^{-1}.
\end{aligned}$$

Thus, we can obtain

$$\beta = \frac{r}{\langle \hat{X}, Y - \hat{X} \rangle}. \tag{12}$$

Since $\langle \hat{X}, Y - \hat{X} \rangle$ is with the same order of magnitude as

$$E(\hat{X})E(Y - \hat{X}) = E(Y)E(Y - \hat{X}) \approx E(Y)E(Y - X),$$

where $X$ define the groundtruth, and $E(Y - X)$ is with the same order of magnitude as $v = \sqrt{E\left((Y - X)^2\right)}$, then we have

$$\beta \approx \hat{c}\frac{r}{mnE(Y)}v^{-1}. \tag{13}$$

This derivation can be easily generalized to tensor case in a similar fashion, and thus we can use $\beta = cv^{-1}$ to set $\beta$.

# 3. More demonstrations on HSI denoising

In this section, we will depict more MSI denoising effects of the competing methods to further demonstrate the superiority of the proposed method.

## 3.1. More simulated MSI denoising results

In this section, we present more simulated MSI denoising results on Columbia Datasets. Table 1 lists the average performance over different scenes with several noise settings of all methods. From these quantitative comparison, the advantage of the proposed method can be evidently observed. Among almost all noise settings, our method can significantly outperform other competing methods with respective to all evaluation measures, e.g., the PSNR obtained by our method are more than 1.5 larger and ERGAS is around 10 smaller than the second best methods.

To further depict the denoising performance of our method, we show in Fig. 1-3 two bands in several MSIs that centered at 400nm (the darker one) and 700nm (the brighter one), respectively. From the figure, it is easy to observe that the proposed method evidently performs better than other competing ones, both in the recovery of finer-grained textures and coarser-grained structures. Especially, when the band energy is low, most competing methods begin to fail, while our method still performs consistently well in such harder cases.

## 3.2. More real MSI denoising demonstration

**The natural scene data set:** We have used a natural scene data set [1][1] with some MSIs from real-world scenes in our real data experiment. This dataset comprises 15 rural scenes (containing rocks, trees, leaves, grass, earth, etc.) and 15 urban scenes (containing walls, roofs, windows, plants, indoor, etc.). All of them are illuminated by the direct sunlight between mid-morning to mid-afternoon. As the images are taken from a fairly far distance and the energy is spread over all bands, these MSIs contain degree of Gaussian noises and amounts of sharp noises. For such noise, in our experiments, we first apply a variance-stabilizing (VST) [3] to the noised MSI, and then perform all competing methods, and finally apply an inverse VST to the denoised data to obtain the final reconstruction.

The experimental results show that our method can consistently ameliorate the image quality contained in these MSIs. For easy observation we illustrates an example image located at a band of a rural MSI in Fig. 4. It can be easily observed that the image restored by our method properly removes the noise while finely preserves the structure underlying the MSI, while the results obtained by most of other competing methods either contain evident blurry or remain more unexpected sharp noises.

# References


[1] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster. The frequency of metamerism in natural scenes. *Optics and Photonics News*, 18(1):47–47, 2007.

[2] P. Gong, C. Zhang, Z. Lu, J. Z. Huang, and J. Ye. A general iterative shrinkage and thresholding algorithm for non-convex regularized optimization problems. In *ICML*, 2013.

[3] M. Markku and F. Alessandro. Optimal inversion of the generalized anscombe transformation for poisson-gaussian noise. *IEEE Trans. Image Processing*, 22(1):91–103, 2013.

[4] L. Mirsky. A trace inequality of john von neumann. *Monatshefte für Mathematik*, 79(4):303–306, 1975.


