HYPERSPECTRAL AND MULTISPECTRAL DATA FUSION BY A REGULARIZATION CONSIDERING

Saori Takeyama[†], *Shunsuke Ono*[†], *and Itsuo Kumazawa*[†]

[†]Tokyo Institute of Technology

ABSTRACT

A hyperspectral (HS) image has high spectral resolution information but low spatial resolution information. To get an HS image of high spatial and spectral resolution (high-spatial HS image), fusion techniques are actively studied, which synthesize an HS image of low spatial and high spectral resolution and a multispectral (MS) image. The techniques can generate a high-spatial HS image by exploiting the high spectral and spatial resolution information of HS and MS images, respectively. However, the methods do not evaluate the edge similarity between generated HS and observed MS images, and do not denoise the MS image. As a result, when an observed MS image is noisy, these methods produce artifacts and spectral distortion. To tackle this problem, we propose a new HS and MS data fusion method using a hybrid spatio-spectral total variation (HSSTV), which is a regularization for HS image restoration. The method not only generates a high-spatial HS image but also denoises a given MS image, so that we obtain a high-spatial HS image even if the observed images are contaminated by noise. In the experiments, we demonstrate the advantages of our method over existing fusion methods and the effectiveness of HSSTV for MS image restoration.

Index Terms— hyperspectral image, multispectral image, image fusion, primal-dual splitting

1. INTRODUCTION

Hyperspectral (HS) imaging is a powerful technology as observing the material characteristic beyond one's vision. An HS image has invisible and narrow interval spectral information, and is used in many applications in a wide range of fields, e.g., agriculture, medical science and remote sensing [1, 2]. Since high spatial resolution information localizes the object, and high spectral resolution information stands for the true material property, the applications require an HS image of high spatial and spectral resolution (a high-spatial HS image). However, capturing a high-spatial HS image is a very challenging task for tradeoffs between both resolutions.

HS and multispectral (MS) data fusion merges images of different spatial and spectral resolutions, i.e., an HS image of low spatial resolution (a low-spatial HS image) and an MS image, to generate a high-spatial HS image. Here, an MS image has high spatial but low spectral resolution. These techniques can solve the above problem, so that it has been actively studied [3].

The methods of HS and MS data fusion create a high-spatial HS image by utilizing spatial information in the MS image and spectral information in the low-spatial HS image. In [4, 5], the methods are based on Gram-Schmidt (GS) or multiresolution analysis (MRA), and fusion low-spatial HS and MS images bandwise. However, they do not use a-priori knowledge on HS images (especially spectral piecewise smoothness), so that the results have artifacts. On the other hands, some fusion methods use a-priori knowledge to generate a high-spatial HS image. As a result, they achieve high-quality fusion, which are based on unmixing [6–9], Bayesian probability [10, 11], and both [12]. However, since the methods do not denoise the MS image, the results have artifacts and spectral distortion when the MS image is noisy.

To overcome the above problems, we propose a novel method of HS and MS data fusion. The method fuses them by solving a newly-formulated convex optimization problem. It simultaneously generates a high-spatial HS image and denoises an MS image for high-quality fusion. In addition, the proposed method can restore detailed structure, because it evaluates edge similarity between them. We provide an algorithm to solve the problem based on a primal-dual splitting method [13]. Our experiments illustrate the strength of the proposed method by comparison with existing methods.

2. PROPOSED METHOD

2.1. Observation Model

Let $\bar{\mathbf{u}} \in \mathbb{R}^{NB_h}$ be a true high-spatial HS image, where N is the number of pixel, and B_h is a spectral bands number of the HS image. For HS and MS fusion, we assume observation models of a low-spatial HS image $\mathbf{v}_h$ and an observed MS image $\mathbf{v}_m$ as follows:

$$\mathbf{v}_h = \mathbf{S}\mathbf{B}\bar{\mathbf{u}} + \mathbf{n}_h \in \mathbb{R}^{\frac{NB_h}{r}}, \tag{1}$$

$$\mathbf{v}_m = \mathbf{R}\bar{\mathbf{u}} + \mathbf{n}_m \in \mathbb{R}^{NB_m}, \tag{2}$$

where $\mathbf{S} \in \mathbb{R}^{\frac{NB_h}{r} \times NB_h}$ is a downsampling matrix with a downsampling rate of r that divides N, $\mathbf{B}$ is a blur matrix, $\mathbf{n}_h$ and $\mathbf{n}_m$ are additive white Gaussian noises with standard deviation σ_h and σ_m, respectively, and B_m is a number of spectral bands of the MS image. Since an MS image contains less noise than an HS image, we assume $\sigma_h > \sigma_m$. Moreover, $\mathbf{R} \in \mathbb{R}^{NB_m \times NB_h}$ is a matrix regarding the property of MS sensor. Specifically, $\mathbf{R}$ calculates partial weighted average along the spectral direction. In general, since HS and MS images are contaminated by noise, the setting is realistic.

