



**HAL**  
open science

## **LRTCFPan: Low-Rank Tensor Completion Based Framework for Pansharpening**

Zhong-Cheng Wu, Ting-Zhu Huang, Liang-Jian Deng, Jie Huang, Jocelyn Chanussot, Gemine Vivone

► **To cite this version:**

Zhong-Cheng Wu, Ting-Zhu Huang, Liang-Jian Deng, Jie Huang, Jocelyn Chanussot, et al.. LRTCFPan: Low-Rank Tensor Completion Based Framework for Pansharpening. *IEEE Transactions on Image Processing*, 2023, 32, pp.1640-1655. 10.1109/TIP.2023.3247165 . hal-04473781

**HAL Id: hal-04473781**

**<https://hal.science/hal-04473781>**

Submitted on 22 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# LRTCFFPan: Low-Rank Tensor Completion Based Framework for Pansharpening

Zhong-Cheng Wu, Ting-Zhu Huang, *Member, IEEE*, Liang-Jian Deng, *Member, IEEE*, Jie Huang, *Member, IEEE*, Jocelyn Chanussot, *Fellow, IEEE*, Gemine Vivone, *Senior Member, IEEE*

**Abstract**—Pansharpening refers to the fusion of a low spatial-resolution multispectral image with a high spatial-resolution panchromatic image. In this paper, we propose a novel low-rank tensor completion (LRTC)-based framework with some regularizers for multispectral image pansharpening, called LRTCFFPan. The tensor completion technique is commonly used for image recovery, but it cannot directly perform the pansharpening or, more generally, the super-resolution problem because of the formulation gap. Different from previous variational methods, we first formulate a pioneering image super-resolution (ISR) degradation model, which equivalently removes the downsampling operator and transforms the tensor completion framework. Under such a framework, the original pansharpening problem is realized by the LRTC-based technique with some deblurring regularizers. From the perspective of regularizer, we further explore a local-similarity-based dynamic detail mapping (DDM) term to more accurately capture the spatial content of the panchromatic image. Moreover, the low-tubal-rank property of multispectral images is investigated, and the low-tubal-rank prior is introduced for better completion and global characterization. To solve the proposed LRTCFFPan model, we develop an alternating direction method of multipliers (ADMM)-based algorithm. Comprehensive experiments at reduced-resolution (i.e., simulated) and full-resolution (i.e., real) data exhibit that the LRTCFFPan method significantly outperforms other state-of-the-art pansharpening methods. The code is publicly available at: [https://github.com/zhongchengwu/code\\_LRTCFFPan](https://github.com/zhongchengwu/code_LRTCFFPan).

**Index Terms**—Low-rank tensor completion (LRTC), Dynamic detail mapping (DDM), Tubal rank, Alternating direction method of multipliers (ADMM), Pansharpening, Super-resolution.

## I. INTRODUCTION

High-resolution multispectral (HR-MS) remote sensing images play a crucial role in many practical applications, e.g., change detection [1], target recognition [2], and classification [3]. Because of some physical constraints on the signal-to-noise ratio [4], many sensors onboard satellite platforms, such

This research is supported by NSFC (Grant Nos. 12171072, 12271083), Key Projects of Applied Basic Research in Sichuan Province (Grant No. 2020YJ0216), Natural Science Foundation of Sichuan Province (Grant No. 2022NSFSC0501), and National Key Research and Development Program of China (Grant No. 2020YFA0714001). (Corresponding authors: Ting-Zhu Huang; Liang-Jian Deng.)

Z. C. Wu, T. Z. Huang, L. J. Deng and J. Huang are with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, China (e-mails: wuzhch97@163.com; tingzhuhuang@126.com; liangjian.deng@uestc.edu.cn; huangjie\_uestc@uestc.edu.cn).

J. Chanussot is with Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France, also with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100045, China (e-mail: jocelyn.chanussot@grenoble-inp.fr).

G. Vivone is with the Institute of Methodologies for Environmental Analysis, CNR-IMAA, 85050 Tito Scalo, Italy (e-mail: gemine.vivone@imaa.cnr.it).

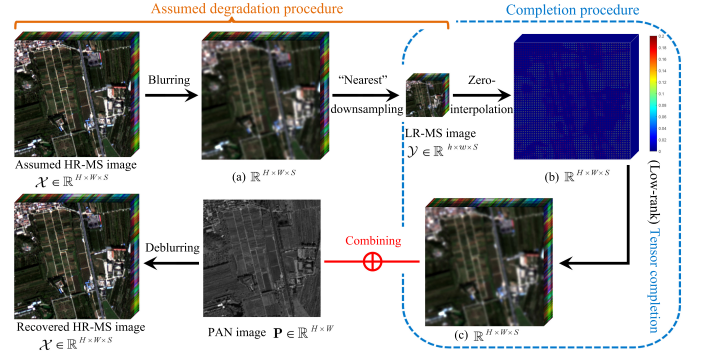


Fig. 1. The whole procedure of the proposed LRTCFFPan, which is a low-rank tensor completion (LRTC)-based framework with the deblurring regularizer.

as Gaofen-2 (GF-2), QuickBird (QB), and WorldView-3 (WV-3), acquire a low spatial-resolution multispectral (LR-MS) image while capturing higher spatial information into a gray-scaled panchromatic (PAN) image through another sensor. Pansharpening refers to the spatial-spectral fusion of the LR-MS image and the corresponding PAN image, aiming to yield an underlying HR-MS image. To clearly illustrate the proposed LRTCFFPan model, the whole procedure is depicted in Fig. 1.

Different methodologies have recently been developed to address the pansharpening problem. The most classical category is the component substitution (CS)-based methods. Some exemplary methods mainly include the principal component analysis (PCA) [5] method, the intensity-hue-saturation (IHS) [6] method, the Gram-Schmidt adaptive (GSA) [7] method, the band-dependent spatial-detail (BDS) [8] method, and the partial replacement adaptive component substitution (PRACS) [9] method. In these methods, the spatial component of the LR-MS image is separated by spectral transformation and substituted with the PAN image. Generally, the CS-based methods are appealing for their reduced computational burden, but they inevitably cause severe spectral distortion [10]. Another widely used category is the multi-resolution analysis (MRA)-based methods. These methods inject the spatial details extracted from the PAN image via multi-scale decomposition into the upsampled LR-MS image. The instances of this class are the “à-trous” wavelet transform (ATWT) [11] method, the additive wavelet luminance proportional (AWLP) [12] method, and the smoothing filter-based intensity modulation (SFIM) [13] method. Compared with CS methods, the MRA methods are characterized by higher spectral coherence while reducing spatial preservation. Overall, both the CS and MRA methods have

robust performance along different datasets. Furthermore, they usually do not require intensive tuning of parameters and have a lower computational complexity. Therefore, these methods are commonly used for benchmarking in pansharpening.

More recently, deep learning (DL) has been rapidly developed for computer vision applications [14]–[17]. Many convolutional neural network (CNN)-based approaches, e.g., [18]–[23], have been designed for pansharpening, showing excellent capabilities for feature extraction and nonlinear mapping learning [24], and getting better performance than traditional methods. However, these CNN-based methods generally require a lot of computational resources and training data [25], which severely limits their computational efficiency, generalization ability, and model interpretability.

Variational optimization-based implementations [26]–[29] are in-between the CS/MRA and CNN-based methods, generally realizing a trade-off between performance and efficiency. The variational methods are characterized by high generalization and model interpretability [24]. These methods, e.g., [25], [30]–[36], consider the pansharpening problem as an ill-posed inverse problem constructing the link among the LR-MS image, the PAN image, and the underlying HR-MS image, thus formulating an optimization model. The promising results have been generated by adopting traditional image super-resolution (ISR) degradation model, as in [24], [37], [38], especially when the characteristics of the MS sensors are considered, e.g., [24], [36]. However, due to the coupling of the ill-posed blurring and downsampling problems, many super-resolution models either exhibit the unnecessary solving complexity for decoupling, e.g., [24], or result in the unintuitive mixture of unfolding-based and tensor-based modeling, e.g., [39].

In this paper, we propose a novel variational pansharpening method, i.e., the low-rank tensor completion (LRTC)-based framework with the deblurring regularizer, called LRTCFFPan. More specifically, the proposed model consists of three folds. Firstly, we formulate a new ISR degradation model, thus theoretically decoupling and converting the original pansharpening problem into the LRTC-based framework, which directly eliminates the downsampling operator before regularization. Secondly, motivated by both the high-pass modulation (HPM) scheme and the local similarity of remote sensing images, we develop a new local-similarity-based dynamic detail mapping (DDM) regularizer, which is imposed on the LRTC-based framework to dynamically capture the high-frequency information of the PAN image. Furthermore, the low-tubal-rank characteristic is investigated, and the low-tubal-rank prior is introduced for better completion and global characterization. Under the ADMM framework, the proposed LRTCFFPan model is efficiently solved. Extensive experiments confirm the superiority of the proposed LRTCFFPan method over other classical and state-of-the-art pansharpening methods.

The *contributions* of this paper are summarized as follows:

- We formulate a novel ISR degradation model, allowing the LRTC-based framework with the deblurring regularizer for pansharpening. Such a strategy directly eliminates the downsampling operator and provides a valuable perspective for the pansharpening task.

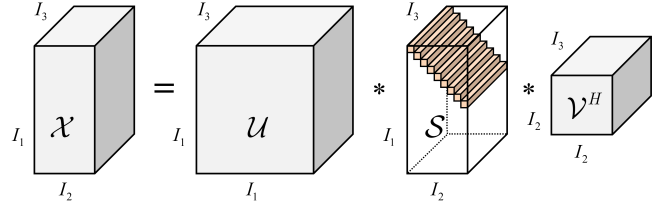


Fig. 2. The graphical illustration of the t-SVD of tensor  $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ .

- We design a local-similarity-based DDM regularizer to better characterize the spatial structure information of the PAN image. Within such a regularizer, we also explore a new procedure for estimating injection coefficients.
- We investigate the low-tubal-rank characteristic of multi-spectral images and impose the low-tubal-rank prior on the LRTC-based framework, aiming for better completion and global characterization.

The remainder of the paper is organized as follows. The notations and preliminaries are introduced in Section II. The related works and the proposed model are described in Section III. The proposed algorithm is provided in Section IV. The numerical experiments are performed in Section V. Finally, the conclusion is drawn in Section VI.

## II. NOTATIONS AND PRELIMINARIES

### A. Notations

Scalars, vectors, matrices, and tensors are denoted by lowercase letters, e.g.,  $a$ , lowercase bold letters, e.g.,  $\mathbf{a}$ , uppercase bold letters, e.g.,  $\mathbf{A}$ , and calligraphic letters, e.g.,  $\mathcal{A}$ , respectively. For a third-order tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , we employ  $\mathcal{A}(:, :, i)$  or  $\mathbf{A}^{(i)}$  for its  $i$ -th frontal slice,  $\mathcal{A}(i, j, :)$  for its  $(i, j)$ -th tube, and  $\mathcal{A}(i, j, k)$  or  $a_{i,j,k}$  for its  $(i, j, k)$ -th element. The Frobenius norm of  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  is defined as  $\|\mathcal{A}\|_F := \sqrt{\sum_{i,j,k} |a_{i,j,k}|^2}$ . Besides, we use  $\bar{\mathcal{A}}$  for the discrete Fourier transformation (DFT) on all the tubes of  $\mathcal{A}$ . Relying upon the MATLAB command, we have  $\bar{\mathcal{A}} = \text{fft}(\mathcal{A}, [], 3)$ . Conversely,  $\mathcal{A}$  can be obtained from  $\bar{\mathcal{A}}$  via the inverse DFT along each tube, i.e.,  $\mathcal{A} = \text{ifft}(\bar{\mathcal{A}}, [], 3)$ .

### B. Preliminaries

For clarity, we provide some definitions and theorems, and briefly introduce the LRTC basics.

**Definition II.1 (Tensor convolution (t-Conv)).** Given a third-order tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  and a convolution kernel tensor  $\mathcal{B} \in \mathbb{R}^{m \times m \times I_3}$ , where set  $\{\mathbf{B}^{(i)}\}_{i=1}^{I_3}$  indicates various kernels along the spectral dimension. Then, the t-Conv between  $\mathcal{A}$  and  $\mathcal{B}$  yields a tensor  $\mathcal{A} \bullet \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , whose  $i$ -th frontal slice is defined by

$$(\mathcal{A} \bullet \mathcal{B})(:, :, i) := \mathbf{A}^{(i)} \otimes \mathbf{B}^{(i)},$$

where  $\otimes$  represents the spatial convolution operator.

**Theorem 1 (Tensor singular value decomposition (t-SVD) [40]).** Let  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  be a third-order tensor, then it can be factorized as

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^H,$$

167 where  $*$  is the tensor-tensor product (t-product) operator,  $\mathcal{U} \in \mathbb{R}^{I_1 \times I_1 \times I_3}$   
 168 and  $\mathcal{V} \in \mathbb{R}^{I_2 \times I_2 \times I_3}$  are orthogonal tensors,  $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$   
 169 is an f-diagonal tensor, and  $(\cdot)^H$  represents the conjugate transpose operator. See [40], [41] for more details.

171 The graphical illustration of the t-SVD is shown in Fig. 2.

172 **Definition II.2 (Tensor multi-rank and tubal rank [42]).** Let  
 173  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  be a third-order tensor, then the tensor multi-  
 174 rank is a vector  $\text{rank}_m(\mathcal{A}) \in \mathbb{R}^{I_3}$  with its  $i$ -th entry being the  
 175 rank of the  $i$ -th frontal slice of  $\bar{\mathcal{A}}$ , where  $\bar{\mathcal{A}} = \text{fft}(\mathcal{A}, [ ], 3)$ .  
 176 The tubal rank, denoted as  $\text{rank}_t(\mathcal{A})$ , is defined as the number  
 177 of nonzero singular tubes of  $\mathcal{S}$ , that is,

$$\text{rank}_t(\mathcal{A}) := \#\{i, \mathcal{S}(i, i, :) \neq \mathbf{0}\},$$

178 where  $\mathcal{S}$  is provided by the t-SVD  $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^H$ .

In particular, the inverse DFT  $\bar{\mathcal{S}} = \text{ifft}(\bar{\mathcal{S}}, [ ], 3)$  gives the following equation

$$\bar{\mathcal{S}}(i, i, 1) = \frac{1}{I_3} \sum_{k=1}^{I_3} \bar{\mathcal{S}}(i, i, k),$$

179 where  $\bar{\mathcal{S}}(:, :, k)$  is the singular value matrix of the  $k$ -th frontal  
 180 slice of  $\bar{\mathcal{A}}$ . That is,  $\text{rank}_t(\mathcal{A}) = \max(\text{rank}_m(\bar{\mathcal{A}}))$ .

181 **Definition II.3 (Tensor singular value [43]).** Given a third-  
 182 order tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , then the singular values of  $\mathcal{A}$   
 183 are defined as the diagonal elements of  $\mathcal{S}(i, i, 1)$ , where  $\mathcal{S}$  is  
 184 provided by the t-SVD  $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^H$ .

185 Therefore,  $\text{rank}_t(\mathcal{A})$  is equivalent to the number of non-  
 186 zero tensor singular values of  $\mathcal{A}$ , and its non-convex approxi-  
 187 mation can be given via the following Definition II.4.

188 **Definition II.4 (Log tensor nuclear norm [39]).** For a tensor  
 189  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , the log tensor nuclear norm is defined as the  
 190 log-sum of the singular values of all the frontal slices of  $\bar{\mathcal{A}}$ ,  
 191 i.e.,

$$\|\mathcal{A}\|_{lt} := \frac{1}{I_3} \sum_{k=1}^{I_3} \sum_{i=1}^t \log(\bar{\mathcal{S}}(i, i, k) + \epsilon),$$

192 where  $\bar{\mathcal{S}} = \text{fft}(\mathcal{S}, [ ], 3)$ , in which  $\mathcal{S}$  is provided by the t-SVD  
 193  $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^H$ ,  $t$  is the  $\text{rank}_t(\mathcal{A})$ , and  $\epsilon$  is a small positive  
 194 value enforcing a non-zero input.

