

l_0-l_1 Hybrid Total Variation Regularization and Its Applications on Hyperspectral Image Mixed Noise Removal and Compressed Sensing

Minghua Wang, *Student Member, IEEE* Qiang Wang, *Member, IEEE* Jocelyn Chanussot, *Fellow, IEEE*, and Danfeng Hong, *Member, IEEE*

The total variation (TV) regularization has been widely used in various applications related to hyperspectral (HS) signal and image processing, owing to its potential in modeling the underlying smoothness of HS data. However, most existing TV norms usually tend to generate spatial over-smoothing or artifacts. To this end, we propose a novel l_0-l_1 hybrid TV (l_0-l_1 HTV) regularization with the applications to HS mixed noise removal and compressed sensing. More specifically, l_0-l_1 HTV can be regarded as a globally and locally integrated TV regularizer, where the l_0 gradient constraint is incorporate into the l_1 spatial-spectral TV (l_1 -SSTV). l_1 -SSTV is capable of exploiting the local structure information across both spatial and spectral domains, while the l_0 gradient can promote a globally spectral-spatial smoothness by directly controlling the number of non-zero gradients of HS images. This efficient combination considers more comprehensive prior knowledge of HS images, yielding sharper edge preservation and resolving the above drawbacks of existing pure TV norms. More significantly, l_0-l_1 HTV can be easily injected into HS-related processing models and an effective algorithm based on the alternating direction method of multipliers (ADMM) is developed to solve the optimization problems. Extensive experiments conducted on several HS datasets substantiate the superiority and effectiveness of the proposed method in comparison with many state-of-the-art methods.

keywords— l_0-l_1 hybrid total variation, l_0 gradient, l_1 SSTV, hyperspectral image denoising, compressed sensing, ADMM.

I. INTRODUCTION

HYPERSPECTRAL imagery acquires hundreds of spectral bands ranging from ultraviolet to infrared wavelength for the same scene on the Earth surface. Compared with other traditional imagery techniques, hyperspectral image (HSI) provides more abundant spatial and spectral information and thus has its wide range of applications in agriculture, military, environment monitoring and food industry [1]–[3].

Manuscript received June 13, 2020; revised July 16 and October 21, 2020; accepted January 11, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61876054, in part by the Delegation Generale de l'Armement (Project ANR-DGA APHYPIS) under Grant ANR-16 ASTR-0027-01, and in part by the support provided by China Scholarship Council (CSC). (Corresponding author: Minghua Wang.)

M. Wang is with Department of Control Science and Engineering, Harbin Institute of Technology, Harbin 150001, PR China (Email: minghuawang1993@163.com).

Q. Wang is with Department of Control Science and Engineering, Harbin Institute of Technology, Harbin 150001, PR China (Email: wangqiang@hit.edu.cn).

J. Chanussot is with Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, Grenoble, 38000, France and Chinese Academy of Sciences, Aerospace Information research Institute, Beijing, PR China (Email: jocelyn.chanussot@grenoble-inp.fr).

D. Hong is with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Wessling, Germany, and also with the Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France. (e-mail: danfeng.hong@dlr.de)

However, due to various factors of imaging system and environment, the observed HSIs are unavoidably degraded by noises in real-world scenes. Moreover, in some high-resolution or long-distance transmission hyperspectral imaging systems, the increasing amount of 3-D hyperspectral data causes the high cost on imaging, storage or transmission. HSI compressed sensing (CS) is designed to reduce the computation of imaging and the final recovered HSI images are estimated from a few measurements. As essential preprocessing steps, these HSI restorations have direct influence on the performances of further HSI image processing including classification [4]–[6], unmixing [7], [8], super resolution [9], [10], feature learning [11], [12] and target detection [13]–[15]. Therefore, it is critical to design effective techniques for HSI restoration.

In general, the optimization problem of HSI restoration methods consists of data fidelity and regularization terms. The former term is derived from a forward model of the object observation. The regularization term can be interpreted as the prior information from underlying properties on recovered hyperspectral images. One of prevalent regularization techniques for image restoration is based on total variation (TV). The TV regularization was introduced to remove noise and preserve image edges for natural image denoising [16]–[18] and CS reconstruction [19]–[21]. Yuan *et al.* [22] explored local spatial prior and developed a spectral-spatial adaptive TV (SSAHTV) model for HSI denoising. But SSAHTV just reduces Gaussian noise and loses useful image detail information. Li *et al.* [23] minimized the total variation of the abundance fractions subject for HSI CS and unmixing. Kuiteing *et al.* [24] proposed an iterative TV method for HSI CS. However, these TV norms strongly enhance spatial piecewise smoothness, which ignores HSI spectral correlation and leads to spatial over-smoothing.

To overcome the drawbacks of the above TV norms, spectral correlation of 3-D HSIs is taken into account and many new TV norms are proposed. Ono *et al.* [25] proposed an Arranged Structure tensor Total Variation (ASTV) for multichannel image restoration. But this regularization based on the singular value decomposition of the low-rank matrix spends much computational time. Chang *et al.* [26] proposed an anisotropic spectral-spatial total variation (ASSTV) regularization to promote the smoothness across both the spectral and the spatial differences. Unfortunately, ASSTV causes spectral smoothing due to its strong suppression of direct spectral difference. Aggarwal and Majumdar [27] combined 2-D TV along the spatial dimension and 1-D TV along the spectral dimension and proposed a spatio-spectral total variation (SSTV) model. Although SSTV can remove several types of noises, it produces obvious artifacts in the final recovered HSI images.

Takeyama *et al.* [28] proposed a hybrid spatio-spectral total variation (HSSTV) which is designed to fuse two kinds of local differences with a balancing weight. Specifically, HSSTV1 and HSSTV2 utilize local l_1 spatial norm and l_2 spatial norm, respectively. However, authors need to analyze and determine an extra parameter, and unfortunately lose the consideration of HSI global spectral information.

By taking advantage of another HSI prior: low-rank (LR), TV regularization is merged with low-rank models. LR representation reveals that the spectral bands of HSI lie in the same low-rank space and the HSI data can be represented by a linear combination of finite endmembers. Zhang *et al.* [29] rearranged a 3-D HSI data as a 2-D matrix and proposed a low-rank matrix approximation (LRMR) method. Subsequently, He *et al.* [30] proposed a band-by-band total variation regularized low-rank Matrix Factorization (LRTV) method and Wu *et al.* [31] fused the band-by-band TV regularization into the weighted nuclear norm minimization (WNMM). Golbabaei *et al.* [32] proposed a convex function that penalizes both the trace norm and the band-by-band TV norm of the low-rank HSI matrix for HSI CS. However, this band-by-band TV regularization handles each HSI band separately and thus yields unsatisfying spectral distortions. Wu *et al.* [33] incorporated the structure tensor total variation (STV) into the above WNNM to proposed a STWNNM model. To exploit spatio-spectral total variation (SSTV), Wang *et al.* [34] proposed a novel low-rank constraint and SSTV regularization (LSSTV) model and Wang *et al.* [35] integrated this kind of SSTV into multi-directional low-rank frameworks. Chen *et al.* [36] proposed a weighted group sparsity regularized $l_{2,1}$ norm to exploit the shared sparse pattern of the difference image for different bands, and combined this regularization into the well-known LR Tucker decomposition. By using a 3-D anisotropic total variation (3DATV), He *et al.* [37] combined it with the local low-rank matrix and proposed an LLRSSTV method for HSI mixed noise removal. Wang *et al.* [38] merged this 3DATV into a low-rank tensor Tucker decomposition model. Peng *et al.* [39] calculated the sparsity on the subspace of gradient maps and proposed an Enhanced 3-D Total Variation (E-3DTV) regularization for HSI denoising and CS. Wang *et al.* [40] propose a novel tensor-based approach named JTenRe3-DTV for HSI CS, by jointing the tensor Tucker decomposition and 3-D total variation. Nevertheless, all the above TVs lean upon l_1 norms and penalize large image gradient magnitudes. In real cases, HSI edges can be adversely affected.

To effectively sharpen image edges, Xu *et al.* [41] firstly proposed a new l_0 gradient minimization for natural image smoothing. This minimization globally controlling how many non-zero gradients of results to recover prominent structure, has been used in image denoising [41], image deblurring [42] and image inpainting [43]. Xu *et al.* [42] proposed an effective method based on a generalized l_0 sparse expression for motion deblurring. Xue *et al.* [43] extended the above l_0 gradient into a low gradient and converted this regularization into an LR framework. We [44] developed the l_0 gradient to explore smoothness across both the spatial and the spectral domains and combined it with the LR tensor Tucker decomposition

model. Xiong *et al.* [45] embedded the l_0 gradient regularization into a low-rank tensor factorization. However, these LR based methods are computationally expensive since they require singular value decomposition. Utilizing either spatial smoothness or spatial-spectral smoothness will inevitably cause over-smoothing or artifacts, respectively.

To consider more prior HSI structure information and introduce less computational burden, we combine spatial-spectral TV and l_0 gradient regularization. The fore-mentioned l_0 gradient has a fatal flaw: a user-give parameter penalizing the degree of smoothness has no physical meaning. Therefore, Ono [46] proposed a novel l_0 gradient projection employing a parameter with a clear meaning: the l_0 gradient value of the output image, to control the degree of smoothness. This l_0 gradient projection variant, l_0 hyperspectral TV (l_0 hTV) is extended for HSI restoration and is effective for extracting image edges information. But it is difficult for l_0 hyperspectral TV to distinguish outliers like sparse noise when these outliers locate on HSI image edges. Base on the above-mentioned discussion, we proposed a l_0 - l_1 Hybrid Total Variation (l_0 - l_1 HTV) technique for HSI mixed noise removal and compressed sensing. Compared with existing techniques, the main contributions of this study are summarized as follows:

- (1) We extend the l_0 gradient minimization into HSI restoration, named as l_0 hTV, for directly controlling the number of non-zero image gradients and establishing a physical relationship with the l_0 gradient value of the recovered result. l_0 hTV characterizes the global spectral-spatial smoothness to complement image details like sharp edges and valuable texture.
- (2) To takes full consideration of the local spatial-spectral structure and the global spectral-spatial information of HSI, we design a globally and locally integrated TV regularizer, where l_0 hTV is added on SSTV. SSTV captures the local information across spatial and spectral domains to recover the overall structure of HSIs. This TV regularizer compensates mutually to conquer the shortcomings of artifacts or over-smoothing.
- (3) We integrate SSTV and l_0 hTV of HSI into a consolidated framework of l_0 - l_1 HTV, which can be formulated for two HSI tasks: HSI noise removal and compressed sensing. The Alternating Direction Method of Multipliers (ADMM) is adapted to solve the optimization problems. Experimental results with both simulated and real data demonstrate the superiority of l_0 - l_1 HTV. In particular, the classification accuracy of the results by l_0 - l_1 HTV is improved significantly.

The remainder of this paper is organized as follows. Some notations and preliminaries of tensors are introduced in Sec.II. Sec.III presents the study of l_0 gradient minimization, as well as the new l_0 - l_1 HTV. Then, Sec.IV and Sec.V provides HSI denoising and compressed sensing models by involving the l_0 - l_1 HTV regularization, respectively. The experimental results and analysis are reported in Sec.VI. Finally, the conclusions are drawn in Sec.VII

II. NOTATIONS AND PRELIMINARIES

In this paper, scalars, vectors, matrices and tensors are denoted by lowercase letters, e.g. x , boldface lowercase letters, e.g. \mathbf{x} , boldface capital letters, e.g. \mathbf{X} , and Euler script letters, e.g., \mathcal{X} respectively. For a three-way hyperspectral image tensor $\mathcal{X} \in \mathbb{R}^{h \times v \times z}$, we denote its (h_i, v_i, z_i) -element as $\mathcal{X}_{h_i, v_i, z_i}$ and use notation $\mathcal{X}(i, :, :)$, $\mathcal{X}(:, i, :)$ and $\mathcal{X}(:, :, i)$ to represent the i^{th} horizontal, lateral and frontal slices, respectively. The l_1 norm is calculated as $\|\mathcal{X}\|_1 = \sum_{h_i, v_i, z_i} |\mathcal{X}_{h_i, v_i, z_i}|$ and the Frobenius norm is written as $\|\mathcal{X}\|_F = \sqrt{\sum_{h_i, v_i, z_i} |\mathcal{X}_{h_i, v_i, z_i}|^2}$. The l_1 norm of a vector denotes $\|\mathbf{x}\|_1 = \sum_i |\mathbf{x}_i|$ and the l_2 norm of a vector is $\|\mathbf{x}\|_2 = \sqrt{\sum_i |\mathbf{x}_i|^2}$. A matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$ can be vectorized as $\mathbf{x} = \text{vec}(\mathbf{X}) \in \mathbb{R}^t$, where $t = m * n$ and $\text{vec}(\mathbf{X})$ stacks the columns of \mathbf{X} .

Definition 1 (Mixed $l_{1,0}$ pseudo-norm [46]). Let \mathbf{y} be a vector of \mathbb{R}^m , and index sets $\theta_1, \dots, \theta_i, \dots, \theta_n (1 \leq n \leq m)$ satisfies

- Each θ_i is a subset of $1, \dots, m$,
- $\theta_i \cap \theta_l = \emptyset$ for any $i \neq l$,
- $\cup_{i=1}^n \theta_i = 1, \dots, m$,

the mixed $l_{1,0}$ pseudo-norm of y is defined as:

$$\|\mathbf{y}\|_{1,0}^\theta = \|(\|\mathbf{y}_{\theta_1}\|_1, \dots, \|\mathbf{y}_{\theta_i}\|_1, \dots, \|\mathbf{y}_{\theta_n}\|_1)\|_0, \quad (1)$$

where \mathbf{y}_{θ_i} denotes a sub-vector of \mathbf{y} with its entries specified by θ_i .

Definition 2 (Indicator function $I(\mathbf{y})$ of $l_{1,0}$ mixed pseudo-norm [46])

$$I_{\|\mathcal{B}\mathbf{y}\|_{1,0}^\theta}(y) = \begin{cases} 0, & \|\mathcal{B}\mathbf{y}\|_{1,0}^\theta \leq \gamma, \\ \infty, & \text{otherwise,} \end{cases} \quad (2)$$

where \mathcal{B} is a given operator. In this paper, \mathcal{B} is a diagonal tensor with 0 or 1 entries.

Proposition 1 (Projection onto $l_{1,0}$ mixed pseudo-norm ball with binary mask [46]): Set $\mathbf{y} \in \mathbb{R}^m$ as a known vector, set γ as a non-negative integer. Let \mathbf{W} be a known diagonal binary matrix, and let $\theta_1, \dots, \theta_i, \dots, \theta_n (1 \leq n \leq m)$ be index sets satisfying the conditions from **definition 1**. Without loss of generality, $\mathbf{W}\mathbf{y} = (\mathbf{y}_{\theta_1}^T \dots \mathbf{y}_{\theta_n}^T)^T$ are assumed. $\mathbf{y}_{\theta_1} \dots \mathbf{y}_{\theta_n}$ are sorted in descending order according to $\|\mathbf{y}_{\theta_1}\|_2 > \|\mathbf{y}_{\theta_2}\|_2 > \dots > \|\mathbf{y}_{\theta_n}\|_2$, where $\mathbf{y}_{\theta_1} \dots \mathbf{y}_{\theta_n}$ are obtained with the new order, and the original index sets i have \overrightarrow{i} that corresponds to it one by one.