2.2. Problem Formulation

We propose an HS and MS fusion problem using a hybrid spatio-spectral total variaton (HSSTV) [14]. Here, HSSTV is a regularization function for HS image restoration, and we proposed it at ICASSP 2017. The regularization simultaneously evaluates both spatio-spectral piecewise smoothness and direct spatial piecewise smoothness, and it is defined as follows:

$$\mathrm{HSSTV}(\mathbf{u}, \omega) := \|\mathbf{A}_\omega \mathbf{u}\|_{1,p} := \left\| \begin{bmatrix} \mathbf{D}\mathbf{D}_b \\ \omega\mathbf{D} \end{bmatrix} \mathbf{u} \right\|_{1,p},$$

The work was partially supported by JSPS Grants-in-Aid (18H05413, 18J20290, 17K12710) and JST-PRESTO.

978-1-5386-4658-8/18/$31.00 ©2019 IEEE

Algorithm 1: PDS method for Prob. (3)

```
input : $\mathbf{u}^{(0)}, \mathbf{y}_1^{(0)}, \mathbf{y}_2^{(0)}, \mathbf{y}_3^{(0)}, \gamma_1, \gamma_2, \mathbf{v}_h, \mathbf{v}_m, \varepsilon, \eta$
while A stopping criterion is not satisfied do do
    $\mathbf{u}_h^{(n+1)} = \mathrm{prox}_{\gamma_1, \iota_{[\underline{\mu_{\mathbf{u}_h}}, \overline{\mu_{\mathbf{u}_h}}]^{NB_h}}}(\mathbf{u}_h^{(n)} - \gamma_1(\mathbf{A}_{\omega_h}^\top \mathbf{y}_1^{(n)} + \mathbf{D}^\top \mathbf{y}_2^{(n)} + \mathbf{B}^\top \mathbf{S}^\top \mathbf{y}_4^{(n)}))$;
    $\mathbf{u}_m^{(n+1)} = \mathrm{prox}_{\gamma_1, \iota_{[\underline{\mu_{\mathbf{u}_m}}, \overline{\mu_{\mathbf{u}_m}}]^{NB_m}}}(\mathbf{u}_m^{(n)} - \gamma_1(-\mathbf{M}^\top \mathbf{D}^\top \mathbf{y}_2 + \mathbf{A}_{\omega_m}^\top \mathbf{y}_3^{(n)} + \mathbf{y}_5^{(n)}))$;
    $\mathbf{y}_1^{(n)} \leftarrow \mathbf{y}_1^{(n)} + \gamma_2 \mathbf{A}_{\omega_h}(2\mathbf{u}_h^{(n+1)} - \mathbf{u}_h^{(n)})$;
    $\mathbf{y}_2^{(n)} \leftarrow \mathbf{y}_2^{(n)} + \gamma_2(\mathbf{D}(2\mathbf{u}_h^{(n+1)} - \mathbf{u}_h^{(n)}) - \mathbf{DM}(2\mathbf{u}_m^{(n+1)} - \mathbf{u}_m^{(n)}))$;
    $\mathbf{y}_3^{(n)} \leftarrow \mathbf{y}_3^{(n)} + \gamma_2 \mathbf{A}_{\omega_m}(2\mathbf{u}_m^{(n+1)} - \mathbf{u}_m^{(n)})$;
    $\mathbf{y}_4^{(n)} \leftarrow \mathbf{y}_4^{(n)} + \gamma_2 \mathbf{SB}(2\mathbf{u}_h^{(n+1)} - \mathbf{u}_h^{(n)})$;
    $\mathbf{y}_5^{(n)} \leftarrow \mathbf{y}_5^{(n)} + \gamma_2(2\mathbf{u}_m^{(n+1)} - \mathbf{u}_m^{(n)})$;
    $\mathbf{y}_1^{(n+1)} = \mathbf{y}_1^{(n)} - \gamma_2 \, \mathrm{prox}_{\frac{1}{\gamma_2}, \|\cdot\|_{1,p}}\left(\frac{\mathbf{y}_1^{(n)}}{\gamma_2}\right)$;
    $\mathbf{y}_2^{(n+1)} = \mathbf{y}_2^{(n)} - \gamma_2 \, \mathrm{prox}_{\frac{\lambda}{\gamma_2}, \|\cdot\|_{1,2}}\left(\frac{\mathbf{y}_2^{(n)}}{\gamma_2}\right)$;
    $\mathbf{y}_3^{(n+1)} = \mathbf{y}_3^{(n)} - \gamma_2 \, \mathrm{prox}_{\frac{1}{\gamma_2}, \|\cdot\|_{1,p}}\left(\frac{\mathbf{y}_3^{(n)}}{\gamma_2}\right)$;
    $\mathbf{y}_4^{(n+1)} = \mathbf{y}_4^{(n)} - \gamma_2 \, \mathrm{prox}_{\frac{1}{\gamma_2}, \iota_{\mathcal{B}_{2,\varepsilon}^{\mathbf{v}_h}}}\left(\frac{\mathbf{y}_4^{(n)}}{\gamma_2}\right)$;
    $\mathbf{y}_5^{(n+1)} = \mathbf{y}_5^{(n)} - \gamma_2 \, \mathrm{prox}_{\frac{1}{\gamma_2}, \iota_{\mathcal{B}_{2,\eta}^{\mathbf{v}_m}}}\left(\frac{\mathbf{y}_5^{(n)}}{\gamma_2}\right)$;
    $n \leftarrow n + 1$;
```