195 **Theorem 2 (Tensor singular value thresholding (t-SVT) [44]).** For any  $\tau > 0$ , and let  $\mathcal{Y} = \mathcal{U} * \mathcal{S} * \mathcal{V}^H$  be the t-SVD  
 196 of tensor  $\mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , a closed-form minimizer of

$$\arg \min_{\mathcal{X}} \tau \|\mathcal{X}\|_{lt} + \frac{1}{2} \|\mathcal{X} - \mathcal{Y}\|_F^2$$

198 is given by the t-SVT as  $\text{Prox}_\tau^\epsilon(\mathcal{Y})$ , which is defined by

$$\text{Prox}_\tau^\epsilon(\mathcal{Y}) := \mathcal{U} * \mathcal{S}_\tau^\epsilon * \mathcal{V}^H,$$

199 where  $\mathcal{S}_\tau^\epsilon = \text{ifft}(\bar{\mathcal{S}}_\tau^\epsilon, [ ], 3)$ . Let  $\bar{\mathcal{S}} = \text{fft}(\mathcal{S}, [ ], 3)$ , the  
 200 elements of  $\bar{\mathcal{S}}_\tau^\epsilon$  obey

$$\bar{\mathcal{S}}_\tau^\epsilon(i, j, k) = \begin{cases} 0, & \text{if } c_2 \leq 0, \\ \frac{c_1 + \sqrt{c_2}}{2}, & \text{if } c_2 > 0, \end{cases}$$

201 where  $c_1 = |\bar{\mathcal{S}}(i, j, k)| - \epsilon$  and  $c_2 = c_1^2 - 4(\tau - \epsilon|\bar{\mathcal{S}}(i, j, k)|)$ .

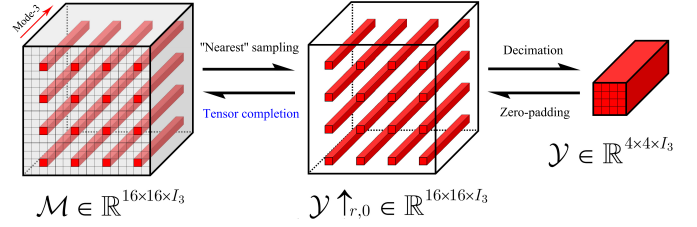


Fig. 3. A deeper perspective on the “nearest” downsampling operator, which is widely adopted [37], [39]. The scale factor  $r$  is equal to 4, and  $\mathcal{M} \downarrow_r = \mathcal{Y}$ . Moreover,  $\mathcal{Y} \uparrow_{r,0}$  denotes the result of using the scale factor  $r$  to perform zero-interpolation for  $\mathcal{Y}$ .

In what follows, we also briefly introduce the LRTC basics. The LRTC aims to recover the missing entries (values of 0) from an observed incomplete tensor by exploiting various low-rank priors, such as the Tucker rank [45], the multi-rank [46], and the fibered rank [44]. Mathematically, the general rank-minimization tensor completion model is formulated as

$$\min_{\mathcal{X}} \text{rank}(\mathcal{X}) \quad \text{s.t.} \quad \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{Y}, \quad (1)$$

where  $\mathcal{X}$  is the underlying tensor,  $\mathcal{Y}$  is the observed tensor,  $\Omega$  is the index set indicating available entries, and  $\mathcal{P}_\Omega(\cdot)$  is the projection function keeping the entries of  $\mathcal{X}$  in  $\Omega$  while forcing all the other values to zeros, i.e.,

$$(\mathcal{P}_\Omega(\mathcal{X}))_{i_1, i_2, \dots, i_N} := \begin{cases} x_{i_1, i_2, \dots, i_N}, & \text{if } (i_1, i_2, \dots, i_N) \in \Omega, \\ 0, & \text{otherwise.} \end{cases}$$

**Remark II.1.** According to the requirements of the projection function in (1), variables  $\mathcal{X}$  and  $\mathcal{Y}$  must have the same size, and their elements in the set  $\Omega$  must be numerically equivalent. However, any two images involved in the pansharpening task typically do not satisfy the prerequisites. Consequently, the LRTC cannot be awkwardly applied to the pansharpening task.

### III. RELATED WORKS AND PROPOSED MODEL

Three images are involved in pansharpening, including the underlying HR-MS image  $\mathcal{X} \in \mathbb{R}^{H \times W \times S}$ , the LR-MS image  $\mathcal{Y} \in \mathbb{R}^{h \times w \times S}$ , and the PAN image  $\mathbf{P} \in \mathbb{R}^{H \times W}$ . Additionally,  $H = h \times r$  and  $W = w \times r$  hold, where  $r$  is the scale factor.

#### A. Related Works

1) *Spectral Perspective:* Since the LR-MS image can be regarded as the degraded version of the underlying HR-MS image, the primary objective of the pansharpening methods is to construct the degradation model between them. Similar to the single image super-resolution problem [47], [48], there also exists an acknowledged and widely used degradation model for pansharpening, which is formulated by

$$\mathcal{Y} = (\mathcal{X} \bullet \mathcal{B}) \downarrow_r + \mathcal{N}_0, \quad (2)$$

where  $\bullet$  is the defined t-Conv operator,  $\downarrow_r$  denotes the “nearest” downsampling with the scale factor  $r$ , and  $\mathcal{N}_0$  indicates an additive zero-mean Gaussian noise. Such a degradation model has extensively been adopted in the field of pansharpening, significantly contributing to the variational optimization-based pansharpening methods, such as [24], [37], [49].

233 2) *Spatial Perspective*: As an ill-posed imaging inverse  
 234 problem, the ISR degeneration model (2) makes it challenging  
 235 to accurately reconstruct the underlying HR-MS image. Con-  
 236 sequently, the pansharpening problem requires establishing  
 237 another relationship between the underlying HR-MS image  
 238 and the PAN image, thereby leveraging the spatial prior  
 239 information of the latter. Considering the difficulty of nonlin-  
 240 ear mapping, the multi-resolution analysis (MRA) framework  
 241 [10], [18], [24], [50] has emerged as a powerful tool for  
 242 learning the spatial information of the PAN image. Formally,  
 243 the MRA framework is

$$\mathcal{X} = \hat{\mathcal{Y}} + \mathcal{G} \cdot (\hat{\mathcal{P}} - \hat{\mathcal{P}}_{LP}), \quad (3)$$

244 where  $\hat{\mathcal{Y}} \in \mathbb{R}^{H \times W \times S}$  denotes the interpolated version of  $\mathcal{Y}$ ,  
 245  $\hat{\mathcal{P}} \in \mathbb{R}^{H \times W \times S}$  is the replicated or histogram-matched  
 246 version of  $\mathbf{P}$ ,  $\hat{\mathcal{P}}_{LP} \in \mathbb{R}^{H \times W \times S}$  is the low-pass filtered  
 247 version of  $\hat{\mathcal{P}}$ ,  $\mathcal{G}$  is the injection coefficient, and  $\cdot$  is the  
 248 Hadamard product. Two common options for defining the  
 249 coefficient are  $\mathcal{G} = 1$  (i.e., the additive injection  
 250 scheme) and  $\mathcal{G} = \hat{\mathcal{Y}} / \hat{\mathcal{P}}_{LP}$  (i.e., the high-pass  
 251 modulation (HPM) scheme), where  $\cdot /$  denotes the  
 252 element-wise division. Benefiting from the greater  
 253 flexibility in configuring the local weights, the HPM  
 254 scheme is generally superior to the additive one and is  
 successfully introduced into the variational pansharpening  
 methods, e.g., [24], [51].

### 255 B. Proposed Model

256 As previously described, the coupled formulation between  
 257 blurring and downsampling typically causes two drawbacks:  
 258 1) the unnecessary solving complexity for decoupling,  
 259 and 2) the inconsistency in modeling form. To alleviate  
 260 these limitations, we consider developing a new ISR  
 261 degeneration model by investigating the downsampling  
 262 operator. As illustrated in Fig. 3, the “nearest”  
 263 downsampling  $\downarrow_r$  can actually be refined into a  
 264 two-stage operator, i.e., “nearest” sampling and  
 265 decimation, and the former is a sampling mode for the  
 266 LR-MS image. Accordingly, when the form of  $\mathcal{M} \downarrow_r = \mathcal{Y}$   
 267 is established and the LR-MS image is preprocessed,  
 268 the inverse problem of “nearest” downsampling can be  
 269 modeled by tensor completion. Inspired by it, we easily  
 270 modify the original ISR degeneration model (2) based  
 on the fact that there exists a zero-mean Gaussian  
 noise  $\mathcal{N}_1$  such that  $\mathcal{N}_1 \downarrow_r = \mathcal{N}_0$ , leading to

$$\begin{aligned} \mathcal{Y} &= (\mathcal{X} \bullet \mathcal{B}) \downarrow_r + \mathcal{N}_0 = (\mathcal{X} \bullet \mathcal{B}) \downarrow_r + \mathcal{N}_1 \downarrow_r \\ &= (\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) \downarrow_r. \end{aligned} \quad (4)$$

271 Consequently, the new ISR degeneration model can be  
 272 represented as  $\mathcal{Y} = (\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) \downarrow_r$ , which assumes  
 273 that the LR-MS image is the blurred, noisy, then  
 274 downsampled version of the underlying HR-MS image.  
 275 When the LR-MS image is further processed, the  
 276 degradation model can be equivalently rewritten as  
 the following projection-based form

$$\mathcal{P}_\Omega(\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) = \mathcal{Y} \uparrow_{r,0}, \quad (5)$$

277 where  $\mathcal{Y} \uparrow_{r,0} \in \mathbb{R}^{H \times W \times S}$  is the preprocessed image.  
 278 Relying upon the projection-based formulation, the  
 279 downsampling operator  $\downarrow_r$  is eliminated, and only  
 the  $\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1$  is maintained.

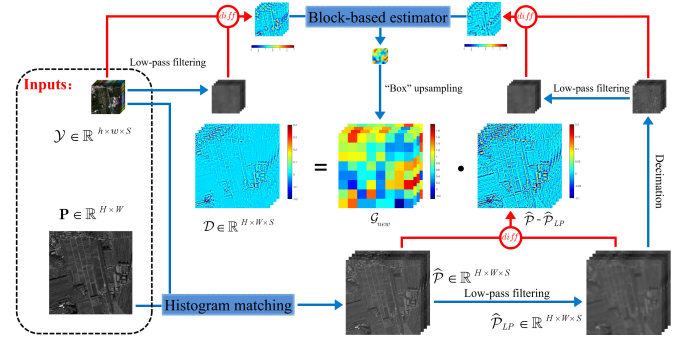


Fig. 4. The graphical illustration of estimating the modulated image  $\mathcal{D}$  (i.e.,  $\mathcal{G}_{new} \cdot (\hat{\mathcal{P}} - \hat{\mathcal{P}}_{LP})$ ) on a reduced-resolution Guangzhou image (source: GF-2). Symbols  $\cdot$  and  $\text{diff}$  denote the Hadamard product and the pixel-wise difference, respectively. The block size is  $8 \times 8$ , and the low-pass filters are available<sup>2</sup>.

280 To generate the underlying HR-MS image, we can formulate  
 281 the following rank-minimization problem

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1} \quad & \text{rank}(\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) \\ \text{s.t.} \quad & \mathcal{P}_\Omega(\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) = \mathcal{Y} \uparrow_{r,0}, \end{aligned} \quad (6)$$

282 where  $\text{rank}(\cdot)$  indicates the tensor rank to be determined.  
 283 Since model (6) is obviously ill-posed, the regularizer that  
 284 can leverage the spatial information of the PAN image is  
 required.

285 To explore a superior regularizer, the HPM model of (3)  
 286 is further improved. Despite the significant merits of the  
 287 HPM model, the coefficient  $\mathcal{G}$ , i.e.,  $\hat{\mathcal{Y}} / \hat{\mathcal{P}}_{LP}$ , generally  
 288 demonstrates unstable computational accuracy and  
 289 hypersensitivity, which are explained by the nonuniqueness  
 290 of  $\hat{\mathcal{Y}}$  and the oversensitivity of  $\hat{\mathcal{P}}_{LP}$  for different  
 291 low-pass filters. Moreover, although  $\hat{\mathcal{Y}}$  is originally  
 292 adopted to approximate the low-frequency information  
 293 of  $\mathcal{X}$ , the chaotic relationship is inevitably caused  
 294 owing to  $\mathcal{Y} = (\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) \downarrow_r$ . To address these  
 295 deficiencies, we consider directly computing the  
 296 low-frequency information of  $\mathcal{X}$  by  $\mathcal{X} \bullet \mathcal{B}$  and  
 developing a novel strategy for estimating the  
 coefficient. Resultantly, we have

$$\mathcal{X} - \mathcal{X} \bullet \mathcal{B} + \mathcal{N}_2 = \mathcal{G}_{new} \cdot (\hat{\mathcal{P}} - \hat{\mathcal{P}}_{LP}), \quad (7)$$

297 where  $\mathcal{N}_2$  is a Gaussian error,  $\hat{\mathcal{P}}^1$  is the histogram-  
 298 matched  $\mathbf{P}$ , and  $\mathcal{G}_{new}$  is the new coefficient determined  
 299 in Section III-C. For simplicity, model (7) can  
 compactly be expressed as

$$\mathcal{X} - \mathcal{X} \bullet \mathcal{B} + \mathcal{N}_2 = \mathcal{D}, \quad (8)$$

300 where  $\mathcal{D} = \mathcal{G}_{new} \cdot (\hat{\mathcal{P}} - \hat{\mathcal{P}}_{LP})$  is the pre-modulated  
 301 image. Furthermore, considering the similarity of the  
 302 local spatial details, we conduct model (8) on each  
 303 image patch to learn more accurate coefficients (see  
 304 Fig. 4), thus completely forming the local-similarity-  
 305 based DDM regularizer. Equipped with such a  
 regularizer, the rank-minimization model (6) is  
 improved as

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1} \quad & \text{rank}(\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) + \lambda_1 \|\mathcal{X} - \mathcal{X} \bullet \mathcal{B} - \mathcal{D}\|_F^2 \\ \text{s.t.} \quad & \mathcal{P}_\Omega(\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) = \mathcal{Y} \uparrow_{r,0}. \end{aligned} \quad (9)$$

<sup>1</sup> $\hat{\mathbf{P}}^{(i)} = (\text{Std}(\mathbf{Y}^{(i)}) / \text{Std}(\mathbf{P}))(\mathbf{P} - \text{Mean}(\mathbf{P})) + \text{Mean}(\mathbf{Y}^{(i)})$ , where  $\text{Mean}(\cdot)$  and  $\text{Std}(\cdot)$  are the mean and standard deviation operators.

