Multispectral Image Super-Resolution via RGB Image Fusion and Radiometric Calibration

Zhi-Wei Pan and Hui-Liang Shen

***Abstract*— Multispectral imaging is of wide application for its capability in acquiring the spectral information of scenes. Due to hardware limitation, multispectral imaging device usually cannot achieve high-spatial resolution. To address the issue, this paper proposes a multispectral image super-resolution algorithm, referred as SRIF, by fusing the low-resolution multispectral image and the high-resolution (HR) RGB image. It deals with the general circumstance that image intensity is linear to scene radiance for multispectral imaging devices while is nonlinear and unknown for most RGB cameras. The SRIF algorithm first solves the inverse camera response function and a spectral sensitivity function of RGB camera, and establishes the linear relationship between multispectral and RGB images. Then the unknown HR multispectral image is efficiently reconstructed according to the linear image degradation models. Meanwhile, the edge structure of the reconstructed HR multispectral image is kept in accordance with that of the RGB image using a weighted total variation regularizer. The effectiveness of the SRIF algorithm is evaluated on both public datasets and our image set. Experimental results validate that the SRIF algorithm outperforms the state-of-the-arts in terms of both reconstruction accuracy and computational efficiency.**

***Index Terms*— Multispectral imaging, super-resolution, camera response function, spectral sensitivity function, weighted total variation, image reconstruction, image fusion.**

I. Introduction

MULTISPECTRAL imaging acquires spectral information of scenes and has been a promising tool for applications in biomedicine [1], remote sensing [2], color reproduction [3], and etc. Although multispectral imaging can achieve high spectral resolution, it has the limitation of low spatial resolution when compared with general RGB cameras [4]. The objective of this work is to reconstruct an high-resolution (HR) multispectral image by fusing a low-resolution (LR) multispectral image and an HR RGB image of the same scene.

Fig. 1 shows the hybrid imaging system employed in this work. The multispectral imaging device, which will also be referred as multispectral camera hereafter, consists of a monochrome camera and a liquid crystal tunable filter (LCTF) module. The multispectral camera acquires multiple LR band images at different wavelengths. The HR RGB image of the same scene can be acquired by replacing the multispectral camera with an RGB camera. As a multispectral imaging device usually aims at acquiring faithful spectral information, it always uses high-quality monochrome camera whose intensity response is linear to the scene radiance entering the imaging system. For most RGB cameras, however, the image intensity is nonlinear to the scene radiance.

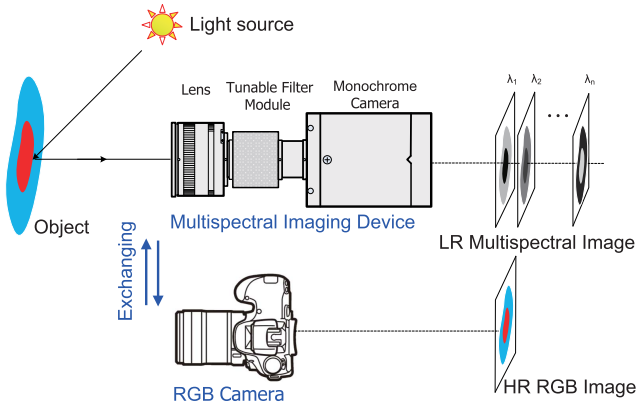


Fig. 1. Schematic diagram of the hybrid imaging system. The multispectral imaging device, which consists of a monochrome camera and a tunable filter module, acquires LR multispectral images. HR RGB images can be acquired by replacing the multispectral imaging device with the RGB camera while keeping the same viewpoint.

The acquired images and the target image to be reconstructed are related by physical imaging models. The LR multispectral image is spatially degraded by Gaussian blur and downsampling operations, provided that the monochrome camera has the linear response function. The Gaussian blur models the point spread function (PSF) of image sensor in high spatial resolution. The RGB image is spectrally degraded with the effect of spectral sensitivity function (SSF) and also subject to the effect of camera response function (CRF). The SSF integrates spectral reflectance into RGB triplets. The CRF relates scene radiance to image intensity, and can be either linear or nonlinear to meet different scenarios.