where $\mathbf{D} := (\mathbf{D}_v^\top \mathbf{D}_h^\top)^\top$ is a spatial difference operator, $\mathbf{D}_v$, $\mathbf{D}_h$ and $\mathbf{D}_b$ are a vertical, horizontal and spectral difference operator, respectively, and $\| \cdot \|_{1,p}$ is the mixed $\ell_{1,p}$ norm (in [14], $p = 1$ or 2). The parameter ω balances between a spatio-spectral and a direct spatial difference operator, and by setting it suitable value, it can restore a higher quality HS image than existing regularization for HS image restoration.

Based on the model in Sec. 2.1, an HS and MS fusion problem using HSSTV is formulated by

$$\begin{aligned} &\min_{\mathbf{u}_h, \mathbf{u}_m} \mathrm{HSSTV}(\mathbf{u}_h, \omega_h) + \lambda \|\mathbf{Du}_h - \mathbf{DMu}_m\|_{1,2} \\ &+ \mathrm{HSSTV}(\mathbf{u}_m, \omega_m) \\ &\text{s.t.} \left[\begin{array}{l} \mathbf{SBu}_h \in \mathcal{B}_{2,\varepsilon}^{\mathbf{v}_h} := \{\mathbf{x} \in \mathbb{R}^{\frac{NB_h}{r}} \,|\, \|\mathbf{x} - \mathbf{v}_h\| \leq \varepsilon\}, \\ \mathbf{u}_m \in \mathcal{B}_{2,\eta}^{\mathbf{v}_m} := \{\mathbf{x} \in \mathbb{R}^{NB_m} \,|\, \|\mathbf{x} - \mathbf{v}_m\| \leq \eta\}, \\ \mathbf{u}_h \in [\underline{\mu_{\mathbf{u}_h}}, \overline{\mu_{\mathbf{u}_h}}]^{NB_h}, \\ \mathbf{u}_m \in [\underline{\mu_{\mathbf{u}_m}}, \overline{\mu_{\mathbf{u}_m}}]^{NB_m}, \end{array} \right. \end{aligned} \tag{3}$$

where $\| \cdot \|_{1,2}$ is the mixed $\ell_{1,2}$ norm, and $\mathbf{M} \in \mathbb{R}^{NB_h \times NB_m}$ is a linear operator that expands the MS image along the spectral direction. The problem not only generates a high-spatial HS image $\mathbf{u}_h$ but also denoises an MS image $\mathbf{u}_m$, leading to high-quality fusion.

The first and third terms in Prob. (3) are HSSTV. The problem uses HSSTV for both denoising an MS image and generating a high-spatial HS image. It is because an MS image has a spatio-spectral piecewise smoothness similar to an HS image. Since the balance between both piecewise smoothness would be different, the ω_h and ω_m are independently set. The second term in Prob. (3) evaluates edge similarity between HS and MS images by taking differences between edge images of generated HS and denoised MS images. Here, we assume that the non-zero differences of the high-spatial HS image are a sparse edge image, and the edges of the high-spatial HS and MS images locate in the same position.

The first and second constraints in (3) serves as data-fidelity to the low-spatial HS image $\mathbf{v}_h$ and the observed MS image $\mathbf{v}_m$. The set $\mathcal{B}_{2,\varepsilon}^{\mathbf{v}_h}$ is defined as the $\mathbf{v}_h$-centered ℓ_2-norm ball with the radius $\varepsilon > 0$, and $\mathcal{B}_{2,\eta}^{\mathbf{v}_m}$ is also defined as the $\mathbf{v}_m$-centered ℓ_2-norm ball with the radius $\eta > 0$. In the hard constraints type of problem, the parameter settings are relatively easy because they have a clear meaning. It is mentioned by [14–18]. The third and forth constraints in (3) are the dynamic range of HS and MS images, respectively.

2.3. Optimization

We adopt a primal-dual splitting method [13] to solve Prob. (3). It is because the problem is a highly nonsmooth convex optimization problem. The algorithm can solve convex optimization problems of the form:

$$\min_{\mathbf{x}} g(\mathbf{x}) + h(\mathbf{Lx}), \tag{4}$$

where $\mathbf{L}$ is a linear operator. In addition, g and h are proper lower semicontinuous convex functions and are assumed to be *proximable*. Here, *proximable* means that the proximity operators[1] [19] of g and h are computable.