<sup>2</sup><http://openremotesensing.net/knowledgebase/a-critical-comparison-among-pansharpening-algorithms/>

307 Regarding the above model (9), the low-rank characteristic  
 308 of variable  $\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1$  needs to be investigated. Among the  
 309 traditional and classical tensor decompositions, the CANDE-  
 310 COMP/PARAFAC (CP) one [52], Tucker one [53], and tensor  
 311 singular value decomposition (t-SVD) [40] have been widely  
 312 applied to the hyperspectral super-resolution problem [54],  
 313 [55]. Corresponding to these decompositions, the CP rank,  
 314 Tucker rank, and tubal rank have also been introduced into the  
 315 tensor completion problem [56]–[58]. However, the existence  
 316 of the optimal CP-rank approximation cannot be assured [59].  
 317 Moreover, since the  $\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1$  for pansharpening is merely  
 318 the multispectral image, the low-Tucker-rank property is rela-  
 319 tively insignificant, especially along the spectral dimension.  
 320 Accordingly, we investigate the tubal-rank rather than other  
 321 characteristics of multispectral images. From Fig. 5(c) and (f),  
 322 we observe that  $\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1$  has a significant low-rankness,  
 323 revealing the validity of the low-tubal-rank prior. Additionally,  
 324 Fig. 5(a) and (d) depict that the underlying HR-MS image  $\mathcal{X}$   
 325 can also exhibit the low-tubal-rank property, which implies  
 326 that the global low-tubal-rank prior can be imposed on the  
 327 underlying HR-MS image to penalize the ill-posed deconvolu-  
 328 tion problem. By combining two corresponding low-tubal-  
 329 rank regularizers, model (9) can be transformed into the final  
 330 LRTC-based framework, i.e., LRTCFFpan, as follows,

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1} \quad & \text{rank}_{\mathfrak{t}}(\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) + \lambda_1 \|\mathcal{X} - \mathcal{X} \bullet \mathcal{B} - \mathcal{D}\|_F^2 \\ & + \lambda_2 \text{rank}_{\mathfrak{t}}(\mathcal{X}) \\ \text{s.t.} \quad & \mathcal{P}_{\Omega}(\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) = \mathcal{Y} \uparrow_{r,0}. \end{aligned} \quad (10)$$

331 Since directly solving rank minimization is NP-hard, we give  
 332 the non-convex approximation of model (10) by

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1} \quad & \|\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1\|_{lt} + \lambda_1 \|\mathcal{X} - \mathcal{X} \bullet \mathcal{B} - \mathcal{D}\|_F^2 \\ & + \lambda_2 \|\mathcal{X}\|_{lt} \\ \text{s.t.} \quad & \mathcal{P}_{\Omega}(\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) = \mathcal{Y} \uparrow_{r,0}. \end{aligned} \quad (11)$$

333 Let  $\mathcal{T} = \mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1$ , model (11) can be further converted to

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{T}} \quad & \|\mathcal{X}\|_{lt} + \lambda_1 \|\mathcal{X} - \mathcal{X} \bullet \mathcal{B} - \mathcal{D}\|_F^2 + \lambda_2 \|\mathcal{X} \bullet \mathcal{B} - \mathcal{T}\|_F^2 \\ & + \lambda_3 \|\mathcal{T}\|_{lt} \\ \text{s.t.} \quad & \mathcal{P}_{\Omega}(\mathcal{T}) = \mathcal{Y} \uparrow_{r,0}, \end{aligned} \quad (12)$$

334 where  $\mathcal{D} = \mathcal{G}_{new} \cdot (\hat{\mathcal{P}} - \hat{\mathcal{P}}_{LP})$  is computed before regulariza-  
 335 tion, and  $\lambda_l, l = 1, 2, 3$ , are positive regularization parameters.

### 336 C. Estimating Coefficient $\mathcal{G}_{new}$

337 According to (7), we easily have the following equation

$$\left( (\mathcal{X} - \mathcal{X} \bullet \mathcal{B} + \mathcal{N}_2) \bullet \mathcal{B} \right) \downarrow_r = \left( \mathcal{G}_{new} \cdot (\hat{\mathcal{P}} - \hat{\mathcal{P}}_{LP}) \bullet \mathcal{B} \right) \downarrow_r. \quad (13)$$

338 When  $\mathbf{G}_{new}^{(i)}, i = 1, 2, \dots, S$ , are constant matrices, the above  
 339 equation (13) is equivalent to

$$\begin{aligned} & (\mathcal{X} \bullet \mathcal{B}) \downarrow_r + (\mathcal{N}_2 \bullet \mathcal{B}) \downarrow_r - (\mathcal{X} \bullet \mathcal{B} \bullet \mathcal{B}) \downarrow_r \\ & = \mathcal{G}_{new} \downarrow_r \cdot \left( (\hat{\mathcal{P}} \bullet \mathcal{B}) \downarrow_r - (\hat{\mathcal{P}}_{LP} \bullet \mathcal{B}) \downarrow_r \right). \end{aligned} \quad (14)$$

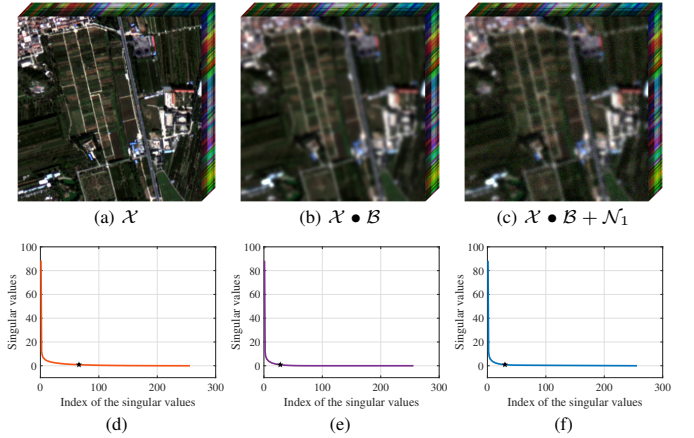


Fig. 5. The illustration of the low-tubal-rank characteristic on the reduced-resolution Guangzhou image (sensor: GF-2). The first row is (a) the HR-MS image modeled as  $\mathcal{X} \in \mathbb{R}^{256 \times 256 \times 4}$ , (b) the low-pass filtered image, and (c) the filtered image with Gaussian noise of standard deviation level 0.01. The (d), (e), and (f) illustrate the singular value curves of (a), (b), and (c), respectively. The approximated tubal ranks [44] are marked by black stars.

340 Since  $\mathcal{X}$  is unavailable, we assume that there exists a Gaussian  
 341 error  $\mathcal{E}$  such that the following equation holds

$$(\mathcal{X} \bullet \mathcal{B} \bullet \mathcal{B}) \downarrow_r = (\mathcal{X} \bullet \mathcal{B} + \mathcal{E}) \downarrow_r \bullet \mathcal{B}. \quad (15)$$

342 Subsequently, equation (14) can be rewritten as

$$\begin{aligned} & (\mathcal{X} \bullet \mathcal{B}) \downarrow_r + (\mathcal{N}_2 \bullet \mathcal{B}) \downarrow_r - (\mathcal{X} \bullet \mathcal{B} + \mathcal{E}) \downarrow_r \bullet \mathcal{B} \\ & = \mathcal{G}_{new} \downarrow_r \cdot \left( (\hat{\mathcal{P}} \bullet \mathcal{B}) \downarrow_r - (\hat{\mathcal{P}}_{LP} \bullet \mathcal{B}) \downarrow_r \right). \end{aligned} \quad (16)$$

343 For  $\mathcal{E}_1, \mathcal{E}_2 \in \mathbb{R}^{H \times W \times S}$  and  $\mathcal{E}_3 \in \mathbb{R}^{h \times w \times S}$ , we further define

$$\Gamma_{\mathcal{E}_1, \mathcal{E}_2} := (\mathcal{X} \bullet \mathcal{B} + \mathcal{E}_1) \downarrow_r - (\mathcal{X} \bullet \mathcal{B} + \mathcal{E}_2) \downarrow_r \bullet \mathcal{B} \quad (17)$$

344 and

$$\Upsilon_{\mathcal{E}_3} := (\hat{\mathcal{P}} \bullet \mathcal{B}) \downarrow_r - \mathcal{E}_3. \quad (18)$$

345 Ultimately, coefficients  $\mathbf{G}_{new}^{(i)}, i = 1, 2, \dots, S$ , can be esti-  
 346 mated by

$$\begin{aligned} \mathbf{G}_{new}^{(i)} &= \frac{\sum_{k=1}^w \sum_{j=1}^h \left( (\Gamma_{\mathcal{N}_2 \bullet \mathcal{B}, \mathcal{E}})^{(i)} \cdot (\Upsilon_{(\hat{\mathcal{P}}_{LP} \bullet \mathcal{B}) \downarrow_r})^{(i)} \right)_{j,k} \mathbf{1}}{\left\| (\Upsilon_{(\hat{\mathcal{P}}_{LP} \bullet \mathcal{B}) \downarrow_r})^{(i)} \right\|_F^2} \\ &\approx \frac{\sum_{k=1}^w \sum_{j=1}^h \left( (\mathcal{Y} - \mathcal{Y} \bullet \mathcal{B})^{(i)} \cdot (\Upsilon_{(\hat{\mathcal{P}}_{LP} \bullet \mathcal{B}) \downarrow_r})^{(i)} \right)_{j,k} \mathbf{1}}{\left\| (\Upsilon_{(\hat{\mathcal{P}}_{LP} \bullet \mathcal{B}) \downarrow_r})^{(i)} \right\|_F^2} \\ &\approx \frac{\sum_{k=1}^w \sum_{j=1}^h \left( (\mathcal{Y} - \mathcal{Y} \bullet \mathcal{B})^{(i)} \cdot (\Upsilon_{(\hat{\mathcal{P}}_{LP}) \downarrow_r \bullet \mathcal{B}})^{(i)} \right)_{j,k} \mathbf{1}}{\left\| (\Upsilon_{(\hat{\mathcal{P}}_{LP}) \downarrow_r \bullet \mathcal{B}})^{(i)} \right\|_F^2}, \end{aligned} \quad (19)$$

347 where  $\mathbf{1}$  is the all-ones matrix, and the  $\Upsilon_{(\hat{\mathcal{P}}_{LP}) \downarrow_r \bullet \mathcal{B}}$  is adopted  
 348 to maintain consistency with the  $(\mathcal{X} \bullet \mathcal{B} + \mathcal{E}) \downarrow_r \bullet \mathcal{B}$  in (15).  
 349 When  $\mathcal{N}_1 \rightarrow 0, \mathcal{N}_2 \rightarrow 0, \mathcal{G}_{new} \rightarrow 1$ , but  $\mathcal{E} \rightarrow 0$ , the negative  
 350 impact from  $\mathcal{E}$  can be appropriately weakened.

#### IV. PROPOSED ALGORITHM

##### A. Algorithm

For optimizing the proposed LRTCFFan model, we develop an efficient ADMM-based algorithm. By introducing auxiliary variables  $\mathcal{Q}$ ,  $\mathcal{R}$  and  $\mathcal{Z}$ , we can rewrite (12) as the following constrained problem

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{T}} \quad & \|\mathcal{Q}\|_{lt} + \lambda_1 \|\mathcal{R} - \mathcal{Z} - \mathcal{D}\|_F^2 + \lambda_2 \|\mathcal{Z} - \mathcal{T}\|_F^2 + \lambda_3 \|\mathcal{T}\|_{lt} \\ \text{s.t.} \quad & \mathcal{P}_\Omega(\mathcal{T}) = \mathcal{Y} \uparrow_{r,0}, \quad \mathcal{Q} = \mathcal{X}, \quad \mathcal{R} = \mathcal{X}, \quad \mathcal{Z} = \mathcal{X} \bullet \mathcal{B}. \end{aligned} \quad (20)$$

The augmented Lagrangian function of (20) is

$$\begin{aligned} \mathcal{L}(\mathcal{X}, \mathcal{T}, \mathcal{Q}, \mathcal{R}, \mathcal{Z}) = & \|\mathcal{Q}\|_{lt} + \lambda_1 \|\mathcal{R} - \mathcal{Z} - \mathcal{D}\|_F^2 \\ & + \lambda_2 \|\mathcal{Z} - \mathcal{T}\|_F^2 + \lambda_3 \|\mathcal{T}\|_{lt} + \iota(\mathcal{T}) + \frac{\eta_1}{2} \left\| \mathcal{X} - \mathcal{Q} + \frac{\Lambda_1}{\eta_1} \right\|_F^2 \\ & + \frac{\eta_2}{2} \left\| \mathcal{X} - \mathcal{R} + \frac{\Lambda_2}{\eta_2} \right\|_F^2 + \frac{\eta_3}{2} \left\| \mathcal{X} \bullet \mathcal{B} - \mathcal{Z} + \frac{\Lambda_3}{\eta_3} \right\|_F^2, \end{aligned} \quad (21)$$

where  $\Lambda_l$ ,  $l = 1, 2, 3$ , are the Lagrange multipliers,  $\eta_l$ ,  $l = 1, 2, 3$ , are positive penalty parameters, and  $\iota(\mathcal{T})$  is an indicator function defined as

$$\iota(\mathcal{T}) := \begin{cases} 0, & \text{if } \mathcal{P}_\Omega(\mathcal{T}) = \mathcal{Y} \uparrow_{r,0}, \\ \infty, & \text{otherwise.} \end{cases} \quad (22)$$

Afterwards, model (20) can be solved by alternatively minimizing the following simpler subproblems:

1)  $\mathcal{X}$ -subproblem: By fixing  $\mathcal{T}$ ,  $\mathcal{Q}$ ,  $\mathcal{R}$ ,  $\mathcal{Z}$ , and  $\Lambda_l$ , the  $\mathcal{X}$ -subproblem can be given as

$$\begin{aligned} \min_{\mathcal{X}} \quad & \frac{\eta_1}{2} \left\| \mathcal{X} - \mathcal{Q} + \frac{\Lambda_1}{\eta_1} \right\|_F^2 + \frac{\eta_2}{2} \left\| \mathcal{X} - \mathcal{R} + \frac{\Lambda_2}{\eta_2} \right\|_F^2 \\ & + \frac{\eta_3}{2} \left\| \mathcal{X} \bullet \mathcal{B} - \mathcal{Z} + \frac{\Lambda_3}{\eta_3} \right\|_F^2. \end{aligned} \quad (23)$$

According to the modulation transfer function (MTF)-matched filters [60], the  $\mathbf{B}^{(i)}$ ,  $i = 1, 2, \dots, S$ , can be configured with different blurring kernels [36]. Accordingly, we can rearrange problem (23) as the frontal slice-based expression, i.e.,

$$\begin{aligned} \min_{\mathcal{X}} \quad & \frac{\eta_1}{2} \sum_{i=1}^S \left\| \mathbf{X}^{(i)} - \mathbf{Q}^{(i)} + \frac{\Lambda_1^{(i)}}{\eta_1} \right\|_F^2 \\ & + \frac{\eta_2}{2} \sum_{i=1}^S \left\| \mathbf{X}^{(i)} - \mathbf{R}^{(i)} + \frac{\Lambda_2^{(i)}}{\eta_2} \right\|_F^2 \\ & + \frac{\eta_3}{2} \sum_{i=1}^S \left\| \mathbf{X}^{(i)} \otimes \mathbf{B}^{(i)} - \mathbf{Z}^{(i)} + \frac{\Lambda_3^{(i)}}{\eta_3} \right\|_F^2, \end{aligned} \quad (24)$$

which is equivalent to

$$\begin{aligned} \min_{\mathcal{X}} \quad & \sum_{i=1}^S \left( \frac{\eta_1}{2} \left\| \mathbf{X}^{(i)} - \mathbf{Q}^{(i)} + \frac{\Lambda_1^{(i)}}{\eta_1} \right\|_F^2 \right. \\ & + \frac{\eta_2}{2} \left\| \mathbf{X}^{(i)} - \mathbf{R}^{(i)} + \frac{\Lambda_2^{(i)}}{\eta_2} \right\|_F^2 \\ & \left. + \frac{\eta_3}{2} \left\| \mathbf{X}^{(i)} \otimes \mathbf{B}^{(i)} - \mathbf{Z}^{(i)} + \frac{\Lambda_3^{(i)}}{\eta_3} \right\|_F^2 \right). \end{aligned} \quad (25)$$

##### Algorithm 1 The ADMM-based LRTCFFan Solver

**Input:**  $\mathcal{Y}$ ,  $\mathbf{P}$ ,  $\lambda_l$ ,  $\eta_l$ ,  $r = 4$ , and  $\epsilon = 2 \times 10^{-5}$ .

**Initialization:**

- 1:  $\mathcal{X} \leftarrow 0, \mathcal{T} \leftarrow 0, \mathcal{Q} \leftarrow 0, \mathcal{R} \leftarrow 0, \mathcal{Z} \leftarrow 0$ , and  $\Lambda_l \leftarrow 0$ .
  - 2:  $\mathcal{D} \leftarrow \mathcal{G}_{new} \cdot (\widehat{\mathbf{P}} - \widehat{\mathbf{P}}_{LP})$ .
  - 3: **while** not converged **do**
  - 4: Record the last-update result  $\mathcal{X}_{last}$ .
  - 5: Update  $\mathcal{X}$  via (27)-(28).
  - 6: Update  $\mathcal{T}$  via (30).
  - 7: Update  $\mathcal{Q}$  via (32).
  - 8: Update  $\mathcal{R}$  via (34).
  - 9: Update  $\mathcal{Z}$  via (36).
  - 10: Update Lagrange multipliers  $\Lambda_l$  via (37).
  - 11: Check the convergence criterion:
  - 12:  $\|\mathcal{X} - \mathcal{X}_{last}\|_F / \|\mathcal{X}_{last}\|_F < \epsilon$ .
  - 13: **end while**
- Output:** The HR-MS image  $\mathcal{X}$ .

