Rotation Invariant Spatio-Spectral Total Variation for Hyperspectral Image Denoising

Shingo Takemoto* and Shunsuke Ono*
* Tokyo Institute of Technology, Japan
E-mail: takemoto.s.af@m.titech.ac.jp; ono@c.titech.ac.jp
Tel/Fax: +80-45-924-5089

Abstract—We propose a novel regularization function, named Rotation-Invariant Spatio-Spectral Total Variation (RISSTV), for hyperspectral (HS) image denoising. Spatio-Spectral Total Variation (SSTV) and its extended methods, defined using the second-order spatio-spectral differences, are known as popular regularization approaches that capture the HS image-specific spatio-spectral piecewise smoothness for effectively removing noise from HS images. However, these methods lack rotation-invariance and tend to corrupt round structures and oblique edges on HS images. To address this issue, we assign rotation-invariance to TV and establish RISSTV, which can accurately recover the detailed structures in HS images. Furthermore, we formulate the HS image denoising problem as a convex optimization problem that includes RISSTV and develop an efficient algorithm based on a preconditioned primal-dual splitting method to solve this problem efficiently. Finally, we demonstrate the effectiveness of RISSTV compared with existing HS image regularization models through mixed noise removal experiments.

I. INTRODUCTION

Hyperspectral (HS) images have rich spectral information in a wide range of wavelengths from the near-infrared to the ultraviolet. This information can visualize materials and phenomena that cannot be distinguished by RGB images, making HS imaging a key technology for applications in diverse fields, such as agriculture, mineralogy, astronomy, and biotechnology [1]–[4]. However, HS images are unavoidably contaminated by various types of noise in the measurement process, such as thermal noise, quantization noise, and shot noise. Such noise significantly degrades the performance of subsequent processing, including unmixing [5], [6], classification [7]–[9], and anomaly detection [10], [11]. Therefore, HS image denoising is an essential preprocessing task [12]–[14].

In the domain of HS image denoising, the Spatio-Spectral Total Variation (SSTV) model [15] stands out as a powerful regularization approach that captures underlying properties on HS images. SSTV is defined by the ℓ_1-norm of the second-order spatio-spectral differences, where the first-order spatial differences are calculated after the spectral ones are calculated, thus removing noise while preserving the HS image-specific spatio-spectral consistency. For this reason, SSTV has been widely used in many state-of-the-art HS image denoising methods [16]–[22]. One of the successful extensions of SSTV is Hybrid Spatio-Spectral Total Variation (HSSTV) [19] proposed by Takeyama *et al.*, which integrates the first-order spatial differences into SSTV. Directly promoting of spatial piecewise

smoothness removes similar noise (called artifacts) between adjacent bands that SSTV tends to retain. However, SSTV and its extended methods, including HSSTV, can corrupt the round structures and oblique edges in HS images due to their lack of rotation-invariance. This raises the question: *Could we design SSTV-type functions with the rotation-invariance so that noise can be removed while preserving the more detailed structure of HS images?*

In this paper, we propose a denoising method for HS images using a newly introduced *Rotation Invariant Spatio-Spectral Total Variation* model (RISSTV). RISSTV is built upon a theoretical framework proposed by Condat [23] that significantly improves the rotation-invariance of TV for grayscale images. The rotation-invariance is an important property for restoring round structures and oblique edges because the TV value is independent of direction. The main contributions of this article are listed below:

- We design a novel regularization formulation, namely RISSTV. RISSTV is designed to combine two types of TVs, consisting of the second-order spatio-spectral differences and the first-order spatial differences, thus effectively removing noise and suppressing artifacts. Furthermore, since the two types of TVs are defined to be rotation-invariant, RISSTV can preserve the detailed structure on HS images, including round structures and oblique edges, while removing noise.
- We formulate the HS image denoising problem as a constrained convex optimization problem that includes RISSTV. In this formulation, the data-fidelity term and the term characterizing sparse noise are imposed as convex constraints rather than added as part of the objective function, which makes parameter tuning easier.
- To solve our optimization problem for HS image denoising, we develop an efficient algorithm based on a preconditioned primal-dual splitting method (P-PDS) [24] with an operator norm-based stepsize selection method [25]. Unlike the standard PDS [26], [27], our proposed algorithm can automatically determine the appropriate stepsizes based on the problem structure.

Finally, we demonstrate the effectiveness of the proposed method by comparing it with state-of-the-art HS image regularization models through HS image denoising experiments.