For the following problem

$$\min_{\mathbf{y}^*} \|\mathbf{y}^* - \mathbf{y}\|^2 \text{ s.t. } \|\mathbf{W}\mathbf{y}^*\|_{1,0}^\theta \leq \gamma, \quad (3)$$

one of the optimal solution is given by:

$$\mathbf{y}^* = \begin{cases} \mathbf{y}, & \|\mathbf{W}\mathbf{y}\|_{1,0}^\theta \leq \gamma, \\ (\tilde{\mathbf{y}}_{\theta_1}^T \dots \tilde{\mathbf{y}}_{\theta_n}^T)^T + (\mathbf{I} - \mathbf{W})\mathbf{y}, & \|\mathbf{W}\mathbf{y}\|_{1,0}^\theta > \gamma, \end{cases} \quad (4)$$

where

$$\tilde{\mathbf{y}}_{\theta_i}^T = \begin{cases} \mathbf{y}_{\theta_i}^T, & \text{if } i \in \{1, \dots, \gamma\}, \\ 0, & \text{if } i \in \{\gamma + 1, \dots, n\}. \end{cases} \quad (5)$$

Definition 3 (Mode- k_1, k_2 unfolding of a tensor [47]): For a m -way tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_m}$, this operator is denoted by $\mathcal{X}_{(k_1, k_2)} \in \mathbb{R}^{n_{k_1} \times k_2 \times \prod_{i \neq k_1, k_2} n_i}$, whose frontal slices are the lexicographic ordering of the mode- k_1, k_2 slice of \mathcal{X} .

III. l_0 - l_1 HYBRID TOTAL VARIATION

A. l_0 Hyperspectral TV

Xu et al. [42] first proposed the l_0 gradient model to preserve natural image edges. For a natural 2-D image $\mathbf{X} \in \mathbb{R}^{m \times n}$, the l_0 Gradient is defined as [42]:

$$\begin{aligned} l_0\text{TV}(\mathbf{X}) &= \sum_{i=1}^{m+n} C(\|\text{vec}(D_h \mathbf{X})\|_1 + \|\text{vec}(D_v \mathbf{X})\|_1) \\ &= \sum_i^n C(|\mathbf{X}_{i+1,j} - \mathbf{X}_{i,j}| + |\mathbf{X}_{i,j+1} - \mathbf{X}_{i,j}|) \end{aligned} \quad (6)$$

where D_h and D_v are linear difference operators along image horizontal and vertical directions. $C(\mathbf{X})$ is a binary function counting the number of non-zero image gradients and $C(\mathbf{X}_{i,j}) = 1$, if $\mathbf{X}_{i,j} \neq 0$; $C(\mathbf{X}_{i,j}) = 0$, otherwise.

For a HSI $\mathcal{X} \in \mathbb{R}^{h \times v \times z}$, $\sum_{k=1}^z l_0\text{TV}(\mathcal{X}^k)$ can be a band-by-band model by the simplest way to extend. Although the spatial smoothness of each HSI band is utilized separately, HSI spectral correlation is neglected. Simultaneously, different HSI bands are acquired from the same scene, the spatial information of all spectral bands should be fused and formulated. $l_0\text{hTV}$ is defined as

$$l_0\text{hTV}(\mathcal{X}) = \sum_i^h \sum_j^v \sum_k^z C(\sum_k (|\mathcal{X}_{i+1,j,k} - \mathcal{X}_{i,j,k}| + |\mathcal{X}_{i,j+1,k} - \mathcal{X}_{i,j,k}|)), \quad (7)$$

where boundary values of gradients are written as: $\mathcal{X}_{i,j+1,k} - \mathcal{X}_{i,j,k} = 0$, if $j = v$, and $\mathcal{X}_{i+1,j,k} - \mathcal{X}_{i,j,k} = 0$, if $i = h$. $l_0\text{hTV}(\mathcal{X})$ actually counting the non-zero gradients in the spatial dimension with the assistance of all spectral bands greatly benefits for capturing the spatial edges of the hyperspectral image.

The mixed $l_{1,0}$ pseudo-norm defined by **Definition 1**, another formulation is adopted to express $l_0\text{hTV}$:

$$l_0\text{hTV}(\mathcal{X}) = \|BD\mathcal{X}\|_{1,0}^\theta, \quad (8)$$

where operator D is an operator to fuse the differences along the horizontal and vertical directions into one tensor. Operator B is an operator that forces boundary values of gradients to be zero when $i = h$ and $j = v$. A simple example is given as follows. $D_h\mathcal{X}$ and $D_v\mathcal{X}$ denote the difference along image horizontal and vertical directions, respectively. $D_h\mathcal{X}$ and $D_v\mathcal{X}$ are stored into one tensor $D\mathcal{X}$, i.e., $D\mathcal{X}(:, :, 1, :) = D_h\mathcal{X}$ and $D\mathcal{X}(:, :, 2, :) = D_v\mathcal{X}$. The boundary values of $BD\mathcal{X}(:, :, 1, 1)$, $BD\mathcal{X}(:, :, 1, 2)$ and $BD\mathcal{X}(:, :, 2, 1)$, $BD\mathcal{X}(:, :, 2, 2)$ are set as 0 when $i = h = 3$ and $j = v = 3$, respectively.

Example 1: A tensor $\mathcal{X} \in \mathbb{R}^{3 \times 3 \times 2}$ and its $BD\mathcal{X} \in \mathbb{R}^{3 \times 3 \times 2 \times 2}$ can be expressed as:

$$\mathcal{X}(:,:,1) = \begin{bmatrix} \mathcal{X}_{111} & \mathcal{X}_{121} & \mathcal{X}_{131} \\ \mathcal{X}_{211} & \mathcal{X}_{221} & \mathcal{X}_{231} \\ \mathcal{X}_{311} & \mathcal{X}_{321} & \mathcal{X}_{331} \end{bmatrix} \quad (9)$$

$$\mathcal{X}(:,:,2) = \begin{bmatrix} \mathcal{X}_{112} & \mathcal{X}_{122} & \mathcal{X}_{132} \\ \mathcal{X}_{212} & \mathcal{X}_{222} & \mathcal{X}_{232} \\ \mathcal{X}_{312} & \mathcal{X}_{322} & \mathcal{X}_{332} \end{bmatrix} \quad (10)$$

$$(BD\mathcal{X})((:,:,1,1)) = \begin{bmatrix} \mathcal{X}_{121} - \mathcal{X}_{111} & \mathcal{X}_{131} - \mathcal{X}_{121} & 0 \\ \mathcal{X}_{221} - \mathcal{X}_{211} & \mathcal{X}_{231} - \mathcal{X}_{221} & 0 \\ \mathcal{X}_{321} - \mathcal{X}_{311} & \mathcal{X}_{331} - \mathcal{X}_{321} & 0 \end{bmatrix} \quad (11)$$

$$(BD\mathcal{X})((:,:,1,2)) = \begin{bmatrix} \mathcal{X}_{122} - \mathcal{X}_{112} & \mathcal{X}_{132} - \mathcal{X}_{122} & 0 \\ \mathcal{X}_{222} - \mathcal{X}_{212} & \mathcal{X}_{232} - \mathcal{X}_{222} & 0 \\ \mathcal{X}_{322} - \mathcal{X}_{312} & \mathcal{X}_{332} - \mathcal{X}_{322} & 0 \end{bmatrix} \quad (12)$$

$$(BD\mathcal{X})((:,:,2,1)) = \begin{bmatrix} \mathcal{X}_{211} - \mathcal{X}_{111} & \mathcal{X}_{221} - \mathcal{X}_{121} & \mathcal{X}_{231} - \mathcal{X}_{131} \\ \mathcal{X}_{311} - \mathcal{X}_{211} & \mathcal{X}_{321} - \mathcal{X}_{221} & \mathcal{X}_{331} - \mathcal{X}_{231} \\ 0 & 0 & 0 \end{bmatrix} \quad (13)$$

$$(BD\mathcal{X})((:,:,2,2)) = \begin{bmatrix} \mathcal{X}_{212} - \mathcal{X}_{112} & \mathcal{X}_{222} - \mathcal{X}_{122} & \mathcal{X}_{232} - \mathcal{X}_{132} \\ \mathcal{X}_{312} - \mathcal{X}_{212} & \mathcal{X}_{322} - \mathcal{X}_{222} & \mathcal{X}_{332} - \mathcal{X}_{232} \\ 0 & 0 & 0 \end{bmatrix} \quad (14)$$

With **Definition 1**, here $n = 9$ and each \mathbf{y}_{θ_i} is determined by mode-1,2 unfolding $(BD\mathcal{X})_{(1,2)} \in \mathbb{R}^{9 \times 2 \times 2}$, for example, $\mathbf{y}_{\theta_1} = [\mathcal{X}_{121} - \mathcal{X}_{111}, \mathcal{X}_{122} - \mathcal{X}_{112}, \mathcal{X}_{211} - \mathcal{X}_{111}, \mathcal{X}_{212} - \mathcal{X}_{112}]^T$. It is easily proven that $l_0\text{hTV}(\mathcal{X})$ and $\|BD\mathcal{X}\|_{1,0}^\theta$ are mathematically equivalent.

B. l_0-l_1 Hybrid Total Variation Regularization

Fig. 1 represents the visual illustration of the l_0-l_1 HTV regularization. As shown at the top of Fig. 1, the global spatial smoothness of all spectral bands is restrained by $l_0\text{hTV}(\mathcal{X}) = \|BD\mathcal{X}\|_{1,0}^\theta$, where γ denotes the l_0 gradient value of the output image. As the number of iteration increases, the image edges of HSI gradually becomes clearer. Meanwhile, if l_0 norm is excessively penalized, the final results will appear over-smoothing phenomenon. At the bottom of Fig. 1, the SSTV handling local spatial-spectral differences combines 2-D spatial TV (defined by $2\text{DTV} = \|D_h\mathcal{X}\|_1 + \|D_v\mathcal{X}\|_1$) with 1-D spectral TV. Although most of structure information is restored, SSTV unavoidably caused artifacts. For more comprehensively exploiting the global spatial and the local spatial-spectral information, $l_0\text{hTV}(\mathcal{X})$ and SSTV are incorporated as a globally and locally integrated regularizer and, with **Definition 2**, the proposed l_0-l_1 Hybrid Total Variation(l_0-l_1 HTV) is defined as:

$$\begin{aligned} l_0-l_1\text{HTV}(\mathcal{X}) &= l_0\text{hTV}(\mathcal{X}) + \text{SSTV}(\mathcal{X}), \\ &= I_{\|B\cdot\|_{1,0}^\theta}(BD\mathcal{X}) + (\|D_h\mathcal{X}D_z\|_1 + \|D_v\mathcal{X}D_z\|_1), \end{aligned} \quad (15)$$

where D_h and D_v are two operators to calculate differences along the horizontal and vertical directions. D_z is a 1-D finite differencing operator to extract the spectral signature of each pixel. In practical applications, it is hard to find a

proper parameter λ for minimizing $\lambda l_0\text{hTV}(\mathcal{X})$ since this user-given parameter λ has no relationship with the degree of the restored image \mathcal{X} . Therefore, $\|BD\mathcal{X}\|_{1,0}^\theta \leq \gamma$ is used for HSI restoration and γ directly represents the l_0 gradient value of the output image.

IV. l_0-l_1 HYBRID TOTAL VARIATION FOR HSI DENOISING

A clear HSI cube $\mathcal{X} \in \mathbb{R}^{h \times v \times z}$ is corrupted by mixed noise including Gaussian noise, salt and pepper noise, deadlines, and strips. Gaussian noise can be denoted by \mathcal{N} and the others are unified as sparse noise \mathcal{S} . Thus an observed 3D hyperspectral image $\mathcal{T} \in \mathbb{R}^{h \times v \times z}$ is represented as:

$$\mathcal{T} = \mathcal{X} + \mathcal{S} + \mathcal{N}. \quad (16)$$

With Eq. (15) in mind, the proposed l_0-l_1 HTV for removing HSI mixed noise is formulated as:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{S}} & \|\mathcal{T} - \mathcal{X} - \mathcal{S}\|_F^2 + \lambda \|\mathcal{S}\|_1 + \lambda_{tv} (\|D_h\mathcal{X}D_z\|_1 \\ & + \|D_v\mathcal{X}D_z\|_1) + I_{\|B\cdot\|_{1,0}^\theta}(BD\mathcal{X}). \end{aligned} \quad (17)$$

where λ and λ_{tv} are two positive constants to balance the contributions of each term. The first term $\|\mathcal{T} - \mathcal{X} - \mathcal{S}\|_F^2$ is a data fidelity term. $\|\mathcal{S}\|_1$ is the l_1 norm of the sparse noise \mathcal{S} to remove salt and pepper noise, dead-lines and strips. $\|D_h\mathcal{X}D_z\|_1 + \|D_v\mathcal{X}D_z\|_1$ is used to characterize the local smoothness in the spatial and spectral directions. $I_{\|B\cdot\|_{1,0}^\theta}(BD\mathcal{X})$ is the mixed $l_{1,0}$ pseudo-norm to enhance image edges and remove Gaussian noise.

Algorithm 1 l_0-l_1 HTV for HSI mixed noise removal

Input: $\mathcal{T}, \lambda, \lambda_{tv}, \gamma$

- 1: **Initialization:** $\mathcal{X} = \mathcal{S} = 0, \mathcal{V} = \mathcal{P} = \mathcal{Q} = \mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{Y}_3 = 0, i = 0, \epsilon = 10^{-4} * hv$
- 2: **while** $0 \leq i \leq maxiter$ or $|\|BD\mathcal{X}\|_{1,0}^\theta - \gamma| > \epsilon$
- 3: $i = i + 1$;
- 4: update \mathbf{x} by solving Eq. (26):

$$(1 + \beta D^* D + \beta D_1^* D_1 + \beta D_2^* D_2) \mathbf{x} = \mathbf{t} - \mathbf{s}$$
- 5: $+ \beta(D_1^* \mathbf{p} - D_1^* \mathbf{y}_1 + D_2^* \mathbf{q} - D_2^* \mathbf{y}_2 + D^* \mathbf{v} - D^* \mathbf{y}_3)$
- 6: update \mathcal{X} : $\mathcal{X} = \text{rearrange}(\mathbf{x})$
- 7: update \mathcal{S}, \mathcal{P} and \mathcal{Q} by solving Eq. (30), (31) and (32):

$$\mathcal{S} = S_{\lambda/2}[\mathcal{T} - \mathcal{X}]$$
- 8: $\mathcal{P} = S_{\lambda_{tv}/2}[D_h\mathcal{X}D_z + \mathcal{Y}_1]$
- 9: $\mathcal{Q} = S_{\lambda_{tv}/2}[D_v\mathcal{X}D_z + \mathcal{Y}_2]$
- 10: update \mathcal{V} by solving Eq. (33):

$$\mathcal{V}' = B(D\mathcal{X} + \mathcal{Y}_3),$$
- 11: Sort $\mathcal{V}'_{\theta_1}, \dots, \mathcal{V}'_{\theta_n}$ in descending order in terms of l_2 norms,
- 12: Set $\mathcal{V}'_{\theta_{\gamma+1}} = 0, \dots, \mathcal{V}'_{\theta_n} = 0$ and then obtain a updated \mathcal{V}' ,
- 13: $\mathcal{V} = \mathcal{V}' + (I - B)(D\mathcal{X} + \mathcal{Y}_3)$
- 14: update the multipliers

$$\mathcal{Y}_1 = \mathcal{Y}_1 + (D_h\mathcal{X}D_z - \mathcal{P})$$
- 15: $\mathcal{Y}_2 = \mathcal{Y}_2 + (D_v\mathcal{X}D_z - \mathcal{Q})$
- 16: $\mathcal{Y}_3 = \mathcal{Y}_3 + (D\mathcal{X} - \mathcal{V})$
- 17: **end while**

Output: \mathcal{X} .