Therefore, the original minimization problem (23) can be separated into  $S$  independent problems as follows,

$$\begin{aligned} \min_{\mathbf{X}^{(i)}} \quad & \frac{\eta_1}{2} \left\| \mathbf{X}^{(i)} - \mathbf{Q}^{(i)} + \frac{\Lambda_1^{(i)}}{\eta_1} \right\|_F^2 + \frac{\eta_2}{2} \left\| \mathbf{X}^{(i)} - \mathbf{R}^{(i)} + \frac{\Lambda_2^{(i)}}{\eta_2} \right\|_F^2 \\ & + \frac{\eta_3}{2} \left\| \mathbf{X}^{(i)} \otimes \mathbf{B}^{(i)} - \mathbf{Z}^{(i)} + \frac{\Lambda_3^{(i)}}{\eta_3} \right\|_F^2, \quad i = 1, 2, \dots, S. \end{aligned} \quad (26)$$

Under the condition of periodic boundary, the closed-form solution of the  $i$ -th problem is given by

$$\mathbf{X}^{(i)} \leftarrow \mathcal{F}^{-1} \left( \Sigma ./ \left( \eta_3 \mathcal{F}(\mathbf{B}^{(i)}) \cdot \mathcal{F}(\mathbf{B}^{(i)})^\dagger + \eta_1 + \eta_2 \right) \right) \quad (27)$$

with

$$\begin{aligned} \Sigma = & \eta_1 \mathcal{F}(\mathbf{Q}^{(i)}) + \eta_2 \mathcal{F}(\mathbf{R}^{(i)}) - \mathcal{F}(\Lambda_1^{(i)}) - \mathcal{F}(\Lambda_2^{(i)}) \\ & + \left( \eta_3 \mathcal{F}(\mathbf{Z}^{(i)}) - \mathcal{F}(\Lambda_3^{(i)}) \right) \cdot \mathcal{F}(\mathbf{B}^{(i)})^\dagger, \end{aligned} \quad (28)$$

where  $\mathcal{F}(\cdot)$  and  $\mathcal{F}^{-1}(\cdot)$  are the 2-D fast Fourier transform (FFT) and its inverse operator, respectively, and  $\dagger$  denotes the complex conjugate.

2)  $\mathcal{T}$ -subproblem: Similarly, the  $\mathcal{T}$ -subproblem is

$$\min_{\mathcal{T}} \quad \lambda_2 \|\mathcal{Z} - \mathcal{T}\|_F^2 + \lambda_3 \|\mathcal{T}\|_{lt} + \iota(\mathcal{T}). \quad (29)$$

Based on Theorem 2 and the definition of indicator function  $\iota(\mathcal{T})$ , we have

$$\mathcal{T} \leftarrow \mathcal{P}_{\Omega^c} \left( \text{Prox}_{\frac{\lambda_3}{2\lambda_2}}^\epsilon(\mathcal{Z}) \right) + \mathcal{Y} \uparrow_{r,0}, \quad (30)$$

where  $\Omega^c$  indicates the complementary set of  $\Omega$ .

3)  $\mathcal{Q}$ -subproblem: By fixing the other estimated directions for alternating, we obtain the  $\mathcal{Q}$ -subproblem as

$$\min_{\mathcal{Q}} \quad \|\mathcal{Q}\|_{lt} + \frac{\eta_1}{2} \left\| \mathcal{X} - \mathcal{Q} + \frac{\Lambda_1}{\eta_1} \right\|_F^2. \quad (31)$$

Based on Theorem 2 again, we can immediately get

$$\mathcal{Q} \leftarrow \text{Prox}_{\frac{\epsilon}{\eta_1}}^\epsilon \left( \mathcal{X} + \frac{\Lambda_1}{\eta_1} \right). \quad (32)$$

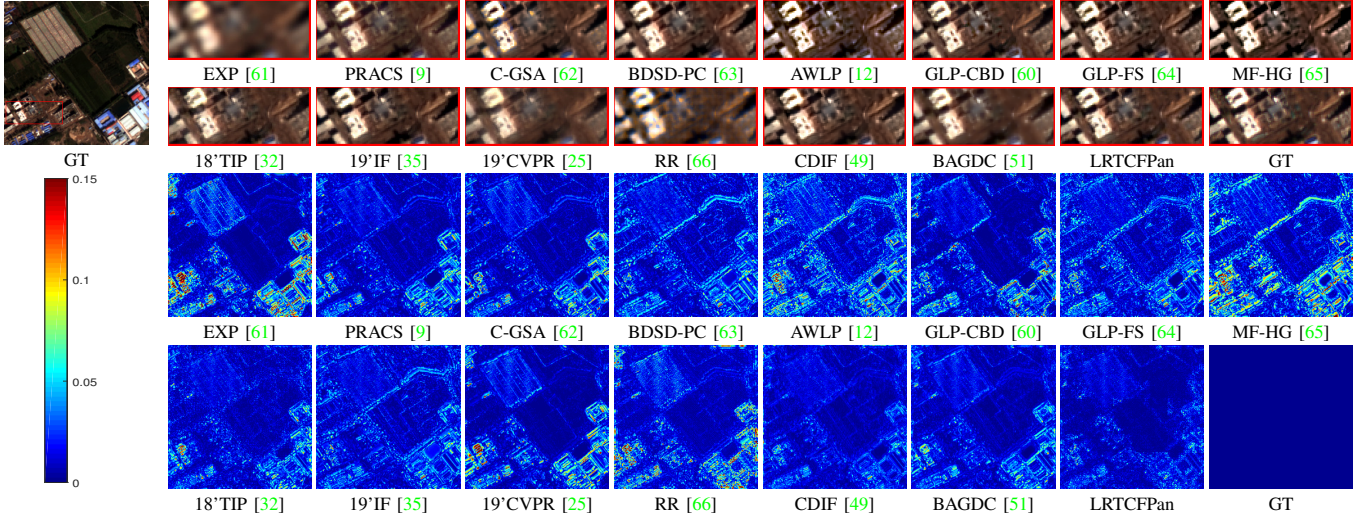


Fig. 6. The fusion results on the reduced-resolution Guangzhou dataset (source: GF-2). The first two rows: the visual inspection of the ground-truth (GT) image and the close-ups of the fused images. The last two rows: the residual maps using the GT image as a reference.

4)  $\mathcal{R}$ -subproblem: The  $\mathcal{R}$ -subproblem is

$$\min_{\mathcal{R}} \lambda_1 \|\mathcal{R} - \mathcal{Z} - \mathcal{D}\|_F^2 + \frac{\eta_2}{2} \left\| \mathcal{X} - \mathcal{R} + \frac{\Lambda_2}{\eta_2} \right\|_F^2, \quad (33)$$

which has the closed-form solution as follows,

$$\mathcal{R} \leftarrow \frac{2\lambda_1(\mathcal{Z} + \mathcal{D}) + \eta_2\mathcal{X} + \Lambda_2}{2\lambda_1 + \eta_2}. \quad (34)$$

5)  $\mathcal{Z}$ -subproblem: The  $\mathcal{Z}$ -subproblem is

$$\min_{\mathcal{Z}} \lambda_1 \|\mathcal{R} - \mathcal{Z} - \mathcal{D}\|_F^2 + \frac{\eta_3}{2} \left\| \mathcal{X} \bullet \mathcal{B} - \mathcal{Z} + \frac{\Lambda_3}{\eta_3} \right\|_F^2 + \lambda_2 \|\mathcal{Z} - \mathcal{T}\|_F^2. \quad (35)$$

Correspondingly, the closed-form solution is given by

$$\mathcal{Z} \leftarrow \frac{2\lambda_1(\mathcal{R} - \mathcal{D}) + 2\lambda_2\mathcal{T} + \eta_3\mathcal{X} \bullet \mathcal{B} + \Lambda_3}{2(\lambda_1 + \lambda_2) + \eta_3}. \quad (36)$$

Under the ADMM framework, the Lagrangian multipliers  $\Lambda_l$ ,  $l = 1, 2, 3$ , can be directly updated by

$$\begin{pmatrix} \Lambda_1 \\ \Lambda_2 \\ \Lambda_3 \end{pmatrix} \leftarrow \begin{pmatrix} \Lambda_1 \\ \Lambda_2 \\ \Lambda_3 \end{pmatrix} + \begin{pmatrix} \eta_1 & 0 & 0 \\ 0 & \eta_2 & 0 \\ 0 & 0 & \eta_3 \end{pmatrix} \begin{pmatrix} \mathcal{X} - \mathcal{Q} \\ \mathcal{X} - \mathcal{R} \\ \mathcal{X} \bullet \mathcal{B} - \mathcal{Z} \end{pmatrix}. \quad (37)$$

The solving pseudocode for the proposed LRTCFFan model is summarized in Algorithm 1.

### B. Computational Complexity Analysis

The complexity of Algorithm 1 mainly involves computing the FFT, the inverse FFT (IFFT), and the SVD. More specifically, the computational complexity of updating  $\mathcal{X}$  is  $\mathcal{O}(HWS \log(HW))$ . The computational complexity of updating  $\mathcal{T}$  and  $\mathcal{Q}$  is  $\mathcal{O}(HWS(\log(S) + \min(H, W)))$ . Since  $\log(S) + \min(H, W) \gg \log(HW)$ , more computational resources are generally consumed for solving the  $\mathcal{T}$  and  $\mathcal{Q}$  subproblems. Furthermore, the computational complexity of updating  $\mathcal{R}$ ,  $\mathcal{Z}$ , and  $\Lambda_l$  ( $l = 1, 2, 3$ ) is  $\mathcal{O}(HWS)$ . Therefore, the total computational complexity for each iteration in Algorithm 1 is  $\mathcal{O}(HWS(\log(HWS) + \min(H, W)))$ .

## V. EXPERIMENTAL RESULTS

To validate the superiority of the proposed LRTCFFan method, we conduct comprehensive numerical experiments on several commonly used datasets<sup>1</sup>, including the Guangzhou dataset (source: GF-2), the Indianapolis dataset (source: QB), and the Rio dataset (source: WV-3). The scale factors for all the datasets are 4, i.e.,  $r = 4$ . Numerically, all experimental data are pre-normalized into  $[0, 1]$ . All the experiments are implemented in MATLAB (R2018a) on a computer with 16Gb of RAM and an Intel(R) Core(TM) i5-4590 CPU: @3.30 GHz. For each sensor, e.g., GF-2, QB, and WV-3,  $S + 1$  low-pass filters are required for configuring the  $\mathcal{B}^{(i)}$ ,  $i = 1, 2, \dots, S$  (i.e., the blurring kernels of the MS image), and the  $(\cdot)_{LP}$  (i.e., the blurring kernel of the PAN image). According to [60], the kernels designed to match the modulation transfer functions (MTFs) of MS and PAN sensors are advisable. More specifically, these  $S + 1$  blurring kernels are assumed to be Gaussian-shaped with size of  $41 \times 41$  having  $S + 1$  standard deviations. When applied to a specific sensor, the  $S + 1$  standard deviations can be determined accordingly.

The compared methods include EXP [61], PRACS [9], C-GSA [62], BDS-PC [63], AWLP [12], GLP-CBD [60], GLP-FS [64], MF-HG [65], 18'TIP [32], 19'IF [35], 19'CVPR [25], RR [66], CDIF [49], and BAGDC [51]. It is worth remarking that the source codes of the competitors are available at either the website<sup>2</sup> or the authors' homepages. The hyper-parameters adopted in these variational optimization-based methods, i.e., the 18'TIP, the 19'IF, the 19'CVPR, the RR, the CDIF, and the BAGDC, are configured within a specific range suggested by their authors to achieve high performance.

When evaluated at reduced-resolution (i.e., simulated) data, six popular metrics, i.e., the peak signal-to-noise ratio (PSNR), the structural similarity index measure (SSIM) [67], the spectral angle mapper (SAM) [68], the spatial correlation coefficient (SCC) [12], the relative dimensionless global error

<sup>1</sup><http://www.digitalglobe.com/samples?search=Imagery>

<sup>2</sup><http://openremotesensing.net/kb/codes/pansharpening/>



TABLE I  
THE QUALITY METRICS ON 82 IMAGES WITH A PAN SIZE OF  $256 \times 256$  FROM THE REDUCED-RESOLUTION GUANGZHOU DATASET (SOURCE: GF-2).  
(BOLD: BEST; UNDERLINE: SECOND BEST)

Method	PSNR	SSIM	SAM	SCC	ERGAS	Q4	Runtime[s]
EXP [61]	31.094 ± 2.125	0.794 ± 0.060	2.007 ± 0.361	0.911 ± 0.029	2.645 ± 0.394	0.794 ± 0.043	<b>0.01</b>
PRACS [9]	33.973 ± 1.862	0.896 ± 0.027	1.883 ± 0.317	0.953 ± 0.021	1.894 ± 0.283	0.887 ± 0.033	0.07
C-GSA [62]	33.944 ± 2.113	0.895 ± 0.031	1.910 ± 0.396	0.950 ± 0.021	1.924 ± 0.358	0.889 ± 0.036	0.29
BDS-PC [63]	33.882 ± 2.086	0.894 ± 0.030	1.844 ± 0.327	0.953 ± 0.018	1.911 ± 0.323	0.893 ± 0.029	<u>0.04</u>
AWLP [12]	33.504 ± 2.012	0.870 ± 0.035	2.164 ± 0.454	0.946 ± 0.018	1.919 ± 0.288	0.870 ± 0.035	0.08
GLP-CBD [60]	33.423 ± 1.862	0.886 ± 0.030	1.763 ± 0.343	0.944 ± 0.023	1.981 ± 0.310	0.888 ± 0.030	24.91
GLP-FS [64]	33.984 ± 1.770	0.892 ± 0.028	1.804 ± 0.319	0.953 ± 0.018	1.838 ± 0.264	0.890 ± 0.035	0.07
MF-HG [65]	33.772 ± 1.853	0.894 ± 0.027	1.787 ± 0.310	0.951 ± 0.015	1.910 ± 0.242	0.886 ± 0.036	<u>0.04</u>
18'TIP [32]	34.014 ± 1.797	0.888 ± 0.026	1.623 ± 0.298	0.952 ± 0.020	1.820 ± 0.282	0.890 ± 0.032	35.40
19'IF [35]	33.411 ± 1.819	0.885 ± 0.029	1.719 ± 0.315	0.947 ± 0.022	1.952 ± 0.349	0.879 ± 0.041	11.81
19'CVPR [25]	33.176 ± 2.198	0.877 ± 0.037	1.737 ± 0.322	0.946 ± 0.018	2.114 ± 0.332	0.870 ± 0.027	9.66
RR [66]	32.668 ± 1.835	0.835 ± 0.044	2.357 ± 0.443	0.921 ± 0.033	1.986 ± 0.340	0.832 ± 0.055	15.47
CDIF [49]	<u>35.312 ± 2.087</u>	<u>0.917 ± 0.025</u>	<u>1.508 ± 0.292</u>	<u>0.965 ± 0.015</u>	<u>1.594 ± 0.293</u>	<u>0.925 ± 0.021</u>	25.58
BAGDC [51]	33.930 ± 1.653	0.890 ± 0.024	2.033 ± 0.359	0.953 ± 0.018	1.895 ± 0.232	0.892 ± 0.027	0.67
LRTCFFan	<b>35.918 ± 2.087</b>	<b>0.921 ± 0.022</b>	<b>1.391 ± 0.274</b>	<b>0.968 ± 0.014</b>	<b>1.496 ± 0.275</b>	<b>0.926 ± 0.039</b>	29.41
<b>Ideal value</b>	<b>+∞</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>-</b>

