Spatio-Spectral Structure Tensor Total Variation for Hyperspectral Image Denoising and Destriping

Shingo Takemoto, Student Member, IEEE, Shunsuke Ono, Member, IEEE

Abstract—This paper proposes a novel regularization method, named Spatio-Spectral Structure Tensor Total Variation (S3TTV), for denoising and destriping of hyperspectral (HS) images. HS images are inevitably contaminated by various types of noise, during acquisition process, due to the measurement equipment and the environment. For HS image denoising and destriping tasks, Spatio-Spectral Total Variation (SSTV), defined using second-order spatio-spectral differences, is widely known as a powerful regularization approach that models the underlying spatio-spectral properties. However, since SSTV refers only to adjacent pixels/bands, semi-local spatial structures are not preserved during denoising process. To address this problem, we newly design S3TTV, defined by the sum of the nuclear norms of matrices consisting of second-order spatio-spectral differences in small spectral blocks (we call these matrices as spatio-spectral structure tensors). The proposed regularization method simultaneously models the spatial piecewise-smoothness, the spatial similarity between adjacent bands, and the spectral correlation across all bands in small spectral blocks, leading to effective noise removal while preserving the semi-local spatial structures. Furthermore, we formulate the HS image denoising and destriping problem as a convex optimization problem involving S3TTV and develop an algorithm based on a preconditioned primal-dual splitting method to solve this problem efficiently. Finally, we demonstrate the effectiveness of S3TTV by comparing it with existing methods, including state-of-the-art ones through denoising and destriping experiments.

Index Terms—Hyperspectral image, denoising, destriping, spatio-spectral regularization, total variation, structure tensor

I. INTRODUCTION

HYPERSPECTRAL (HS) imaging measures a wide spectrum of light ranging from the ultraviolet to the near-infrared. The rich spectral information of HS images, with more than one hundred bands, can distinguish materials and phenomena that the human eye and existing RGB cameras cannot. This capability has been applied in diverse fields including agriculture, mineralogy, astronomy, and biotechnology [1]–[4]. However, observed HS images are inevitably contaminated by various types of noise—including thermal noise, quantization noise, shot noise, and stripe noise—due to factors caused by measurement equipment and environment, such as photon effects, atmospheric absorption, dark currents, and sensor disturbances [5]–[7]. Since such noise significantly degrades the performance of subsequent processing such as unmixing [8], [9], classification [10]–[12], and anomaly detection [13], [14]. HS image denoising is an essential preprocessing step for the applications.

To recover desirable HS images from degraded observations, many existing methods design some regularization functions that mathematically model the inherent spatio-spectral structure in HS images and then solve optimization problems that incorporate the regularization functions. As a representative class of regularization methods for HS images, total variation (TV) type regularization methods have been developed. This approach originated from edge-preserving natural image denoising [15], [16], and the pioneering TV-type method for HS image denoising is Spectral-Spatial Adaptive Hyperspectral TV (SASAHTV) [17]. SASAHTV models the spatial piecewise-smoothness of HS images as the sparsity of first-order spatial differences between adjacent pixels. Furthermore, Anisotropic Spectral Spatial TV (ASSSTV) [18] extends this model by incorporating first-order spectral differences into the SASAHTV regularization function, capturing the spectral piecewise-smoothness in addition to the spatial piecewise-smoothness of HS images. However, SASAHTV and ASSSTV cause 1) over-smoothing by minimizing \( \ell_1 \)-type norms of first-order differences, and 2) corruption of the semi-local spatial structures during the denoising process because they only refer to adjacent pixels/bands.

To mitigate these weaknesses, some regularization methods have combined ASSSTV with an additional regularization function, specifically a low-rank (LR)-type regularization function [19]–[21]. The LR-type regularization functions are designed to capture the spectral correlation across all bands due to the fact that HS images are composed of a few variety of material-specific spectra. However, these methods have two other challenges: the mischaracterization of stripe noise (see Figs. 13–14 and 15 in Sec. IV) and the difficulty of setting the parameter that balances the two different regularization functions. Therefore, our paper focuses on enhancing single TV-type regularization methods.

One promising single TV-type regularization method that overcomes over-smoothing is Spatio-Spectral TV (S3TV) [22]. S3TV is defined by the \( \ell_1 \) norm of second-order spatio-spectral differences, i.e., first-order spatial differences of spectral ones to avoid over-smoothing. In addition, S3TV captures not only the spatial piecewise-smoothness, but also the spatial similarity between adjacent bands by its formulation. For these reasons, S3TV has been widely used in state-of-the-art HS
image denoising methods [23]–[27]. Takeyama et al. proposed Hybrid Spatio-Spectral TV (HSSTV) [25], which enhances SSTV by incorporating SSAHTV. To recover more detailed spatial structures of HS images, Graph Spatio-Spectral TV (GSSTV) [27] was proposed, which weighs the spatial difference operator of SSTV based on a graph reflecting the spatial structures of an HS image. To directly control the degree of the smoothness, Wang et al. proposed the \( \ell_0-\ell_1 \) hybrid TV (\( \ell_0-\ell_1 \) HTV) [26], which incorporates \( \ell_0 \)-type constraints of the spatial differences (originally proposed for color image processing [28]) into SSTV. However, SSTV and its extension methods cannot preserve the semi-local spatial structures of HS images during the denoising process because they only refer to adjacent pixels/bands.