For any $\mathbf{y}^{(0)}$, the algorithm is given by

$$\left\lfloor \begin{array}{l} \mathbf{u}^{(n+1)} = \mathrm{prox}_{\gamma_1 g}(\mathbf{u}^{(n)} - \gamma_1 \mathbf{L}^\top \mathbf{y}^{(n)}), \\ \mathbf{y}^{(n+1)} = \mathrm{prox}_{\gamma_2 h^*}(\mathbf{y}^{(n)} + \gamma_2 \mathbf{L}(2\mathbf{u}^{(n+1)} - \mathbf{u}^{(n)})), \end{array} \right.$$

where $\gamma_1, \gamma_2 > 0$ are stepsizes of the primal-dual splitting method, and need to satisfy $\gamma_1 \gamma_2 \|\mathbf{L}\|_{op}^2 \leq 1$ ($\| \cdot \|_{op}$ is the operator norm) to converge the algorithm. The function h^* is the convex conjugate of h, and the proximity operator of h^* is available via that of h [20, Theorem 14.3(ii)] as follows:

$$\mathrm{prox}_{\gamma h^*}(\mathbf{x}) = \mathbf{x} - \gamma \, \mathrm{prox}_{\frac{1}{\gamma} h}(\tfrac{1}{\gamma} \mathbf{x}). \tag{5}$$

To solve Prob. (3) using the primal-dual splitting method, we reformulate Prob. (3) into Prob. (4).

First, since the four hard constraint in Prob. (3) are put into the objective function, we introduce the indicator function, which is defined by

$$\iota_C(\mathbf{x}) := \begin{cases} 0, & \text{if } \mathbf{x} \in C, \\ \infty, & \text{otherwise,} \end{cases}$$

where C is a nonempty closed convex set. With this in mind, Prob. (3) can be rewritten as

$$\begin{aligned} &\min_{\mathbf{u}_h, \mathbf{u}_m} \|\mathbf{A}_{\omega_h} \mathbf{u}_h\|_{1,p} + \lambda \|\mathbf{Du}_h - \mathbf{DMu}_m\|_{1,2} + \|\mathbf{A}_{\omega_m} \mathbf{u}_m\|_{1,p} \\ &+ \iota_{\mathcal{B}_{2,\varepsilon}^{\mathbf{v}_h}}(\mathbf{SBu}_h) + \iota_{\mathcal{B}_{2,\eta}^{\mathbf{v}_m}}(\mathbf{u}_m) + \iota_{[\underline{\mu_{\mathbf{u}_h}}, \overline{\mu_{\mathbf{u}_h}}]^{NB_h}}(\mathbf{u}_h) \\ &+ \iota_{[\underline{\mu_{\mathbf{u}_m}}, \overline{\mu_{\mathbf{u}_m}}]^{NB_m}}(\mathbf{u}_h). \end{aligned} \tag{6}$$

As the definition of the indicator function, Prob. (3) and Prob. (6) are equivalent.

---

[1] The proximity operator of index $\gamma > 0$ of a proper lower semicontinuous convex function f is defined by $\mathrm{prox}_{\gamma f}(\mathbf{x}) := \underset{\mathbf{y}}{\mathrm{argmin}} \, f(\mathbf{y}) + \frac{1}{2\gamma} \|\mathbf{y} - \mathbf{x}\|^2$.

Table 1. Quality measures for $\sigma_m = 0.04$ (left) and $\sigma_m = 0.06$ (right) averaged all HS images.

	$\sigma_m = 0.04$				$\sigma_m = 0.08$			
	PSNR	SAM	ERGAS	$Q2^n$	PSNR	SAM	ERGAS	$Q2^n$
MTF-GLP [5]	22.80	10.45	7.557	0.4937	22.28	11.09	7.905	0.4649
CNMF [6]	26.40	9.139	6.104	0.6264	24.25	10.68	7.211	0.5273
HySure [12]	28.65	8.354	7.511	0.6836	27.93	7.676	7.448	0.6550
MAPSMM [10]	24.06	10.39	14.00	0.4973	23.49	11.12	14.53	0.4658
proposed (HTV, $p = 1$**)**	29.79	5.103	4.459	0.7110	28.64	5.277	4.931	0.6790
proposed (HTV, $p = 2$**)**	29.89	**5.020**	**4.386**	0.7163	**28.70**	**5.209**	**4.866**	0.6832
proposed (HSSTV, $p = 1$**)**	29.86	5.774	5.062	0.7179	28.56	6.062	5.569	0.6829
proposed (HSSTV, $p = 2$	**29.95**	5.711	5.006	**0.7231**	28.64	6.002	5.513	**0.6880**