II. PRELIMINARIES

A. Notations

Throughout this paper, we denote vectors and matrices by the boldface lowercase letters (e.g., $\mathbf{x}$) and boldface capital letters (e.g., $\mathbf{X}$), respectively. We treat an HS image, denoted by $\mathbf{u}$ with N_1 vertical pixels, N_2 horizontal pixels, and N_3 bands. We denotes the total number of cube data elements by $N = N_1 N_2 N_3$. For a matrix data $\mathbf{x} \in \mathbb{R}^{N_1 N_2}$, the value of the location (n_1, n_2) in the domain $\{1, \ldots, N_1\} \times \{1, \ldots, N_2\}$ is denoted by $[\mathbf{X}]_{n_1, n_2}$. The ℓ_1-norm and the ℓ_2-norm of a vector $\mathbf{x} \in \mathbb{R}^N$ are defined as $\|\mathbf{x}\|_1 := \sum_{n=1}^{N} |x_n|$ and $\|\mathbf{x}\|_2 := \sqrt{\sum_{n=1}^{N} x_n^2}$, respectively, where x_n represents the n-th entry of $\mathbf{x}$. For an HS image $\mathbf{u} \in \mathbb{R}^N$, let $\mathbf{D}_v \in \mathbb{R}^{N \times N}$, $\mathbf{D}_h \in \mathbb{R}^{N \times N}$, and $\mathbf{D}_s \in \mathbb{R}^{N \times N}$ be the forward difference operators in the horizontal, vertical, and spectral directions, respectively, and the boundary condition is the Neumann boundary. Here, spatial difference operator is denoted by $\mathbf{D} := \begin{pmatrix} \mathbf{D}_v^\top & \mathbf{D}_h^\top \end{pmatrix}^\top \in \mathbb{R}^{2N \times N}$. Other notations will be introduced as needed.

B. Hybrid Spatio-Spectral Total Variation (HSSTV) [19]

For an HS image $\mathbf{u} \in \mathbb{R}^N$, HSSTV [19] is defined as

$$\text{HSSTV}(\mathbf{u}) := \left\| \begin{matrix} \mathbf{D}\mathbf{D}_s \mathbf{u} \\ \omega \mathbf{D} \mathbf{u} \end{matrix} \right\|_{1,p}, \qquad (1)$$

where $\| \cdot \|_{1,p}$ is the mixed $\ell_{1,p}$ norm, and $\omega \geq 0$ is a hyperparameter, p is assumed to be 1 or 2, i.e., the ℓ_1 norm ($\| \cdot \|_{1,1} = \| \cdot \|_1$) or the mixed $\ell_{1,2}$ norm, respectively. In HSSTV, the second-order spatio-spectral differences and the first-order spatial differences capture the spatio-spectral piecewise smoothness and directly the spatial piecewise smoothness, respectively. In addition, the first-order spatial differences suppress noise-like artifacts that are produced by only promoting the spatio-spectral piecewise smoothness.

C. Condat's rotation-invariant TV [23]

A new TV with rotation-invariance proposed by Condat [23] for a grayscale image $\mathbf{x} \in \mathbb{R}^{N_1 N_2}$ is defined in the dual domain as follows: for all $n_1 = 1, \ldots, N_1$, $n_2 = 1, \ldots, N_2$,

$$\text{TV}_c(\mathbf{x}) := \max_{\mathbf{y} \in \mathbb{R}^{2N_1 N_2}} \langle \mathbf{D}_g \mathbf{x}, \mathbf{y} \rangle \text{ s.t. } \begin{cases} \|[\mathbf{L}_\updownarrow \mathbf{y}]_{n_1,n_2}\|_1 \leq 1, \\ \|[\mathbf{L}_\leftrightarrow \mathbf{y}]_{n_1,n_2}\|_1 \leq 1, \\ \|[\mathbf{L}_\bullet \mathbf{y}]_{n_1,n_2}\|_1 \leq 1, \end{cases}$$
$$(2)$$

where $\mathbf{D}_g \in \mathbb{R}^{2N_1 N_2 \times N_1 N_2}$ is the spatial difference operator for the grayscale image and $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$ is the dual image pair. For the definition of the three operators $\mathbf{L}_\updownarrow$, $\mathbf{L}_\leftrightarrow$, $\mathbf{L}_\bullet$, see Eqs. (9)-(14) in [23]. The dual images $[\mathbf{y}_1]_{n_1,n_2}$, $[\mathbf{y}_2]_{n_1,n_2}$, like $[\mathbf{D}_v \mathbf{x}]_{n_1,n_2}$, $[\mathbf{D}_h \mathbf{x}]_{n_1,n_2}$, can be located at the points $(n_1 + \frac{1}{2}, n_2)$, $(n_1, n_2 + \frac{1}{2})$, respectively. The three operators $\mathbf{L}_\updownarrow$, $\mathbf{L}_\leftrightarrow$, $\mathbf{L}_\bullet$ correct this half-pixel shift and interpolate bilinearly the dual image pair $\mathbf{y}$ onto the grids $(n_1 + \frac{1}{2}, n_2)$, $(n_1, n_2 + \frac{1}{2})$, (n_1, n_2), respectively. The dual image pair $\mathbf{y}$ satisfies the constraints on their respective grids, then Condat's TV has the rotation-invariant property. The dual formulation Eq. (2) can be rewritten into the equivalent primal formulation as

$$\text{TV}_c(\mathbf{x}) = \min_{\mathbf{w}_\updownarrow, \mathbf{w}_\leftrightarrow, \mathbf{w}_\bullet \in \mathbb{R}^{2N_1 N_2}} \|\mathbf{w}_\updownarrow\|_{1,2} + \|\mathbf{w}_\leftrightarrow\|_{1,2} + \|\mathbf{w}_\bullet\|_{1,2}$$
$$\text{s.t. } \mathbf{L}_\updownarrow^\top \mathbf{w}_\updownarrow + \mathbf{L}_\leftrightarrow^\top \mathbf{w}_\leftrightarrow + \mathbf{L}_\bullet^\top \mathbf{w}_\bullet = \mathbf{D}_g \mathbf{x}, \qquad (3)$$