Other single TV-type regularization methods that preserve the semi-local spatial structures are the Structure tensor TV (STV)-type regularization methods [29]–[32]. These regularization functions are defined as the sum of nuclear norms of structure tensors consisting of matrices of local differences in small spectral blocks containing small spatial areas and all bands. The original STV [29] and Structure Tensor Total variation-regularized Weighted Nuclear Norm Minimization (STWNNM) [30] capture the semi-local spatial piecewise-smoothness, but do not model the spectral property of HS images due to consisting only of spatial differences. The regularization function of Arranged Structure tensor TV (ASTV) [31] characterizes the spectral correlations across all bands by modifying the ordering of spatial differences in the structure tensors. In addition, Spatio-Spectral Structure Tensor (SSST) [32] explicitly exploits the spectral piecewise-smoothness of HS images by including not only spatial differences but also spectral ones in its formulation. However, the existing STV-type regularization methods minimize the nuclear norms of “first-order” differences, leading to spatial or spectral over-smoothing, as in SSAHTV.

From the discussion so far, SSTV and STV-type regularization methods are powerful approaches that capture the underlying spatial and spectral characteristics of HS images. However, they have their own limitations: SSTV cannot preserve the semi-local spatial structures and STV-type regularization methods cause over-smoothing. Now a natural question arises: Can we design a regularization function with the four requirements: avoid over-smoothing, preserve semi-local spatial structures, effectively remove stripe noise, and be a single regularization function?

Inspired by both SSTV and the above family of STV, we propose a denoising method for HS images using a newly introduced spatio-spectral structure tensor total variation (S\(_3\)TTV) model. The main contributions of this article are summarized as follows.

1) We design a novel formulation of an STV-type regularization method, namely S\(_3\)TTV. S\(_3\)TTV is defined as the sum of the nuclear norms of newly designed structure tensors (called spatio-spectral structure tensor), each of which consists of second-order spatio-spectral differences within a small spectral block. S\(_3\)TTV can be seen as a hybrid regularization function that is constructed by integrating SSTV and STV into a convex function, and thus can fully capture the spatial piecewise-smoothness, the spatial similarity between adjacent bands, and the spectral correlation across all bands of HS images. This leads to effective noise removal while preserving the semi-local spatial structures of HS images without over-smoothing.

2) We formulate the mixed noise removal problem as a constrained convex optimization problem involving S\(_3\)TTV. By including the functions that characterize Gaussian, sparse, and stripe noise in the optimization problem, the proposed method effectively removes the three types of noise. In this formulation, we model data-fidelity and noise terms as hard constraints instead of adding them to the objective function. This type of constrained formulation decouples interdependent hyperparameters into independent ones, making parameter setting easier. Such advantages are also addressed, for example, in [33]–[38]. Furthermore, since the objective function has only a single regularization function, the proposed method does not require a balance parameter.

3) To solve our optimization problem for the HS image denoising, we develop an efficient algorithm based on a Preconditioned Primal-Dual Splitting method (P-PDS) [39]. Unlike other popular algorithms used in existing HS image denoising methods, such as an alternating direction method of multipliers [40] and PDS [41]. P-PDS can automatically determine the appropriate steplengths based on the problem structure [39], [43]. Experimental results show the superiority of the proposed method to existing methods including state-of-the-art ones. The comparison of the features of the existing and proposed methods is summarized in Fig. I.

The paper is organized as follows. In Sec. II, we introduce the mathematical tools required for the proposed method. Sec. III provides the proposed HS image denoising method involving S\(_3\)TTV. The experimental results are reported in Sec. IV. Finally, we give concluding remarks in Sec. V. The preliminary version of this paper, without considering stripe noise, mathematical details, comprehensive experimental comparison, or deeper discussion, has appeared in conference proceedings [44].

II. PRELIMINARIES

A. Notations

Throughout this paper, we denote vectors and matrices by boldface lowercase letters (e.g., \( \mathbf{x} \)) and boldface capital letters (e.g., \( \mathbf{X} \)). We treat an HS image, denoted by \( \mathbf{u} \) with \( N_1 \) vertical pixels, \( N_2 \) horizontal pixels, and \( N_3 \) bands. We denote the total number of elements in the HS image by \( N = N_1 N_2 N_3 \). For a matrix data \( \mathbf{X} \in \mathbb{R}^{N_1 \times N_2} \), the value at the location \((i,j)\) is denoted by \( [\mathbf{X}]_{i,j} \). The \( \ell_1 \)-norm and the \( \ell_2 \)-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 \( \mathbf{x} \). The nuclear norm of a matrix, which is the sum of all the singular values, is denoted by \( \| \mathbf{X} \|_* \). For an HS image \( \mathbf{u} \in \mathbb{R}^N \), let \( \mathbf{D}_s \in \mathbb{R}^{N \times N}, \mathbf{D}_h \in \mathbb{R}^{N \times N}, \) and \( \mathbf{D}_s \in \mathbb{R}^{N \times N} \) be the forward difference operators along
the horizontal, vertical, and spectral directions, respectively, with the periodic boundary condition. Here, a spatial difference operator is denoted by \( D \) := \( (D_u, D_h, D_s) \) \( \in \mathbb{R}^{2N \times N} \). Using \( D_u, D_h, \) and \( D_s \), we denote the second-order spatio-spectral differences by \( D_u D_h D_s u \in \mathbb{R}^N \) and \( D_h D_s u \in \mathbb{R}^N \). Other notations will be introduced as needed.

### B. Proximal Tools

In this chapter, we introduce basic proximal tools that play a central role in the optimization part of our method.