This paper proposes a multispectral image super-resolution algorithm, which is referred as SRIF, to reconstruct the target HR multispectral data with unknown SSF and CRF via RGB image fusion and radiometric calibration. In the algorithm, the inverse CRF (iCRF) and SSF are solved alternately by

Manuscript received February 9, 2018; revised August 15, 2018 and September 20, 2018; accepted November 12, 2018. Date of publication November 26, 2018; date of current version December 3, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61371160, in part by the Zhejiang Provincial Key Research and Development Project under Grant 2017C01044, and in part by the Fundamental Research Funds for the Central Universities under Grant 2017XZZX009-01. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Keigo Hirakawa. *(Corresponding author: Hui-Liang Shen.)*

The authors are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: pankdda@zju.edu.cn; shenhl@zju.edu.cn).

Digital Object Identifier 10.1109/TIP.2018.2881911

keeping the relationship between the acquired multispectral and RGB images. The iCRF and SSF are computed under the constraints of monotonicity and smoothness, respectively. The linear relationship among pixels in multispectral image will remain the same in the RGB image linearized by the iCRF.

The reconstruction of HR multispectral image is formulated as an optimization problem. The cost function contains two data-fitting terms and an edge-preserving prior regularizer. The data-fitting terms model the spatial and spectral degradations. The prior regularizer, in the form of weighted total variation (wTV), incorporates the edge structure information of the RGB image into the reconstructed multispectral image, based on the assumption that the edge structures of RGB image and band images are highly correlated. The reconstruction is performed on a low-dimensional subspace of the LR multispectral image. The multispectral image reconstruction problem is finally solved by the alternating direction method of multipliers (ADMM) algorithm [5] in an iterative manner. The subproblems in each iteration are simple and can be solved in closed form.

To summarize, the main contributions of this work are as follows:

- A super-resolution algorithm is proposed to reconstruct HR multispectral image under the guidance of an HR RGB image. Compared with the state-of-the-arts, this algorithm deals with the more general circumstance that, in addition to the unknown SSF, the CRF of RGB camera is nonlinear and also unknown.
- The iCRF and SSF are solved alternatively under appropriate constraints. By recovering an intrinsic linear relationship from the RGB image, the initial iCRF can be estimated robustly without knowing SSF.
- The reconstruction of HR multispectral image is guided by the edge structure of RGB image via the wTV regularization. The edges can be well preserved even when the original LR band images are heavily corrupted by noise.
- The subproblems in the ADMM iterations are all solved in closed form to improve computation efficiency. This makes our algorithm suitable for practical applications.

A. Organization and Notation

The rest of this paper is organized as follows. Section II reviews the related work on multispectral image reconstruction and radiometric calibration. Section III introduces our SRIF algorithm that consists of radiometric calibration (i.e., iCRF and SSF computation) and HR multispectral image reconstruction. The implementation of our algorithms is elaborated in Section IV. Section V shows the experimental results, and finally Section VI concludes the paper.

For clarity, Table I lists the main notations and their corresponding meanings used in this work.

TABLE I
MAIN NOTATIONS USED IN THIS WORK

Notations	Meanings
$\widetilde{\mathbf{Y}}$, $\mathbf{Y}$	LR multispectral image cube and matrix
$\widetilde{\mathbf{Z}}$, $\mathbf{Z}$	HR RGB image cube and matrix
$\widetilde{\mathbf{X}}$, $\mathbf{X}$	Reconstructed HR multispectral image cube and matrix
$\overline{\mathbf{Z}}$	Linearized HR RGB image matrix
$\mathbf{B}$, $\mathbf{S}$	Blur matrix, downsampling matrix
$\mathbf{R}$	SSF matrix
$\mathbf{\Psi}$	Low-dimensional space basis of multispectral image
$\mathbf{C}$	Coefficient matrix
$f(\cdot), g(\cdot)$	CRF, iCRF
M, m	Numbers of rows of $\widetilde{\mathbf{Z}}$ and $\widetilde{\mathbf{Y}}$
N, n	Numbers of columns of $\widetilde{\mathbf{Z}}$ and $\widetilde{\mathbf{Y}}$
L	Number of total band images

II. RELATED WORK

In the community of remote sensing and computer vision, the spatial super-resolution of multispectral images is achieved by image fusion. Pansharpening has been widely employed to fuse the LR band images with an HR panchromatic (PAN) image [6]. With only one HR image available, however, pansharpening can produce obvious spectral distortion when compensating spatial details. To deal with this limitation, most recent image fusion algorithms employ HR RGB image to improve both spatial resolution and spectral accuracy [4], [7]–[11]. Our work also belongs to this category.