where $\| \cdot \|_{1,2}$ is the mixed $\ell_{1,2}$ norm grouping the vertical and horizontal directions. The three vectors $\mathbf{w}_\updownarrow$, $\mathbf{w}_\leftrightarrow$, $\mathbf{w}_\bullet$ are viewed as gradients on the grids $(n_1 + \frac{1}{2}, n_2)$, $(n_1, n_2 + \frac{1}{2})$, (n_1, n_2), respectively. For a more compact form, let the linear operator $\mathbf{L}_g = \begin{pmatrix} \mathbf{L}_\updownarrow^\top & \mathbf{L}_\leftrightarrow^\top & \mathbf{L}_\bullet^\top \end{pmatrix}^\top \in \mathbb{R}^{6N_1 N_2 \times 2N_1 N_2}$, and the vector $\mathbf{w} = \begin{pmatrix} \mathbf{w}_\updownarrow & \mathbf{w}_\leftrightarrow & \mathbf{w}_\bullet \end{pmatrix}^\top \in \mathbb{R}^{6N_1 N_2 \times 2N_1 N_2}$. Then Eq. (3) can be written as

$$\text{TV}_c(\mathbf{x}) = \min_{\mathbf{w} \in \mathbb{R}^{6N_1 N_2}} \|\mathbf{w}\|_{1,2} \quad \text{s.t. } \mathbf{L}_g^\top \mathbf{w} = \mathbf{D}_g \mathbf{x}. \qquad (4)$$

III. PROPOSED METHOD

A. Rotation Invariant Spatio-Spectral Total Variation

Combining the second-order spatio-spectral differences with the first-order spatial differences and having the rotation-invariant property, our RISSTV is defined as follows:

$$\text{RISSTV}(\mathbf{u}) := \min_{\mathbf{w}_1, \mathbf{w}_2} \omega \|\mathbf{w}_1\|_{1,2} + \|\mathbf{w}_2\|_{1,2}$$
$$\text{s.t. } \begin{cases} \mathbf{L}^\top \mathbf{w}_1 = \mathbf{D}\mathbf{u}, \\ \mathbf{L}^\top \mathbf{w}_2 = \mathbf{D}\mathbf{D}_s \mathbf{u}, \end{cases} \qquad (5)$$

where $\omega \geq 0$, $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^{6N}$ are auxiliary variables, $\mathbf{L} \in 2N \times 6N$ is linear operator formed by arranging N_3 diagonals of $\mathbf{L}_g$ defined in Eq. (4). Mainly, TV of the second-order differences related to the second term and second constraint in Eq. (5) captures the HS image-specific spatio-spectral piecewise smoothness. Associated with the first term and the first constraint, TV of the first-order spatial differences directly promotes the spatial piecewise smoothness and plays an important role in suppressing the noise-like artifacts produced by only imposing TV of the second-order spatio-spectral differences. These two types of TVs are an extension of Condat's idea and have rotation-invariance, so RISSTV can recover round structures and oblique edges in HS images more accurately than existing TV-based methods.

The parameter ω controls the relative importance of TV of the first-order spatial differences to one of the second-order spatio-spectral differences. If ω is larger, i.e., we make the direct spatial smoothness stronger on a restored HS image, RISSTV would cause over-smoothing of the detailed structures. Therefore, ω should be set to less than one. In fact, a good choice of ω is empirically around 0.01 to 0.05 for various HS images.

B. HS Image Denoising by RISSTV

An observed HS image $\mathbf{v} \in \mathbb{R}^N$ contaminated by mixed noise is modeled by

$$\mathbf{v} = \bar{\mathbf{u}} + \bar{\mathbf{s}} + \mathbf{n}, \qquad (6)$$

where $\bar{\mathbf{u}}$ is a clean HS image, $\bar{\mathbf{s}}$ is a sparse noise that models outliers, and $\mathbf{n}$ represents random noise, respectively.

Based on the above observation model, we formulate the HS image denoising problem that handles RISSTV as a constrained convex optimization problem with the following form:

$$\min_{\mathbf{u},\mathbf{w}_1,\mathbf{w}_2,\mathbf{s}} \omega\|\mathbf{w}_1\|_{1,2} + \|\mathbf{w}_2\|_{1,2} \quad \text{s.t.} \begin{cases} \mathbf{u} \in R_{\underline{\mu},\bar{\mu}}, \\ \mathbf{s} \in B_{1,\eta}, \\ \mathbf{u} + \mathbf{s} \in B_{2,\varepsilon}^{\mathbf{v}}, \\ \mathbf{L}^\top\mathbf{w}_1 = \mathbf{D}\mathbf{u}, \\ \mathbf{L}^\top\mathbf{w}_2 = \mathbf{D}\mathbf{D}_s\mathbf{u}, \end{cases}$$
(7)

where

$$R_{\underline{\mu},\bar{\mu}} := \{\mathbf{x} \in \mathbb{R}^N | \underline{\mu} \le x_i \le \bar{\mu}\,(i = 1,\ldots,N)\}, \quad (8)$$

$$B_{1,\eta} := \{\mathbf{x} \in \mathbb{R}^N | \|\mathbf{x}\|_1 \le \eta\}, \quad (9)$$

$$B_{2,\varepsilon}^{\mathbf{v}} := \{\mathbf{x} \in \mathbb{R}^N | \|\mathbf{x} - \mathbf{v}\|_2 \le \varepsilon\}. \quad (10)$$