TABLE II  
THE QUALITY METRICS ON 42 IMAGES WITH A PAN SIZE OF  $256 \times 256$  FROM THE REDUCED-RESOLUTION INDIANAPOLIS DATASET (SOURCE: QB).  
(BOLD: BEST; UNDERLINE: SECOND BEST)

Method	PSNR	SSIM	SAM	SCC	ERGAS	Q4	Runtime[s]
EXP [61]	28.038 ± 2.710	0.682 ± 0.075	8.280 ± 1.453	0.771 ± 0.026	11.927 ± 1.387	0.595 ± 0.081	<b>0.01</b>
PRACS [9]	31.029 ± 2.203	0.829 ± 0.034	8.058 ± 1.502	0.898 ± 0.023	8.499 ± 0.694	0.786 ± 0.104	0.07
C-GSA [62]	32.057 ± 2.138	0.861 ± 0.027	<u>7.143 ± 1.244</u>	0.910 ± 0.020	7.530 ± 0.665	0.835 ± 0.099	0.29
BDS-PC [63]	31.920 ± 2.130	0.855 ± 0.028	7.801 ± 1.457	0.906 ± 0.019	7.648 ± 0.630	0.832 ± 0.096	<u>0.04</u>
AWLP [12]	31.506 ± 2.278	0.845 ± 0.032	8.172 ± 1.566	0.903 ± 0.017	8.037 ± 0.790	0.813 ± 0.093	0.07
GLP-CBD [60]	31.774 ± 2.173	0.857 ± 0.028	7.241 ± 1.289	0.906 ± 0.018	7.711 ± 0.624	0.833 ± 0.088	24.46
GLP-FS [64]	31.689 ± 2.058	0.850 ± 0.028	7.614 ± 1.358	0.905 ± 0.020	7.776 ± 0.588	0.822 ± 0.100	0.07
MF-HG [65]	31.161 ± 2.159	0.835 ± 0.034	7.782 ± 1.434	0.890 ± 0.018	8.485 ± 0.766	0.804 ± 0.091	<u>0.04</u>
18'TIP [32]	31.228 ± 2.211	0.824 ± 0.035	8.730 ± 1.593	0.899 ± 0.016	8.415 ± 0.705	0.798 ± 0.083	34.42
19'IF [35]	31.512 ± 2.061	0.844 ± 0.030	8.329 ± 1.530	0.903 ± 0.018	7.901 ± 0.641	0.822 ± 0.099	11.76
19'CVPR [25]	30.224 ± 2.477	0.798 ± 0.044	7.828 ± 1.364	0.879 ± 0.016	9.440 ± 0.938	0.743 ± 0.093	8.32
RR [66]	30.453 ± 2.556	0.814 ± 0.049	7.807 ± 1.296	0.859 ± 0.021	8.602 ± 0.968	0.779 ± 0.073	20.64
CDIF [49]	<u>32.485 ± 2.078</u>	<u>0.866 ± 0.026</u>	<u>7.247 ± 1.297</u>	<u>0.919 ± 0.018</u>	<u>7.223 ± 0.619</u>	<u>0.852 ± 0.089</u>	29.05
BAGDC [51]	30.822 ± 2.227	0.800 ± 0.040	8.500 ± 1.492	0.884 ± 0.016	8.828 ± 0.668	0.776 ± 0.103	0.84
LRTCFFan	<b>32.727 ± 2.132</b>	<b>0.873 ± 0.025</b>	<b>7.032 ± 1.264</b>	<b>0.922 ± 0.016</b>	<b>6.964 ± 0.596</b>	<b>0.861 ± 0.092</b>	29.38
<b>Ideal value</b>	<b>+∞</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>-</b>

440 in synthesis (ERGAS) [69], and the  $Q2^n$  [70], are adopted.  
 441 When evaluated at full-resolution (i.e., real) data, the quality  
 442 with no reference (QNR) [71] metric, which consists of the  
 443 spectral distortion index (i.e.,  $D_\lambda$ ) and the spatial distortion  
 444 index (i.e.,  $D_s$ ), is employed.

#### 445 A. Qualitative Comparison

446 1) *Reduced-Resolution Data Experiment*: To qualitatively  
 447 evaluate the performance of the proposed LRTCFFan method,  
 448 we first conduct the numerical experiments on the reduced-  
 449 resolution images, which are simulated from the real-world  
 450 Guangzhou (sensor: GF-2), Indianapolis (sensor: QB), and Rio  
 451 (sensor: WV-3) datasets. According to Wald's protocol [72],

the simulated HR-MS image, the simulated LR-MS image, 452  
 and the simulated PAN image can be considered as the blurred 453  
 and downsampled versions of the underlying HR-MS image, 454  
 the real LR-MS image, and the real PAN image, respectively. 455  
 Since the ISR degradation model  $\mathcal{Y} = (\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1) \downarrow_r$  assumes 456  
 that the real LR-MS image is the blurred and downsampled 457  
 version of the underlying HR-MS image when noise-free, 458  
 the real LR-MS image is actually assigned as the simulated 459  
 HR-MS image without additional processing. Considering the 460  
 page layout, we present only the visual comparative results 461  
 of a 4-bands (i.e., the simulated Guangzhou data) experiment 462  
 and an 8-bands (i.e., the simulated Rio data) experiment. By 463  
 the RGB rendering, the corresponding results are depicted in 464

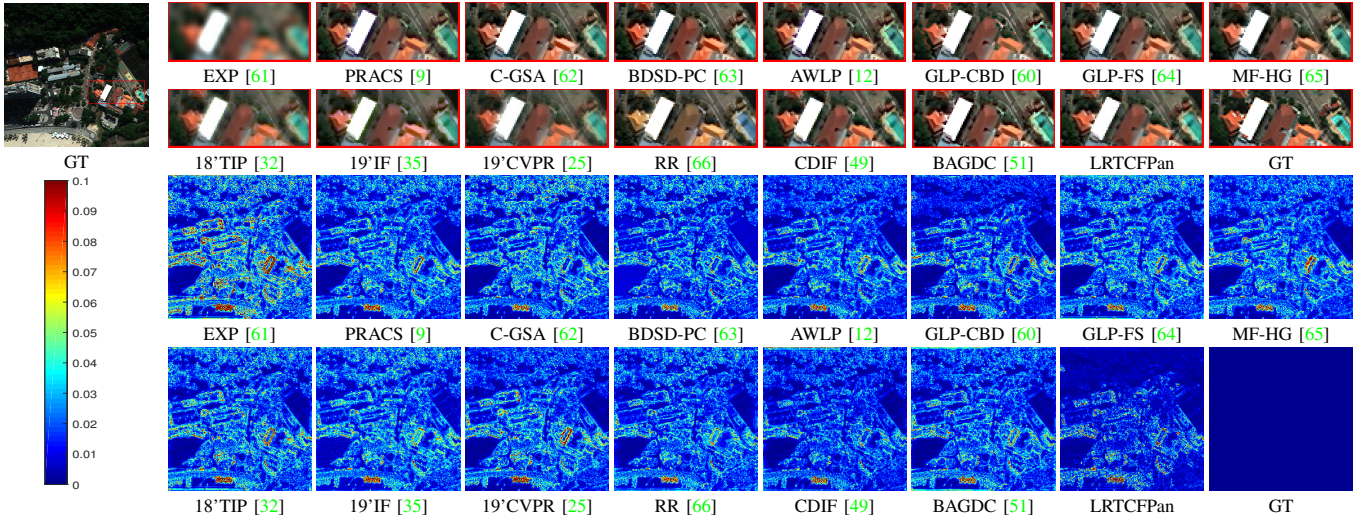


Fig. 7. The fusion results on the reduced-resolution Rio dataset (source: WV-3). The first two rows: the visual inspection of the ground-truth (GT) image and the close-ups of the fused images. The last two rows: the residual maps using the GT image as a reference.

465 Figs. 6-7. Compared with the GT image, Fig. 6 reveals that  
 466 the GLP-CBD, the GLP-FS, the CDIF, and our LRTCFFPan  
 467 methods obtain the better performance from both spectral  
 468 and spatial perspectives. However, other comparators achieve  
 469 inferior performance considering the overall or local feature  
 470 evaluation. It is worth underlining that clearer details do not  
 471 always mean superior performance, e.g., the images recovered  
 472 by the C-GSA, the AWLP, and the MF-HG methods. That  
 473 is because the details exceeding those of the GT image are  
 474 regarded as errors. The performance of Fig. 7 is similar to that  
 475 of Fig. 6. More specifically, the C-GSA, the GLP-CBD, the  
 476 19'IF, the CDIF, and the proposed LRTCFFPan methods achieve  
 477 better visual performance. Nonetheless, the other compared  
 478 methods reflect varying levels of color deviation and spatial  
 479 blurring. From the corresponding residual images of Figs. 6-7,  
 480 we can further confirm that the proposed LRTCFFPan method is  
 481 superior to other methods, clarifying its significant advantages.  
 482

483 2) *Full-Resolution Data Experiment*: To corroborate the re-  
 484 sults obtained at reduced resolution, the proposed LRTCFFPan  
 485 method is further evaluated at the real experimental images,  
 486 which are cropped from the real datasets, including the real-  
 487 world Guangzhou (sensor: GF-2), Indianapolis (sensor: QB),  
 488 and Rio (sensor: WV-3) datasets. Subsequently, the visual  
 489 performance is displayed in Figs. 8-9. In this case, the visual  
 490 comparison requires the PAN image as the spatial reference,  
 491 whereas the LR-MS image (or the recovered image of the  
 492 EXP method) is the spectral reference. According to Fig. 8,  
 493 although many compared approaches, e.g., the PRACS, the  
 494 AWLP, the GLP-FS, the MF-HG, and the 19'IF, obtain clearer  
 495 details, the inferior spectral fidelity is caused. Moreover, the  
 496 C-GSA, the GLP-CBD, the 18'TIP, and the CDIF methods  
 497 generate abnormal colors, structures, or artifacts. In contrast,  
 498 the LRTCFFPan and the BDS-D-PC methods show the better  
 499 trade-off between spatial sharpening and spectral consistency.  
 500 From Fig. 9, we can observe that only the C-GSA, the BDS-D-  
 501 PC, the 19'IF, and the LRTCFFPan methods can reconstruct the  
 502 right shape and color of the acquired car. Especially, only the

LRTCFFPan method can recover the correct direction of the  
 shadow of the car. Therefore, the effectiveness and superiority  
 of the LRTCFFPan method are corroborated at full resolution.

### B. Quantitative Comparison

To quantitatively compare the LRTCFFPan method with  
 other methods, we provide the average numerical metrics  
 of 82, 42, 15, 15, 15, and 42 images, which are selected  
 from the simulated Guangzhou (sensor: GF-2), the simulated  
 Indianapolis (sensor: QB), the simulated Rio (sensor: WV-3),  
 the real-world Guangzhou (sensor: GF-2), the real-world Indi-  
 anapolis (sensor: QB), and the real-world Rio (sensor: WV-3)  
 datasets, respectively. The statistical values of all the metrics  
 (means and related standard deviations) and the computational  
 times are shown in Tables I, II, III, IV-(a), IV-(b), and V.  
 Notably, the variational methods, i.e., the 18'TIP, the 19'IF,  
 the 19'CVPR, the RR, the CDIF, the BAGDC, and the LRTCFFPan,  
 are implemented using only one set of parameters for all  
 the experiments of the same dataset. Consequently, better  
 performance also implies higher robustness of the parameters.  
 From the results, we observe that the proposed LRTCFFPan  
 method generally achieves better average values than the other  
 methods, demonstrating its numerical superiority.

### C. Discussions

1) *Parameter Analysis*: In Algorithm 1, seven hyperparam-  
 eters are theoretically involved, including the regularization  
 parameters (i.e.,  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$ ), the penalty parameters (i.e.,  
 $\eta_1$ ,  $\eta_2$ , and  $\eta_3$ ), and the blocksize of the block-based DDM  
 regularizer. Among them,  $\lambda_3$  and  $\eta_1$  control the low-tubal-rank  
 properties of  $\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1$  and  $\mathcal{X}$ , respectively. Empirically,  
 $\lambda_3$  and  $\eta_1$  can be pre-determined within a small range, e.g.,  
 $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$ . Similarly, the blocksize can also be  
 selected from  $\{8 \times 8, 10 \times 10\}$ , showing promising results in  
 almost all the experiments. Afterwards, the remaining param-  
 eters, i.e.,  $\lambda_1$ ,  $\lambda_2$ ,  $\eta_2$ , and  $\eta_3$ , are searched by jointly reaching  
 the optimal SAM, SCC, ERGAS, and  $Q^{2n}$  metrics. For

TABLE III  
THE QUALITY METRICS ON 15 IMAGES WITH A PAN SIZE OF  $256 \times 256$  FROM THE REDUCED-RESOLUTION RIO DATASET (SOURCE: WV-3).  
(BOLD: BEST; UNDERLINE: SECOND BEST)

Method	PSNR	SSIM	SAM	SCC	ERGAS	Q8	Runtime[s]
EXP [61]	27.409 ± 1.281	0.678 ± 0.054	7.472 ± 1.144	0.835 ± 0.044	8.441 ± 0.954	0.678 ± 0.034	<b>0.02</b>
PRACS [9]	30.615 ± 1.263	0.844 ± 0.028	7.704 ± 1.245	0.923 ± 0.018	5.871 ± 0.624	0.843 ± 0.012	0.16
C-GSA [62]	31.245 ± 1.051	0.853 ± 0.027	7.888 ± 1.408	0.928 ± 0.016	5.567 ± 0.548	0.862 ± 0.026	0.53
BDS-PC [63]	31.521 ± 1.106	0.873 ± 0.021	7.443 ± 1.143	0.933 ± 0.015	5.313 ± 0.535	0.879 ± 0.018	<u>0.08</u>
AWLP [12]	31.182 ± 1.189	0.874 ± 0.020	7.109 ± 1.016	0.930 ± 0.016	5.412 ± 0.585	0.871 ± 0.007	0.18
GLP-CBD [60]	31.131 ± 1.235	0.879 ± 0.019	6.608 ± 0.891	0.929 ± 0.017	5.549 ± 0.545	0.877 ± 0.003	52.30
GLP-FS [64]	31.102 ± 1.070	0.861 ± 0.025	7.308 ± 1.230	0.930 ± 0.016	5.499 ± 0.538	0.865 ± 0.017	0.14
MF-HG [65]	30.884 ± 1.200	0.865 ± 0.026	7.067 ± 1.166	0.925 ± 0.018	5.664 ± 0.614	0.863 ± 0.011	0.25
18'TIP [32]	29.786 ± 1.178	0.812 ± 0.031	7.227 ± 1.124	0.912 ± 0.020	6.373 ± 0.685	0.825 ± 0.012	73.95
19'IF [35]	30.088 ± 1.108	0.841 ± 0.024	7.855 ± 1.173	0.921 ± 0.016	5.831 ± 0.574	0.840 ± 0.015	23.74
19'CVPR [25]	30.157 ± 1.413	0.838 ± 0.033	6.680 ± 1.034	0.920 ± 0.021	6.159 ± 0.718	0.829 ± 0.017	17.10
RR [66]	30.972 ± 1.103	0.870 ± 0.019	7.043 ± 1.018	0.928 ± 0.017	5.317 ± 0.583	0.867 ± 0.017	54.48
CDIF [49]	<u>31.808 ± 1.395</u>	<u>0.883 ± 0.020</u>	<u>6.260 ± 0.851</u>	<u>0.938 ± 0.014</u>	<u>5.010 ± 0.522</u>	<u>0.891 ± 0.013</u>	81.18
BAGDC [51]	30.881 ± 0.921	0.874 ± 0.018	7.276 ± 1.051	0.928 ± 0.015	5.388 ± 0.579	0.872 ± 0.018	1.16
LRTCFFan	<b>32.251 ± 1.333</b>	<b>0.891 ± 0.018</b>	<b>6.132 ± 0.880</b>	<b>0.945 ± 0.015</b>	<b>4.834 ± 0.576</b>	<b>0.901 ± 0.004</b>	57.15
<b>Ideal value</b>	<b><math>+\infty</math></b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>	-