For any \( \gamma > 0 \), the proximity operator of \( f \) is defined by

\[
\text{prox}_{\gamma f}(x) := \arg \min_{y \in \mathbb{R}^N} f(y) + \frac{1}{2\gamma} \|x - y\|^2_2.
\] (1)

The Fenchel–Rockafellar conjugate function \( f^* \) of the function \( f \) is defined by

\[
f^*(x) := \sup_{y} \langle x, y \rangle - f(y),
\] (2)

where \( \langle \cdot, \cdot \rangle \) is the Euclidean inner product. Thanks to a generalization of Moreau’s identity \[45\], the proximity operator of \( f^* \) is calculated as

\[
\text{prox}_{\gamma f^*}(x) = x - \gamma \text{prox}_{\frac{1}{\gamma} f}(\frac{1}{\gamma} x).
\] (3)

The indicator function of a nonempty closed convex set \( C \subset \mathbb{R}^N \), denoted by \( \iota_C \), is defined as

\[
\iota_C(x) := \begin{cases} 
0, & \text{if } x \in C, \\
\infty, & \text{otherwise.}
\end{cases}
\] (4)

The proximity operator of \( \iota_C \) is equivalent to the projection onto \( C \), as given by

\[
\text{prox}_{\iota_C}(x) = P_C(x) := \arg \min_{y \in C} \|y - x\|_2.
\] (5)

### C. Preconditioned Primal-Dual Splitting Method (P-PDS)

Consider the following generic form of convex optimization problems:

\[
\begin{align*}
\min_{x_1, \ldots, x_M, y_1^1, \ldots, y_M^M} & \sum_{i=1}^{N} \sum_{j=1}^{M} f_i(x_i) + \sum_{j=1}^{M} g_j(y_j) \\
\text{s.t.} & \begin{cases} 
y_1^1 = \sum_{i=1}^{N} A_{1,i} x_i, \\
\vdots 
y_M^M = \sum_{i=1}^{N} A_{M,i} x_i,
\end{cases}
\end{align*}
\] (6)

where \( f_i(i = 1, \ldots, N) \) and \( g_j(j = 1, \ldots, M) \) are lower semi-continuous proper convex functions, \( x_i \in \mathbb{R}^{n_i} \) \( (i = 1, \ldots, N) \) are primal variables, \( y_j \in \mathbb{R}^{n_j} \) \( (j = 1, \ldots, M) \) are dual variables, and \( A_{j,i} \in \mathbb{R}^{n_j \times n_i} \) \((i = 1, \ldots, N, j = 1, \ldots, M)\) are linear operators.

The standard PDS \[41\], \[42\] and P-PDS \[39\], on which our algorithm is based, solve Prob. (6) by the following iterative procedures:

\[
\begin{align*}
x_{i}^{(t+1)} & \leftarrow \text{prox}_{\iota_{1,1} \cdot f_1} (x_{i}^{(t)} - \gamma_1, 1, (\sum_{j=1}^{M} A^T_{1,j} y_{j}^{(t)}) ), \\
\vdots 
x_{N}^{(t+1)} & \leftarrow \text{prox}_{\iota_{1,1} \cdot f_N} (x_{N}^{(t)} - \gamma_1, 1, (\sum_{j=1}^{M} A^T_{1,j} y_{j}^{(t)}) ), \\
x_{i}^{(t+1)} & \leftarrow 2x_{i}^{(t+1)} - x_{i}^{(t)} (\forall i = 1, \ldots, N), \\
y_{1}^{(t+1)} & \leftarrow \text{prox}_{\iota_{2,1} \cdot g_1} (y_{1}^{(t)} - \gamma_2, 1, (\sum_{i=1}^{N} A_{1,i} x_{i}^{(t)}) ), \\
\vdots 
y_{M}^{(t+1)} & \leftarrow \text{prox}_{\iota_{2,1} \cdot g_M} (y_{M}^{(t)} - \gamma_2, 1, (\sum_{i=1}^{N} A_{M,i} x_{i}^{(t)}) ),
\end{align*}
\] (7)

where \( \gamma_{1,i}(i = 1, \ldots, N) \) and \( \gamma_{2,j}(j = 1, \ldots, M) \) are the stepsize parameters. The standard PDS needs to adjust the appropriate stepsize parameters to satisfy the convergence conditions. On the other hand, P-PDS can automatically determine the stepsize parameters that guarantee convergence \[39\], \[43\]. According to \[39\], this paper designs the stepsize parameters as

\[1\]

In \[39\], the stepsize parameters are defined per element not just per variable, because the reciprocals of the row/column absolute sums of the linear operator are generally different for each row/column. However, in this paper, the resulting row/column absolute sum of linear operators is the same for each variable. Therefore, we define the variable-wise stepsize parameters in Eq. (5) for simplicity.
Fig. 1. Illustration of the proposed method, i.e., $S_3$TTV.

as follows:

$$\gamma_{1,i} = \frac{1}{\sum_{j=1}^{M} \sum_{i=1}^{m_i} \|A_{j,i}\|_1}, \quad (\forall i = 1, \ldots, N),$$

$$\gamma_{2,j} = \frac{1}{\sum_{i=1}^{N} \sum_{k=1}^{n_i} \|A_{j,i,k}\|_1}, \quad (\forall j = 1, \ldots, M). \quad (8)$$

### III. PROPOSED METHOD

In the following, we first describe the design of the $S_3$TTV regularization function. Next, we consider a situation where an HS image is contaminated with mixed noise and introduce the corresponding observation model. Based on this model, we formulate an HS image denoising problem as a constrained optimization problem involving the $S_3$TTV regularization function. Finally, we derive an algorithm based on P-PDS to efficiently solve the optimization problem. A schematic diagram of $S_3$TTV is shown in Fig. [1].