[1] `http://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectralimagesofnaturalscenes02`

Table 1. Average performance of 9 competing methods with respect to 4 PQIs. For both settings, the results are obtained by averaging through the 32 scenes and the varied parameters.

| $v$ | Method | PSNR | SSIM | FSIM | ERGAS |
|---|---|---|---|---|---|
| 0.10 | *Nosiy image* | 20.00±0.00 | 0.14±0.06 | 0.65±0.10 | 575.76±164.34 |
| | BwK-SVD | 30.24±1.23 | 0.60±0.04 | 0.89±0.03 | 173.91±39.36 |
| | BwBM3D | 37.21±2.82 | 0.92±0.03 | 0.95±0.01 | 79.13±23.97 |
| | 3DK-SVD | 32.66±1.30 | 0.76±0.02 | 0.93±0.02 | 134.14±26.41 |
| | LRTA | 36.93±2.83 | 0.89±0.05 | 0.95±0.01 | 82.23±26.50 |
| | PARAFAC | 33.53±3.96 | 0.84±0.10 | 0.92±0.04 | 126.07±62.80 |
| | ANLM3D | 36.86±2.87 | 0.93±0.04 | 0.96±0.01 | 85.23±21.36 |
| | TDL | 39.47±2.35 | 0.95±0.02 | 0.97±0.01 | 60.41±14.48 |
| | BM4D | 39.76±2.23 | 0.94±0.02 | 0.97±0.01 | 58.51±14.80 |
| | ITSReg | **41.24 ± 2.58** | **0.95 ± 0.04** | **0.98 ± 0.01** | **50.59 ± 30.16** |
| 0.15 | *Nosiy image* | 16.48±0.00 | 0.07±0.04 | 0.53±0.11 | 863.64±246.49 |
| | BwK-SVD | 28.65±1.33 | 0.52±0.05 | 0.84±0.03 | 208.74±46.49 |
| | BwBM3D | 34.97±2.74 | 0.88±0.05 | 0.93±0.01 | 101.92±29.81 |
| | 3DK-SVD | 31.04±1.59 | 0.71±0.03 | 0.91±0.02 | 160.81±30.60 |
| | LRTA | 34.88±2.78 | 0.85±0.07 | 0.93±0.02 | 103.93±32.98 |
| | PARAFAC | 32.80±3.54 | 0.80±0.09 | 0.91±0.04 | 135.47±61.75 |
| | ANLM3D | 35.11±2.85 | 0.90±0.05 | 0.95±0.02 | 103.26±25.97 |
| | TDL | 37.05±2.15 | 0.91±0.02 | 0.95±0.02 | 79.58±18.97 |
| | BM4D | 37.45±2.13 | 0.90±0.03 | 0.95±0.01 | 76.41±16.58 |
| | ITSReg | **39.33 ± 2.48** | **0.94 ± 0.04** | **0.97 ± 0.01** | **62.35 ± 17.30** |
| 0.20 | *Nosiy image* | 13.98±0.00 | 0.05±0.03 | 0.45±0.12 | 1151.56±328.82 |
| | BwK-SVD | 27.52±1.35 | 0.45±0.06 | 0.80±0.03 | 237.65±53.13 |
| | BwBM3D | 33.71±2.88 | 0.86±0.06 | 0.92±0.02 | 118.32±36.64 |
| | 3DK-SVD | 29.97±1.76 | 0.68±0.04 | 0.89±0.02 | 181.08±34.53 |
| | LRTA | 33.43±2.78 | 0.82±0.08 | 0.92±0.02 | 122.75±38.36 |
| | PARAFAC | 31.34±2.80 | 0.73±0.07 | 0.89±0.03 | 157.44±59.16 |
| | ANLM3D | 33.84±2.75 | 0.87±0.05 | 0.93±0.02 | 118.68±29.64 |
| | TDL | 35.35±1.99 | 0.87±0.03 | 0.93±0.02 | 96.76±23.55 |
| | BM4D | 35.80±2.05 | 0.86±0.03 | 0.94±0.01 | 92.43±19.74 |
| | ITSReg | **37.70 ± 2.42** | **0.91 ± 0.05** | **0.96 ± 0.01** | **75.13 ± 20.70** |
| 0.25 | *Nosiy image* | 12.04±0.00 | 0.03±0.02 | 0.39±0.12 | 1439.39±410.92 |
| | BwK-SVD | 26.62±1.32 | 0.40±0.06 | 0.77±0.04 | 263.79±59.54 |
| | BwBM3D | 32.54±2.79 | 0.83±0.06 | 0.90±0.02 | 134.85±39.85 |
| | 3DK-SVD | 29.22±1.91 | 0.65±0.04 | 0.88±0.02 | 196.95±38.38 |
| | LRTA | 32.29±2.76 | 0.79±0.09 | 0.90±0.02 | 139.87±43.47 |
| | PARAFAC | 30.16±2.33 | 0.66±0.06 | 0.87±0.03 | 179.15±59.78 |
| | ANLM3D | 32.82±2.61 | 0.83±0.05 | 0.91±0.02 | 132.50±32.25 |
| | TDL | 33.93±1.79 | 0.83±0.04 | 0.91±0.03 | 113.74±27.35 |
| | BM4D | 34.49±1.97 | 0.83±0.04 | 0.92±0.02 | 107.35±22.65 |
| | ITSReg | **36.26 ± 2.32** | **0.89 ± 0.06** | **0.95 ± 0.02** | **88.36 ± 23.95** |
| 0.30 | *Nosiy image* | 10.46±0.00 | 0.02±0.02 | 0.35±0.11 | 1727.34±493.14 |
| | BwK-SVD | 25.84±1.25 | 0.36±0.06 | 0.73±0.04 | 288.82±65.91 |
| | BwBM3D | 31.56±2.71 | 0.81±0.06 | 0.89±0.03 | 150.33±42.33 |
| | 3DK-SVD | 28.67±2.04 | 0.62±0.04 | 0.86±0.02 | 209.91±42.66 |
| | LRTA | 31.39±2.74 | 0.76±0.10 | 0.89±0.03 | 155.14±48.02 |
| | PARAFAC | 28.93±1.95 | 0.58±0.06 | 0.85±0.03 | 205.17±62.13 |
| | ANLM3D | 31.97±2.48 | 0.79±0.05 | 0.90±0.02 | 145.40±34.65 |
| | TDL | 32.74±1.61 | 0.78±0.05 | 0.89±0.03 | 130.58±32.20 |
| | BM4D | 33.41±1.91 | 0.79±0.05 | 0.90±0.02 | 121.33±25.29 |
| | ITSReg | **34.39 ± 1.87** | **0.83 ± 0.08** | **0.93 ± 0.02** | **110.31 ± 30.38** |