By letting

$$g : \mathbb{R}^{N(B_h+B_m)} \to \mathbb{R}^2 : (\mathbf{u}_h, \mathbf{u}_m) \mapsto$$
$$(\iota_{[\underline{\mu_{\mathbf{u}_h}}, \overline{\mu_{\mathbf{u}_h}}]^{NB_h}}(\mathbf{u}_h), \iota_{[\underline{\mu_{\mathbf{u}_m}}, \overline{\mu_{\mathbf{u}_m}}]^{NB_m}}(\mathbf{u}_m)),$$
$$h : \mathbb{R}^{((6+\frac{1}{r})B_h+5B_m)N} \to \mathbb{R} \cup \{\infty\} : (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4, \mathbf{y}_5) \mapsto$$
$$\|\mathbf{y}_1\|_{1,p} + \lambda\|\mathbf{y}_2\|_{1,2} + \|\mathbf{y}_3\|_{1,p} + \iota_{\mathcal{B}_{2,\varepsilon}^{\mathbf{v}_h}}(\mathbf{y}_4) + \iota_{\mathcal{B}_{2,\eta}^{\mathbf{v}_m}}(\mathbf{y}_5),$$
$$\mathbf{L} : \mathbb{R}^{N(B_h+B_m)} \to \mathbb{R}^{((6+\frac{1}{r})B_h+5B_m)N} : (\mathbf{u}_h, \mathbf{u}_m) \mapsto$$
$$(\mathbf{A}_{\omega_h}\mathbf{u}_h, \mathbf{D}\mathbf{u}_h - \mathbf{D}\mathbf{M}\mathbf{u}_m, \mathbf{A}_{\omega_m}\mathbf{u}_m, \mathbf{S}\mathbf{B}\mathbf{u}_h, \mathbf{u}_m),$$

Prob. (6) come down to Prob. (4). By using (5), the algorithm for solving Prob. (3) is derived as Algorithm 1.

We explain the proximity operators in Alg. 1. First, the proximity operator of the indicator function of C is equivalent to the metric projection on to C, which are defined by

$$\mathrm{prox}_{\gamma\iota_C}(\mathbf{x}) := \min_{\mathbf{y}} \|\mathbf{x} - \mathbf{y}\| \text{ s.t. } \mathbf{x} \in C =: P_C(\mathbf{x}).$$

The proximity operators in steps 2, 3, 12, and 13 can be computed as follows: for $i = 1, \ldots, NB$

$$[\mathrm{prox}_{\gamma\iota_{[\underline{\mu_\cdot}, \overline{\mu_\cdot}]^{NB}}}(\mathbf{x})]_i = [P_{[\underline{\mu_\cdot}, \overline{\mu_\cdot}]^{NB}}(\mathbf{x})]_i$$
$$= \min\{\max\{x_i, \underline{\mu_\cdot}\}, \overline{\mu_\cdot}\},$$
$$\mathrm{prox}_{\gamma\iota_{\mathcal{B}_{2,\varepsilon}^{\mathbf{v}_\cdot}}}(\mathbf{x}) = P_{\mathcal{B}_{2,\varepsilon}^{\mathbf{v}_\cdot}}(\mathbf{x}) = \begin{cases} \mathbf{x}, & \text{if } \mathbf{x} \in \mathcal{B}_{2,\varepsilon}^{\mathbf{v}_\cdot}, \\ \mathbf{v}_\cdot + \frac{\varepsilon(\mathbf{x}-\mathbf{v}_\cdot)}{\|\mathbf{x}-\mathbf{v}_\cdot\|_2}, & \text{otherwise.} \end{cases}$$

Meanwhile, the proximity operators of the ℓ_1 norm and the mixed $\ell_{1,2}$ norm in steps 9, 10, and 11 are reduced to simple soft-thresholding type operations: for $\gamma > 0$ and for $i = 1, \ldots, 2NB$,

$$[\mathrm{prox}_{\gamma\|\cdot\|_1}(\mathbf{x})]_i = \mathrm{sgn}(x_i) \max\{|x_i| - \gamma, 0\},$$
$$[\mathrm{prox}_{\gamma\|\cdot\|_{1,2}}(\mathbf{x})]_i = \max\left\{1 - \gamma\left(\textstyle\sum_{j=0}^{1} x_{\tilde{i}+jNB_\cdot}^2\right)^{-\frac{1}{2}}, 0\right\} x_i,$$

where sgn is the sign function, and $\tilde{i} := ((i-1) \mod NB_\cdot) + 1$.

3. EXPERIMENTS

We demonstrate the advantages of the proposed method over existing HS and MS fusion methods. In the experience, first, we generated both low-spatial HS and MS images using a true high HS image based on (1) and (2), where $\mathbf{n}_h$ and $\mathbf{n}_m$ are additive white Gaussian noises with standard deviation σ_h and σ_m, respectively. At this time, we adopted 15 HS images from the *SpecTIR* [21], *MultiSpec* [22], *GIC* [23], *a Moffett field dataset* and *Chikusei* [24] as the true HS image, which are normalized its dynamic range into $[0, 1]$. In addition, the downsampling rate of $\mathbf{S}$ was set as $r = 4$, the kernel of $\mathbf{B}$ was a 9×9 Gaussian blur matrix, and $\mathbf{R}$ was a partial average matrix, which divides bands of an HS image into eight parts, and averages it in each group.