TABLE IV  
THE QUANTITATIVE RESULTS FOR ALL THE COMPARED METHODS ON (A) 15 IMAGES FROM THE FULL-RESOLUTION GUANGZHOU DATASET (SOURCE: GF-2) AND (B) 15 IMAGES FROM THE FULL-RESOLUTION INDIANAPOLIS DATASET (SOURCE: QB). THE SIZE OF THE PAN IMAGE IS  $400 \times 400$ .  
(BOLD: BEST; UNDERLINE: SECOND BEST)

Method	(a) Full-Resolution Guangzhou Dataset			(b) Full-Resolution Indianapolis Dataset			Time[s]
	$D_\lambda$	$D_s$	QNR	$D_\lambda$	$D_s$	QNR	
EXP [61]	<b>0.002 ± 0.001</b>	0.163 ± 0.040	0.836 ± 0.040	<b>0.003 ± 0.001</b>	0.121 ± 0.022	0.877 ± 0.022	<b>0.03</b>
PRACS [9]	0.054 ± 0.018	0.063 ± 0.028	0.886 ± 0.041	0.038 ± 0.020	0.083 ± 0.038	0.883 ± 0.052	0.19
C-GSA [62]	0.100 ± 0.036	0.099 ± 0.045	0.812 ± 0.070	0.080 ± 0.058	0.137 ± 0.081	0.798 ± 0.118	0.68
BDS-PC [63]	0.066 ± 0.029	0.077 ± 0.041	0.863 ± 0.065	0.029 ± 0.026	0.068 ± 0.031	0.906 ± 0.049	<u>0.07</u>
AWLP [12]	0.086 ± 0.071	0.090 ± 0.081	0.836 ± 0.132	0.061 ± 0.025	0.068 ± 0.035	0.876 ± 0.055	0.28
GLP-CBD [60]	0.078 ± 0.038	0.053 ± 0.043	0.874 ± 0.075	0.038 ± 0.029	0.048 ± 0.029	0.917 ± 0.055	62.19
GLP-FS [64]	0.090 ± 0.033	0.075 ± 0.053	0.843 ± 0.078	0.063 ± 0.027	0.069 ± 0.037	0.873 ± 0.058	0.14
MF-HG [65]	0.110 ± 0.058	0.106 ± 0.081	0.799 ± 0.119	0.072 ± 0.033	0.073 ± 0.033	0.861 ± 0.059	0.09
18'TIP [32]	0.070 ± 0.042	0.050 ± 0.037	0.884 ± 0.069	0.060 ± 0.060	0.058 ± 0.056	0.889 ± 0.105	111.31
19'IF [35]	0.167 ± 0.063	0.158 ± 0.092	0.706 ± 0.125	0.147 ± 0.072	0.201 ± 0.104	0.688 ± 0.141	34.14
19'CVPR [25]	<u>0.006 ± 0.002</u>	0.101 ± 0.028	0.893 ± 0.028	<u>0.013 ± 0.007</u>	0.071 ± 0.014	0.916 ± 0.020	40.83
RR [66]	0.107 ± 0.047	0.128 ± 0.049	0.781 ± 0.081	0.089 ± 0.066	0.112 ± 0.064	0.812 ± 0.111	44.57
CDIF [49]	0.032 ± 0.016	<u>0.040 ± 0.018</u>	<u>0.929 ± 0.022</u>	0.026 ± 0.013	<u>0.031 ± 0.007</u>	<u>0.943 ± 0.015</u>	90.74
BAGDC [51]	0.037 ± 0.025	0.042 ± 0.021	0.923 ± 0.040	0.031 ± 0.027	0.036 ± 0.023	0.935 ± 0.047	2.85
LRTCFFan	0.043 ± 0.018	<b>0.019 ± 0.010</b>	<b>0.939 ± 0.021</b>	0.020 ± 0.009	<b>0.025 ± 0.011</b>	<b>0.955 ± 0.018</b>	77.73
<b>Ideal value</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>	-

538 brevity, Fig. 10 presents the performance of varying  $\lambda_1$ ,  $\lambda_2$ ,  
539  $\eta_2$ , and  $\eta_3$  on the reduced-resolution Guangzhou data (source:  
540 GF-2). Obviously,  $\lambda_1 = 5 \times 10^{-2}$ ,  $\lambda_2 = 1.8 \times 10^1$ ,  $\eta_2 = 8.1$ ,  
541 and  $\eta_3 = 1.8$  are the best parameters for configuration. By  
542 adopting the same strategy on different datasets, all parameter  
543 configurations can be obtained and provided in Table VI.

544 2) *Algorithm Convergence*: Since the log tensor nuclear  
545 norm of Definition II.4 is non-convex, the convergence of the  
546 proposed ADMM-based LRTCFFan algorithm cannot be theo-  
547 retically guaranteed. As depicted in Fig. 11, we numerically  
548 illustrate the algorithm convergence on the reduced-resolution  
549 Guangzhou (sensor: GF-2), Indianapolis (sensor: QB), and

Rio (sensor: WV-3) datasets. For a better presentation, the  
550 maximum number of iterations is empirically set to 200.  
551 In any considered case, the value of the objective function  
552 becomes stable as the iteration number increases, implying  
553 the numerical convergence behavior of Algorithm 1.  
554

555 3) *Ablation Study*: For deeper insights into the LRTCFFan  
556 model, we further conduct the ablation study of model (12)  
557 on the reduced-resolution Guangzhou image (sensor: GF-2).  
558 The following three sub-models are generated to independently  
559 verify the contributions of the two low-tubal-rank priors and  
560 the proposed local-similarity-based DDM regularizer.

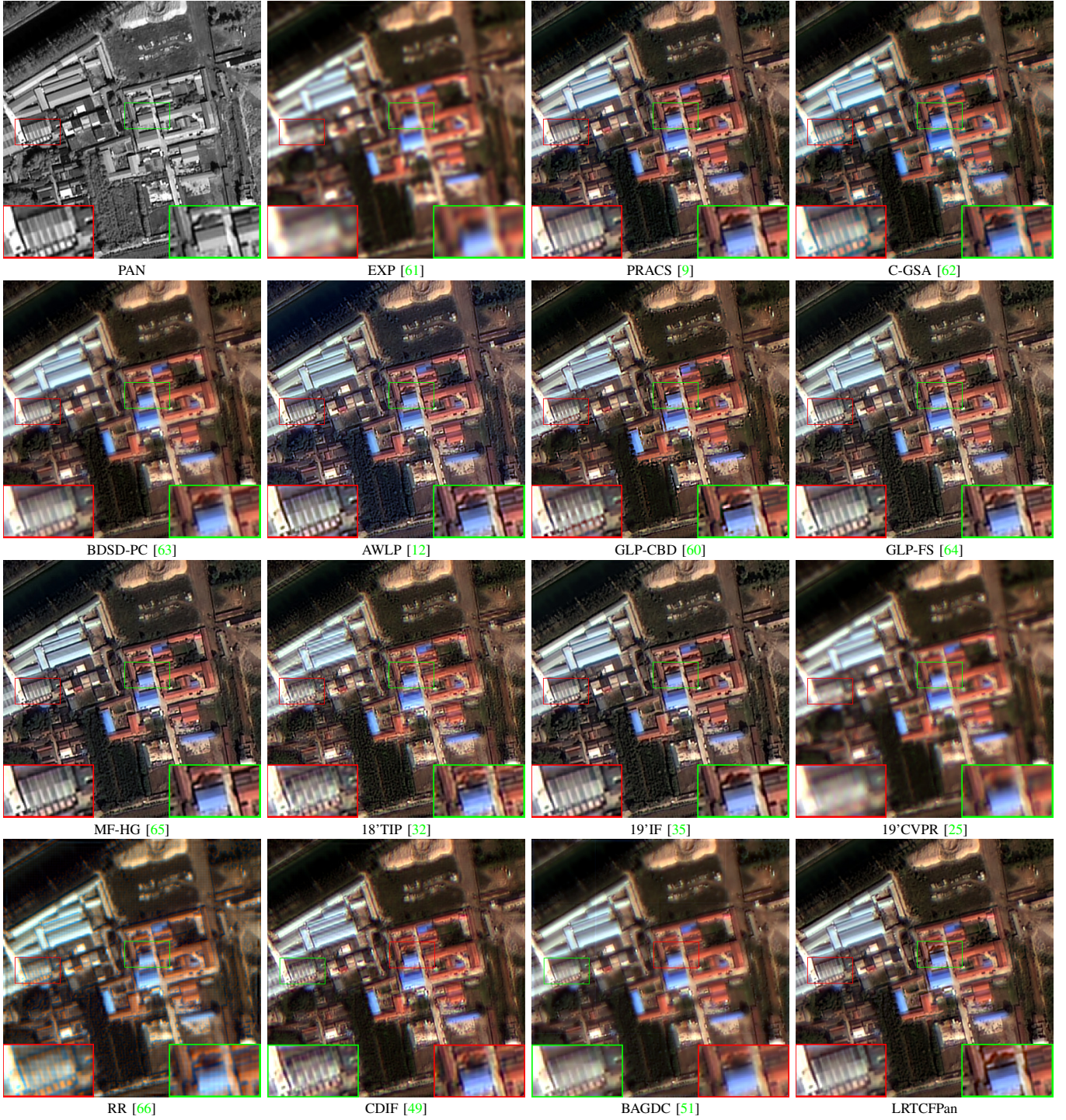


Fig. 8. The RGB compositions of the fused images on the full-resolution Guangzhou dataset (source: GF-2). The size of the PAN image is  $400 \times 400$ . The close-ups are depicted in the bottom corners of the images.

Submodel-I:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{T}} \|\mathcal{X}\|_{lt} + \lambda_1 \|\mathcal{X} - \mathcal{X} \bullet \mathcal{B} - \mathcal{D}\|_F^2 + \lambda_2 \|\mathcal{X} \bullet \mathcal{B} - \mathcal{T}\|_F^2 \\ \text{s.t. } \mathcal{P}_\Omega(\mathcal{T}) = \mathcal{Y} \uparrow_{r,0}, \end{aligned}$$

Submodel-II:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{T}} \|\mathcal{T}\|_{lt} + \lambda_1 \|\mathcal{X} - \mathcal{X} \bullet \mathcal{B} - \mathcal{D}\|_F^2 + \lambda_2 \|\mathcal{X} \bullet \mathcal{B} - \mathcal{T}\|_F^2 \\ \text{s.t. } \mathcal{P}_\Omega(\mathcal{T}) = \mathcal{Y} \uparrow_{r,0}, \end{aligned}$$

Submodel-III:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{T}} \|\mathcal{X}\|_{lt} + \lambda_1 \|\mathcal{X} \bullet \mathcal{B} - \mathcal{T}\|_F^2 + \lambda_2 \|\mathcal{T}\|_{lt} \\ \text{s.t. } \mathcal{P}_\Omega(\mathcal{T}) = \mathcal{Y} \uparrow_{r,0}. \end{aligned}$$

After all optimal parameter configurations are satisfied, the quantitative results of these models are reported in Table VII. As observed, the models employing the local-similarity-based DDM regularizer (i.e., Submodel-I and Submodel-II) perform better, implying the remarkable effectiveness of the regularizer.

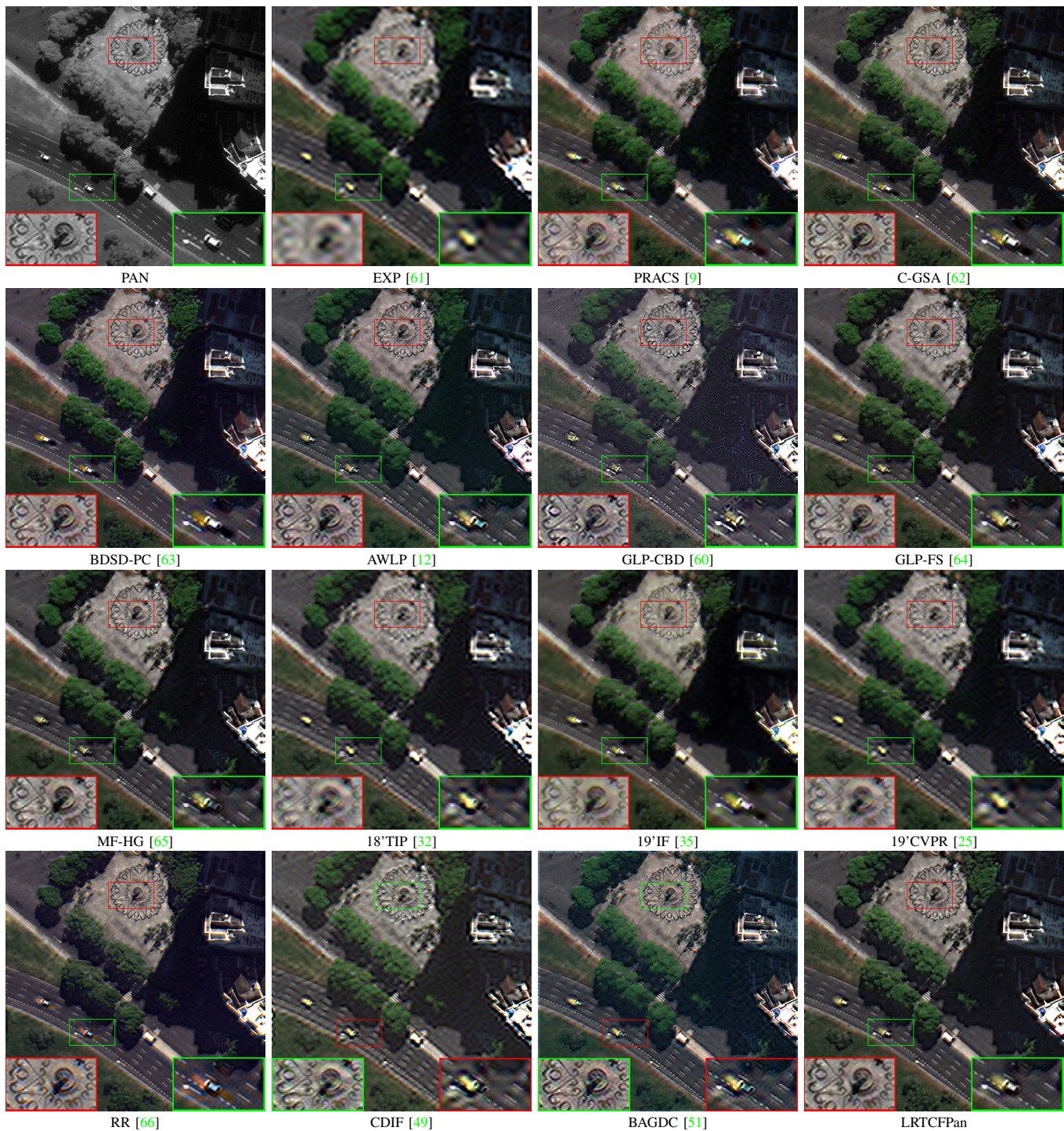


Fig. 9. The RGB compositions of the fused images on the full-resolution Rio dataset (source: WV-3). The size of the PAN image is  $400 \times 400$ . The close-ups are depicted in the bottom corners of the images.

566 Moreover, two low-tubal-rank priors also realize incremental  
 567 performance improvements. Accordingly, the three regulariz-  
 568 ers collectively contribute to the LRTCfPan model.