#### A. Spatio-Spectral Structure Tensor Total Variation ($S_3$TTV)

Before describing the proposed regularization function, we introduce the notion of **spatio-spectral structure tensor**\(^2\). First, for a given HS image $u$, we calculate the second-order spatial-spectral differences $D_x D_y u$ and $D_y D_x u$. Next, we extract small spectral blocks by cropping the second-order spatio-spectral differences to the size $N_1 \times N_2 (N_1 << N_1, N_2 << N_2)$ for all bands\(^3\). Then, the $b$-th spatio-spectral structure tensor $L_b^{(b)}$ is defined by vectorizing the second-order spatio-spectral differences in the $b$-th small spectral block by band and arranging them in parallel as follows:

$$L_b^{(b)} := \begin{bmatrix} [D_x D_y u]_1^{(b)} & [D_x D_y u]_2^{(b)} & \cdots & [D_x D_y u]_L^{(b)} \\ [D_y D_x u]_1^{(b)} & [D_y D_x u]_2^{(b)} & \cdots & [D_y D_x u]_L^{(b)} \end{bmatrix} \in \mathbb{R}^{N_1' N_2' \times 2 N_3}. \quad (9)$$

where $[D_x D_y u]_i^{(b)} \in \mathbb{R}^{N_1' N_2'}$ and $[D_y D_x u]_i^{(b)} \in \mathbb{R}^{N_1' N_2'}$ are the second-order spatio-spectral differences of the $i$-th band in the $b$-th small spectral block. Since HS images have the strong correlation across all bands, $[D_x D_y u]_1^{(b)}, \ldots, [D_x D_y u]_L^{(b)}$ and $[D_y D_x u]_1^{(b)}, \ldots, [D_y D_x u]_L^{(b)}$ are similar vectors, respectively, i.e., the columns of $L_b^{(b)}$ are approximately linearly dependent.

The flow of constructing the spatio-spectral structure tensor is depicted in the middle right of Fig. [1].

To capture the spatial piecewise-smoothness, the spatial similarity between adjacent bands, and the spectral correlation of an HS image, we propose a regularization function using the spatio-spectral structure tensors as follows:

$$S_3$$TTV$(u) := \sum_{b=1}^{B} \|L_b^{(b)}\|_*, \quad (10)$$

where $B$ is the number of the extracted small spectral blocks. We call this function as **Spatio-Spectral Structure Tensor Total Variation ($S_3$TTV)**. Here, $L_b^{(b)}$ is represented with an operator $P_b \in \mathbb{R}^{2 N_1' N_2' \times 2 N_1 N_2 N_3}$ that extracts the $b$-th small spectral block as

$$L_b^{(b)} = P_b D D_x u. \quad (11)$$

Rather than directly suppressing some norm of the spatial differences, $S_3$TTV suppresses that of the spatial differences of the spectral differences (i.e., the second-order spatio-spectral differences), which preserves spatial similarity between adjacent bands. In addition, since the second-order difference vectors of each band are horizontally aligned in the structure
tensors, $S_3$TTV can capture strong correlation across all bands when the sum of singular values of $L_u^{(b)}$ is reduced. This is because the stronger the correlation across all bands in HS images is, the more approximately linearly dependent the columns of $L_u^{(b)}$ are.

B. HS Image Denoising Problem by $S_3$TTV

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

$$v = \bar{u} + \bar{s} + \bar{t} + n,$$

where $\bar{u}$ is a clean HS image, $\bar{s}$ is sparse noise, $\bar{t}$ is stripe noise, and $n$ is Gaussian noise, respectively.

Based on the above observation model, we formulate an HS image denoising problem involving $S_3$TTV as a constrained convex optimization problem with the following form:

$$\min_{u, s, t \in \mathbb{R}^N} S_3^{TTV}(u) \text{ s.t. }\begin{cases} s \in B_{1,\alpha}, \\ t \in B_{1,\beta}, \\ D_s t = 0, \\ u + s + t \in B_{2,\varepsilon}^y, \\ u \in R_{\bar{u},\bar{\mu}}, \end{cases}$$

where

$$B_{1,\alpha} := \{x \in \mathbb{R}^N | \|x\|_1 \leq \alpha\},$$

$$B_{1,\beta} := \{x \in \mathbb{R}^N | \|x\|_1 \leq \beta\},$$

$$B_{2,\varepsilon}^y := \{x \in \mathbb{R}^N | \|x - v\|_2 \leq \varepsilon\},$$

$$R_{\bar{u},\bar{\mu}} := \{x \in \mathbb{R}^N | \bar{\mu} \leq x_i \leq \bar{\mu} \ (i = 1, \ldots, N)\}.$$ (17)

The first constraint characterizes sparse noise $s$ with the zero-centered $\ell_1$-ball of the radius $\alpha > 0$. The second constraint controls the intensity of stripe noise $t$ and the third constraint captures the vertical flatness property by imposing zero to the vertical gradient of $t$. These constraints effectively characterize stripe noise $[38]$. The fourth constraint serves as data-fidelity with the $\bar{v}$-centered $\ell_2$-ball of the radius $\varepsilon > 0$. The fifth constraint is a box constraint with $\bar{\mu} < \bar{\mu}$ which represents the dynamic range of $u$. For normalized HS images, we can set $\bar{\mu} = 0$ and $\bar{\mu} = 1$.

Using the first, second, and fourth constraints instead of adding terms to the objective function makes it much easier to adjust the parameters $\alpha$, $\beta$, and $\varepsilon$. This is because by expressing multiple terms as constraints, rather than adding them to the objective function, the hyperparameters associated with each term are converted to be independent of each other, and appropriate parameters can be determined without interdependence. Such advantage has been addressed, e.g., in $[33]-[37]$.