The first constraint is a box constraint with $\underline{\mu} < \bar{\mu}$ which represents the dynamic range of $\mathbf{u}$. The second constraint serves as data-fidelity with the $\mathbf{v}$-centered ℓ_2-ball of the radius $\varepsilon > 0$. The third constraint characterizes sparse noise with the zero-centered ℓ_1-ball of the radius $\eta > 0$. Because ε and η can be determined based solely on the intensity of each noise (independently of the other terms in the objective function), such a data fidelity constraint makes the adjustment of the hyperparameters much easier than when a data fidelity term is added to the objective function. These advantages are also addressed, for example, in [28]–[32].

C. Optimization

Using indicator functions[1] of $\{\mathbf{0}\}$, $B_{2,\varepsilon}^{\mathbf{v}}$, $B_{1,\eta}$, and $R_{\underline{\mu},\bar{\mu}}$, we rewrite Prob. (7) into an equivalent form:

$$\min_{\substack{\mathbf{u},\mathbf{w}_1,\mathbf{w}_2,\mathbf{s} \\ \mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3}} \omega\|\mathbf{w}_1\|_{1,2} + \|\mathbf{w}_2\|_{1,2} + \iota_{R_{\underline{\mu},\bar{\mu}}}(\mathbf{u}) + \iota_{B_{1,\eta}}(\mathbf{s})$$

$$+ \iota_{\{\mathbf{0}\}}(\mathbf{y}_1) + \iota_{\{\mathbf{0}\}}(\mathbf{y}_2) + \iota_{B_{2,\varepsilon}^{\mathbf{v}}}(\mathbf{y}_3)$$

$$\text{s.t.} \begin{cases} \mathbf{y}_1 = \mathbf{D}\mathbf{u} - \mathbf{L}^\top\mathbf{w}_1, \\ \mathbf{y}_2 = \mathbf{D}\mathbf{D}_s\mathbf{u} - \mathbf{L}^\top\mathbf{w}_2, \\ \mathbf{y}_3 = \mathbf{u} + \mathbf{s}. \end{cases}$$
(11)

Prob. (11) can be solved by P-PDS [24]. We show the detailed algorithm in Alg. 1. The proximity operators[2] of $\iota_{R_{\underline{\mu},\bar{\mu}}}$, $\iota_{\{\mathbf{0}\}}$,

---

Algorithm 1 P-PDS-based solver for (11)

Input: $\mathbf{u}^{(0)}, \mathbf{s}^{(0)}, \mathbf{w}_1^{(0)}, \mathbf{w}_2^{(0)}, \mathbf{y}_1^{(0)}, \mathbf{y}_2^{(0)}, \mathbf{y}_3^{(0)}$
Output: $\mathbf{u}^{(t)}$
1: **while** A stopping criterion is not satisfied **do**
2: $\quad \mathbf{u}^{(t+1)} \leftarrow$
$\quad\quad P_{R_{\underline{\mu},\bar{\mu}}}\left(\mathbf{u}^{(t)} - \gamma_{\mathbf{u}}\left(\mathbf{D}^\top\mathbf{y}_1^{(t)} + \mathbf{D}_s^\top\mathbf{D}^\top\mathbf{y}_2^{(t)} + \mathbf{y}_3^{(t)}\right)\right)$
3: $\quad \mathbf{s}^{(t+1)} \leftarrow P_{B_{1,\eta}}\left(\mathbf{s}^{(t)} - \gamma_{\mathbf{s}}\mathbf{y}_3^{(t)}\right)$
4: $\quad \mathbf{w}_1^{(t+1)} \leftarrow \text{prox}_{\gamma_{\mathbf{s}}\omega,\|\cdot\|_{1,2}}\left(\mathbf{w}_1^{(t)} + \gamma_{\mathbf{w}_1}\mathbf{L}\mathbf{y}_1^{(t)}\right)$
5: $\quad \mathbf{w}_2^{(t+1)} \leftarrow \text{prox}_{\gamma_{\mathbf{s}},\|\cdot\|_{1,2}}\left(\mathbf{w}_2^{(t)} + \gamma_{\mathbf{w}_2}\mathbf{L}\mathbf{y}_2^{(t)}\right)$
6: $\quad \mathbf{u}' \leftarrow 2\mathbf{u}^{(t+1)} - \mathbf{u}^{(t)};$
7: $\quad \mathbf{s}' \leftarrow 2\mathbf{s}^{(t+1)} - \mathbf{s}^{(t)};$
8: $\quad \mathbf{w}_1' \leftarrow 2\mathbf{w}_1^{(t+1)} - \mathbf{w}_1^{(t)};$
9: $\quad \mathbf{w}_2' \leftarrow 2\mathbf{w}_2^{(t+1)} - \mathbf{w}_2^{(t)};$
10: $\quad \mathbf{y}_1^{(t+1)} \leftarrow \mathbf{y}_1^{(t)} + \gamma_{\mathbf{y}_1}\left(\mathbf{D}\mathbf{u}' - \mathbf{L}^\top\mathbf{w}_1'\right)$
11: $\quad \mathbf{y}_2^{(t+1)} \leftarrow \mathbf{y}_2^{(t)} + \gamma_{\mathbf{y}_2}\left(\mathbf{D}\mathbf{D}_s\mathbf{u}' - \mathbf{L}^\top\mathbf{w}_2'\right)$
12: $\quad \mathbf{y}_3' \leftarrow \mathbf{y}_3^{(t)} + \gamma_{\mathbf{y}_3}\left(\mathbf{u}' + \mathbf{s}'\right);$
13: $\quad \mathbf{y}_3^{(t+1)} \leftarrow \mathbf{y}_3' - \gamma_{\mathbf{y}_3}P_{B_{2,\varepsilon}^{\mathbf{v}}}\left(\frac{1}{\gamma_{\mathbf{y}_3}}\mathbf{y}_3'\right);$
14: $\quad t \leftarrow t + 1;$
15: **end while**