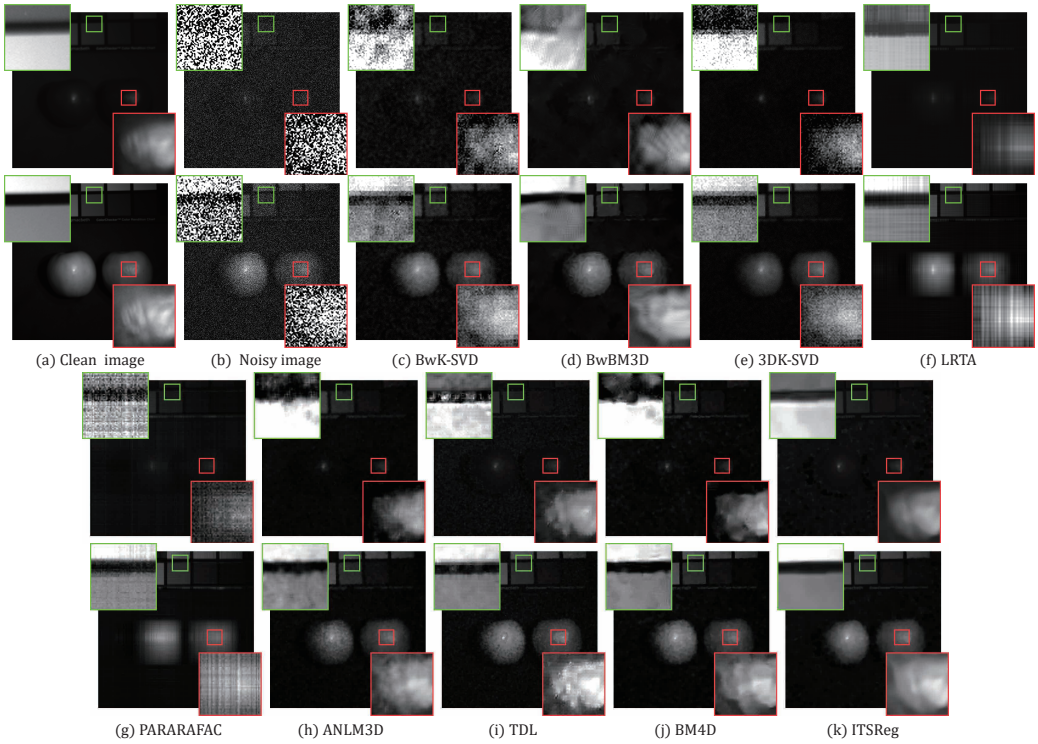


Figure 1. (a) The image at two bands (400nm and 700nm) of *real and fake apples ms*; (b) The noisy images corrupted by Gaussian noise with variance $v = 0.2$, (c)-(k) The restored image obtained by the 9 utilized MSI denoising methods. Two demarcated areas in each image are amplified at a 4 times larger scale and the same degree of contrast for easy observation of details.