C. Optimization

We develop an efficient solver for Prob. (13) based on P-PDS $[39]$. Using the indicator functions $\ell(0)$, $\ell_{B^y_{2,\varepsilon}}$, $\ell_{B_{1,\alpha}}$, and $\ell_{R_{\bar{u},\bar{\mu}}}$, we rewrite Prob. (13) into an equivalent form:

$$\min_{y_{1,1}, \ldots, y_{1,b}, y_2, y_3} \sum_{b=1}^B \|y_{1,b}\|_s + \ell(0)(y_2) + \ell_{B^y_{2,\varepsilon}}(y_3)$$

$$+ \ell_{B_{1,\alpha}}(s) + \ell_{B_{1,\beta}}(t) + \ell_{R_{\bar{u},\bar{\mu}}}(u),$$

s.t. $$\begin{align*}
 y_{1,1} &= P_1 DD_s u, \\
 \vdots \\
 y_{1,B} &= P_B DD_s u, \\
 y_2 &= D_s t, \\
 y_3 &= u + s + t.
\end{align*}$$

Then, by defining,

$$f_1(u) := \ell_{R_{\bar{u},\bar{\mu}}}(u),$$

$$f_2(s) := \ell_{B_{1,\alpha}}(s),$$

$$f_3(t) := \ell_{B_{1,\beta}}(t),$$

$$g_1(y_{1,1}) := \|y_{1,1}\|_s, \ldots, g_B(y_{1,B}) := \|y_{1,B}\|_s,$$

$$g_{B+1}(y_2) := \ell(0)(y_2),$$

$$g_{B+2}(y_3) := \ell_{B^y_{2,\varepsilon}}(y_3),$$

Prob. (13) is reduced to Prob. (5). Therefore, P-PDS is applicable to Prob. (13).

We show the detailed algorithm in Alg. 1. The proximity
operators of \( \ell_{B_{\mu,\gamma}} \), \( \ell_{B_{\gamma}} \), and \( \ell_{B_{\gamma,\nu}} \) are calculated by

\[
[\text{prox}_{\gamma_{B_{\mu,\gamma}}} (x)]_i = \begin{cases} 
    \mu_i, & \text{if } x_i < \mu_i, \\
    \mu_i, & \text{if } x_i > \mu_i, \\
    x_i, & \text{otherwise},
\end{cases}
\]

(20)

\[
\text{prox}_{\gamma_{B_{\gamma}}} (x) = 0,
\]

(21)

\[
\text{prox}_{\gamma_{B_{\gamma,\nu}}} (x) = \begin{cases} 
    x, & \text{if } x \in B_{\gamma,\nu}, \\
    v + \epsilon (v - x) / \|v - x\|_2, & \text{otherwise},
\end{cases}
\]

(22)

The proximity operators of \( \ell_{B_{\mu,\gamma}}(s) \) and \( \ell_{B_{\gamma}}(t) \) can be efficiently computed by a fast \( \ell_1 \)-ball projection algorithm \cite{46}. The proximity operator for the nuclear norm \( \|\cdot\|_\star \) is calculated by

\[
\text{prox}_{\|\cdot\|_\star} (X) = U \Sigma^\dagger V^\top, \\
\Sigma_{\gamma} = \text{diag}(\max\{\rho_1 - \gamma, 0\}, \cdots, \max\{\rho_Q - \gamma, 0\}),
\]

(23)

where \( Q \) is the number of nonzero singular values, i.e., \( \min\{N_1' N_2', 2 N_3\} \), and the singular value decomposition of \( X \) is \( U \text{ diag}(\rho_1, \cdots, \rho_Q) V^\top \).

Based on Eq. (8), the stepsize parameters \( \gamma_{u_1}, \gamma_{s_1}, \gamma_t, \gamma_{y_1,1}, \cdots, \gamma_{y_1,8}, \gamma_{y_2} \), and \( \gamma_3 \) are given as

\[
\gamma_u = \frac{1}{8B + 1}, \quad \gamma_s = 1, \quad \gamma_t = \frac{1}{3}, \\
\gamma_{y_1,1} = \cdots = \gamma_{y_1,8} = \frac{1}{4}, \\
\gamma_{y_2} = \frac{1}{2}, \quad \gamma_3 = \frac{1}{3},
\]

(24)

IV. Experiments

To demonstrate the effectiveness of \( S_3\text{TTV} \), we conducted mixed noise removal experiments on HS image contaminated with simulated or real noise. We compared \( S_3\text{TTV} \) with three types of methods; SSTV-based methods, i.e., SSTV \cite{22}, HSSTV \cite{25}, and \( l_0-l_1\text{HTV} \) \cite{26}, STV-based methods, i.e., \( STV \) \cite{29} and SST \cite{32}, and TV-LR hybrid methods, i.e., LRTDTV \cite{20} and TPTV \cite{21}. Here, HSSTV with \( \ell_1 \)-norm and \( \ell_{1,2} \)-norm are denoted by HSSTV1 and HSSTV2, respectively. For a fair comparison, the regularization functions of the P-PDS applicable methods, i.e., SSTV, HSSTV1, HSSTV2, \( l_0-l_1\text{HTV} \), STV, and SST, were replaced with the \( S_3\text{TTV} \) regularization function in Prob. (13), and we solve each problem by P-PDS. For LRTDTV and TPTV, we used implementation codes published by the authors\(^4\).

A. Simulated HS Image Experiments

We adopt two HS image datasets which have different structures information.

1) Jasper Ridge: This HS image was captured using an Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor in a rural area of California, USA. Jasper Ridge consists of a large river in the center and fine structure on the left and right. The resolution of the original data is \( 512 \times 614 \) pixels with 224 spectral bands per pixel. After removing several noisy bands

\(^4\)The LRTDTV and TPTV implementation codes are available at https://github.com/chuchulyf/ETPTV, respectively.
Fig. 2. Denoising results for Jasper Ridge with the 131st band in Case 1. Upper row images are the restored images by each method. Lower row images are the absolute difference between the original image and each restored image (the range of the only observed noisy image in (b) is [0, 1]). (a) Ground-truth. (b) Observed noisy image. (c) SSTV. (d) HSSTV1. (e) HSSTV2. (f) \(l_0-l_1\)HTV. (g) STV. (h) SSST. (i) LRTDTV. (j) TPTV. (k) \(S_3\)TTV (ours).

and cropping the original data, we obtained the HS image with 100 × 100 pixels and 198 bands.