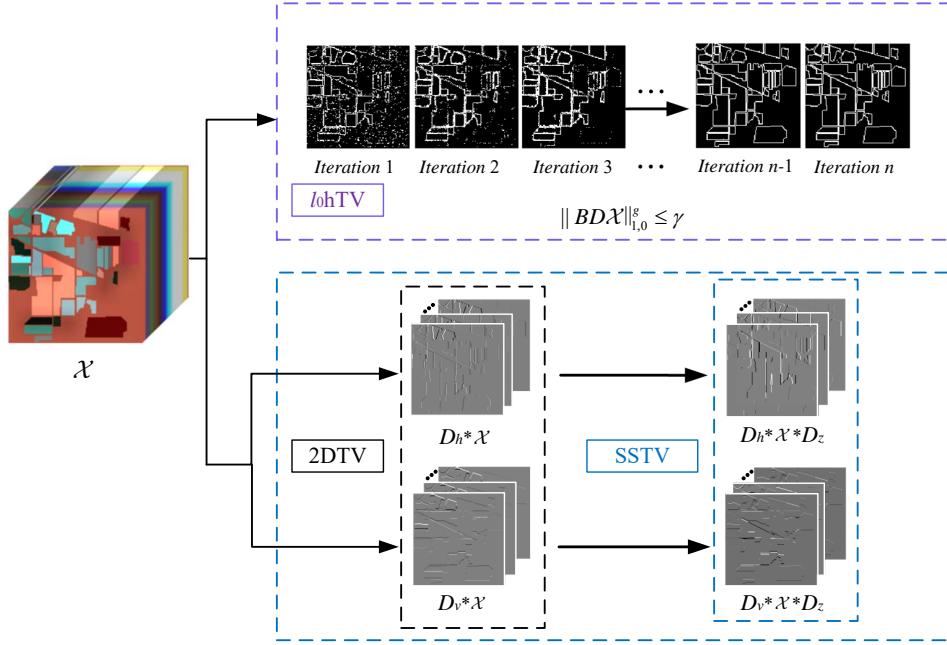


Fig. 1: Illustration of the proposed l_0 - l_1 hybrid total variation regularization.

To solve the above HSI denoising problem, three auxiliary variable \mathcal{V} , \mathcal{P} and \mathcal{Q} are introduced and the Eq. (17) is reformulated as follows:

$$\begin{aligned} & \min_{\mathcal{X}, \mathcal{S}, \mathcal{V}, \mathcal{P}, \mathcal{Q}} \|\mathcal{T} - \mathcal{X} - \mathcal{S}\|_F^2 + \lambda \|\mathcal{S}\|_1 + \lambda_{tv}(\|\mathcal{P}\|_1 + \|\mathcal{Q}\|_1) \\ & \quad + I_{\|\mathcal{B}\cdot\|_{1,0}^\theta}(\mathcal{V}), \\ & \text{s.t. } \mathcal{P} = D_h \mathcal{X} D_z, \mathcal{Q} = D_v \mathcal{X} D_z, \mathcal{V} = D \mathcal{X}. \end{aligned} \quad (18)$$

The ADMM method [48], [49] is applied for the aforementioned problem and the unconstrained function of Eq.(18) is rewritten as:

$$\begin{aligned} & \min_{\mathcal{X}, \mathcal{S}, \mathcal{V}, \mathcal{P}, \mathcal{Q}} \|\mathcal{T} - \mathcal{X} - \mathcal{S}\|_F^2 + \lambda \|\mathcal{S}\|_1 + \lambda_{tv}(\|\mathcal{P}\|_1 + \|\mathcal{Q}\|_1) \\ & + I_{\|\mathcal{B}\cdot\|_{1,0}^\theta}(\mathcal{V}) + \beta \|\mathcal{P} - D_h \mathcal{X} D_z - \mathcal{Y}_1\|_F^2 \\ & + \beta \|\mathcal{Q} - D_v \mathcal{X} D_z - \mathcal{Y}_2\|_F^2 + \beta \|\mathcal{V} - D \mathcal{X} - \mathcal{Y}_3\|_F^2. \end{aligned} \quad (19)$$

Eq. (19) is divided into five sub-problems:

$$\begin{aligned} 1) \quad & \min_{\mathcal{X}} \|\mathcal{T} - \mathcal{X} - \mathcal{S}\|_F^2 + \beta \|\mathcal{P} - D_h \mathcal{X} D_z - \mathcal{Y}_1\|_F^2 \\ & + \beta \|\mathcal{Q} - D_v \mathcal{X} D_z - \mathcal{Y}_2\|_F^2 + \beta \|\mathcal{V} - D \mathcal{X} - \mathcal{Y}_3\|_F^2, \end{aligned} \quad (20)$$

$$2) \quad \min_{\mathcal{S}} \lambda \|\mathcal{S}\|_1 + \|\mathcal{T} - \mathcal{X} - \mathcal{S}\|_F^2, \quad (21)$$

$$3) \quad \min_{\mathcal{P}} \lambda_{tv} \|\mathcal{P}\|_1 + \beta \|\mathcal{P} - D_h \mathcal{X} D_z - \mathcal{Y}_1\|_F^2 \quad (22)$$

$$4) \quad \min_{\mathcal{Q}} \lambda_{tv} \|\mathcal{Q}\|_1 + \beta \|\mathcal{Q} - D_v \mathcal{X} D_z - \mathcal{Y}_2\|_F^2. \quad (23)$$

$$5) \quad \min_{\mathcal{V}} I_{\|\mathcal{B}\cdot\|_{1,0}^\theta}(\mathcal{V}) + \beta \|\mathcal{V} - D \mathcal{X} - \mathcal{Y}_3\|_F^2. \quad (24)$$

The \mathcal{X} -related sub-problem is rewritten into its vector:

$$\begin{aligned} \mathbf{x} &= \arg \min_{\mathbf{x}} f(\mathbf{x}) \\ &= \arg \min_{\mathbf{x}} \|\mathbf{t} - \mathbf{x} - \mathbf{s}\|_F^2 + \beta \|\mathbf{p} - D_1 \mathbf{x} - \mathbf{y}_1\|_F^2 \\ & \quad + \beta \|\mathbf{q} - D_2 \mathbf{x} - \mathbf{y}_2\|_F^2 + \beta \|\mathbf{v} - D \mathbf{x} - \mathbf{y}_3\|_F^2, \end{aligned} \quad (25)$$

where $D_1 = D_h \otimes D_z$, $D_2 = D_v \otimes D_z$, and \otimes denotes the Kronecker product. Set the partial derivative of $f(\mathbf{x})$ as zero, and yield the following linear equation:

$$\begin{aligned} & (1 + \beta D^* D + \beta D_1^* D_1 + \beta D_2^* D_2) \mathbf{x} = \\ & \mathbf{t} - \mathbf{s} + \beta(D_1^* \mathbf{p} - D_1^* \mathbf{y}_1 + D_2^* \mathbf{q} - D_2^* \mathbf{y}_2 + D^* \mathbf{v} - D^* \mathbf{y}_3). \end{aligned} \quad (26)$$

Eq. (26) can be solved by the Preconditioned Conjugate Gradient (PCG) [50] or the algorithm for sparse linear equations and Least Squares (LSQR) [51], [52].

Eq. (21), (22) and (23) about \mathcal{S} , \mathcal{P} and \mathcal{Q} are the same form and respectively rewritten as:

$$\min_{\mathcal{S}} \lambda \|\mathcal{S}\|_1 + \|\mathcal{S} - (\mathcal{T} - \mathcal{X})\|_F^2, \quad (27)$$

$$\min_{\mathcal{P}} \lambda_{tv} \|\mathcal{P}\|_1 + \beta \|\mathcal{P} - (D_h \mathcal{X} D_z + \mathcal{Y}_1)\|_F^2 \quad (28)$$

$$\min_{\mathcal{Q}} \lambda_{tv} \|\mathcal{Q}\|_1 + \beta \|\mathcal{Q} - (D_v \mathcal{X} D_z + \mathcal{Y}_2)\|_F^2, \quad (29)$$

where sign operator $sgn(x)$ is defined as

$$sgn(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$$

By using the operator $S_\varepsilon[x] = sgn(x) \cdot \max(|x| - \varepsilon, 0)$, the solutions are obtained as:

$$\mathcal{S} = S_{\lambda/2}[\mathcal{T} - \mathcal{X}] \quad (30)$$

$$\mathcal{P} = S_{\lambda_{tv}/2}[D_h \mathcal{X} D_z + \mathcal{Y}_1] \quad (31)$$

$$\mathcal{Q} = S_{\lambda_{tv}/2}[D_v \mathcal{X} D_z + \mathcal{Y}_2] \quad (32)$$

The sub-problem of \mathcal{V} is equivalent to the following constrained minimization problem:

$$\begin{aligned} \min_{\mathcal{V}} & \|\mathcal{V} - D\mathcal{X} - \mathcal{Y}_3\|_F^2 \\ \text{s.t. } & \|B\mathcal{V}\|_{1,0}^\theta \leq \gamma, \end{aligned} \quad (33)$$

The parameter β is omitted because it has no impact on the optimization. **Proposition 1** is adopted to solve Eq. (33). Firstly, \mathcal{V}' is initialized to $\mathcal{V}' = B(D\mathcal{X} + \mathcal{Y}_3)$ and at the i -th iteration, $\mathcal{V}'^{(i)}$ is sorted in descending order. $\mathcal{V}'_{\theta_{\gamma+1}} = 0, \dots, \mathcal{V}'_{\theta_n}$ are set as zero vectors, which effectively removes most of noise while preserving sharper image edges. Simultaneously, artifacts of the denoised result are avoided since the information from all spectral bands is used to promote spatial smoothness. The detail of the l_0-l_1 HTV method for HSI mixed noise removal is summarized in Algorithm 1.

V. l_0-l_1 HYBRID TOTAL VARIATION FOR HSI COMPRESSED SENSING

HSI compressed sensing aims at reconstructing a 3-D HSI data \mathcal{X} from compressive measurements $\mathbf{y} \in \mathbb{R}^m$. In real acquisition, disturbed by Gaussian noise \mathbf{n} , the compressive measurements \mathbf{y} can be obtained by:

$$\mathbf{y} = \Psi(\mathcal{X}) + \mathbf{n}. \quad (34)$$

Ψ is instantiated as $\Psi = \mathbf{D} \cdot \mathbf{H} \cdot \mathbf{P}$, where \mathbf{D} is a random down sampling operator, \mathbf{H} is a random permutation matrix, \mathbf{P} is the WalshHadamard transform and the mapping of Ψ is $\mathbb{R}^{h \times v \times z} \rightarrow \mathbb{R}^m$. This comprssed operator has been successfully applied for various CS cases [32], [53], [54]. However, it is an ill-posed inverse problem to recover \mathcal{X} from Eq. (34). As the prior information is efficient to find the optimal solution, l_0-l_1 HTV regularization is adopted and the proposed HSI CS model is formulated as:

$$\begin{aligned} \min_{\mathcal{X}} & \|\mathbf{y} - \Psi(\mathcal{X})\|_F^2 + \lambda_{tv}(\|D_h \mathcal{X} D_z\|_1 + \|D_v \mathcal{X} D_z\|_1) \\ & + I_{\|B\cdot\|_{1,0}^\theta}(BD\mathcal{X}). \end{aligned} \quad (35)$$

By introducing three auxiliary variables \mathcal{P} , \mathcal{Q} and \mathcal{V} , we rewrite the Eq. (35) as follows:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{V}, \mathcal{P}, \mathcal{Q}} & \|\mathbf{y} - \Psi(\mathcal{X})\|_F^2 + \lambda_{tv}(\|\mathcal{P}\|_1 + \|\mathcal{Q}\|_1) + I_{\|B\cdot\|_{1,0}^\theta}(\mathcal{V}), \\ \text{s.t. } & \mathcal{P} = D_h \mathcal{X} D_z, \mathcal{Q} = D_v \mathcal{X} D_z, \mathcal{V} = D\mathcal{X}. \end{aligned} \quad (36)$$

Based on the ADMM method [48], the unconstrained function of Eq.(36) is given as:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{V}, \mathcal{P}, \mathcal{Q}} & \|\mathbf{y} - \Psi(\mathcal{X})\|_F^2 + \lambda_{tv}(\|\mathcal{P}\|_1 + \|\mathcal{Q}\|_1) + I_{\|B\cdot\|_{1,0}^\theta}(\mathcal{V}) \\ & + \beta\|\mathcal{P} - D_h \mathcal{X} D_z - \mathcal{Y}_1\|_F^2 + \beta\|\mathcal{Q} - D_v \mathcal{X} D_z - \mathcal{Y}_2\|_F^2 \\ & + \beta\|\mathcal{V} - D\mathcal{X} - \mathcal{Y}_3\|_F^2. \end{aligned} \quad (37)$$

Eq. (37) is divided into four sub-problems:

$$\begin{aligned} 1) & \min_{\mathcal{X}} \|\mathbf{y} - \Psi(\mathcal{X})\|_F^2 + \beta\|\mathcal{P} - D_h \mathcal{X} D_z - \mathcal{Y}_1\|_F^2 \\ & + \beta\|\mathcal{Q} - D_v \mathcal{X} D_z - \mathcal{Y}_2\|_F^2 + \beta\|\mathcal{V} - D\mathcal{X} - \mathcal{Y}_3\|_F^2, \end{aligned} \quad (38)$$

$$2) \min_{\mathcal{P}} \lambda_{tv} \|\mathcal{P}\|_1 + \beta\|\mathcal{P} - D_h \mathcal{X} D_z - \mathcal{Y}_1\|_F^2 \quad (39)$$

$$3) \min_{\mathcal{Q}} \lambda_{tv} \|\mathcal{Q}\|_1 + \beta\|\mathcal{Q} - D_v \mathcal{X} D_z - \mathcal{Y}_2\|_F^2. \quad (40)$$

$$4) \min_{\mathcal{V}} I_{\|B\cdot\|_{1,0}^\theta}(\mathcal{V}) + \beta\|\mathcal{V} - D\mathcal{X} - \mathcal{Y}_3\|_F^2. \quad (41)$$

Algorithm 2 l_0-l_1 HTV for HSI compressed sensing

Input: $\mathbf{y}, \lambda_{tv}, \gamma$

- 1: **Initializtion:** $\mathcal{X} = 0, \mathcal{V} = \mathcal{P} = \mathcal{Q} = \mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{Y}_3 = 0, i = 0, \epsilon = 10^{-4} * hv$
- 2: **while** $0 \leq i \leq maxiter$ or $|\|BD\mathcal{X}\|_{1,0}^\theta - \gamma| > \epsilon$
- 3: $i = i + 1;$
- 4: update \mathbf{x} by solving Eq. (43):

$$(1 + \beta D^* D + \beta D_1^* D_1 + \beta D_2^* D_2) \mathbf{x} = \mathbf{y} - \Psi(\mathbf{x})$$
- 5: $+ \beta(D_1^* \mathbf{p} - D_1^* \mathbf{y}_1 + D_2^* \mathbf{q} - D_2^* \mathbf{y}_2 + D^* \mathbf{v} - D^* \mathbf{y}_3)$
- 6: update \mathcal{X} : $\mathcal{X} = \text{rearrange}(\mathbf{x})$
- 7: update \mathcal{P} and \mathcal{Q} by solving (44) and (45):

$$\mathcal{P} = S_{\lambda_{tv}/2}[D_h \mathcal{X} D_z + \mathcal{Y}_1]$$

$$\mathcal{Q} = S_{\lambda_{tv}/2}[D_v \mathcal{X} D_z + \mathcal{Y}_2]$$
- 10: update \mathcal{V} by solving Eq. (33):

$$\mathcal{V}' = B(D\mathcal{X} + \mathcal{Y}_3),$$

$$\text{Sort } \mathcal{V}'_{\theta_1}, \dots, \mathcal{V}'_{\theta_n} \text{ in descending order in terms of } l_2 \text{ norms,}$$

$$\text{Set } \mathcal{V}'_{\theta_{\gamma+1}} = 0, \dots, \mathcal{V}'_{\theta_n} = 0 \text{ and then obtain a updated } \mathcal{V}',$$
- 14: $\mathcal{V} = \mathcal{V}' + (I - B)(D\mathcal{X} + \mathcal{Y}_3)$
- 15: update the multipliers

$$\mathcal{Y}_1 = \mathcal{Y}_1 + (D_h \mathcal{X} D_z - \mathcal{P})$$

$$\mathcal{Y}_2 = \mathcal{Y}_2 + (D_v \mathcal{X} D_z - \mathcal{Q})$$

$$\mathcal{Y}_3 = \mathcal{Y}_3 + (D\mathcal{X} - \mathcal{V})$$

16: **end while**

Output: \mathcal{X} .