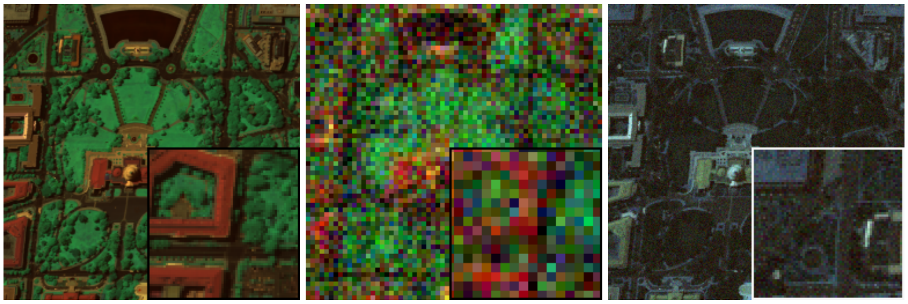
Second, we estimated a high-spatial HS image from the pair by each method. In this experiments, we utilize MTF-GLP [5], CNMF [6], HySure [12] and MAPSMM [10] as copared methods. Here, with respect to MTF-GLP and MAPSMM, all parameters were set to the recommended values in a MATLAB toolbox of HS and MS fusion [3]. In the case of HySure, we decided that its parameter regarding data fidelity $\lambda_\phi = 0.5(\sigma_h + \sigma_m)$, and for CNMF, the parameter was set to $\theta_h = \sigma_h$ and $\theta_m = \sigma_m$ for fair comparison. The other parameters in HySure and CNMF were set to the values of the same toolbox. For the proposed method, we decided that the parameters $\varepsilon = \|\mathbf{v}_h - \mathbf{S}\mathbf{B}\bar{\mathbf{u}}_h\|_2$, $\eta = \|\mathbf{v}_m - \mathbf{R}\bar{\mathbf{u}}_h\|_2$, $\omega_h = 0.02$, $\omega_m = 0.1$, and $\lambda = 0.3$. Besides, for verification of the availability of HSSTV regarding an MS image denoising, we experimented by the method replacing HSSTV of the third term in Prob. (3) with hyperspectral total variation (HTV) [25], i.e., $\mathrm{HSSTV}(\mathbf{u}_m, \omega_m)$ is replaced $\|\mathbf{D}\mathbf{u}_m\|_{1,2}$. The settings of max iteration number and the stopping criterion of Alg. 1 were 5000 and $\|\mathbf{u}_h^{(n)} - \mathbf{u}_h^{(n+1)}\|_2 / \|\mathbf{u}_h^{(n)}\|_2 < 1.0 \times 10^{-4}$, respectively.

Finally, we evaluated the estimated high HS images based on four standard quality measures: PSNR[dB], the Spectral Angle Mapper (SAM) [26], Erreur Relative Globale Adimensionnelle de synthèse (ERGAS) [27] and $Q2^n$ [28]. Here, to define SAM, and ERGAS, we assume that spectral and spatial vectors of $\mathbf{u}_h$ are $\mathbf{u}_h^i = [u_h^i, u_h^{i+N}, \ldots, u_h^{i+(B_h-1)N}] \in \mathbb{R}^B$ $(i = 1, \ldots, N)$ and $\mathbf{u}_h^{j\,*} = [u_h^{N(j-1)+1}, u_h^{N(j-1)+2}, \ldots, u_h^{N(j-1)+N}] \in \mathbb{R}^N$ $(j = 1, \ldots, B_h)$, respectively. The spectral and spatial vectors of $\bar{\mathbf{u}}_h$ are assumed alike. Then, PSNR, SAM and ERGAS are defined by

$$\mathrm{PSNR}(\mathbf{u}_h, \bar{\mathbf{u}}_h) = 10 \log_{10} \frac{NB_h}{\|\mathbf{u}_h - \bar{\mathbf{u}}_h\|_2}$$
$$\mathrm{SAM}(\mathbf{u}_h, \bar{\mathbf{u}}_h) = \frac{1}{N} \sum_{i=1}^{N} \arccos\left(\frac{\mathbf{u}_i^\top \bar{\mathbf{u}}_i}{\|\mathbf{u}_i\|_2 \|\bar{\mathbf{u}}_i\|_2}\right)$$
$$\mathrm{ERGAS}(\mathbf{u}_h, \bar{\mathbf{u}}_h) = \frac{100}{r} \sqrt{\frac{1}{B} \sum_{j=1}^{B} \frac{\|\mathbf{u}_j^* - \bar{\mathbf{u}}_j^*\|_2^2}{(\frac{1}{p}\mathbf{1}^\top \mathbf{u}_j^*)^{2'}}},$$

where $\mathbf{1} = [1, \ldots, 1] \in \mathbb{R}^N$. Note that the higher value of PSNR and $Q2^n$ are, the more similar $\mathbf{u}_h$ and $\bar{\mathbf{u}}_h$, and the lower value of SAM and ERGAS are, the more similar both images.