$\iota_{B_{2,\varepsilon}^{\mathbf{v}}}$, and $\|\cdot\|_{1,2}$ are calculated by

$$[\text{prox}_{\gamma\iota_{R_{\underline{\mu},\bar{\mu}}}}(\mathbf{x})]_i = [P_{R_{\underline{\mu},\bar{\mu}}}(\mathbf{x})]_i = \begin{cases} \underline{\mu}, & \text{if } x_i < \underline{\mu}, \\ \bar{\mu}, & \text{if } x_i > \bar{\mu}, \\ x_i, & \text{otherwise}, \end{cases} \quad (12)$$

$$\text{prox}_{\gamma\iota_{\{\mathbf{0}\}}}(\mathbf{x}) = \mathbf{0}, \quad (13)$$

$$\text{prox}_{\gamma\iota_{B_{2,\varepsilon}^{\mathbf{v}}}}(\mathbf{x}) = P_{B_{2,\varepsilon}^{\mathbf{v}}}(\mathbf{x}) = \begin{cases} \mathbf{x}, & \text{if } \mathbf{x} \in B_{2,\varepsilon}^{\mathbf{v}}, \\ \mathbf{v} + \frac{\varepsilon(\mathbf{x}-\mathbf{v})}{\|\mathbf{x}-\mathbf{v}\|_2}, & \text{otherwise}, \end{cases}$$
(14)

$$[\text{prox}_{\gamma\|\cdot\|_{1,2}}(\mathbf{x})]_i = \max\left\{1 - \frac{\gamma}{\sqrt{x_{dN+\tilde{i}}^2 + x_{(d+1)N+\tilde{i}}^2}}, 0\right\}x_i, \quad (15)$$

where $\tilde{i} := ((i-1) \mod N) + 1$, $d := (i - \tilde{i})/N$. The sum of $x_{dN+\tilde{i}}^2$ and $x_{(d+1)N+\tilde{i}}^2$ is intended to be the sum of elements in vertical and horizontal directions. The proximity operators of $\iota_{B_{1,\eta}}(\mathbf{s})$ can be efficiently computed by a fast ℓ_1-ball projection algorithm [33].

We set the stepsize parameters as $\gamma_{\mathbf{u}} = \frac{1}{41}$, $\gamma_{\mathbf{s}} = 1$, $\gamma_{\mathbf{w}_1} = \gamma_{\mathbf{w}_2} = \frac{1}{4}$, $\gamma_{\mathbf{y}_1} = \gamma_{\mathbf{y}_2} = \gamma_{\mathbf{y}_3} = \frac{1}{4}$ according to P-PDS [25].

IV. Experiments

To demonstrate the effectiveness of RISSTV, we conducted HS image denoising experiments where we compared RISSTV with three types of state-of-the-art methods. The first are TV-based methods: SSTV [15], HSSTV [19], and l_0-l_1HTV [20]. The second are tensor-based methods: Decomposable Nonlocal

---

[1]The indicator function of a closed convex set C is defined by $\iota_C(\mathbf{x}) := 0$, if $\mathbf{x} \in C$; ∞, otherwise.

[2]The proximity operator of index $\gamma > 0$ of a proper lower semicontinuous convex function f is defined by $\text{prox}_{\gamma f}(\mathbf{x}) := \arg\min_{\mathbf{y}} f(\mathbf{y}) + \frac{1}{2\gamma}\|\mathbf{x}-\mathbf{y}\|_2^2$. The proximity operator of ι_C is the projection onto C, given by $\text{prox}_{\iota_C}(\mathbf{x}) = \arg\min_{\mathbf{y} \in C}\|\mathbf{y} - \mathbf{x}\|_2 := P_C(\mathbf{x})$.

TABLE I
MPSNRs AND MSSIMs OF ALL DENOISING RESULTS.