The \mathcal{X} -related sub-problem is rewritten into its vector:

$$\begin{aligned} \mathbf{x} &= \arg \min_{\mathbf{x}} f(\mathbf{x}) \\ &= \arg \min_{\mathbf{x}} \|\mathbf{y} - \Psi(\mathbf{x})\|_F^2 + \beta\|\mathbf{p} - D_1 \mathbf{x} - \mathbf{y}_1\|_F^2 \\ &+ \beta\|\mathbf{q} - D_2 \mathbf{x} - \mathbf{y}_2\|_F^2 + \beta\|\mathbf{v} - D \mathbf{x} - \mathbf{y}_3\|_F^2, \end{aligned} \quad (42)$$

The partial derivative of $f(\mathbf{x})$ is set as zero, and optimizing \mathbf{x} is equivalent to solve the following linear equation:

$$\begin{aligned} & (1 + \beta D^* D + \beta D_1^* D_1 + \beta D_2^* D_2) \mathbf{x} = \\ & \mathbf{y} - \Psi(\mathbf{x}) + \beta(D_1^* \mathbf{p} - D_1^* \mathbf{y}_1 + D_2^* \mathbf{q} - D_2^* \mathbf{y}_2 + D^* \mathbf{v} - D^* \mathbf{y}_3). \end{aligned} \quad (43)$$

Eq. (43) is also solved by the Preconditioned Conjugate Gradient (PCG) [50] or the algorithm for sparse linear equations and Least Squares (LSQR) [52]. Based on the operator $S_\varepsilon[x] = \text{sgn}(x) * \max(|x| - \varepsilon, 0)$, the solutions of (39) and

TABLE I: The added mixed noises in nine cases.

Noise Case	Mixed Noise
Case 1	Gaussian (SNR = 10 ~ 20 dB)
Case 2	Gaussian (SNR = 10 ~ 20 dB)+salt and pepper(10%) in all bands
Case 3	Gaussian (SNR = 10 ~ 20 dB)+salt and pepper(20%) in all bands + four vertical and horizontal deadlines in each of four random bands
Case 4	Gaussian (mean SNR = 12.12 dB, and noise variance is calculated by the Eq.(50))
Case 5	Gaussian (mean SNR = 12.12 dB, and noise variance is calculated by the Eq.(50)) +salt and pepper(10%) in all bands
Case 6	Gaussian (mean SNR = 12.12 dB, and noise variance is calculated by the Eq.(50))+salt and pepper(20%) in all bands + eight vertical and horizontal deadlines in each of eight random bands
Case 7	Gaussian (zero-mean, and noise variance is different for different bands and is selected from 0.1~0.2 randomly)
Case 8	Gaussian (zero-mean, and noise variance is different for different bands and is selected from 0.1~0.2 randomly) +salt and pepper(selected from 0 to 20% randomly) in all bands
Case 9	Gaussian (zero-mean, and noise variance is different for different bands and is selected from 0.1~0.2 randomly) +salt and pepper(selected from 0 to 20% randomly) in all bands + dead-lines in 61~70 bands (with the number of stripes randomly selected from 1 to 10 and the width of the dead-lines randomly generated from 1 to 3)

(40) are calculated as:

$$\mathcal{P} = S_{\lambda_{tv}/2}[D_h \mathcal{X} D_z + \mathcal{Y}_1] \quad (44)$$

$$\mathcal{Q} = S_{\lambda_{tv}/2}[D_v \mathcal{X} D_z + \mathcal{Y}_2] \quad (45)$$

Extracting all items containing \mathcal{V} from Eq. (37), similar to the process of solving \mathcal{V} in the HSI denoising model. The detail of the l_0-l_1 HTV method for HSI CS is summarized in Algorithm 2.

VI. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, extensive experiments on HSI denoising and HSI compressed sensing are conducted to evaluate the performances of the proposed method. In the denoising experiments, we select four types of methods: TV based method, i.e., 2DTV, ASSTV [26], SSTV [27], HSSTV [28], E-3DTV [39] and LR based method, i.e., LRMR [29]. In the CS experiments, l_0-l_1 HTV is compared with including 2DTV, ASSTV [26], SSTV [27] and LR based method, i.e., LR and joint-sparse recovery (SparCS) [55]. Three TV norms are replaced in Eq. (35) and also solved by ADMM.

In simulated experiments, we employ the Mean Peak Signal to Noise Ratio (MPSNR), the Mean Structural Similarity Index (MSSIM), the Erreur Relative Globale Adimensionnelle de Synthese (ERGAS) and the mean spectral angle distance (MSAD) to evaluate HSI restoration qualities:

$$\text{MPSNR} = \frac{1}{z} \sum_{i=1}^z \text{PSNR}(\mathcal{X}(:,:,i), \bar{\mathcal{X}}(:,:,i)), \quad (46)$$

$$\text{MSSIM} = \frac{1}{z} \sum_{i=1}^z \text{SSIM}(\mathcal{X}(:,:,i), \bar{\mathcal{X}}(:,:,i)). \quad (47)$$

$$\text{ERGAS} = \sqrt{\frac{1}{z} \sum_{i=1}^z \text{mse}(\mathcal{X}(:,:,i), \bar{\mathcal{X}}(:,:,i))}, \quad (48)$$

$$\text{MSAD} = \frac{1}{hv} \sum_{i=1}^{hv} \frac{180}{\pi} \times \arccos \frac{(\mathbf{X}_{(3)}(:,:,i))^T \cdot (\bar{\mathbf{X}}_{(3)}(:,:,i))}{\|\mathbf{X}_{(3)}(:,:,i)\| \cdot \|\bar{\mathbf{X}}_{(3)}(:,:,i)\|},$$

where $\mathcal{X}(:,:,i)$ and $\bar{\mathcal{X}}(:,:,i)$ denote the i th band of the reference and recovered images, and $\mathbf{X}_{(3)}(:,:,i)$ and $\bar{\mathbf{X}}_{(3)}(:,:,i)$ are the i th spectral signatures of two images.

A. HSI Mixed Noise Removal Experiments

Two HSI datasets are adopted including the Washington DC Mall (WDC) dataset and the Indian synthetic (Indian-s) data. The WDC dataset was captured by the hyperspectral digital imagery collection experiment. A sub-region image in size of $200 \times 200 \times 160$ holds complex structure and texture information. The Indian synthetic (Indian-s) data has been considered in [29], [33], [38], which is generated by USGS splib06 using the ground truth of the Indian Pines dataset. The Indian-s data contains 145×145 pixels and 175 bands, with evident local smoothness information.

The noise of HSI data generally manifests as a mixture of Gaussian noise, salt and pepper noise, dead-line noise and strips. The details of noise cases are given in Table I. Three different kinds of Gaussian noise are taken into account. In cases 1~3, Gaussian noise is randomly added to all bands of HSIs. The noisy intensity of different bands is different and the SNR values of Gaussian noise vary from 10~20 dB randomly. For cases 4~6, the variance σ_z^2 of Gaussian noise is defined as [46]

$$\sigma_z^2 = \delta^2 \frac{\exp \left\{ -(z - Z/2)^2 / (2\eta^2) \right\}}{\sum_{z=1}^Z \exp \left\{ -(z - Z/2)^2 / (2\eta^2) \right\}} \quad (50)$$

where δ and η are two adjustable parameters and the mean noise SNR is set as 12.12 dB.

1) WDC dataset: In terms of visual quality, Fig. 2 shows band 70 of WDC data denoised by eight different methods under Case 5. ASSTV fails to remove all noise completely. Although noise is largely suppressed by 2DTV, its result yields some blurry areas and loses much detail information. HSSTV and E-3DTV perform comparatively better than SSTV and LRMR in mixed noise removal, but their results exist more or less undesired artifacts in the green rectangle area. As

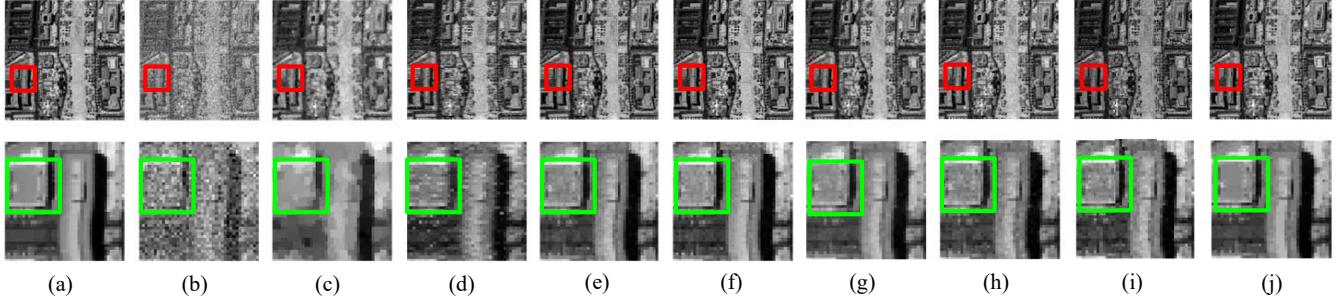


Fig. 2: Denoised results by all the compared methods for WDC dataset: (a) original band 70, (b) simulated noise band of case 5, (c) 2DTV, (d) ASSTV, (e) SSTV, (f) LRMR, (g) HSSTV1, (h) HSSTV2, (i) E-3DTV, (j) l_0-l_1 HTV.

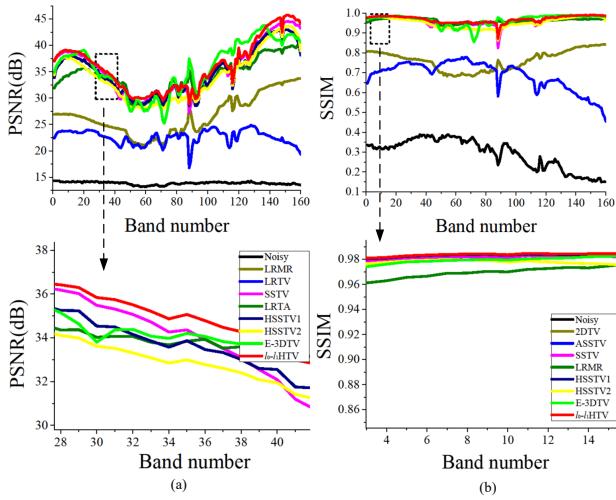


Fig. 3: PSNR and SSIM values of each band in Case 5 for WDC dataset: (a) PSNR, (b) SSIM.

observed, the proposed l_0-l_1 HTV produces superior denoised results without any over-smoothing or artifacts and meanwhile preserves clearer structure and texture information like four vertical boundaries in the middle of recovered image.

To quantitatively compare the performances of these methods, we list the assessment results in Table II, with the best ones marked in bold and the second ones underlined. Consistent with visual comparison, the proposed method achieves the highest evaluation scores among seven denoised approaches. SSTV, HSSTV, and E-3DTV outperform LRMR under cases 1~6 since these state-of-the-art TV-based methods ponder more local smoothness information. E-3DTV and LRMR have similar denoised results under cases 7~9. Nevertheless, LRMR explores the low-rank property of HSI and enables to handle complex Gaussian noise, leading to higher MSSIM values than HSSTV and E-3DTV. As ASSTV utilizes local information along three directions of HSI, ASSTV gains higher scores than 2DTV. Fig. 3 presents the PSNR and SSIM values of each band in Case 5 for the WDC dataset. Due to the same curves of l_0-l_1 HTV, LRMR, HSSTV, and E-3DTV, the curves in the black dashed frames are enlarged. Obviously, the proposed method shown in red provides the best performances.

2) Indian-s dataset: Fig. 4 shows the denoised perfor-

TABLE II: Quantitative results of all the methods under different mixed noise cases for WDC dataset.

Noise Case	Index	Noisy	WDC Dataset							
			2DTV	ASSTV	SSTV	LRMR	HSSTV1	HSSTV2	E-3DTV	l_0-l_1 HTV
Case 1	MPSNR	27.22	29.96	32.95	37.84	36.85	38.30	38.09	38.18	38.89
	MSSIM	0.7973	0.8779	0.9476	0.9776	0.9797	0.9809	0.9794	<u>0.9819</u>	0.9834
	EGRAS	11.9726	7.6773	5.5183	3.3347	3.2881	<u>2.9490</u>	3.0544	3.0633	2.7484
	MSAD	10.6152	7.5971	4.7666	3.1015	3.6719	2.8716	2.9465	3.0112	2.6935
Case 2	MPSNR	14.11	26.14	22.76	37.15	36.00	37.17	36.71	<u>38.13</u>	38.27
	MSSIM	0.3140	0.7781	0.7005	0.9767	0.9752	0.9776	0.9763	<u>0.9804</u>	0.9813
	EGRAS	46.9504	11.1352	17.8865	3.5510	3.6063	3.2904	3.4705	<u>3.1541</u>	2.9667
	MSAD	36.9851	11.6210	14.4817	3.3381	4.0624	3.2414	3.4029	<u>3.0299</u>	2.9030
Case 3	MPSNR	11.30	25.04	19.85	36.04	35.15	33.16	33.53	<u>36.06</u>	36.82
	MSSIM	0.1756	0.6995	0.5508	0.9689	0.9679	0.9575	0.9564	<u>0.9702</u>	0.9750
	EGRAS	46.9504	12.7272	25.3070	3.7629	4.1288	4.8508	4.7852	<u>3.9624</u>	3.4041
	MSAD	43.6584	13.3058	19.1071	3.7519	4.6473	4.7344	4.7852	<u>3.8913</u>	3.4023
Case 4	MPSNR	24.34	27.40	31.26	36.40	35.13	36.35	<u>36.86</u>	36.19	37.00
	MSSIM	0.7008	0.8029	0.9099	0.9674	0.9689	0.9713	<u>0.9747</u>	0.9698	0.9734
	EGRAS	15.2816	10.3531	6.9993	3.8969	3.6063	3.6474	<u>3.6205</u>	3.8483	3.5670
	MSAD	14.1859	10.0657	6.1004	3.6553	4.0624	3.5669	<u>3.4223</u>	3.7338	3.3771
Case 5	MPSNR	13.85	25.73	22.60	35.58	34.24	35.60	34.32	<u>35.89</u>	36.25
	MSSIM	0.2919	0.7277	0.6873	0.9622	0.9618	0.9632	0.9519	<u>0.9655</u>	0.9670
	EGRAS	47.7630	11.8439	18.0731	4.1833	4.4109	4.2951	4.7135	<u>4.0372</u>	3.8344
	MSAD	37.6949	12.1861	14.8037	3.9773	4.9412	4.1192	4.5168	<u>3.8409</u>	3.6591
Case 6	MPSNR	11.16	24.56	19.73	34.52	33.28	32.50	32.08	33.67	35.03
	MSSIM	0.1662	0.6796	0.5383	0.9541	<u>0.9539</u>	0.9428	0.9376	0.9481	0.9605
	EGRAS	65.9574	13.4884	25.4364	4.6480	4.9389	5.5669	5.8622	5.7319	4.3581
	MSAD	44.1772	14.0139	19.4124	4.4672	5.4491	5.3261	5.6505	5.2271	4.1878
Case 7	MPSNR	16.48	20.43	25.67	30.10	31.17	<u>31.27</u>	29.43	31.14	31.38
	MSSIM	0.3543	0.5043	0.8107	0.8853	<u>0.9178</u>	0.9170	0.8741	0.9049	0.9198
	EGRAS	34.9656	23.1351	12.3470	7.3824	6.2509	<u>6.2464</u>	7.7847	6.6029	6.2090
	MSAD	32.4990	22.9750	10.5510	7.9050	7.2501	<u>6.6695</u>	8.3656	6.9146	6.6289
Case 8	MPSNR	12.58	15.39	22.20	29.41	29.68	29.00	28.92	<u>29.71</u>	30.07
	MSSIM	0.2220	0.3104	0.6337	0.8708	0.9054	0.8825	0.8785	0.8726	<u>0.8951</u>
	EGRAS	58.4574	43.0059	18.9829	8.0110	7.4304	8.2410	8.3480	8.9605	7.3081
	MSAD	41.8123	34.0091	15.9943	8.4385	<u>8.2500</u>	8.2950	8.4206	8.3783	7.5561
Case 9	MPSNR	12.50	15.30	22.14	29.30	29.57	28.50	28.35	<u>29.63</u>	30.00
	MSSIM	0.2199	0.3763	0.6315	0.8160	0.9033	0.8600	0.8554	0.8632	<u>0.8949</u>
	EGRAS	58.3482	42.5358	18.8190	7.9522	<u>7.5299</u>	8.9066	9.0515	9.5305	7.2896
	MSAD	42.1183	33.6726	16.0386	8.4101	<u>8.3486</u>	8.8264	8.9889	8.6956	7.5511

mances of compared methods under Case 9. As seen in Fig. 4 (a), Indian-s data has much sharper edges than the WDC dataset. Similarly, ASSTV still remains the amount of noise and 2DTV generates serious over-smoothing. ASSTV and SSTV fail to remove strips completely. Unlike SSTV, LRMR restores a better image without lots of artifacts since LRMR explores the low-rank property of HSI and enables to handle complex Gaussian noise, leading to higher MSSIM values than HSSTV and E-3DTV. However, HSSTV just exploits the local differences of HSI, which causes wrong boundary judgment. Although E-3DTV outperforms the above-mentioned algorithms, there exist fewer artifacts in the smoothing areas since E-3DTV only utilizes the local spatial information and the sparsity insight of 3DTV. The white area on the left reveals l_0-l_1 HTV obtains the best smoothness and the sharpest edges. l_0-l_1 HTV is superior to