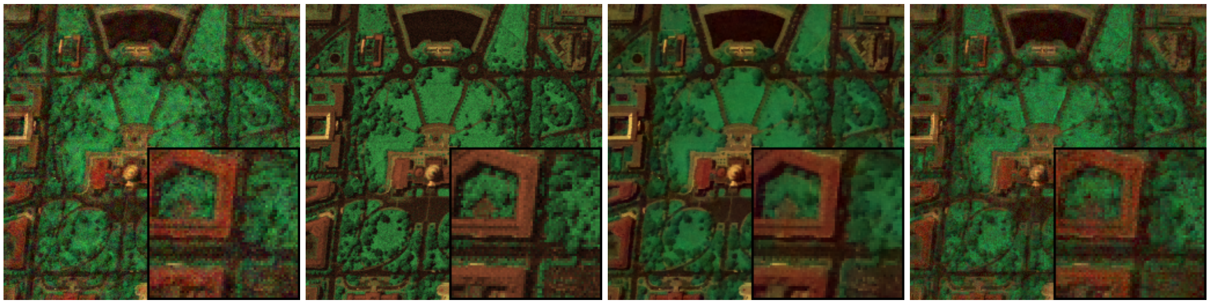
Table 1 shows PSNR[dB], SAM, ERGAS and $Q2^n$ of the



True HS image

Observed HS image

Observed MS image

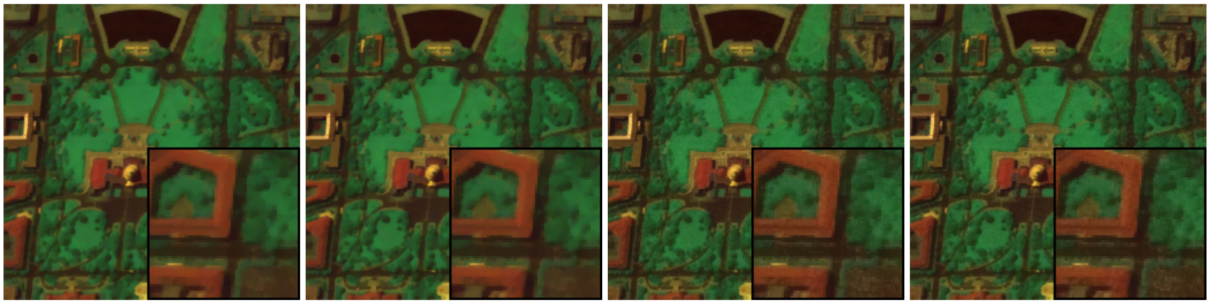


MTF-GLP $(19.74, 12.63, 10.32, 0.3373)$

CNMF $(22.97, 9.713, 8.825, 0.4808)$

HySure $(24.65, 8.670, 22.40, 0.4339)$

MAPSMM $(21.64, 10.26, 29.24, 0.2890)$



Proposed (HTV, $p = 1$) $(25.34, 7.520, 8.349, 0.9072)$

Proposed (HTV, $p = 2$) $(\mathbf{25.36}, \mathbf{7.500}, \mathbf{8.340}, \mathbf{0.9100})$

Proposed (HSSTV, $p = 1$) $(24.55, 8.369, 9.480, 0.9020)$

Proposed (HSSTV, $p = 2$) $(24.61, 8.316, 9.462, 0.9058)$

Fig. 1. Resulting HS images with four quality measure (PSNR, SAM, ERGAS, $Q2^n$) (DC, $\sigma_m = 0.04$).

estimated HS images by the existing and proposed methods for $(\sigma_h, \sigma_m) = (0.1,\ 0.04)$ and $(0.1,\ 0.06)$. At the time, its values are the average of the results for all HS images. In Table 1, one can see that for all the quality measures and for both noise intensity, the proposed method outperforms all the existing methods. By comparison between the proposed by HTV and HSSTV, in the almost PSNR, SAM and ERGAS cases, HTV is higher quality than HSSTV. We think that since an MS image has weaker spectral correlation than an HS image, it hardly necessary to evaluate the spatio-spectral piecewise smoothness of an MS image. The other hands, in $Q2^n$, the result of HSSTV and $p = 2$ is best. Since $Q2^n$ evaluates the structure information and spectral/spatial distortion, the result of HSSTV have more correct detail than that of HTV.