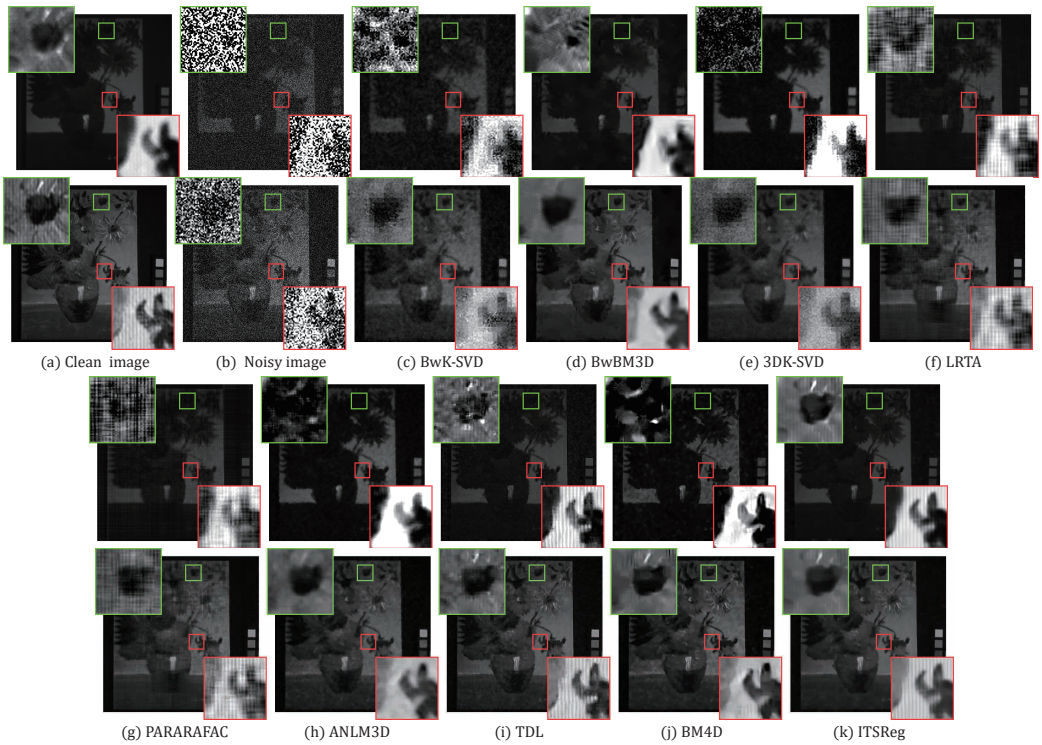


Figure 2. (a) The image at two bands (400nm and 700nm) of *oil painting ms*; (b) The noisy images corrupted by Gaussian noise with variance $v = 0.2$, (c)-(k) The restored image obtained by the 9 utilized MSI denoising methods.