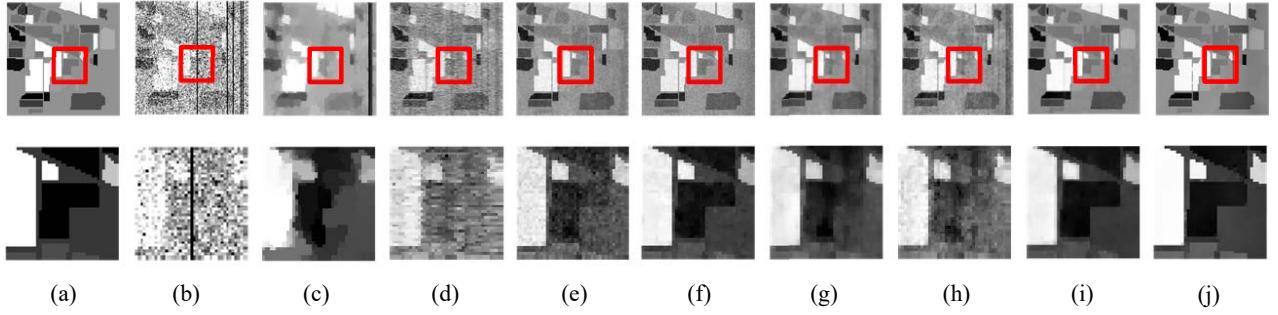


Fig. 4: Denoised results by all the compared methods for Indian-s dataset: (a) original band 65, (b) simulated noise band of case 9, (c) 2DTV, (d) ASSTV, (e) SSTV, (f) LRMR, (g) HSSTV1, (h) HSSTV2, (i) E-3DTV, (j) l_0-l_1 HTV.

TABLE III: Quantitative results of all the methods under different mixed noise cases for Indian-s dataset.

Noise Case	Index	Indian-s Dataset								
		Noisy	2DTV	ASSTV	SSTV	LRMR	HSSTV1	HSSTV2	E-3DTV	l_0-l_1 HTV
Case 1	MPSNR	26.34	31.80	34.04	42.11	41.50	<u>52.20</u>	40.84	42.09	48.36
	MSSIM	0.5125	0.9092	0.7987	0.9706	0.9554	0.9853	0.9723	<u>0.9864</u>	0.9976
	ERGAS	11.4981	5.3534	5.1290	1.6352	1.7970	<u>1.6658</u>	1.9309	2.3059	0.8380
Case 2	MSAD	1.8373	5.6082	4.0230	1.7533	1.8373	1.7766	2.0184	<u>1.6364</u>	0.8329
	MPSNR	14.10	30.55	25.89	40.76	40.18	40.92	39.17	<u>41.73</u>	46.86
	MSSIM	0.1416	0.8869	0.6865	0.9457	0.9264	0.9775	0.9506	<u>0.9882</u>	0.9967
Case 3	ERGAS	45.2641	6.2247	12.7398	1.8783	2.0617	<u>1.8495</u>	2.3058	1.9456	0.9965
	MSAD	34.4532	6.5157	8.9093	2.0228	2.1259	1.8881	2.2816	<u>1.7073</u>	0.9866
	MPSNR	11.38	29.36	22.10	39.33	38.89	38.36	36.12	<u>40.63</u>	45.50
Case 4	MSSIM	0.0734	0.8613	0.5166	0.9457	0.9264	0.9676	0.9142	<u>0.9843</u>	0.9958
	ERGAS	63.1225	7.1611	20.3016	<u>2.2232</u>	2.4029	2.4883	2.3296	2.2755	1.0183
	MSAD	39.9088	7.4905	12.3116	2.3942	2.4565	2.4307	3.1494	<u>1.9189</u>	1.0323
Case 5	MPSNR	23.51	30.79	30.83	39.17	38.84	<u>40.78</u>	39.97	40.07	45.89
	MSSIM	0.4147	0.8799	0.6741	0.9682	0.9081	0.9726	0.9317	<u>0.9742</u>	0.9974
	ERGAS	16.0971	6.0026	8.1258	2.3333	2.5499	<u>1.9414</u>	2.5508	2.6449	1.1403
Case 6	MSAD	14.6163	6.1425	6.0690	2.3942	2.4827	<u>1.9183</u>	2.3623	2.0342	1.0600
	MPSNR	13.62	29.78	25.25	38.14	37.91	38.95	36.76	<u>39.39</u>	44.69
	MSSIM	0.1264	0.8600	0.6348	0.9075	0.8965	0.9632	0.8987	<u>0.9739</u>	0.9940
Case 7	ERGAS	16.0887	6.7552	13.1283	2.6097	2.8033	<u>2.3932</u>	3.2468	2.9820	1.2872
	MSAD	36.5602	6.9482	9.3958	2.6529	2.7355	2.3290	3.0067	<u>2.1509</u>	1.2322
	MPSNR	11.13	28.40	21.67	36.81	36.56	36.54	34.09	<u>37.93</u>	42.85
Case 8	MSSIM	0.0681	0.8275	0.4806	0.8953	0.8715	0.9458	0.8492	<u>0.9695</u>	0.9902
	ERGAS	46.3208	7.9260	20.6494	<u>3.0334</u>	3.4558	3.1618	4.3455	3.0903	1.6262
	MSAD	35.5602	8.1290	12.7695	3.0799	3.1478	<u>2.9339</u>	3.9652	<u>2.5740</u>	1.5174
Case 9	MPSNR	16.07	28.40	27.64	32.04	34.66	36.09	32.53	<u>39.64</u>	40.25
	MSSIM	0.1499	0.7921	0.6297	0.7656	0.8819	0.9468	0.8385	<u>0.9804</u>	0.9980
	ERGAS	36.0327	8.3164	9.9599	5.9355	3.7659	3.4308	5.1724	<u>2.5135</u>	2.1761
Case 10	MSAD	32.9956	9.2680	9.2830	6.6732	4.2082	3.8704	5.8814	<u>2.3930</u>	2.4115
	MPSNR	12.93	27.46	23.17	31.32	33.86	33.19	30.08	<u>36.92</u>	39.37
	MSSIM	0.0956	0.7982	0.5024	0.7403	0.8631	0.9231	0.7656	<u>0.9724</u>	0.9838
Case 11	ERGAS	54.4352	9.0489	18.7659	6.3558	4.1185	<u>4.3218</u>	7.2719	<u>3.2275</u>	2.4449
	MSAD	36.5381	10.0114	12.2504	7.0768	4.5191	5.1050	7.2911	<u>3.1632</u>	2.6785
	MPSNR	12.87	27.19	23.10	31.22	33.78	32.96	29.87	<u>36.80</u>	39.28
Case 12	MSSIM	0.0951	0.7942	0.5014	0.7403	0.8615	0.9208	0.7603	<u>0.9720</u>	0.9852
	ERGAS	54.5384	9.2997	18.8065	6.3824	4.1608	5.3789	7.3462	<u>3.2776</u>	2.4148
	MSAD	36.5962	10.1866	12.2709	7.1077	4.5275	5.1436	7.3547	<u>3.1931</u>	2.6458

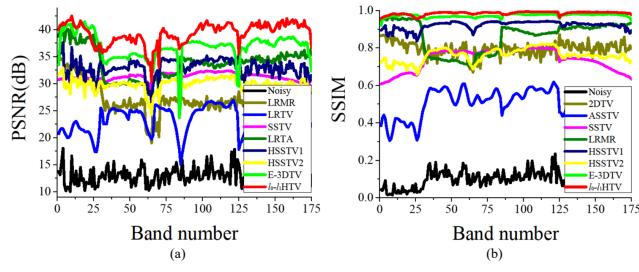


Fig. 5: PSNR and SSIM values of each band in Case 9 for Indian-s dataset: (a) PSNR, (b) SSIM.

the other approaches.

Table II gives the quantitative results of eight denoised algorithms and further proves the best performances of our proposed l_0-l_1 HTV. For this kind of dataset with sharper edges, E-3DTV and HSSTV bring better results than 2DTV,

ASSTV, and LRMR in most of the cases. E-3DTV obtains the second-highest metric values among seven comparing methods. HSSTV deals with complex Gaussian noise cases like cases 8~9 a little worse than LRMR. Due to the good use of local HSI information, HSSTV obtains better MSSIM scores than LRMR. Meanwhile, as we can see in Fig. 5, the PSNR and SSIM values of each band obtained by the proposed l_0-l_1 HTV are much higher than the other state-of-the-art algorithms.

3)Real-world dataset : Indian Pines dataset was acquired by NASA AVIRIS instrument in North-western Indiana, containing 145×145 and 224 spectral reflectance bands. Fig. 6 shows the original band 220 of this scene is corrupted by mixed noise and target objects can barely be recognized, which directly hinders further HSI analysis. The above eight methods are implemented for HSI denoising. 2DTV and ASSTV lose their utility when they encounter heavy noise cases. SSTV and LRMR remove lots of noise, but they produce artifacts. The result denoised by E-3DTV appears blurred image edges and fewer artifacts. In contrast, HSSTV and l_0-l_1 HTV restore the clearest image without artifacts.

In real data experiments, the above image assessments are useless since the reference images are unknown. Meanwhile, preserving Sharp image edges benefits further HSI tasks like classification. Therefore, the supervised classification experiment is designed and the well-known support vector machines (SVM) [56]–[58] is applied to quantitatively compare the performances of different methods. The first and second column of Table IV show 16 different classes, and the number of training (random 10%) and test samples. The accuracy for each class, kappa coefficient, overall accuracy (OA) and average accuracy (AA) are depicted in Table IV and eight denoising techniques can be quantitatively evaluated. Fig. 7 presents the visual comparison of classification results before and after denoising. It is conspicuously seen that classification accuracies are significantly fortified after denoising process. Due to the over-smoothing phenomenon, the results by 2DTV and ASSTV appear large misclassified areas. Therefore, their kappa, OA, and AA scores have risen less. E-3DTV presents higher classification accuracies than LRMR since E-3DTV explores the sparsity of subspace on the gradient maps and is one of the advanced 3DTV variant algorithms. SSTV provides higher metric values, but artifacts cause incorrect classification

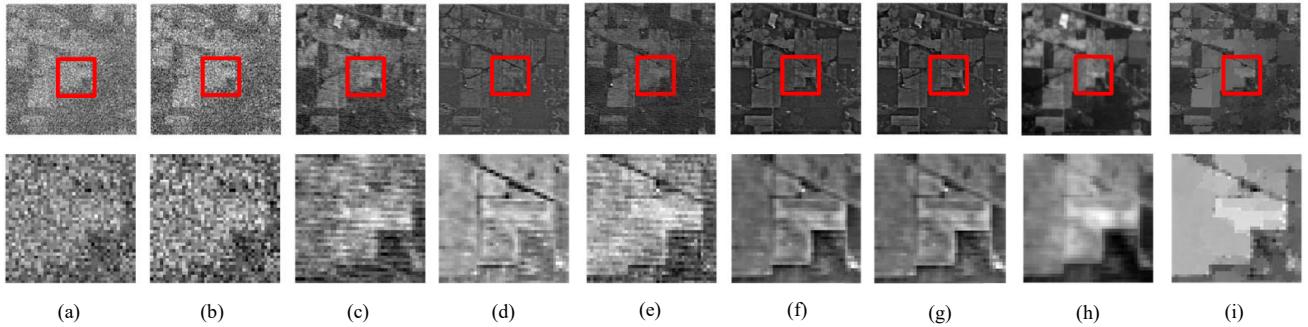


Fig. 6: Denoised results by all the compared methods on the Indian Pine data: (a) original band 220, (b) 2DTV, (c) ASSTV, (d) SSTV, (e) LRMR, (f) HSSTV1, (g) HSSTV2, (h) E-3DTV, (i) $l_0 - l_1$ HTV.

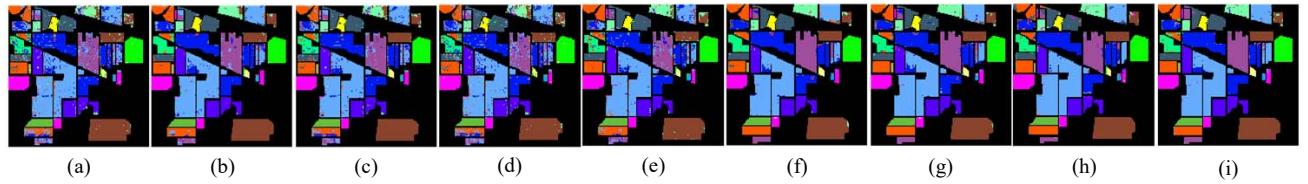


Fig. 7: Classification results by all the compared methods in the Indian Pine data: (a) original, (b) 2DTV, (c) ASSTV, (d) SSTV, (e) LRMR, (f) HSSTV1, (g) HSSTV2, (h) E-3DTV, (i) $l_0 - l_1$ HTV.

TABLE IV: Classification results on Indian Pines data by different methods.