569 4) *Comparison of ISR Degradation Models:* For decoupling  
 570 the original  $\mathcal{Y} = (\mathcal{X} \bullet \mathcal{B}) \downarrow_r + \mathcal{N}_0$ , the variable substitution  
 571 is usually involved, e.g., [24], leading to the following con-  
 572 strained model

$$\min_{\mathcal{X}, \mathcal{Z}} \frac{1}{2} \|\mathcal{Z} \downarrow_r - \mathcal{Y}\|_F^2 \quad \text{s.t.} \quad \mathcal{Z} = \mathcal{X} \bullet \mathcal{B}, \quad (38)$$

whose augmented Lagrangian function is

$$\mathcal{L}(\mathcal{X}, \mathcal{Z}) = \frac{1}{2} \|\mathcal{Z} \downarrow_r - \mathcal{Y}\|_F^2 + \frac{\eta}{2} \left\| \mathcal{X} \bullet \mathcal{B} - \mathcal{Z} + \frac{\Lambda}{\eta} \right\|_F^2. \quad (39)$$

574 However, when the new ISR degradation model  $\mathcal{Y} = (\mathcal{X} \bullet \mathcal{B} +$   
 575  $\mathcal{N}_1) \downarrow_r$  is employed, we only need to consider the augmented  
 576 Lagrangian function as follows,

$$\mathcal{L}(\mathcal{X}, \mathcal{T}) = \frac{1}{2} \|\mathcal{X} \bullet \mathcal{B} - \mathcal{T}\|_F^2 + \iota(\mathcal{T}). \quad (40)$$

573

574

575

576

TABLE V  
THE QUALITY METRICS FOR 42 IMAGES FROM THE FULL-RESOLUTION RIO DATASET (SOURCE: WV-3). THE SIZE OF THE PAN IMAGE IS  $400 \times 400$ . (BOLD: BEST; UNDERLINE: SECOND BEST)

Method	$D_\lambda$	$D_s$	QNR	Time[s]
EXP [61]	<b>0.004</b> $\pm$ <b>0.001</b>	0.105 $\pm$ 0.019	0.892 $\pm$ 0.019	<b>0.06</b>
PRACS [9]	0.018 $\pm$ 0.013	0.054 $\pm$ 0.035	0.928 $\pm$ 0.040	0.51
C-GSA [62]	0.044 $\pm$ 0.038	0.075 $\pm$ 0.064	0.887 $\pm$ 0.086	0.94
BDS-PC [63]	0.020 $\pm$ 0.011	<u>0.044</u> $\pm$ <u>0.021</u>	0.937 $\pm$ 0.029	<u>0.14</u>
AWLP [12]	0.051 $\pm$ 0.057	0.058 $\pm$ 0.072	0.898 $\pm$ 0.101	0.57
GLP-CBD [60]	0.065 $\pm$ 0.084	0.046 $\pm$ 0.037	0.894 $\pm$ 0.100	119.49
GLP-FS [64]	0.045 $\pm$ 0.047	0.056 $\pm$ 0.064	0.904 $\pm$ 0.091	0.26
MF-HG [65]	0.053 $\pm$ 0.050	0.064 $\pm$ 0.055	0.889 $\pm$ 0.087	0.18
18'TIP [32]	0.035 $\pm$ 0.030	0.067 $\pm$ 0.041	0.902 $\pm$ 0.060	212.45
19'IF [35]	0.087 $\pm$ 0.043	0.096 $\pm$ 0.048	0.828 $\pm$ 0.080	55.94
19'CVPR [25]	<u>0.016</u> $\pm$ <u>0.006</u>	0.046 $\pm$ 0.012	<u>0.939</u> $\pm$ <u>0.016</u>	73.26
RR [66]	0.062 $\pm$ 0.052	0.086 $\pm$ 0.077	0.861 $\pm$ 0.103	102.93
CDIF [49]	0.028 $\pm$ 0.009	0.048 $\pm$ 0.016	0.926 $\pm$ 0.018	182.18
BAGDC [51]	0.060 $\pm$ 0.055	0.048 $\pm$ 0.048	0.898 $\pm$ 0.088	4.60
LRTCFFan	0.022 $\pm$ 0.013	<b>0.022</b> $\pm$ <b>0.027</b>	<b>0.956</b> $\pm$ <b>0.036</b>	153.51
Ideal value	<b>0</b>	<b>0</b>	<b>1</b>	-

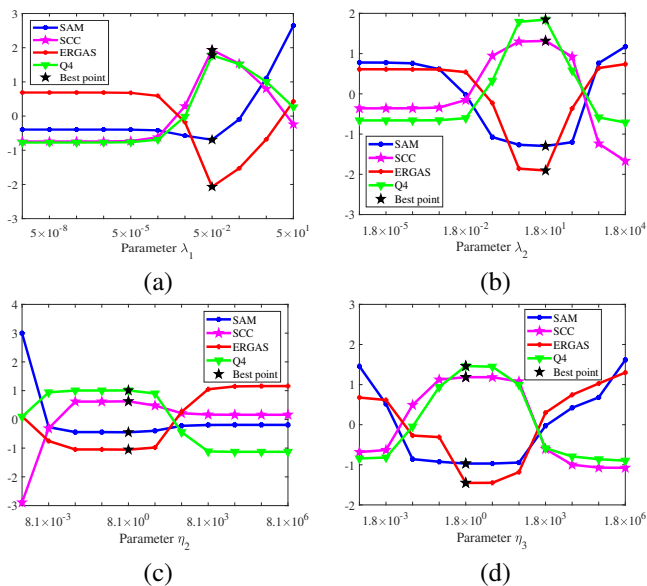


Fig. 10. The SAM, SCC, ERGAS, and Q4 curves for (a)  $\lambda_1$ , (b)  $\lambda_2$ , (c)  $\eta_2$ , and (d)  $\eta_3$  on a reduced-resolution Guangzhou image (sensor: GF-2). To show them with the same range of values, the obtained indexes are post-processed by zero-mean normalization, i.e.,  $(index - \text{Mean}(index)) / \text{Std}(index)$ . Moreover, the means and the standard deviations of the SAM, the SCC, the ERGAS, and the Q4 are provided for four subfigures, i.e., (a)  $2.341 \pm 0.467$ ;  $0.947 \pm 0.017$ ;  $2.897 \pm 0.596$ ;  $0.812 \pm 0.069$ , (b)  $14.913 \pm 9.950$ ;  $0.588 \pm 0.298$ ;  $19.503 \pm 9.366$ ;  $0.267 \pm 0.362$ , (c)  $3.515 \pm 3.335$ ;  $0.904 \pm 0.120$ ;  $13.638 \pm 11.367$ ;  $0.494 \pm 0.438$ , and (d)  $10.008 \pm 8.247$ ;  $0.500 \pm 0.405$ ;  $17.518 \pm 10.917$ ;  $0.358 \pm 0.394$ .

Under the ADMM algorithm framework, the proposed ISR degradation model avoids the computational complexity (i.e.,  $\mathcal{O}(HWS)$ ) of solving  $\frac{1}{2} \|\mathcal{Z} \downarrow_r - \mathcal{Y}\|_F^2$ . As depicted in Fig. 12, the computational times are reduced. Moreover, since the downsampling operator  $\downarrow_r$  is eliminated by the tensor completion step, the matrixization of  $\downarrow_r$  is not included in the resulting model. Consequently, the proposed LRTCFFan model can be formulated in the tensor-based form, which is more physically intuitive than the matrix-based modeling or the mixture of

TABLE VI  
THE HYPER-PARAMETER SETTINGS OF THE PROPOSED MODEL FOR DIFFERENT CASES. (R: REDUCED RESOLUTION; F: FULL RESOLUTION)

Dataset	Case	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\eta_1$	$\eta_2$	$\eta_3$	Blocksize
Guangzhou	R	0.05	18	$10^{-4}$	$10^{-4}$	8.1	1.8	$8 \times 8$
	F	1.00	50	$10^{-4}$	$10^{-4}$	2.1	4.7	$10 \times 10$
Indianapolis	R	0.11	65	$10^{-4}$	$10^{-4}$	1.1	8.7	$8 \times 8$
	F	0.40	75	$10^{-1}$	$10^{-3}$	2.1	6.7	$10 \times 10$
Rio	R	0.14	56	$10^{-4}$	$10^{-4}$	4.2	8.3	$8 \times 8$
	F	1.10	36	$10^{-4}$	$10^{-4}$	6.2	3.3	$10 \times 10$

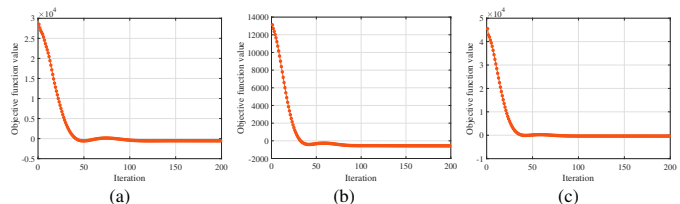


Fig. 11. The curves of the objective function values on the reduced-resolution (a) Guangzhou (sensor: GF-2), (b) Indianapolis (sensor: QB), and (c) Rio (sensor: WV-3) datasets.

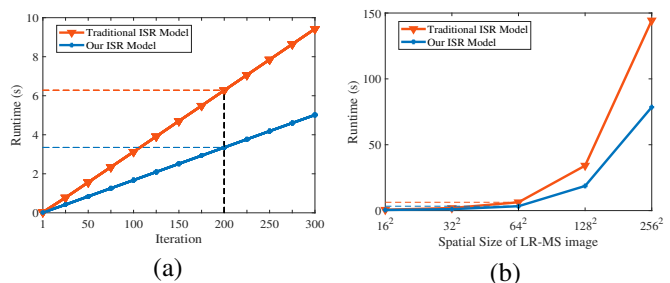


Fig. 12. The comparison of the computational burden between two ISR degeneration models by using two different cases, i.e., (a) the runtime versus the number of iterations when the spatial size of the LR-MS image is  $64 \times 64$  and (b) the runtime versus the spatial size of the LR-MS image when the number of iterations is fixed to 200. The reduced-resolution Guangzhou dataset (source: GF-2) is employed.

unfolding-based and tensor-based modeling, e.g., [39].

5) *Applicable Scope*: Since the proposed LRTCFFan model incorporates the low-tubal-rank prior, we further perform the applicability analysis by investigating the tubal-rank characteristic of numerous multispectral images. For such a statistical analysis, all simulated experimental data, i.e., 82 Guangzhou images (sensor: GF-2), 42 Indianapolis images (sensor: QB), and 15 Rio images (sensor: WV-3), are employed. According to Fig. 13, the corresponding multispectral images demonstrate specific low-rank characteristics. Consequently, the applicability of the proposed LRTCFFan model can be established.

6) *Comparison with CNN-based Method*: In the previous numerical experiments, only the traditional CS, MRA, and variational pansharpening methods are involved. To comprehensively demonstrate the performance, we further compare the proposed LRTCFFan model with the CNN-based DCFNet method [73] on all reduced-resolution data, i.e., 82 Guangzhou images (sensor: GF-2), 42 Indianapolis images (sensor: QB), and 15 Rio images (sensor: WV-3). Particularly, the pretraining datasets of the DCFNet model for the GF-2, QB, and WV-3 cases are the Beijing (sensor: GF-2), Indianapolis (sensor:

TABLE VII  
THE QUANTITATIVE RESULTS OF THE ABLATION EXPERIMENT ON THE REDUCED-RESOLUTION GUANGZHOU DATA (SOURCE: GF-2).  
(BOLD: BEST; UNDERLINE: SECOND BEST)

Configuration	ISR Degradation Model	Low-Rank Prior for $\mathcal{X} \bullet \mathcal{B} + \mathcal{N}_1$	Low-Rank Prior for $\mathcal{X}$	Local-Similarity-Based DDM Regularizer	PSNR	SSIM	SAM	SCC	ERGAS	Q4
EXP [61]	✓	✗	✗	✗	29.3053	0.8016	2.4860	0.9429	3.1620	0.8360
Submodel-I	✓	✗	✓	✓	<u>35.0918</u>	<u>0.9150</u>	<u>2.0104</u>	<u>0.9802</u>	<u>1.6582</u>	<u>0.9359</u>
Submodel-II	✓	✓	✗	✓	34.9260	0.9109	2.0373	0.9794	1.6971	0.9327
Submodel-III	✓	✓	✓	✗	29.9401	0.7899	2.5329	0.9494	2.9143	0.8332
LRTCFFPan	✓	✓	✓	✓	<b>35.1550</b>	<b>0.9155</b>	<b>2.0089</b>	<b>0.9803</b>	<b>1.6470</b>	<b>0.9364</b>
<b>Ideal value</b>	-	-	-	-	$+\infty$	<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>

TABLE VIII  
THE QUALITY METRICS OF DIFFERENT METHODS ON THE REDUCED-RESOLUTION GUANGZHOU (SENSOR: GF-2), INDIANAPOLIS (SENSOR: QB), AND RIO (SENSOR: WV-3) DATASETS. (BOLD: BEST; UNDERLINE: SECOND BEST)

Dataset	Sensor	Method	PSNR	SSIM	SAM	SCC	ERGAS	Q2 <sup>n</sup>
Guangzhou	GF-2	DCFNet [73]	34.695 ± 1.450	0.899 ± 0.018	1.834 ± 0.265	0.957 ± 0.017	1.598 ± 0.179	0.898 ± 0.042
		LRTCFFPan	<b>35.918 ± 2.087</b>	<b>0.921 ± 0.022</b>	<b>1.391 ± 0.274</b>	<b>0.968 ± 0.014</b>	<b>1.496 ± 0.275</b>	<b>0.926 ± 0.039</b>
Indianapolis	QB	DCFNet [73]	31.295 ± 2.231	<b>0.877 ± 0.022</b>	<b>6.002 ± 0.914</b>	0.896 ± 0.018	8.105 ± 0.890	0.848 ± 0.095
		LRTCFFPan	<b>32.727 ± 2.132</b>	0.873 ± 0.025	7.032 ± 1.264	<b>0.922 ± 0.016</b>	<b>6.964 ± 0.596</b>	<b>0.861 ± 0.092</b>
Rio	WV-3	DCFNet [73]	<b>36.692 ± 0.494</b>	<b>0.964 ± 0.006</b>	<b>3.699 ± 0.723</b>	<b>0.982 ± 0.004</b>	<b>2.388 ± 0.625</b>	<b>0.971 ± 0.010</b>
		LRTCFFPan	32.251 ± 1.333	0.891 ± 0.018	6.132 ± 0.880	0.945 ± 0.015	4.834 ± 0.576	0.901 ± 0.004
<b>Ideal value</b>			$+\infty$	<b>1</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>

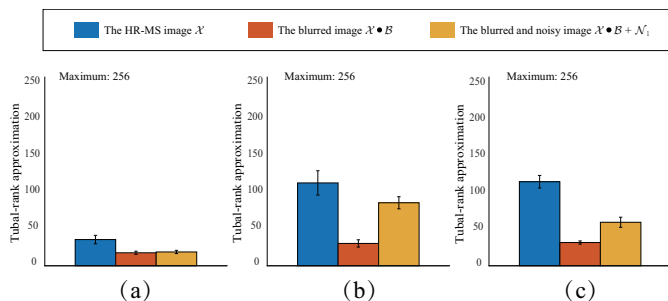


Fig. 13. The statistics of the approximation of the tubal rank on different simulated datasets, including (a) 82 Guangzhou images (sensor: GF-2), (b) 42 Indianapolis images (sensor: QB), and (c) 15 Rio images (sensor: WV-3). The standard deviation of Gaussian noise is 0.01.