Image	$(\sigma, p_{\mathbf{s}})$		TDL [34]	ITSReg [35]	SSTV [18]	HSSTV1 [19]	HSSTV2 [19]	l_0-l_1HTV [20]	LRTDTV [17]	RISSTV
Jasper Ridge	(0.05, 0.05)	MPSNR	21.09	31.65	38.39	38.71	38.36	38.41	37.85	**38.82**
		MSSIM	0.7999	0.8577	0.9475	**0.9567**	0.9535	0.9463	0.9544	0.9540
	(0.1, 0.1)	MPSNR	19.85	27.86	34.03	34.22	33.94	34.29	34.63	**34.68**
		MSSIM	0.6917	0.7694	0.8923	0.9021	0.9005	0.8928	0.9086	**0.9149**
PaviaU	(0.05, 0.05)	MPSNR	22.03	31.38	38.98	39.20	39.07	39.08	34.30	**39.41**
		MSSIM	0.8159	0.8734	0.9585	**0.9638**	0.9616	0.9587	0.9006	0.9634
	(0.1, 0.1)	MPSNR	20.33	27.29	32.76	33.37	33.50	33.20	32.31	**34.04**
		MSSIM	0.7372	0.7969	0.8687	0.8847	0.8859	0.8807	0.8483	**0.9044**

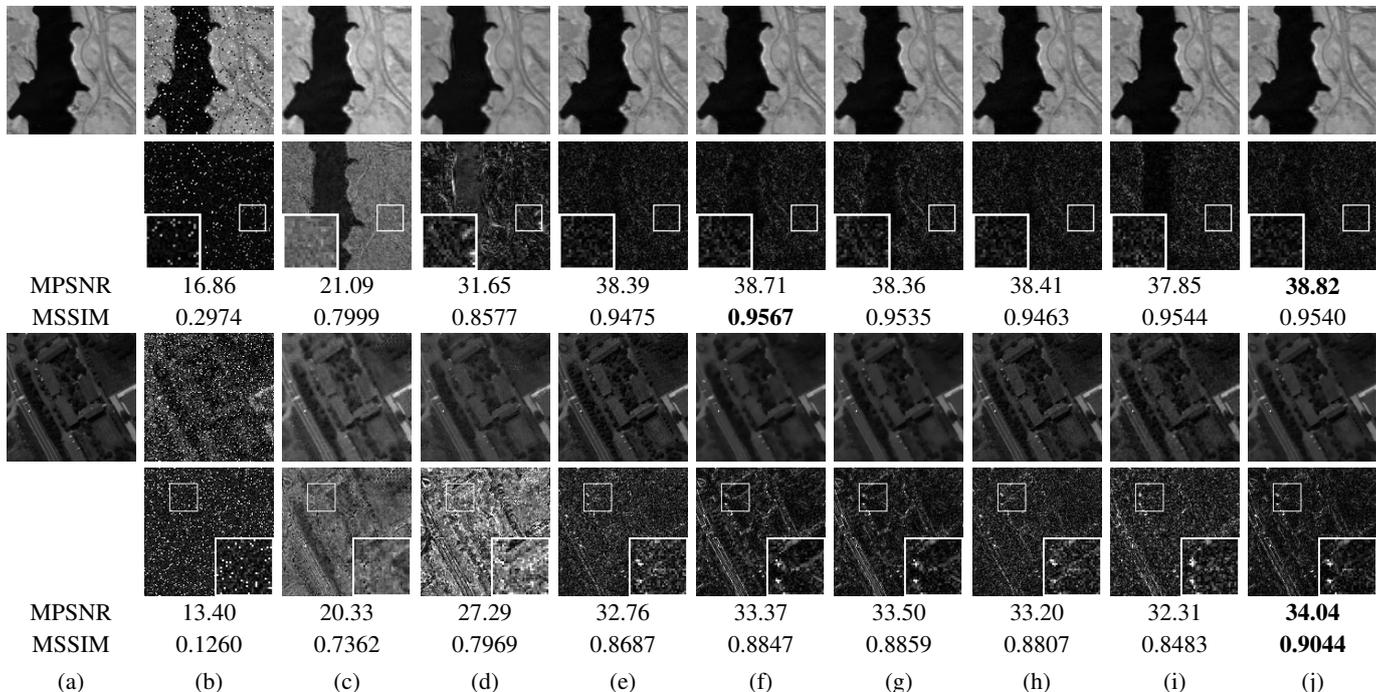|									
MPSNR	16.86	21.09	31.65	38.39	38.71	38.36	38.41	37.85	**38.82**
MSSIM	0.2974	0.7999	0.8577	0.9475	**0.9567**	0.9535	0.9463	0.9544	0.9540
MPSNR	13.40	20.33	27.29	32.76	33.37	33.50	33.20	32.31	**34.04**
MSSIM	0.1260	0.7362	0.7969	0.8687	0.8847	0.8859	0.8807	0.8483	**0.9044**
(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)

Fig. 1. Denoising results. The first and second rows are the results on Jasper Ridge with $\sigma = 0.05$, $p_{\mathbf{s}} = 0.05$ and the third and fourth rows are the results on PaviaU with $\sigma = 0.1$, $p_{\mathbf{s}} = 0.1$. (a) Ground-truth. (b) Observed noisy image. (c) TDL. (d) ITSReg. (e) SSTV. (f) HSSTV1. (g) HSSTV2. (h) l_0-l_1HTV. (i) LRTDTV. (j) RISSTV (ours).