2) Pavia University: This HS image was captured using a Reflective Optics System Imaging Spectrometer (ROSIS) sensor in Pavia, northern Italy. Pavia University consists of complex structures. The resolution of the original data is 610 × 610 pixels with 103 spectral bands per pixel. After removing several noisy bands and cropping the original data, we obtained the HS image with 120 × 120 pixels and 99 bands.

All the intensities of both HS images were normalized within the range \([0, 1]\).

HS images are often degraded by a mixture of various types of noise in real-world scenarios. Thus, in the experiments, we consider the following six cases of noise contamination:

Case 1: The observed HS image is contaminated by white Gaussian noise with the standard deviation \(\sigma = 0.05\) and salt-and-pepper noise with the rate \(p_s = 0.05\).

Case 2: The observed HS image is contaminated by white Gaussian noise with the standard deviation \(\sigma = 0.1\) and salt-and-pepper noise with the rate \(p_s = 0.05\).

Case 3: The observed HS image is contaminated by white Gaussian noise with the standard deviation \(\sigma = 0.05\) and vertical stripe noise whose intensity is uniformly random in the range \([-0.5, 0.5]\) with the rate \(p_t = 0.05\).

Case 4: The observed HS image is contaminated by white Gaussian noise with the standard deviation \(\sigma = 0.1\) and vertical stripe noise whose intensity is uniformly random in the range \([-0.5, 0.5]\) with the rate \(p_t = 0.05\).

Case 5: The observed HS image is contaminated by white Gaussian noise with the standard deviation \(\sigma = 0.05\), salt-and-pepper noise with the rate \(p_s = 0.05\), and vertical stripe noise whose intensity is uniformly random in the range \([-0.5, 0.5]\) with the rate \(p_t = 0.05\).

Case 6: The observed HS image is contaminated by white Gaussian noise with the standard deviation \(\sigma = 0.1\), salt-and-pepper noise with the rates \(p_s = 0.05\), and vertical stripe noise whose intensity is uniformly random in the range \([-0.5, 0.5]\) with the rates \(p_t = 0.05\).

The block size of \(S_3\)TTV was set to \(10 \times 10 \times N_3\). The radii \(\alpha\), \(\beta\), and \(\varepsilon\) were set as follows:

\[
\alpha = \rho \frac{N_p}{2}, \quad \beta = \rho \frac{0.5 N_p (1 - p_s)}{2}, \quad \varepsilon = \rho \sqrt{\sigma^2 N (1 - p_s)},
\]

(25)

where the parameter \(\rho\) was set to 0.95. The stopping criterion of Alg. 1 were set as follows:

\[
\frac{\|u^{(t+1)} - u^{(t)}\|^2}{\|u^{(t)}\|^2} < 1.0 \times 10^{-5}.
\]

(26)
Fig. 3. Denoising and destriping results for Pavia University with the 58th band in Case 3. Upper row images are the restored images by each method. Lower row images are the absolute difference between the original image and each restored image (the range of the only observed noisy image in (b) is [0, 1]). (a) Ground-truth. (b) Observed noisy image. (c) SSTV. (d) HSSTV1. (e) HSSTV2. (f) \(l_0-l_1\)HTV. (g) STV. (h) SSST. (i) LRTDTV. (j) TPTV. (k) S3TTV (ours).

For the quantitative evaluation, we employed the mean peak signal-to-noise ratio (MPSNR):

\[
\text{MPSNR} = \frac{1}{N_3} \sum_{i=1}^{N_3} 10 \log_{10} \frac{N_1 N_2}{\| u_i - \bar{u}_i \|_2^2},
\]

and the mean structural similarity index (MSSIM) [47]:

\[
\text{MSSIM} = \frac{1}{N_3} \sum_{i=1}^{N_3} \text{SSIM}(u_i, \bar{u}_i),
\]

where \(u_i\) and \(\bar{u}_i\) are the \(i\)-th band of the ground-truth HS image \(u\) and the estimated HS image \(\bar{u}\), respectively. Generally, higher MPSNR and MSSIM values are corresponding to better denoising performances.

1) Quantitative Comparison: Tables II and III show MPSNRs and MSSIMs in the experiments on the HS image contaminated with simulated noise. The best and second best results are highlighted in bold and underlined, respectively. STV is worse in all cases. The SSST results show high performance in stripe noise removal, especially in Case 4 for MSSIMs. However, its effectiveness is degraded when the HS image is contaminated with sparse noise. As for LRTDTV, its performance drops when stripe noise is included. The SSTV-type methods, including SSTV, HSSTV1, HSSTV2, and \(l_0-l_1\)HTV, outperform STV, SSST, and LRTDTV. Among the SSTV-type methods, HSSTV1 performs better for MSSIMs of Pavia University. TPTV yields overall higher MPSNRs and MSSIMs than the other existing methods, most notably for the MSSIMs of Jasper Ridge. However, in Cases 3 and 4, MPSNRs of TPTV are worse than those of the SSTV-based methods. On the other hand, S3TTV achieves the best MPSNRs in most cases, and in the two exceptions, it still ranks second highest. Moreover, S3TTV shows a high overall performance independent of the HS images.

2) Visual Quality Comparison: Figs. 2-4 show the results of HS image denoising and destriping. The lower row images are the absolute difference between the original image and each restored image. Fig. 2 shows the denoising results for Jasper Ridge in Case 1, i.e. under contamination by both Gaussian and sparse noise. In the restored images by STV and SSST in (g) and (h), spatial over-smoothing occurs because the nuclear norm of the “first-order” differences is directly suppressed. The restored images by SSTV-based methods (SSTV, HSSTV1, HSSTV2, and \(l_0-l_1\)HTV) and LRTDTV in (c), (d), (e), (f), and (i) have more structure than those by STV and SSST. However, edges and textures appear in these difference images between the ground-truth and these restored images. Especially in the enlarged images, the edges of the road are clearly visible. In
other words, these detailed structures are lost in the restored images. On the other hand, as shown in (j) and (k), TPTV and $S_3$TTV achieve higher restoration performance than the methods presented above.