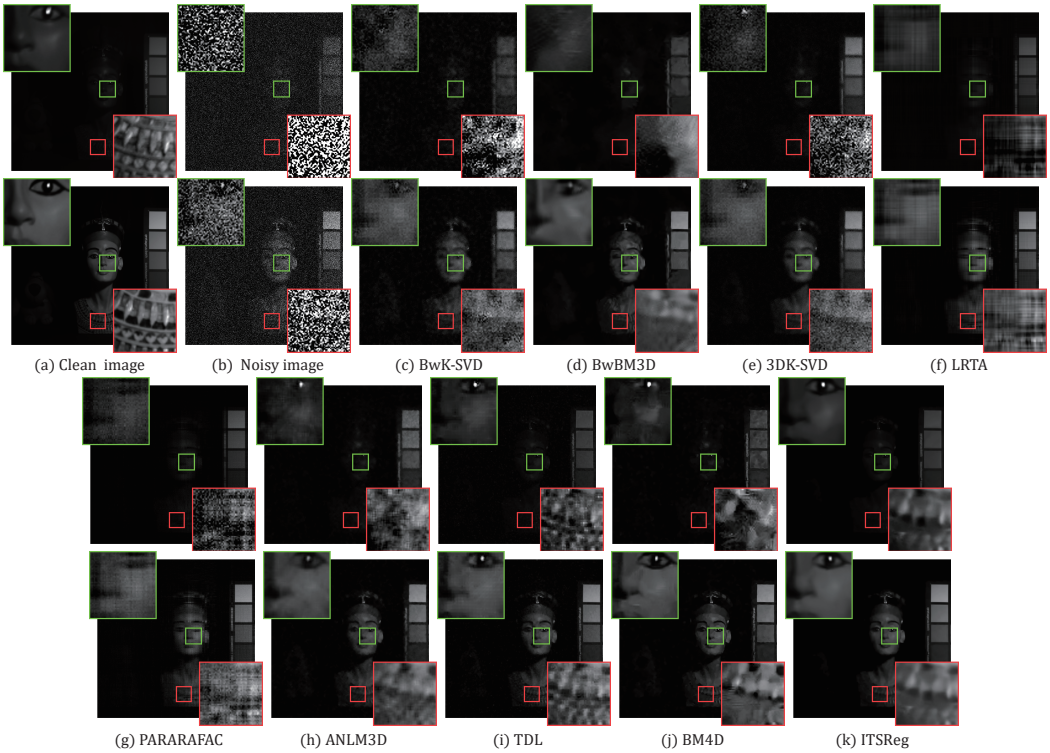


Figure 3. (a) The image at two bands (400nm and 700nm) of *Egyptian statue ms*; (b) The noisy images corrupted by Gaussian noise with variance $v = 0.2$, (c)-(k) The restored image obtained by the 9 utilized MSI denoising methods.

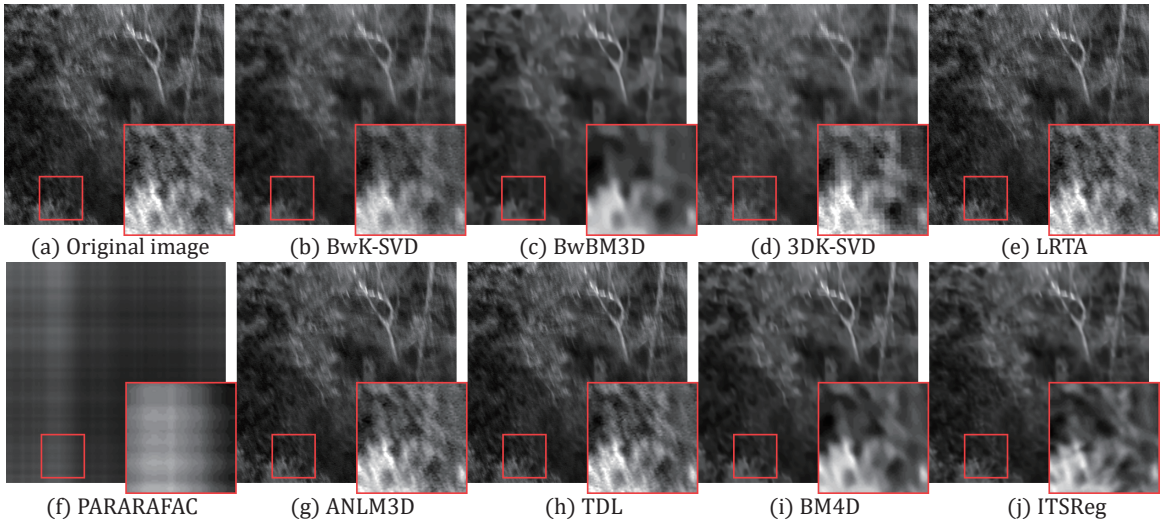


Figure 4. (a) The image located at the $6^{th}$ band in *scene 3* of the natural scene data set; (b)-(j) The restored image obtained by the 9 utilized MSI denoising methods. The demarcated areas in each image are amplified at a 2.5 times larger scale for easy observation of details.