607 QB), and both Rio and Tripoli (sensor: WV-3), respectively.  
 608 The numerical metrics are reported in Table VIII. For the  
 609 WV-3 case, the DCFNet method is significantly superior to  
 610 the LRTCFFPan method, which is reasonable provided that the  
 611 Rio dataset is included in the training data of the former.  
 612 Furthermore, when applied to the Indianapolis dataset (testing  
 613 images), the DCFNet method does not exhibit the advantage  
 614 over the LRTCFFPan method, even if the former is pretrained  
 615 on the Indianapolis dataset. Instead, the DCFNet method is  
 616 inferior to the LRTCFFPan method on the Guangzhou dataset  
 617 owing to its limited generalization ability. Consequently, the  
 618 superior algorithm robustness and generalization capability of  
 619 the LRTCFFPan method are mainly demonstrated, which may  
 620 endow such a method with more practical significance.

## VI. CONCLUSIONS

In this paper, we proposed a novel LRTC-based framework for pansharpening, called LRTCFFPan. Specifically, we first deduced an ISR degradation model, thus eliminating the down-sampling operator and transforming the original pansharpening problem into the LRTC-based framework with the deblurring regularizer. Moreover, we designed a local-similarity-based DDM regularizer, which dynamically and locally integrates the spatial information from the PAN image to the underlying HR-MS image. For better completion and global characterization, two low-tubal-rank constraints are simultaneously imposed. To regularize the proposed model, we developed an efficient ADMM-based algorithm. The numerical experiments demonstrated the superiority of the proposed LRTCFFPan method.

## REFERENCES

- [1] M. Zanetti, F. Bovolo, and L. Bruzzone, "Rayleigh-Rice mixture parameter estimation via EM algorithm for change detection in multispectral images," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5004–5016, 2015. 1
- [2] X. Yu, L. E. Hoff, I. S. Reed, A. M. Chen, and L. B. Stotts, "Automatic target detection and recognition in multiband imagery: A unified ML detection and estimation approach," *IEEE Trans. Image Process.*, vol. 6, no. 1, pp. 143–156, 1997. 1
- [3] P. Zhong and R. Wang, "Learning conditional random fields for classification of hyperspectral images," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1890–1907, 2010. 1
- [4] P. X. Zhuang, Q. S. Liu, and X. H. Ding, "Pan-GGF: A probabilistic method for pan-sharpening with gradient domain guided image filtering," *Signal Process.*, vol. 156, pp. 177–190, 2019. 1
- [5] P. Kwarteng and A. Chavez, "Extracting spectral contrast in landsat thematic mapper image data using selective principal component analysis," *Photogramm. Eng. Remote Sens.*, vol. 55, no. 1, pp. 339–348, 1989. 1
- [6] W. Carper, T. Lillesand, and R. Kiefer, "The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multi-spectral image data," *Photogramm. Eng. Remote Sens.*, vol. 56, no. 4, pp. 459–467, 1990. 1

- [7] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS + Pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, 2007. 1
- [8] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE pansharpening of very high resolution multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 228–236, 2007. 1
- [9] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, 2010. 1, 7, 8, 9, 10, 11, 12, 13
- [10] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. A. Licciardi, R. Restaino, and L. Wald, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, 2014. 1, 4
- [11] M. J. Shensa, "The discrete wavelet transform: Wedding the a trous and Mallat algorithms," *IEEE Trans. Signal Process.*, vol. 40, no. 10, pp. 2464–2482, 1992. 1
- [12] X. Otazu, M. Gonzalezaudicana, O. Fors, and J. Nunez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, 2005. 1, 7, 8, 9, 10, 11, 12, 13
- [13] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, 2000. 1
- [14] J. F. Yang, X. Y. Fu, Y. W. Hu, Y. Huang, X. H. Ding, and J. Paisley, "PanNet: A deep network architecture for pansharpening," in *Int. Conf. Comput. Vision (ICCV)*, 2017, pp. 5449–5457. 2
- [15] J. F. Hu, T. Z. Huang, L. J. Deng, T. X. Jiang, G. Vivone, and J. Chanussot, "Hyperspectral image super-resolution via deep spatio-spectral attention convolutional neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, pp. 1–15, 2021. 2
- [16] Z. Zhu, J. Hou, J. Chen, H. Zeng, and J. Zhou, "Hyperspectral image super-resolution via deep progressive zero-centric residual learning," *IEEE Trans. Image Process.*, vol. 30, pp. 1423–1438, 2020. 2
- [17] T. Huang, W. Dong, J. Wu, L. Li, X. Li, and G. Shi, "Deep hyperspectral image fusion network with iterative spatio-spectral regularization," *IEEE Trans. Comput. Imaging*, vol. 8, pp. 201–214, 2022. 2
- [18] L. J. Deng, G. Vivone, C. Jin, and J. Chanussot, "Detail injection-based deep convolutional neural networks for pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6995–7010, 2020. 2, 4
- [19] P. Adesso, G. Vivone, R. Restaino, and J. Chanussot, "A data-driven model-based regression applied to panchromatic sharpening," *IEEE Trans. Image Process.*, vol. 29, pp. 7779–7794, 2020. 2
- [20] Z. R. Jin, L. J. Deng, T. J. Zhang, and X. X. Jin, "BAM: Bilateral activation mechanism for image fusion," in *ACM Int. Conf. on Multimedia (ACM MM)*, 2021. 2
- [21] Y. D. Wang, L. J. Deng, T. J. Zhang, and X. Wu, "SSconv: Explicit spectral-to-spatial convolution for pansharpening," in *ACM Int. Conf. on Multimedia (ACM MM)*, 2021. 2
- [22] X. Y. Fu, W. Wang, Y. Huang, X. H. Ding, and J. Paisley, "Deep multiscale detail networks for multiband spectral image sharpening," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 5, pp. 2090–2104, 2021. 2
- [23] X. Lu, J. Zhang, D. Yang, L. Xu, and F. Jia, "Cascaded convolutional neural network-based hyperspectral image resolution enhancement via an auxiliary panchromatic image," *IEEE Trans. Image Process.*, vol. 30, pp. 6815–6828, 2021. 2
- [24] Z. C. Wu, T. Z. Huang, L. J. Deng, J. F. Hu, and G. Vivone, "VO+Net: An adaptive approach using variational optimization and deep learning for panchromatic sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022. 2, 3, 4, 12
- [25] X. Y. Fu, Z. H. Lin, Y. Huang, and X. H. Ding, "A variational pansharpening with local gradient constraints," in *IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, 2019, pp. 10265–10274. 2, 7, 8, 9, 10, 11, 12, 13
- [26] F. Fang, F. Li, C. Shen, and G. Zhang, "A variational approach for pansharpening," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2822–2834, 2013. 2
- [27] X. He, L. Condat, J. M. Bioucas-Dias, J. Chanussot, and J. Xia, "A new pansharpening method based on spatial and spectral sparsity priors," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 4160–4174, 2014. 2
- [28] H. A. Aly and G. Sharma, "A regularized model-based optimization framework for pansharpening," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2596–2608, 2014. 2
- [29] C. Chen, Y. Li, W. Liu, and J. Huang, "SIRF: Simultaneous satellite image registration and fusion in a unified framework," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4213–4224, 2015. 2
- [30] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rouge, "A variational model for P+XS image fusion," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 43–58, 2006. 2
- [31] Y. Y. Jiang, X. H. Ding, D. L. Zeng, Y. Huang, and J. Paisley, "Pansharpening with a hyper-Laplacian penalty," in *Int. Conf. Comput. Vision (ICCV)*, 2015, pp. 540–548. 2
- [32] T. Wang, F. Fang, F. Li, and G. Zhang, "High-quality Bayesian pansharpening," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 227–239, 2018. 2, 7, 8, 9, 10, 11, 12, 13
- [33] L. J. Deng, G. Vivone, W. H. Guo, M. Dalla Mura, and J. Chanussot, "A variational pansharpening approach based on reproducible Kernel Hilbert space and Heaviside function," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4330–4344, 2018. 2
- [34] A. M. Teodoro, J. M. Bioucas-Dias, and M. A. Figueiredo, "A convergent image fusion algorithm using scene-adapted Gaussian-mixture-based denoising," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 451–463, 2018. 2
- [35] L. J. Deng, M. Y. Feng, and X. C. Tai, "The fusion of panchromatic and multispectral remote sensing images via tensor-based sparse modeling and hyper-Laplacian prior," *Inf. Fusion*, vol. 52, pp. 76–89, 2019. 2, 7, 8, 9, 10, 11, 12, 13
- [36] Z. C. Wu, T. Z. Huang, L. J. Deng, G. Vivone, J. Q. Miao, J. F. Hu, and X. L. Zhao, "A new variational approach based on proximal deep injection and gradient intensity similarity for spatio-spectral image fusion," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 6277–6290, 2020. 2, 6
- [37] Q. Wei, N. Dobigeon, and J. Y. Tourneret, "Fast fusion of multi-band images based on solving a Sylvester equation," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4109–4121, 2015. 2, 3
- [38] R. W. Dian, S. T. Li, A. J. Guo, and L. Y. Fang, "Deep hyperspectral image sharpening," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 99, pp. 1–11, 2018. 2
- [39] R. W. Dian and S. T. Li, "Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 5135–5146, 2019. 2, 3, 13
- [40] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, 2013. 2, 3, 5
- [41] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Alg. Appl.*, vol. 435, no. 3, pp. 641–658, 2011. 3
- [42] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, "Novel methods for multilinear data completion and de-noising based on tensor-SVD," in *IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, 2014, pp. 3842–3849. 3
- [43] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis with a new tensor nuclear norm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 925–938, 2020. 3
- [44] Y. B. Zheng, T. Z. Huang, X. L. Zhao, T. X. Jiang, T. H. Ma, and T. Y. Ji, "Mixed noise removal in hyperspectral image via low-fibered-rank regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 734–749, 2020. 3, 5
- [45] Y. B. Zheng, T. Z. Huang, T. Y. Ji, X. L. Zhao, T. X. Jiang, and T. H. Ma, "Low-rank tensor completion via smooth matrix factorization," *Appl. Math. Model.*, vol. 70, pp. 677–695, 2019. 3
- [46] X. L. Zhao, W. H. Xu, T. X. Jiang, Y. Wang, and M. Ng, "Deep plug-and-play prior for low-rank tensor completion," *Neurocomputing*, vol. 400, pp. 137–149, 2020. 3
- [47] K. Zhang, W. Zuo, and L. Zhang, "Deep plug-and-play super-resolution for arbitrary blur kernels," in *IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, 2019, pp. 1671–1681. 3
- [48] Q. Song, R. Xiong, D. Liu, Z. Xiong, F. Wu, and W. Gao, "Fast image super-resolution via local adaptive gradient field sharpening transform," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1966–1980, 2018. 3
- [49] J. L. Xiao, T. Z. Huang, L. J. Deng, Z. C. Wu, and G. Vivone, "A new context-aware details injection fidelity with adaptive coefficients estimation for variational pansharpening," *IEEE Trans. Geosci. Remote Sens.*, 2022. 3, 7, 8, 9, 10, 11, 12, 13
- [50] L. Loncan, L. B. Almeida, J. M. Bioucasdias, X. Briottet, et al., "Hyperspectral pansharpening: A Review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 27–46, 2015. 4
- [51] H. Lu, Y. Yang, S. Huang, W. Tu, and W. Wan, "A unified pansharpening model based on band-adaptive gradient and detail correction," *IEEE Trans. Image Process.*, vol. 31, pp. 918–933, 2022. 4, 7, 8, 9, 10, 11, 12, 13
- [52] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *J. Math. Phys.*, vol. 6, no. 1-4, pp. 164–189, 1927. 5



- 811 [53] L. R. Tucker, "Some mathematical notes on three-mode factor analysis,"  
812 *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966. 5
- 813 [54] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W. K. Ma, "Hyper-  
814 spectral super-resolution: A coupled tensor factorization approach,"  
815 *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6503–6517, 2018.  
816 5
- 817 [55] C. Prévost, K. Usevich, P. Comon, and D. Brie, "Hyperspectral super-  
818 resolution with coupled Tucker approximation: Recoverability and SVD-  
819 based algorithms," *IEEE Trans. Signal Process.*, vol. 68, pp. 931–946,  
820 2020. 5
- 821 [56] M. Ashraphijuo and X. Wang, "Fundamental conditions for low-CP-rank  
822 tensor completion," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 2116–2145,  
823 2017. 5
- 824 [57] Y. Liu, Z. Long, H. Huang, and C. Zhu, "Low CP rank and Tucker rank  
825 tensor completion for estimating missing components in image data,"  
826 *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 4, pp. 944–954,  
827 2019. 5
- 828 [58] X. Y. Liu, S. Aeron, V. Aggarwal, and X. Wang, "Low-tubal-rank tensor  
829 completion using alternating minimization," *IEEE Trans. Inf. Theory*,  
830 vol. 66, no. 3, pp. 1714–1737, 2019. 5
- 831 [59] T. G. Kolda, "Orthogonal tensor decompositions," *SIAM J. Matrix Anal.*  
832 *Appl.*, vol. 23, no. 1, pp. 243–255, 2001. 5
- 833 [60] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-  
834 tailored multiscale fusion of high-resolution MS and Pan imagery,"  
835 *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, 2006.  
836 6, 7, 8, 9, 10, 11, 12, 13
- 837 [61] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven  
838 fusion of high spatial and spectral resolution images based on oversam-  
839 pled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol.  
840 40, no. 10, pp. 2300–2312, 2002. 7, 8, 9, 10, 11, 12, 13, 14
- 841 [62] R. Restaino, M. Dalla Mura, G. Vivone, and J. Chanussot, "Context-  
842 adaptive pansharpening based on image segmentation," *IEEE Trans.*  
843 *Geosci. Remote Sens.*, vol. 55, no. 2, pp. 753–766, 2016. 7, 8, 9, 10,  
844 11, 12, 13
- 845 [63] G. Vivone, "Robust band-dependent spatial-detail approaches for  
846 panchromatic sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 57,  
847 no. 9, pp. 6421–6433, 2019. 7, 8, 9, 10, 11, 12, 13
- 848 [64] G. Vivone, R. Restaino, and J. Chanussot, "Full scale regression-based  
849 injection coefficients for panchromatic sharpening," *IEEE Trans. Image*  
850 *Process.*, vol. 27, no. 7, pp. 3418–3431, 2018. 7, 8, 9, 10, 11, 12, 13
- 851 [65] R. Restaino, G. Vivone, M. Dalla Mura, and J. Chanussot, "Fusion  
852 of multispectral and panchromatic images based on morphological  
853 operators," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2882–2895,  
854 2016. 7, 8, 9, 10, 11, 12, 13
- 855 [66] F. Palsson, M. O. Ulfarsson, and J. R. Sveinsson, "Model-based reduced-  
856 rank pansharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 4,  
857 pp. 656–660, 2019. 7, 8, 9, 10, 11, 12, 13
- 858 [67] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image  
859 quality assessment: From error visibility to structural similarity," *IEEE*  
860 *Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004. 7
- 861 [68] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination  
862 among semi-arid landscape endmembers using the spectral angle mapper  
863 (SAM) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci.*  
864 *Workshop*, 1992, vol. 1, pp. 147–149. 7
- 865 [69] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M.  
866 Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006  
867 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45,  
868 no. 10, pp. 3012–3021, 2007. 8
- 869 [70] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality  
870 measurement of pan-sharpened multispectral imagery," *IEEE Geosci.*  
871 *Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, 2004. 8
- 872 [71] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva,  
873 "Multispectral and panchromatic data fusion assessment without refer-  
874 ence," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200,  
875 2008. 8
- 876 [72] G. Vivone, R. Restaino, M. Dalla Mura, G. Licciardi, and J. Chanus-  
877 sot, "Contrast and error-based fusion schemes for multispectral image  
878 pansharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp.  
879 930–934, 2013. 8
- 880 [73] X. Wu, T. Z. Huang, L. J. Deng, and T. J. Zhang, "Dynamic cross  
881 feature fusion for remote sensing pansharpening," in *Int. Conf. Comput.*  
882 *Vision (ICCV)*, 2021, pp. 14687–14696. 13, 14