Fig. 3 shows the denoising results for Pavia University in Case 3, i.e. under contamination by both Gaussian and stripe noise. As indicated by the rightward arrow in (j), the black line exists in the restored HS image by TPTV. More stripe noise remains in the restored HS image by LRTDTV, including that indicated by the leftward arrows. These occur due to the mischaracterization of stripe noise. As for other methods, including our method, stripe noise is adequately removed. This is thanks to the second and third constraints in Eq. (13) that characterize stripe noise. Furthermore, the methods that mainly use the second-order spatio-spectral differences, i.e., SSTV, HSSTV1, HSSTV2, $l_0$-$l_1$HTV, and our method fully recover HS images without over-smoothing.

Fig. 4 shows the denoising results for Jasper Ridge in Case 6, which is the most contaminated case with Gaussian, sparse, and stripe noise. For STV and SSST in (g) and (h), no stripe noise remains in the restored images, but edges and textures are lost along with the noise. SSTV, HSSTV1, HSSTV2, and $l_0$-$l_1$HTV in (c), (d), (e), and (f) restore edges and textures more clearly than STV and SSST, but still remove some edges and textures (see the difference images in (c), (d), (e), and (f) of Fig. 4). This would be due to the fact that these methods suffer more from the limitation of referring only to adjacent pixels/bands. LRTDTV and TPTV almost recover edges and textures in the restored images of (i) and (j), but as in Case 3, stripe noise remains at the positions indicated by the arrows. On the other hand, the difference image of $S_3$TTV in (k) is closer to black than the other methods. Furthermore, no edges or textures are visible in the enlarged image. These suggest that $S_3$TTV removes Gaussian, sparse, and stripe noise most effectively and recovers edges and textures with the highest accuracy.

B. Real HS Image Experiment

We employed the following two datasets:

1) Indian Pines: This HS image was captured using the AVIRIS sensor over the Indian Pines test site in North-western Indiana. The resolution of the original data is $145 \times 145$ pixels, and each pixel has spectral information with 224 bands ranging from 400 nm to 2500 nm. After removing several noisy bands and cropping the original data, we obtained the HS image with $120 \times 120$ pixels and 198 bands.

2) Suwannee: This HS image was captured using a Spec-TIR sensor over the Suwannee River Basin in Florida, USA.
The resolution of the original HS image is 1200×320 pixels, and each pixel has spectral information with 360 bands ranging from 400 nm to 2500 nm. We cropped the HS image to 100×100 pixels and 360 bands.

All the intensities of both HS images were normalized within the range [0,1]. The block size of $S_3$TTV was set to $10 \times 10 \times N_3$. For the radii $\alpha$, $\beta$, and $\varepsilon$, we adjusted them to appropriate values after empirically estimating the intensity of the noise in the real HS image. Specifically, for the Indian Pines, $\alpha$, $\beta$, and $\varepsilon$ were set to 200, 100, and 30, respectively, and for Suwannee, they were set to 800, 5000, and 100, respectively. The stopping criterion of Alg. 1 were set as follows:

$$\frac{\|u^{(t+1)} - u^{(t)}\|_2}{\|u^{(t)}\|_2} < 1.0 \times 10^{-5}.$$  \hfill (29)

Since no reference clean HS image is available, we compare the denoising performance using visual results. Fig. 5 shows the HS image denoising and destriping results for Indian Pines. HSSTV1, STV, and SSST cause over-smoothing in the restored images of (d), (g), and (h). For LRTDTV and TPTV in (i) and (j), noise remains in the restored images. On the other hand, SSTV, HSSTV2, $l_0$-$l_1$HTV, and $S_3$TTV achieve sufficient noise removal while preserving edges and textures. This is due to the characterization of the spatial similarity between adjacent bands of HS images using the second-order spatio-spectral differences. In the enlarged area, SSTV, HSSTV2, and $S_3$TTV restore the edges indicated by the arrows, while $l_0$-$l_1$HTV loses them.

Fig. 6 shows the HS image denoising and destriping results for Suwannee. The restored images by STV, LRTDTV, and
TPTV in (g), (i), and (j) still retain vertical stripe noise. For LRTDTV and TPTV, this is due to the limitation of stripe noise removal, similar to the results seen in the simulation experiments of Figs. 3 and 4. In the case of STV, its regularization method, which mainly captures the spatial piecewise-smoothness in HS images, is insufficient to separate the HS image and vertically smooth stripe noise. On the other hand, SSTV, HSSTV1, HSSTV2, \( l_0-l_1 \)-HTV, SSST, and \( S_3 \)-TVV sufficiently remove stripe noise. This indicates that the second and third constraints characterizing stripe noise in Eq. (13) are effective for real stripe noise. However, the edges of the restored images by HSSTV1 and SSST in (d) and (h) are smoothed in the enlarged areas. Unlike these methods, SSTV, HSSTV2, \( l_0-l_1 \)-HTV, and the proposed method, \( S_3 \)-TVV, achieve the preservation of the narrow river structure while removing noise.

C. Summary

We summarize the insights from the experiments as follows.

1) The simulated HS image data experiments demonstrate that \( S_3 \)-TVV outperforms existing methods in removing mixed noise while preserving edges and textures in HS images with high accuracy. This indicates that \( S_3 \)-TVV is the most effective in removing mixed noise.

2) The real HS image data experiments show that \( S_3 \)-TVV has high performance even when observed HS images are degraded by real noise.