Fig. 1 is the estimated high HS images in the case of $(\sigma_h, \sigma_m) = (0.1, 0.04)$ and *DC*. They are depicted as RGB images, where R, G, and B bands were set to the 30th, 60th and 90th bands of them in the case of the HS images, and these three bands are set as (R, G, B) $= (2, 4, 6)$ bands in the MS image. One can see that the results by existing methods have artifacts and spectral distortion. Specifically, the results using MTF-GLP, MAPSMM, and CNMF have many artifacts, so that the spatial smoothness of them are lost. In the HySure case, the result almost does not have artifacts. However, since the color of the result changes compared with the original HS image, HySure produces the spectral distortion. In contrast, the proposed methods avoid artifacts and spectral distortion. In addition, HSSTV preserves more edge than HTV.

4. CONCLUSION

We have proposed a new method regarding HS and MS data fusion effectively using a-priori knowledge on an HS image. The proposed method simultaneously generate a high-spatial HS image and denoise an observed MS image to make robust HS and MS fusion. In addition, it fully evaluates spatial/spectral piecewise smoothness and edge similarity between the denoising MS image and the generated HS image. As a result, the proposed method can generate a high-spatial HS image without artifacts and spectral distortion. Since the problem of the proposed method is a nonsmooth convex optimization with four hard constraints, a primal-dual splitting method can solve it after suitable deformation. In the experiments, the results show the advantages of the proposed method over existing methods.

5. REFERENCES

[1] C. I. Chang, *Hyperspectral imaging: techniques for spectral detection and classification*, vol. 1, Springer Science & Business Media, 2003.

[2] A. Plaza et al., "Recent advances in techniques for hyperspectral image processing," *Remote sensing of environment*, vol. 113, pp. S110–S122, 2009.

[3] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geosci. and Remote Sens. Magazine*, vol. 5, no. 2, pp. 29–56, 2017.

[4] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of ms + pan data," *IEEE Trans. on Geosci. and Remote Sensing*, vol. 45, no. 10, pp. 3230–3239, 2007.

[5] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "Mtf-tailored multiscale fusion of high-resolution ms and pan imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 72, no. 5, pp. 591–596, 2006.

[6] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. on Geosci. and Remote Sensing*, vol. 50, no. 2, pp. 528–537, 2012.

[7] H. N. Gross and J. R. Schott, "Application of spectral mixture analysis and image fusion techniques for image sharpening," *Remote Sensing of Environment*, vol. 63, no. 2, pp. 85–94, 1998.

[8] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral superresolution by coupled spectral unmixing," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 3586–3594.

[9] Y. Zhang, Y. Gao, Y. Liu, and M. He, "Hyperspectral and multispectral image fusion based on constrained cnmf unmixing," in *2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2015, pp. 1–4.

[10] M. T. Eismann, *Resolution enhancement of hyperspectral imagery using maximum a posteriori estimation with a stochastic mixing model*, Ph.D. dissertation, Dept. Electrical and Computer Engineering, Univ. Dayton, OH, 2004.

[11] Q. Wei, N. Dobigeon, and J. Y. Tourneret, "Fast fusion of multi-band images based on solving a sylvester equation," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4109–4121, 2015.

[12] M. Simões, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. on Geosci. and Remote Sensing*, vol. 53, no. 6, pp. 3373–3388, 2015.

[13] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2010.

[14] S. Takeyama, S. Ono, and I. Kumazawa, "Hyperspectral image restoration by hybrid spatio-spectral total variation," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 4586–4590, 2017.

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

[16] G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu, "Epigraphical projection and proximal tools for solving constrained convex optimization problems," *Signal, Image and Video Process.*, pp. 1–13, 2014.

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

[18] S. Ono, "L_0 gradient projection," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1554–1564, 2017.

[19] J. J. Moreau, "Fonctions convexes duales et points proximaux dans un espace hilbertien," *C. R. Acad. Sci. Paris Ser. A Math.*, vol. 255, pp. 2897–2899, 1962.

[20] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, New York, 2011.

[21] "Spectir," http://www.spectir.com/free-data-samples/.

[22] "Multispec," https://engineering.purdue.edu/ biehl/MultiSpec.

[23] "GIC," http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.

[24] N. Yokoya and A. Iwasaki, "Airborne hyperspectral data over chikusei," Tech. Rep., Space Application Laboratory, University of Tokyo, Japan, May 2016.

[25] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral–spatial adaptive total variation model," *IEEE Trans. on Geosci. and Remote Sensing*, vol. 50, no. 10, pp. 3660–3677, 2012.

[26] F. A. Kruse et al, A. B. Lefkoff, J.W. Boardman, K. B. Heidebrecht, A.T. Shapiro, P. J. Barloon, and A. F. H. Goetz, "The spectral image processing system (sips)-interactive visualization and analysis of imaging spectrometer data," in *AIP Conference Proceedings*, 1993, vol. 283, pp. 192–201.

[27] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?," in *Third conference" Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"*, 2000, pp. 99–103.

[28] A. Garzelli and F. Nencini, "Hypercomplex quality assessment of multi/hyperspectral images," *IEEE Geosci. and Remote Sens. Lett.*, vol. 6, no. 4, pp. 662–665, 2009.