Tensor Dictionary Learning (TDL) [34] and Intrinsic Tensor Sparsity Regularization (ITSReg) [35]. The third is a TV-tensor hybrid method, i.e., Total Variation-Regularized Low-Rank Tensor Decomposition (LRTDTV) [17]. Here, HSSTV with ℓ_1-norm and mixed $\ell_{1,2}$-norm are denoted by HSSTV1 and HSSTV2, respectively. We used *Jasper Ridge*[3] with size $100 \times 100 \times 198$ and *Pavia University*[4] (PaviaU) cropped to size $140 \times 140 \times 98$ as ground-truth HS images. All the intensities were normalized to within the range $[0, 1]$. Noisy HSIs were generated based on (6). In this experiment, we simulated two cases. The first is the low-intensity noise case, where we set the standard deviation of the Gaussian noise σ to 0.05 and the ratio of salt-and-paper noise $p_{\mathbf{s}}$ to 0.05. The second is the high-intensity noise case, where we set σ to 0.1 and $p_{\mathbf{s}}$ to 0.1.

For a fair comparison, we replaced RISSTV in (7) with TV-based methods, i.e., SSTV, HSSTV1, HSSTV2, l_0-l_1HTV,

respectively, and solved each problem by P-PDS. For TDL, ITSReg, and LRTDTV, we used the implementation published by the authors. In RISSTV, we fixed ω to 0.01 for all images and conditions. The radiuses η in (9) and ε in (10) were set to $0.93 \frac{N p_{\mathbf{s}}}{2}$ and $0.93 \sqrt{\sigma^2 N (1 - p_{\mathbf{s}})}$, respectively. We set the stopping criterion of Alg. 1 to $\frac{\|\mathbf{u}^{(t+1)} - \mathbf{u}^{(t)}\|_2}{\|\mathbf{u}^{(t)}\|_2} < 1.0 \times 10^{-5}$. For quality measures, we employed the mean peak signal-to-noise ratio (MPSNR) [dB]: $\frac{1}{N_3} \sum_{i=1}^{N_3} 10 \log_{10}(N_1 N_2 / \|\mathbf{u}_i - \bar{\mathbf{u}}_i\|_2^2)$, and the mean structural similarity index (MSSIM) [36]: $\frac{1}{N_3} \sum_{i=1}^{N_3} \text{SSIM}(\mathbf{u}_i, \bar{\mathbf{u}}_i)$, where $\mathbf{u}_i$ is the i-th band of $\mathbf{u}$. where $\mathbf{u}_i$ and $\bar{\mathbf{u}}_i$ are the i-th band of the ground true HS image $\mathbf{u}$ and the estimated HS image $\bar{\mathbf{u}}$, respectively. Generally, higher MPSNR and MSSIM values are corresponding to better denoising performances.

Table I shows MPSNR [dB] and MSSIM of all the denoising results for each method. The best MPSNR and MSSIM values are highlighted in bold. Our RISSTV performed better than all other methods in most cases. Especially for MPSNR, our RISSTV is superior to all the other methods in all cases.

[3]https://rslab.ut.ac.ir/data
[4]https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes

The first and second rows in Fig. 1 depict the denoising results at $\sigma = 0.05$ and $p_{\mathbf{s}} = 0.05$ in Jasper Ridge with the 99th band, and the third and fourth rows those at $\sigma = 0.1$ and $p_{\mathbf{s}} = 0.1$ in PaviaU with the 20th band. The images in the second and fourth rows show the absolute differences between the original images and each restored image in each result, multiplied by eight for visibility.[5]

In images (c) and (d) restored with the tensor-based methods TDL and ITSReg, respectively, edges and textures are restored to some extent, but the mean pixel values are significantly off due to spectral distortion. More detailed structure is preserved in the restored images by SSTV in (e) and the existing TV-based methods (HSSTV1, HSSTV2, l_0-l_1HTV, LRTDTV) in (f), (g), (h), and (i) than those by TDL and ITSReg. However, in the case of higher noise intensity, noise and artifacts remain in the restored image by the methods using only the second-order spatio-spectral differences, i.e., SSTV, l_0-l_1HTV, and LRTDTV in (e), (h), and (i). The images in (f) and (g) restored by HSSTV1 and HSSTV2 can remove noise more effectively than these methods, but many edges and textures are present in the difference images. On the other hand, as shown in (j), RISSTV achieves the highest restoration performance. In addition, in the enlarged area of the difference images between the ground-truth images and the restored images by RISSTV, edges and textures, oblique edges, do not appear much. These results suggest that RISSTV can remove noise while retaining edges and textures with high accuracy. This may be due to the fact that RISSTV has the rotation-invariant property.

V. Conclusion

We proposed a new HS image regularization model, named RISSTV, for HS image denoising. RISSTV is designed by combining the TVs of the spatial/spatio-spectral differences and making them rotation-invariant, leading to a powerful regularization model that fully recovers the spatio-spectral structures, including round structures and oblique edges, in the target HS image. The HS image denoising problem using RISSTV is formulated as a convex optimization problem and efficiently solved by an efficient algorithm based on a preconditioned primal-dual splitting method. The experimental results on mixed noise removal illustrate the advantage of RISSTV over several existing HS image regularization models.

Acknowledgment

---

[5]The difference images for the observed noisy images or TDL were multiplied by one or three, respectively, in order to avoid excessive white-out.

References

[1] M. Borengasser, W. S. Hungate, and R. Watkins, *Hyperspectral remote sensing: principles and applications.* CRC press, 2007.