Class	Train/Test	Noisy	2DTV	ASSTV	SSTV	LRMR	HSSTV1	HSSTV2	E-3DTV	$l_0 - l_1$ HTV
Alfalfa	5/41	0.6341	0.6585	0.7073	0.7804	0.7317	<u>0.8780</u>	0.8537	0.6585	0.9512
Corn-notill	143/1285	0.8202	0.9253	0.8677	0.9167	0.9105	0.9183	<u>0.9198</u>	0.8996	0.9447
Corn-mintill	83/747	0.7161	0.9076	0.7778	0.9518	0.8955	<u>0.9572</u>	0.9625	0.8983	0.9394
Corn	24/213	0.8028	0.9848	0.8967	0.9154	0.9248	<u>0.9765</u>	0.9390	0.9577	0.9437
Grass-pasture	49/434	0.9377	0.9931	0.9470	0.9723	0.9354	0.9700	0.9677	0.9493	<u>0.9816</u>
Grass-trees	73/657	0.9497	0.9833	0.9696	0.9863	0.9619	<u>0.9897</u>	0.9848	0.9833	0.9985
Grass-pasture-mowed	3/25	0.6400	0.7200	0.7600	0.8400	0.6400	<u>0.8800</u>	0.800	0.5600	0.9600
Hay-winnowed	48/430	0.9906	0.9930	1.0000	0.9953	0.9976	<u>0.9977</u>	<u>0.9977</u>	<u>0.9977</u>	1.0000
Oats	2/18	0.7222	<u>0.9444</u>	1.0000	<u>0.9444</u>	0.8333	0.9359	0.9444	1.0000	1.0000
Soybean-notill	98/874	0.7059	0.8272	0.7780	0.9405	0.8821	<u>0.9359</u>	0.9176	0.8833	0.9245
Soybean-mintill	246/2209	0.8415	0.9140	0.8873	<u>0.9601</u>	0.9388	0.9520	0.9579	0.9511	0.9765
Soybean-clean	60/533	0.8480	0.9587	0.9043	<u>0.9699</u>	0.9324	0.9737	0.9625	0.9043	0.9437
Wheat	21/184	<u>0.9945</u>	0.9946	0.9946	<u>0.9945</u>	0.9945	0.9891	0.9891	0.9946	0.9783
Woods	127/1138	0.9516	<u>0.9824</u>	0.9754	0.9674	0.9455	0.9736	<u>0.9824</u>	0.9807	0.9851
Buildings-Grass-Trees-Drives	39/347	0.6311	0.8559	0.6513	0.8587	0.8040	0.9510	<u>0.9424</u>	0.9049	0.9308
Stone-Steel-Towers	10/83	0.9518	1.0000	0.9518	0.9638	0.9759	1.0000	1.0000	1.0000	1.0000
kappa	-	0.8197	0.9197	0.8677	0.9451	0.9131	<u>0.9510</u>	0.9496	0.9271	0.9594
OA	-	0.8423	0.9293	0.8842	0.9518	0.9237	<u>0.9569</u>	0.9557	0.9360	0.9644
AA	-	0.8212	0.9192	0.8793	0.9349	0.8911	<u>0.9588</u>	0.9451	0.9077	0.9666

results for each class. As kappa, OA, and AA show, two HSSTV methods have similar classification results, which are a little lower than the proposed method. $l_0 - l_1$ HTV produces the highest evaluation scores, with kappa, OA, and AA of 0.9594, 0.9644 and 0.9666.

B. HSI Compressed Sensing Experiments

Three HSI data experiments are employed to validate the effectiveness of $l_0 - l_1$ HTV on HSI CS with seven different sample ratio i.e. 0.05, 0.1, 0.2, 0.4 and 0.6. Three datasets include Pavia University , CAVE Toy and Reno, and their cube used for HSI CS experiments are size of $250 \times 121 \times 130$, $300 \times 300 \times 31$ and $150 \times 150 \times 100$.

1) Pavia University: In terms of visual quality, the representative bands of eight HSI CS results in the sampling

ratio $r = 0.4$ are presented in Fig. 8. As shown in the green square of the enlarged areas, reconstructed images by 2DTV, ASSTV, and HSSTV are smoothed excessively, which causes the loss of street detailed information. As observed from the white areas, SSTV and SparCS lead to more or fewer artifacts. Meanwhile, HSSTV overcomes the artifacts of SSTV and E-3DTV compensates for the defect of over-smoothing. $l_0 - l_1$ HTV provides the best result without any artifact and the comparison of PSNR and SSIM values of all the bands in Fig. 9 also illustrates the superiority of our $l_0 - l_1$ HTV method.

Table V presents the quantitative indices of eight different sampling ratios on the Pavia university dataset. SSTV and SparCS can reconstruct great images in high sampling ratios, but poor results are obtained in low sampling ratios. Nevertheless, the results of 2DTV are just the opposite. As two modified

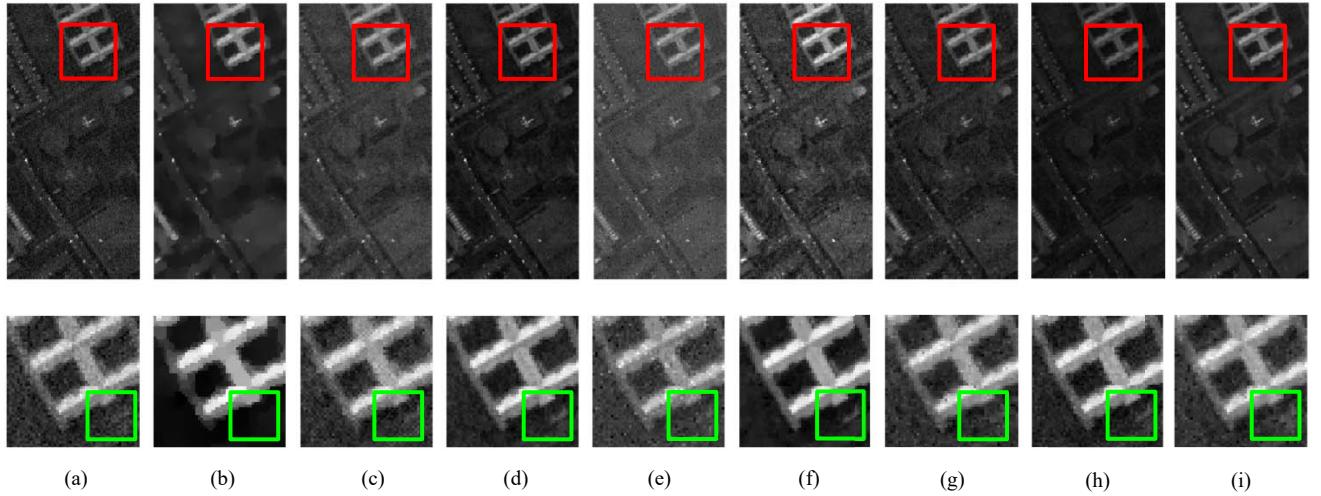


Fig. 8: HSI CS results on the Pavia university dataset and sample ratio is 0.4. (a) Original band 1, (b) 2DTV, (c) ASSTV, (d) SSTV, (e) SparCS, (f) HSSTV1, (g) HSSTV2, (h) E-3DTV, (i) l_0-l_1 HTV.

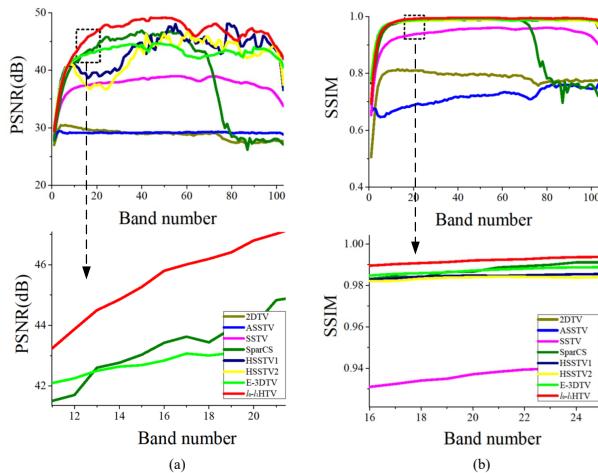


Fig. 9: PSNR and SSIM values of each band for Pavia University under $r = 0.4$: (a) PSNR, (b) SSIM.

TV-based methods, HSSTV and E-3DTV always obtain better results than SSTV, but they still have quality decays with sample ratios decreasing. These indicate the instabilities of the above three methods. l_0-l_1 HTV provides much higher metrics than the other approaches, with lower sampling ratios.

2) Toy: Compared with the above Toy HSIs, more regular clothes texture information exists in the magnified areas of Fig. 10 (a). Unfortunately, ASSTV and SSTV fail to recover the final HSI under low sampling ratio $r = 0.2$. Although a few texture details can be found in the results by 2DTV and SparCS, it is hard to distinguish the colors of texture information. As the zoomed areas of the yellow frame show, much white background information is lost by two HSSTV algorithms. E-3DTV produces a better background result, but there still exist few artifacts. l_0-l_1 HTV method provides the best result with reconstructing most information of the original image. The PSNR and SSIM values of each band in Case 9 for the Indian-s dataset are shown in Fig. 9. We can observed that

TABLE V: Quantitative results of all the methods under different CS cases for Pavia university dataset.

Ratio	Index	Pavia university Dataset							
		2DTV	ASSTV	SSTV	SparCS	HSSTV1	HSSTV2	E-3DTV	l_0-l_1 HTV
0.05	MPSNR	21.48	20.37	21.17	19.71	25.15	22.20	<u>26.89</u>	29.63
	MSSIM	0.4587	0.3174	0.4446	0.3002	0.7009	0.6681	<u>0.7508</u>	0.8492
	ERGAS	20.0014	22.8979	20.9138	25.0856	13.3918	18.5314	<u>10.8490</u>	7.8779
	MSAD	23.8498	27.7240	25.0375	29.2799	14.6353	19.8596	<u>11.7016</u>	8.8695
	MPSNR	23.43	22.04	23.01	20.83	28.57	27.93	<u>30.39</u>	36.33
	MSSIM	0.5374	0.4216	0.5270	0.3597	<u>0.9036</u>	0.8964	0.8613	0.9411
	ERGAS	15.8906	18.8693	16.9183	23.2884	9.2553	9.8463	<u>7.2841</u>	3.7945
	MSAD	18.7013	22.4843	19.9202	20.8280	10.2367	10.9056	<u>7.6556</u>	4.2857
0.1	MPSNR	25.69	25.57	27.79	26.19	<u>36.49</u>	32.09	34.24	41.11
	MSSIM	0.6531	0.5768	0.7175	0.6363	<u>0.9666</u>	0.9560	0.9390	0.9717
	ERGAS	12.2199	12.5174	9.7927	12.5922	<u>4.0403</u>	6.4183	4.7032	2.4319
	MSAD	14.2670	14.6661	11.3162	13.7806	<u>4.3824</u>	7.0270	5.3836	2.5706
0.2	MPSNR	28.80	29.13	37.60	39.23	<u>43.11</u>	42.14	42.82	46.15
	MSSIM	0.7841	0.7160	0.9389	0.9173	0.9803	0.9801	<u>0.9826</u>	0.9844
	ERGAS	8.5431	8.3255	3.3359	4.5526	2.1228	2.3508	<u>1.9911</u>	1.6831
	MSAD	9.8675	9.7278	3.7420	3.9402	2.1740	2.4210	<u>2.1164</u>	1.5689
0.4	MPSNR	31.15	34.69	45.58	47.04	<u>47.25</u>	46.32	46.81	47.92
	MSSIM	0.8533	0.8875	0.9836	0.9875	0.9858	0.9852	<u>0.9909</u>	0.9881
	ERGAS	6.5299	4.4228	1.7491	1.6358	<u>1.4127</u>	1.5741	1.3848	1.4497
	MSAD	7.5053	5.1355	1.6679	1.4617	<u>1.3862</u>	1.5693	1.3894	1.3116

in almost all bands, the metric values of our proposed method (the red curves) are higher than those of the other approaches.

Table VI lists the four metric values of different methods. Under almost all cases, the metric scores obtained by using l_0-l_1 HTV are the highest among all the techniques. 2DTV, ASSTV and SSTV outperform SparCS under most sampling ratios, which demonstrates TV methods are more suitable for recovering HSIs with more texture information than pure low-rank methods. The MSSIM, ERGAS, and MSAD values of HSSTV are the highest when high sample ratio $r = 0.6$, but the metrics of HSSTV decrease more than E-3DTV and l_0-l_1 HTV with decreasing sample ratios.

3) Reno Dataset: Fig. 12 depicts the reconstructed results of band 2 of Reno dataset by different methods, with the random sample ratio $r = 0.1$. 2DTV causes spatial over-smoothing. SparCS exploits the low-rank property of HSI, but SparCS ignores spatial smoothness, which leads to the lack of local structure information. Due to the existence of artifacts, blurry image edges can hardly be distinguished in the result recovered by SSTV. Based on local spatial and spatio-spectral information, HSSTV recovers some clear image edges, but

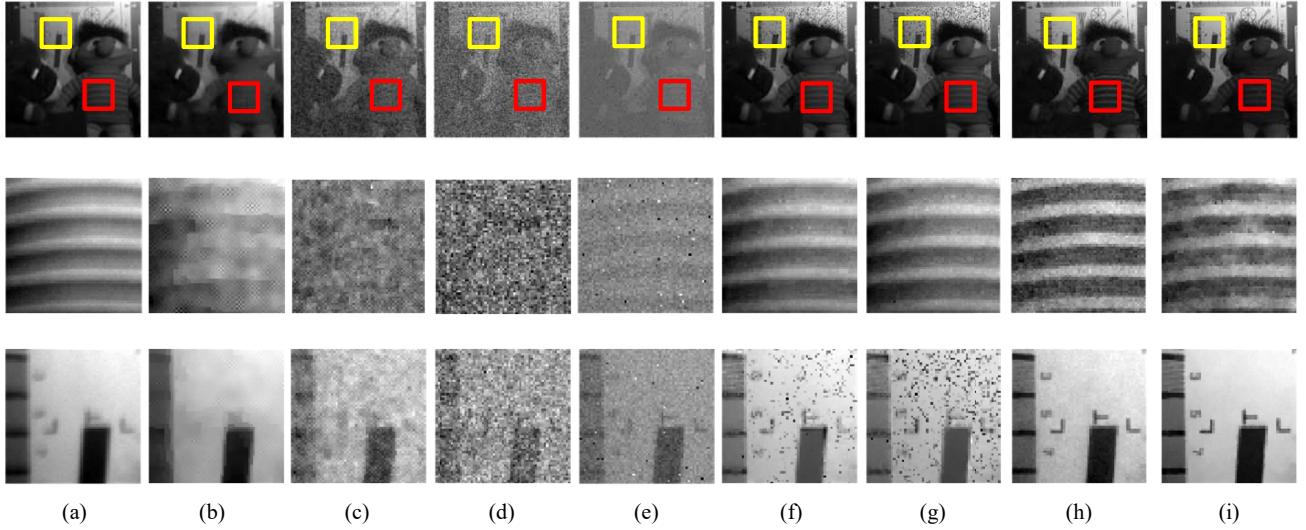


Fig. 10: HSI CS results on the Toy dataset and sample ratio is 0.2. (a) Original band 2, (b) 2DTV, (c) ASSTV, (d) SSTV, (e) SparCS, (f) HSSTV1, (g) HSSTV2, (h)E-3DTV, (i) l_0-l_1 HTV.

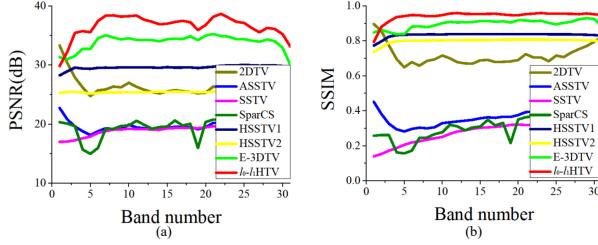


Fig. 11: PSNR and SSIM values of each band for Toy under $r = 0.2$: (a) PSNR, (b) SSIM.

TABLE VI: Quantitative results of all the methods under different CS cases for Toy dataset.