V. Conclusion

In this paper, we have proposed a new regularization method, named \( S_3 \)-TVV, for denoising and destriping of HS images. \( S_3 \)-TVV is defined as the sum of the nuclear norms of matrices consisting of second-order spatio-spectral differences in small spectral blocks, which fully captures the spatial piecewise-smoothness, the spatial similarity between adjacent bands, and the spectral correlation across all bands of HS images. We have formulated the denoising and destriping problem as a constrained convex optimization problem including \( S_3 \)-TVV, and developed the optimization algorithm based on P-PDS. Experiments on HS images with simulated or real noise have demonstrated the superiority of \( S_3 \)-TVV over existing methods. \( S_3 \)-TVV will have a strong impact on the field of remote sensing, including applications using HS images taken in highly degraded measurement environments.

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, pp. 2659–2670, 2020.
[5] H. Shen, X. Li, Q. Cheng, C. Zeng, G. Yang, H. Li, and L. Zhang, “Missing information reconstruction of remote sensing data: A technical review,” IEEE Geosci. Remote Sens. Mag., vol. 3, no. 3, pp. 61–85, 2015.
[6] 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.
[7] 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.
[8] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, “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.
[9] W. Ma, J. M. Bioucas-Dias, T. Chan, N. Gillis, P. Gader, A. J. Plaza, A. Ambikapathi, and C. Chi, “A signal processing perspective on hyperspectral unmixing: Insights from remote sensing,” IEEE Signal Process. Mag., vol. 31, no. 1, pp. 67–81, 2014.
[10] 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.
[11] S. Li, W. Song, L. Fang, Y. Chen, P. Ghaniessi, and J. A. Benediktsson, “Deep learning for hyperspectral image classification: An overview,” IEEE Geosci. Remote Sens. Mag., vol. 7, no. 2, pp. 159–173, 2019.
[12] 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.
[13] H. Su, Z. Wu, H. Zhang, and Q. Du, “Hyperspectral anomaly detection: A survey,” IEEE Geosci. Remote Sens. Mag., vol. 10, no. 1, pp. 64–90, 2022.
[14] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” Physica D: Nonlinear Phenomena, vol. 60, no. 1, pp. 259–268, 1992.
[15] X. Bresson and T. F. Chan, “Fast dual minimization of the vectorial total variation norm and applications to color image processing,” Inverse problems imag., vol. 2, no. 4, pp. 455–484, 2008.
[16] Y. Chang, L. Zhang, and H. Shen, “Hyperspectral image denoising employing a spectral–spatial adaptive total variation model,” IEEE Trans. Geosci. Remote Sens., vol. 50, no. 10, pp. 3660–3677, 2012.
[17] Y. Chang, L. Yan, H. Fang, and C. Luo, “Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping,” IEEE Trans. Image Process., vol. 24, no. 6, pp. 1852–1866, 2015.
[18] W. He, H. Zhang, L. Zhang, and H. Shen, “Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration,” IEEE Trans. Geosci. Remote Sens., vol. 54, no. 1, pp. 178–188, 2016.
[19] 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.
[20] Y. Chen, W. Cao, L. Pang, J. Peng, and X. Cao, “Hyperspectral image denoising via texture-preserved total variation regularizer,” IEEE Trans. Geosci. Remote Sens., vol. 61, no. 1–4, 2013.
[21] H. K. Aggarwal and A. Majumdar, “Hyperspectral image denoising using spatio-spectral total variation,” IEEE Trans. Geosci. Remote Sens., vol. 13, no. 3, pp. 442–446, 2016.
[22] 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.
[23] T. Ince, “Hyperspectral image denoising using group low-rank and spatial-spectral total variation,” IEEE Access, vol. 7, pp. 52095–52109, 2019.
[24] 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.
[25] M. Wang, Q. Wang, J. Chanussot, and D. Hong, “\( l_{0-1} \) 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.
[26] 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.
[27] S. Ono, “\( l_0 \) gradient projection,” IEEE Trans. Image Process., vol. 26, no. 4, pp. 1554–1564, 2017.
[29] S. Lefkimmiatis, A. Roussos, P. Maragos, and M. Unser, “Structure tensor total variation,” *SIAM J. Imag. Sci.*, vol. 8, no. 2, pp. 1090–1122, 2015.

[30] Z. Wu, Q. Wang, J. Jin, and Y. Shen, “Structure tensor total variation-regularized weighted nuclear norm minimization for hyperspectral image mixed denoising,” *Signal Process.*, vol. 131, pp. 202–219, 2017.

[31] S. Ono, K. Shirai, and M. Okuda, “Vectorial total variation based on arranged structure tensor for multichannel image restoration,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2016, pp. 4528–4532.

[32] R. Kurihara, S. Ono, K. Shirai, and M. Okuda, “Hyperspectral image restoration based on spatio-spectral structure tensor regularization,” in *Proc. Eur Signal Process. Conf. (EUSIPCO)*, 2017, pp. 488–492.

[33] 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.

[34] 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.

[35] 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.

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

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

[38] K. Naganuma and S. Ono, “A general destriping framework for remote sensing images using flatness constraint,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022.

[39] 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.

[40] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[41] 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.

[42] 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.

[43] 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.

[44] 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.

[45] P. L. Combettes and N. N. Reyes, “Moreau’s decomposition in banach spaces,” *Math. Program.*, vol. 139, no. 1, pp. 103–114, 2013.

[46] L. Condat, “Fast projection onto the simplex and the $\ell_1$ ball,” *Math. Program.*, vol. 158, no. 1, pp. 575–585, 2016.

[47] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

Shingo Takemoto (S’22) received a B.E. degree and M.E. degree in Information and Computer Sciences in 2021 from Sophia University and in 2023 from Tokyo Institute of Technology, respectively. He is currently pursuing an Ph.D. degree at the Department of Computer Science in the Tokyo Institute of Technology. His current research interests are in signal and image processing and optimization theory.

Mr. Takemoto received the Student Award from IEEE SPS Tokyo Joint Chapter in 2022.