[2] H. Grahn and P. Geladi, *Techniques and applications of hyperspectral image analysis.* John Wiley & Sons, 2007.

[3] P. S. Thenkabail, J. G. Lyon, and A. Huete, *Hyperspectral remote sensing of vegetation.* CRC press, 2016.

[4] B. Lu, P. D. Dao, J. Liu, Y. He, and J. Shang, "Recent advances of hyperspectral imaging technology and applications in agriculture," *Remote Sens.*, vol. 12, no. 16, p. 2659, 2020.

[5] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, 2012.

[6] W. Ma, J. M. Bioucas-Dias, T. Chan, *et al.*, "A signal processing perspective on hyperspectral unmixing: Insights from remote sensing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 67–81, 2014.

[7] P. Ghamisi, J. Plaza, Y. Chen, J. Li, and A. J. Plaza, "Advanced spectral classifiers for hyperspectral images: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 1, pp. 8–32, 2017.

[8] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6690–6709, 2019.

[9] N. Audebert, B. L. Saux, and S. Lefevre, "Deep learning for classification of hyperspectral data: A comparative review," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 159–173, 2019.

[10] S. Matteoli, M. Diani, and J. Theiler, "An overview of background modeling for detection of targets and anomalies in hyperspectral remotely sensed imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2317–2336, 2014.

[11] H. Su, Z. Wu, H. Zhang, and Q. Du, "Hyperspectral anomaly detection: A survey," *IEEE Geos. Remote Sens. Mag.*, vol. 10, no. 1, pp. 64–90, 2022.

[12] H. Shen, X. Li, Q. Cheng, *et al.*, "Missing information reconstruction of remote sensing data: A technical review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 61–85, 2015.

[13] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sens.*, vol. 10, no. 3, 2018.

[14] H. Shen, M. Jiang, J. Li, C. Zhou, Q. Yuan, and L. Zhang, "Coupling model- and data-driven methods for remote sensing image restoration and fusion: Improving physical interpretability," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 231–249, 2022.

[15] H. K. Aggarwal and A. Majumdar, "Hyperspectral image denoising using spatio-spectral total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 442–446, 2016.

[16] H. Fan, C. Li, Y. Guo, G. Kuang, and J. Ma, "Spatial–spectral total variation regularized low-rank tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6196–6213, 2018.

[17] Y. Wang, J. Peng, Q. Zhao, Y. Leung, X. Zhao, and D. Meng, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1227–1243, 2018.

[18] T. Ince, "Hyperspectral image denoising using group low-rank and spatial-spectral total variation," *IEEE Access*, vol. 7, pp. 52 095–52 109, 2019.

[19] S. Takeyama, S. Ono, and I. Kumazawa, "A constrained convex optimization approach to hyperspectral image restoration with hybrid spatio-spectral regularization," *Remote Sens.*, vol. 12, no. 21, 2020.

[20] M. Wang, Q. Wang, J. Chanussot, and D. Hong, "l_0-l_1 hybrid total variation regularization and its applications on hyperspectral image mixed noise removal and compressed sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7695–7710, 2021.

[21] S. Takemoto, K. Naganuma, and S. Ono, "Graph spatio-spectral total variation model for hyperspectral image denoising," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[22] S. Takemoto and S. Ono, "Enhancing spatio-spectral regularization by structure tensor modeling for hyperspectral image denoising," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., (ICASSP)*, 2023, pp. 1–5.

[23] L. Condat, "Discrete total variation: New definition and minimization," *SIAM J. Imag. Sci.*, vol. 10, no. 3, pp. 1258–1290, 2017.

[24] T. Pock and A. Chambolle, "Diagonal preconditioning for first order primal-dual algorithms in convex optimization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2011, pp. 1762–1769.

[25] K. Naganuma and S. Ono, "Variable-wise diagonal preconditioning for primal-dual splitting: Design and applications," *IEEE Trans. Signal Process.*, vol. 71, pp. 3281–3295, 2023.

[26] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, no. 1, pp. 120–145, 2011.

[27] L. Condat, "A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," *J. Optim. Theory Appl.*, vol. 158, no. 2, pp. 460–479, 2013.

[28] M. V. Afonso, J. M. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, 2011.

[29] G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu, "Epigraphical projection and proximal tools for solving constrained convex optimization problems," *Signal Image Video Process.*, vol. 9, no. 8, pp. 1737–1749, 2015.

[30] S. Ono and I. Yamada, "Signal recovery with certain involved convex data-fidelity constraints," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6149–6163, 2015.

[31] S. Ono, "Primal-dual plug-and-play image restoration," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1108–1112, 2017.

[32] S. Ono, "Efficient constrained signal reconstruction by randomized epigraphical projection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., (ICASSP)*, 2019, pp. 4993–4997.

[33] L. Condat, "Fast projection onto the simplex and the ℓ_1 ball," *Math. Program.*, vol. 158, no. 1, pp. 575–585, 2016.

[34] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang, "Decomposable nonlocal tensor dictionary learning for multispectral image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2014, pp. 2949–2956.

[35] Q. Xie, Q. Zhao, D. Meng, *et al.*, "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE conf. comput. vis. pattern recognit. (CVPR)*, 2016, pp. 1692–1700.

[36] W. Zhou, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.