Ratio	Index	Toy Dataset							
		2DTV	ASSTV	SSTV	SparCS	HSSTV1	HSSTV2	E-3DTV	l_0-l_1 HTV
0.05	MPSNR	13.18	12.39	12.91	10.89	14.62	13.36	<u>20.42</u>	23.74
	MSSIM	0.2033	0.1415	0.1299	0.0757	<u>0.3027</u>	0.6318	0.3896	0.6234
	ERGAS	47.0874	51.5439	48.4195	61.4419	39.7529	45.9325	<u>20.5011</u>	13.8817
	MSAD	39.2882	44.2176	40.7690	53.3116	27.8701	33.6779	<u>10.7912</u>	16.6667
0.1	MPSNR	15.83	13.78	14.51	12.82	18.55	16.10	<u>27.66</u>	32.23
	MSSIM	0.2746	0.1716	0.1572	0.1091	<u>0.8110</u>	0.7610	0.7519	0.8800
	ERGAS	34.8259	44.0011	40.2711	49.3815	25.2887	33.5346	<u>8.8715</u>	5.2783
	MSAD	27.0173	36.0451	32.6395	41.5952	17.8904	24.3612	<u>6.47281</u>	4.0486
0.2	MPSNR	26.30	19.65	18.62	19.11	29.60	25.42	<u>33.85</u>	36.55
	MSSIM	0.7171	0.3640	0.2635	0.2877	0.8326	0.7993	<u>0.8957</u>	0.9462
	ERGAS	10.6142	22.4768	25.2059	24.5570	7.0920	11.4444	<u>4.4208</u>	3.3489
	MSAD	7.5504	16.1117	19.5294	41.5952	5.4126	8.6127	<u>3.2960</u>	2.5473
0.4	MPSNR	37.23	31.63	31.15	30.6924	39.84	38.26	<u>40.52</u>	41.23
	MSSIM	0.9480	0.7557	0.7128	0.7133	<u>0.9842</u>	0.9803	0.9673	0.9811
	ERGAS	3.0369	5.6096	6.0629	89464	2.2174	2.6428	<u>2.1044</u>	1.6689
	MSAD	2.2502	4.2858	4.6485	5.4921	1.7069	2.0394	<u>1.9919</u>	1.2115
0.6	MPSNR	43.85	33.97	44.60	35.90	<u>45.27</u>	45.01	44.18	47.92
	MSSIM	0.9884	0.8182	0.9816	0.8457	<u>0.9894</u>	0.9891	0.9822	0.9881
	ERGAS	1.3753	4.2852	1.3168	5.0260	<u>1.1959</u>	1.1866	1.4131	1.4497
	MSAD	1.0637	3.3165	1.0021	3.1134	<u>0.9190</u>	0.9015	1.0589	1.3116

loses lots of pixels in the white areas. It is evident that E-3DTV and l_0-l_1 HTV get the best reconstruction, but l_0-l_1 HTV recovers the complete white areas and the sharpest edges. The PSNR and SSIM of each band restored by different approaches are shown in Fig. 13 and the results demonstrate that l_0-l_1 HTV marked in red achieved the best results in all the bands.

Table VII provides the quantitative results of eight different sampling ratios on the Reno dataset in terms of four evaluation

indices. It can be easily seen that the proposed l_0-l_1 HTV gives a significantly fortified performance under five sampling ratios, as compared with other competing methods. Since important structure details of the final HSIs are barely reconstructed, 2DTV and ASSTV get similar assessment scores. By using more local differences, E-3DTV and HSSTV recover the second-best and third-best results, respectively. Especially, E-3DTV is more beneficial for the CS recovery with lower sampling ratios. However, l_0-l_1 HTV achieves the best reconstructions under all sampling ratios.

In summary, the structure information of three datasets is different, while our proposed method always obtains the best-recovered images. l_0-l_1 HTV can be a useful regularizer to ameliorate the performances for HSI denoising and CS tasks.

TABLE VII: Quantitative results of all the methods under different CS cases for Reno dataset.

Ratio	Index	Reno Dataset							
		2DTV	ASSTV	SSTV	SparCS	HSSTV1	HSSTV2	E-3DTV	l_0-l_1 HTV
0.05	MPSNR	15.76	15.96	16.21	13.56	18.55	17.04	22.80	25.11
	MSSIM	0.2574	0.2743	0.2891	0.1488	<u>0.7336</u>	0.6472	0.6073	0.7724
	ERGAS	33.2416	32.5009	31.5179	45.2039	24.0110	28.4047	14.5842	11.1754
	MSAD	32.1112	31.2797	30.2264	41.6306	20.7043	24.7868	<u>13.3769</u>	9.8166
0.1	MPSNR	16.87	17.62	17.77	14.48	24.67	23.08	<u>27.98</u>	32.05
	MSSIM	0.3240	0.3867	0.3866	0.2189	<u>0.9070</u>	0.8743	0.8876	0.9230
	ERGAS	29.2555	26.8520	26.3792	38.1143	12.1361	14.4872	<u>8.0589</u>	5.0768
	MSAD	27.7898	25.1334	24.8107	41.4849	10.5626	12.6363	<u>6.5857</u>	4.6492
0.2	MPSNR	20.10	22.98	21.84	22.18	35.54	32.76	<u>35.72</u>	38.94
	MSSIM	0.5054	0.6348	0.5921	0.6234	<u>0.9782</u>	0.9744	0.9733	0.9792
	ERGAS	20.1403	14.3473	16.5678	16.5344	3.5288	5.0153	<u>3.3380</u>	2.3849
	MSAD	18.4749	12.8383	15.1398	14.8583	3.1360	4.3568	<u>2.9699</u>	2.1355
0.4	MPSNR	29.54	30.77	36.19	36.84	38.11	36.60	<u>42.82</u>	45.81
	MSSIM	0.8679	0.8669	0.9537	0.9306	0.9887	0.9858	<u>0.9914</u>	0.9961
	ERGAS	6.7873	5.8024	3.2823	5.1925	2.8389	3.3429	<u>1.5316</u>	1.1066
	MSAD	6.1679	5.3277	2.9512	3.6990	2.4193	2.8556	<u>1.3672</u>	0.9465
0.6	MPSNR	33.83	33.69	47.64	47.92	<u>48.34</u>	47.41	47.78	48.81
	MSSIM	0.9413	0.9179	0.9957	0.9890	<u>0.9974</u>	0.9950	0.9964	0.9969
	ERGAS	4.1420	4.1573	0.9421	1.7419	<u>0.9277</u>	1.0385	0.9366	0.8553
	MSAD	3.7852	3.8306	0.8052	0.9994	<u>0.7756</u>	0.8616	0.8017	0.7165

C. Discussion

1) Convergence of the proposed methods: The convergence is verified on HSI denoising and CS experiment datasets. Fig. 14 (a) displays the RalCha of two denoised

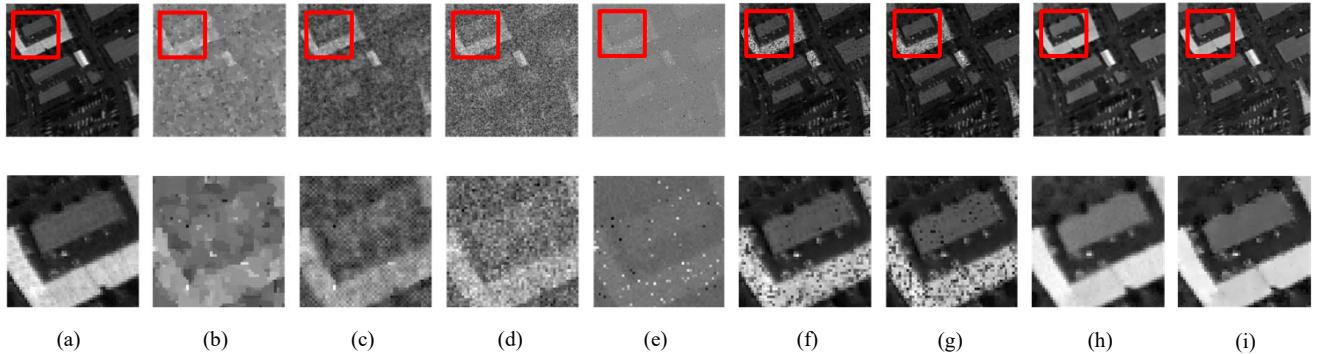


Fig. 12: HSI CS results on the Reno dataset and sample ratio is 0.1. (a) Original band 2, (b) 2DTV, (c) ASSTV, (d) SSTV, (e) SparCS, (f) HSSTV1, (g) HSSTV2, (h) E-3DTV, (i) l_0-l_1 HTV.

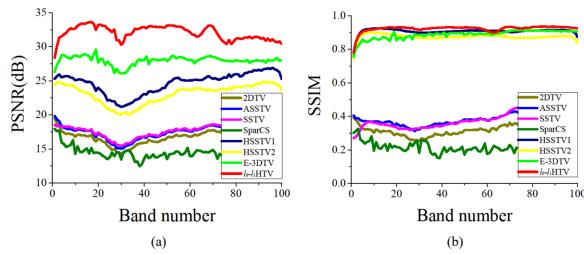


Fig. 13: PSNR and SSIM values of each band for Reno dataset under $r = 0.1$: (a) PSNR, (b) SSIM.

results for WDC and Indian-s datasets under noisy cases 5 and 9, respectively. RalCha is the relative change of two restored results, i.e., $\|\mathcal{X}(i+1) - \mathcal{X}(i)\|_F^2 / \|\mathcal{X}(i)\|_F^2$, where $\mathcal{X}(i)$ is i -th recovered HSI image. Fig. 14 (b) shows the RalCha of two recovered results for PaviaUA and Toy datasets under sampling ratio $r = 0.1$ and $r = 0.2$, respectively. It can be spotted that as the iterations are conducted, the RalCha values of l_0-l_1 HTV for HSI denoising and CS reconstruction monotonically decrease, finally approaching zero.

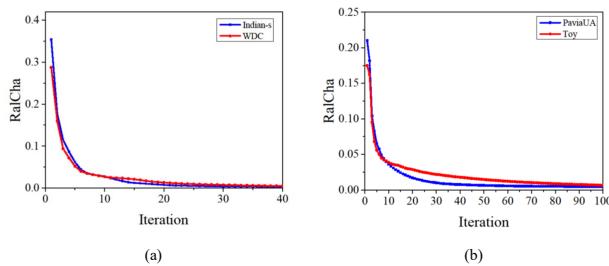


Fig. 14: RelCha values with respect to the number of iterations. (a) HSI denoising, (b) HSI CS.

2) Parameter analysis: In the l_0-l_1 HTV model, four parameters should be carefully identified for the HSI denoising and CS. We analyze the impact of each parameter on the restoration results and discuss how to choose such parameters in our experiments. The denoised results are based on the simulated experiments in WDC and Indian-s datasets. The effect of each parameter on the CS is discussed under sampling ratio $r = 0.1$.

Parameter γ plays an important role to control the number of l_0 gradients. Due to relationship with HSI size, γ is replaced by parameter γ' that is set to various percentages of total HSI size in Fig. 15 (a) and Fig. 16 (a). γ' is chosen from the set [0.1, 1.0] with 0.01 as step size. The Indian-s and Toy datasets have simpler texture information than the WDC and Pavia datasets, thus, γ' values for the Indian-s and Toy datasets are lower than the others.

Parameter λ is related to the strength of sparse noise norm. l_0-l_1 HTV is tested using different values of λ varied among [0.01, 0.1]. Obviously, l_0-l_1 HTV produces an appealing result when λ lies in the range of [0.05, 0.06].

Parameter λ_{tv} restricts the influence of the SSTV regularization on denoising or CS effectiveness. Fig. 15 (c) and Fig. 16 (b) present the robustness of the propose method to the changes of λ_{tv} . When λ_{tv} is given within [0.07, 0.08] for denoising and [8, 10] for CS, respectively, PSNR stabilizes at a high value.

Parameter β is used for the ADMM multipliers. Firstly, β is set as 0.1 for denoising and 0.01 for CS, respectively, and then tuned gradually. β is fixed as 0.3 and 0.05 for denoising and CS, respectively, which achieves the best performances.

3) Computational complexity: The computational complexities of the proposed algorithm for two HSI restorations are discussed in this section and the first three parts for denoising and CS processes are universal: 1) Solving the sub-problem (22), (23) or (39), (40) includes two thresholds and two SSTV operators, which takes about $O(2hvz + 2hvz)$ iterations. 2) For Eq. (24) or (41), the sorting of the L2 norms of hv subvectors spends the most time and requires $O(hv(\log hv))$ time. 3) Most of the computation time for updating the multipliers lies on the SSTV and l_0 HTV operators, which requires about $O(3hvz)$ iterations.

For HSI denoising: 3) The threshold for solving \mathcal{S} -related problem requires $O(hvz)$ and the computation of LSQR algorithm used to update \mathbf{x} costs $O(4hv + 5z)$.

For HSI CS: 3) LSQR algorithm is also utilized to update \mathbf{x} and its computation costs $O(4hv + hvz)$.

Therefore, the computational complexity of the HSI denoising process and CS process are in total $O(8hvz + 5z + hv(\log hv))$ and $O(8hvz + 4hv + hv(\log hv))$, respectively.

The running times of different methods on HSI denoising

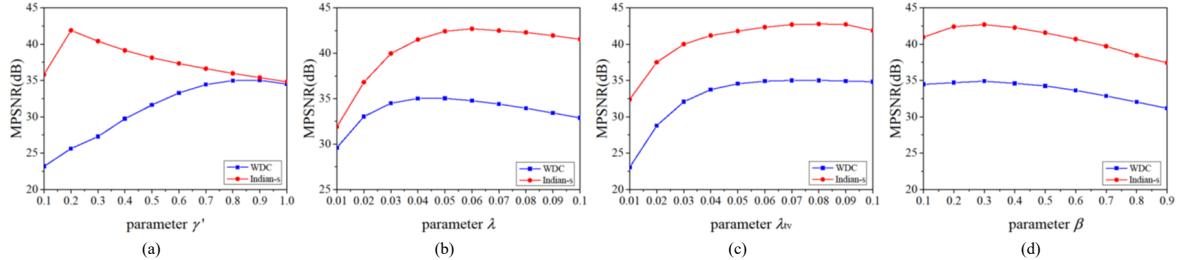


Fig. 15: MPSNR as a function of the parameters (i.e., γ' , λ , λ_{tv} and β) in the proposed method for HSI denoising on the WDC and Indian-s datasets. (a) γ' , (b) λ , (c) λ_{tv} , (d) β .

TABLE VIII: Running time comparison of different methods.

HSI data sets	size	2DTV	ASSTV	SSTV	LRMR	HSSTV1	HSSTV2	E-3DTV	l_0-l_1 HTV
WDC	$200 \times 200 \times 160$	314.5s	109.9s	564.3s	131.6s	1710.7s	1805.6s	390.5s	571.9s
Indian-s	$145 \times 145 \times 175$	189.9s	57.4s	432.3s	99.2s	1054.2s	1061.8s	193.10s	349.6s
Indian Pines	$145 \times 145 \times 220$	235.5s	72.5s	466.5s	115.1s	1124.6s	1286.3s	211.8s	400.6s
HSI data sets	size	2DTV	ASSTV	SSTV	SparCS	HSSTV1	HSSTV2	E-3DTV	l_0-l_1 HTV
Pavia University	$250 \times 121 \times 103$	832.1s	1247.7s	941.3s	1516.9s	1126.1s	1274.9s	288.1s	778.3s
Toy	$300 \times 300 \times 31$	1025.7s	1303.7s	1066.2s	1355.1s	1334.9s	1465.4s	307.4s	819.7s
Reno	$150 \times 150 \times 100$	708.9s	1012.1s	734.7s	1027.5s	1028.4s	1135.9s	267.7s	679.4s

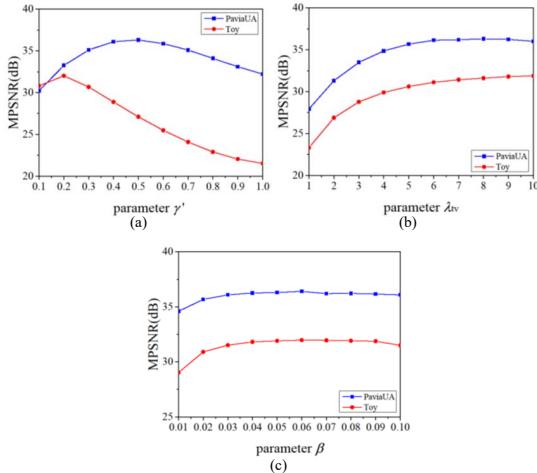


Fig. 16: MPSNR as a function of the parameters (i.e., γ' , λ , λ_{tv} and β) in the proposed method for HSI CS on the PaviaUA and Toy datasets. (a) γ' , (b) λ , (c) λ_{tv} , (d) β .

and CS reconstruction experiments are reported in Table VIII. All of the experiments are conducted in Matlab 2018b on the PC with an Intel i7 CPU at 2.60 GHz and 32 GB of memory. For HSI denoising, ASSTV is the fastest one among all the comparing methods, but most of the mixed noise remains. LRMR is the second-best and hardly handles heavy Gaussian cases. 2DTV and E-3DTV spend similar computational time because the update processing of 2DTV is simple, and E-3DTV converges quickly. HSSTV uses two local differences and has a slow convergence, which leads to an expensive computation cost. In contrast, although l_0-l_1 HTV contains two TV norms: l_0 hTV and SSTV, the running times of l_0-l_1 HTV and SSTV are similar since the pure SSTV model converges slowly. For HSI CS, ASSTV, SparCS, and HSSTV are more

computationally expensive. Thanks to fast convergence speed, E-3DTV is the fastest one among all the CS algorithms. Our proposed method is slightly slower than SSTV. However, l_0-l_1 HTV obtains the best performances in the HSI restoration experiments.

VII. CONCLUSION

In this paper, we have proposed a novel l_0-l_1 HTV regularization for HSI denoising and CS restoration. l_0 hyperspectral TV restrains the spatial smoothness from all spectral bands and SSTV exploits local structure information along spatial and spectral directions. The hybrid TV is designed as a globally and locally integrated regularizer for HSI restorations. Experimental results on HSI mixed noise removal and CS recovery demonstrate the superior performances of our proposed methods.

For further work, two extensions of our regularization technique can be explored: (1) l_0-l_1 HTV can be applied to other HSI processing tasks, like classification. (2) In this study, the simple use of l_0-l_1 HTV has been focused. In fact, this kind of HTV can be injected into low-rank matrix or tensor models.

REFERENCES

- [1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
- [2] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sens.*, vol. 10, 2018.
- [3] B. Rasti, D. Hong, R. Hang, P. Ghamisi, X. Kang, J. Chanussot, and J. Benediktsson, "Feature extraction for hyperspectral imagery: The evolution from shallow to deep: Overview and toolbox," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 4, pp. 60–88, 2020.
- [4] Y. Gu, J. Chanussot, X. Jia, and J. A. Benediktsson, "Multiple kernel learning for hyperspectral image classification: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6547–6565, Nov. 2017.

- [5] D. Hong, X. Wu, P. Ghamisi, J. Chanussot, N. Yokoya, and X. Zhu, “Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 3791–3808, 2020.
- [6] F. Liu and Q. Wang, “A sparse tensor-based classification method of hyperspectral image,” *Signal Process.*, vol. 168, p. 107361, 2020.
- [7] 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. Top. Appl. Earth Obs. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [8] D. Hong, N. Yokoya, J. Chanussot, and X. Zhu, “An augmented linear mixing model to address spectral variability for hyperspectral unmixing,” *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, 2019.
- [9] H. Zhang, L. Zhang, and H. Shen, “A super-resolution reconstruction algorithm for hyperspectral images,” *Signal Process.*, vol. 92, pp. 2082–2096, 2012.
- [10] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, “Nonlocal patch tensor sparse representation for hyperspectral image super-resolution,” *IEEE Trans. Image Process.*, vol. 28, no. 6, 2019.
- [11] D. Hong, N. Yokoya, N. Ge, J. Chanussot, and X. Zhu, “Learnable manifold alignment (lema): A semi-supervised cross-modality learning framework for land cover and land use classification,” *ISPRS J. Photogramm. Remote Sens.*, vol. 147, pp. 193–205, 2019.
- [12] D. Hong, N. Yokoya, J. Chanussot, J. Xu, and X. Zhu, “Cospace: Common subspace learning from hyperspectral-multispectral correspondences,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4349–4359, 2019.
- [13] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, “Joint reconstruction and anomaly detection from compressive hyperspectral images using mahalanobis distance-regularized tensor RPCA,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2919–2930, May 2018.
- [14] X. Wu, D. Hong, J. Tian, J. Chanussot, W. Li, and R. Tao, “Orsim detector: A novel object detection framework in optical remote sensing imagery using spatial-frequency channel features,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, 2019.
- [15] X. Wu, D. Hong, J. Chanussot, Y. Xu, R. Tao, and Y. Wang, “Fourier-based rotation-invariant feature boosting: An efficient framework for geospatial object detection,” *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 302–306, 2020.
- [16] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D*, vol. 60, no. 1-4, pp. 259–268, 1992.
- [17] D. M. Strong, P. Blomgren, and T. F. Chan, “Spatially adaptive local-feature-driven total variation minimizing image restoration,” in *Proc. Conf. Stat. Stochastic Methods Image Process. II*, vol. 3167, 1997, pp. 222–233.
- [18] A. Chopra and H. Lian, “Total variation, adaptive total variation and nonconvex smoothly clipped absolute deviation penalty for denoising blocky images,” *Pattern Recognit.*, vol. 43, no. 8, pp. 2609–2619, 2010.
- [19] D. Needell and R. Ward, “Stable image reconstruction using total variation minimization,” *SIAM J. Imaging Sciences*, vol. 6, no. 2, pp. 1035–1058, 2013.
- [20] C. Poon, “On the role of total variation in compressed sensing,” *SIAM J. on Imaging Sciences*, vol. 8, no. 1, pp. 682–720, 2015.
- [21] D. Needell and R. Ward, “Near-optimal compressed sensing guarantees for total variation minimization,” *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3941–3949, 2013.
- [22] Q. Yuan, 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.
- [23] C. Li, T. Sun, K. F. Kelly, and Y. Zhang, “A compressive sensing and unmixing scheme for hyperspectral data processing,” *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1200–1210, 2011.
- [24] D. T. Eason and M. Andrews, “Total variation regularization via continuation to recover compressed hyperspectral images,” *IEEE trans. image process.*, vol. 24, no. 1, pp. 284–293, 2014.
- [25] S. Ono, K. Shirai, and M. Okuda, “Vectorial total variation based on arranged structure tensor for multichannel image restoration,” in *IEEE Int. Conf. Acoust. Speech Signal Process.* IEEE, 2016, pp. 4528–4532.
- [26] Y. Chang, L. Yan, H. Fang, and H. Liu, “Simultaneous despeckling and denoising for remote sensing images with unidirectional total variation and sparse representation,” *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 6, pp. 1051–1055, 2013.
- [27] 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, Mar. 2016.
- [28] S. Takeyama, S. Ono, and I. Kumazawa, “Hybrid spatio-spectral total variation: A regularization technique for hyperspectral image denoising and compressed sensing,” *arXiv preprint arXiv:1907.13357*, 2019.
- [29] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, “Hyperspectral image restoration using low-rank matrix recovery,” *IEEE Trans. on Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.
- [30] 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, Jan. 2016.
- [31] Z. Wu, Q. Wang, Z. Wu, and Y. Shen, “Total variation-regularized weighted nuclear norm minimization for hyperspectral image mixed denoising,” *J. Electron. Imag.*, vol. 25, no. 1, 2016.
- [32] M. Golbabaei and P. Vandergheynst, “Joint trace/tv norm minimization: A new efficient approach for spectral compressive imaging,” in *IEEE Int. Conf. Image Process.*, 2012, pp. 933–936.
- [33] 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.
- [34] Q. Wang, Z. Wu, J. Jin, T. Wang, and Y. Shen, “Low rank constraint and spatial spectral total variation for hyperspectral image mixed denoising,” *Signal Process.*, vol. 142, pp. 11–26, 2018.
- [35] M. Wang, Q. Wang, J. Chanussot, and D. Li, “Hyperspectral image mixed noise removal based on multidirectional low-rank modeling and spatial-spectral total variation,” *IEEE Trans. Geosci. Remote Sens.*, pp. 1–20, 2020, DOI: 10.1109/TGRS.2020.2993631.
- [36] Y. Chen, W. He, N. Yokoya, and T. Z. Huang, “Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition,” *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3556–3570, 2020.
- [37] W. He, H. Zhang, H. Shen, and L. Zhang, “Hyperspectral image denoising using local low-rank matrix recovery and global spatial-spectral total variation,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 3, pp. 713–729, 2018.
- [38] 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. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 4, pp. 1227–1243, Apr. 2018.
- [39] J. Peng, Q. Xie, Q. Zhao, Y. Wang, L. Yee, and D. Meng, “Enhanced 3tv regularization and its applications on hsi denoising and compressed sensing,” *IEEE Trans. Image Process.*, vol. 29, pp. 7889–7903, 2020.
- [40] Y. Wang, L. Lin, Q. Zhao, T. Yue, D. Meng, and Y. Leung, “Compressive sensing of hyperspectral images via joint tensor tucker decomposition and weighted total variation regularization,” *IEEE Geo. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2457–2461, 2017.
- [41] X. Li, C. L. Y. Xu, and J. Jia, “Image smoothing via ℓ_0 gradient minimization,” *ACM Trans. Graph.*, vol. 30, no. 6, pp. 174:1–174:12, Dec. 2011.
- [42] L. Xu, S. Zheng, and J. Jia, “Unnatural ℓ_0 sparse representation for natural image deblurring,” in *Pro. IEEE conf. comput. vis. pattern recognit. (CVPR)*, 2013, pp. 1107–1114.
- [43] H. Xue, S. Zhang, and D. Cai, “Depth image inpainting: Improving low rank matrix completion with low gradient regularization,” *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4311–4320, 2017.
- [44] M. Wang, Q. Wang, and J. Chanussot, “ ℓ_0 gradient regularized low-rank tensor model for hyperspectral image denoising,” in *10th Workshop Hyperspectral Imag. Signal Process., Evol. Remote Sens. (WHISPERS)*, Sep. 2019, pp. 1–6.
- [45] F. Xiong, J. Zhou, and Y. Qian, “Hyperspectral restoration via ℓ_0 gradient regularized low-rank tensor factorization,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10410–10425, 2019.
- [46] S. Ono, “ ℓ_0 gradient projection,” *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1554–1564, 2017.
- [47] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, T.-X. Jiang, T.-Y. Ji, and T.-H. Ma, “Tensor n -tubal rank and its convex relaxation for low-rank tensor recovery,” *arXiv preprint arXiv:1812.00688*, 2018.
- [48] 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, Jan. 2011.
- [49] J. Kang, D. Hong, J. Liu, G. Baier, N. Yokoya, and B. Demir, “Learning convolutional sparse coding on complex domain for interferometric phase restoration,” *IEEE Trans. Neural Netw. Learn. Syst.*, 2020, doi: 10.1109/TNNLS.2020.2979546.
- [50] C. C. Paige and M. A. Saunders, “LSQR: An algorithm for sparse linear equations and sparse least squares,” *Acm Trans. Math. Softw.*, vol. 8, no. 1, pp. 43–71, 1982.

- [51] D. Hong, N. Yokoya, J. Chanussot, J. Xu, and X. X. Zhu, "Joint and progressive subspace analysis (jpsa) with spatial-spectral manifold alignment for semi-supervised hyperspectral dimensionality reduction," *IEEE Trans. Cybern.*, 2020, dOI: 10.1109/TCYB.2020.3028931.
- [52] R. Barrett, M. Berry, T. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. V. der Vorst, "Templates for the solution of linear systems: Building blocks for iterative methods," *SIAM*, vol. 43, 1994.
- [53] W. Cao, Y. Wang, J. Sun, D. Meng, C. Yang, A. Cichocki, and Z. Xu, "Total variation regularized tensor rpca for background subtraction from compressive measurements," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4075–4090, 2016.
- [54] A. Gogna, A. Shukla, H. Agarwal, and A. Majumdar, "Split bregman algorithms for sparse/joint-sparse and low-rank signal recovery: Application in compressive hyperspectral imaging," in *IEEE Int. Conf. Image Process.* IEEE, 2014, pp. 1302–1306.
- [55] M. Golbabaei and P. Vandergheynst, "Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery," in *IEEE Int. Conf. Acoust. Speech Signal Process.*, 2012, pp. 2741–2744.
- [56] D. Hong, L. Gao, N. Yokoya, J. Yao, J. Chanussot, D. Qian, and B. Zhang, "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, 2020, dOI: 10.1109/TGRS.2020.3016820.
- [57] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, pp. 1778–1790, Aug. 2004.
- [58] D. Hong, L. Gao, J. Yao, B. Zhang, P. Antonio, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, 2020, dOI: 10.1109/TGRS.2020.3015157.



Jocelyn Chanussot (Fellow, IEEE) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from the Universit de Savoie, Annecy, France, in 1998. Since 1999, he has been with Grenoble INP, where he is currently a Professor of signal and image processing. His research interests include image analysis, hyperspectral remote sensing, data fusion, machine learning and artificial intelligence. He has been a visiting scholar at Stanford University (USA), KTH (Sweden) and NUS (Singapore). Since 2013, he is an Adjunct Professor of the University of Iceland. In 2015-2017, he was a visiting professor at the University of California, Los Angeles (UCLA). He holds the AXA chair in remote sensing and is an Adjunct professor at the Chinese Academy of Sciences, Aerospace Information research Institute, Beijing. Dr. Chanussot is the founding President of IEEE Geoscience and Remote Sensing French chapter (2007-2010) which received the 2010 IEEE GRSS Chapter Excellence Award. He has received multiple outstanding paper awards. He was the Vice-President of the IEEE Geoscience and Remote Sensing Society, in charge of meetings and symposia (2017-2019). He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair (2009-2011) and Cochair of the GRS Data Fusion Technical Committee (2005-2008). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006-2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing (2009). He is an Associate Editor for the IEEE Transactions on Geoscience and Remote Sensing, the IEEE Transactions on Image Processing and the Proceedings of the IEEE. He was the Editor-in-Chief of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2011-2015). In 2014 he served as a Guest Editor for the IEEE Signal Processing Magazine. He is a Fellow of the IEEE, a member of the Institut Universitaire de France (2012-2017) and a Highly Cited Researcher (Clarivate Analytics/Thomson Reuters, 2018-2019).



Minghua Wang (Student Member, IEEE) received the B.S. from the School of automation, Harbin Institute of Technology (HIT), Harbin, China, in 2016 and she is currently working toward the Ph.D. degree in Control Science and Engineering. She is also a visiting Ph.D student in the Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France (2019-2020). Her research interests include Hyperspectral image denoising, tensor completion and low-rank representation.



Danfeng Hong (S'16–M'19) received the M.Sc. degree (summa cum laude) in computer vision, College of Information Engineering, Qingdao University, Qingdao, China, in 2015, the Dr. -Ing degree (summa cum laude) in Signal Processing in Earth Observation (SiPEO), Technical University of Munich (TUM), Munich, Germany, in 2019.

Since 2015, he worked as a Research Associate at the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany. Currently, he is a research scientist and leads a Spectral Vision working group at IMF, DLR, and also an adjunct scientist in GIPSA-lab, Grenoble INP, CNRS, Univ. Grenoble Alpes, Grenoble, France.

His research interests include signal / image processing and analysis, hyperspectral remote sensing, machine / deep learning, artificial intelligence and their applications in Earth Vision.



Qiang Wang (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in control science and engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 1998, 2000, and 2004, respectively. Since 2008, he has been a Professor with the Department of Control Science and Engineering, HIT. His research interests include Hyperspectral image denoising, signal/image processing, multisensor data fusion, wireless sensor networks, and intelligent detection technology.