The fusion of multispectral and RGB image can be conveniently formulated in the Bayesian inference framework. The prior knowledge is usually learnt from the scene of interest to improve the accuracy of fused spectral data. For example, the work [12] makes use of a stochastic mixing model (SMM) to learn the prior knowledge of the underlying spectral scene content. In [13], a prior spectrum distribution relating to linear mixing model is introduced. A Gibbs sampling algorithm is used to generate samples asymptotically distributed according to this prior distribution. It should be noted that the extraction of prior information usually requires relatively intensive computation, which will decrease the reconstruction efficiency.

Matrix factorization has been widely employed in image fusion. As band images are highly correlated, they can be decomposed into an orthogonal subspace basis and corresponding coefficients by principal component analysis (PCA) [14]. Inferring in subspace can decrease computational burden and improve solution stability. Based upon spectral unmixing principle [15], K-SVD factorization is adopted in [16] to learn the spectral dictionary from the LR hyperspectral image. The work [17] utilizes the local low-rank property to unmix both the hyperspectral and multispectral images in a coupled fashion. Considering the redundancy and correlation in spectral domain, low-rank matrix factorization is applied in [18] to exploit the latent low-dimensional structure existing in high-dimensional data. In addition, sparse promoting technique is used to yield unique matrix factorization. A set of basis that represents the reflectance spectra is obtained in [19] by applying sparse matrix factorization. The SNNMF algorithm [7] learns a nonnegative spectral basis by solving a constrained sparse representation problem.

Tensor factorization has the potential to fully exploit the inherent spatial-spectral structures during image fusion by

naturally representing the multispectral image as a three dimensional tensor. The work [20] incorporates the non-local spatial self-similarity into sparse tensor factorization and casts the image fusion problem as estimating sparse core tensor and dictionaries of three modes. The CSTF algorithm [11] uses a coupled sparse tensor factorization based approach for image fusion by imposing a sparse constraint on the core tensor.

Regularization techniques have also been employed to produce reasonable approximate solutions when the fusion problems are ill-posed. For example, the work [14] introduces a sparse regularization term based on the overcomplete dictionary learnt from natural images. An edge-preserving regularizer, which is in the form of vector total variation, is used in the HySure algorithm [8] to promote a piecewise-smooth solution. The NSSR algorithm [4] uses a clustering-based regularizer that exploits the spatial correlations among local and nonlocal similar pixels. The CO-CNMF algorithm [10] adopts the sum-of-squared-distances regularizer to yield high-fidelity endmember estimation. Regularization problem is usually solved using an iterative procedure. To decrease computational complexity, the properties of decimation matrix are explored in [21] to derive an analytical solution for the $\ell 2$ norm regularized super-resolution problem. In addition, the work [22] derives an explicit solution to the regularized image fusion problem based on the Sylvester equation. By exploiting the Woodbury formula, a generalized Sylvester equation associated with the multi-band image fusion problem is solved with the R-FUSE algorithm in a more robust and efficient way [9].

Deep learning presents new solutions for the multispectral image super-resolution. The main principle of this category is to learn a mapping function between LR and HR images in a fully supervised manner. The work [23] achieves this goal by training a deep neural network with the modified sparse denoising autoencoder, and regarding pansharpening as a simple image regression problem. PanNet [24] has the ability to preserve both the spectral and spatial information during the learning process, as its network parameters are trained on the high-pass components of the PAN and upsampled LR multispectral images. The work [25] learns the HR images as priors via deep convolutional neural network based learning. The priors are incorporated into the fusion framework to further improve performance. These end-to-end supervised learnings need the availability of diverse HR hyperspectral images to guarantee the accuracy, otherwise the mapping function may cause spectral distortion when applied to the images acquired in different situations. An unsupervised sparse Dirichelet-Net [26] is proposed to solve the fusion problem without the pretraining stage. It consists of two deep learning networks that share the same decoder weights, enabling the extraction of spectral basis and spatial representations in an unsupervised fashion.

Considering that our algorithm involves radiometric calibration, the works relating to iCRF and SSF estimation are further reviewed. The iCRF can be solved parametrically with some prior models. Gamma function [27] is supposed to lack enough freedom to fit the shape of iCRF, while a fifth-order polynomial function [28] and a PCA based model [29] have a good balance between accuracy and flexibility in iCRF estimation. Given a set of images taken with different exposures, a self-calibration algorithm is proposed in [30] to recover the iCRF along with the exposure ratios. By exploring the invariant low-rank structure within these images, iCRF is solved in [28] using a rank minimization approach. As for single image, the work [31] estimates iCRF by linearizing the distributions of measured colors in edge regions. The work [32] directly samples iCRF using temporal color mixtures. The accuracy of iCRF estimation can be considerably improved in our work by using additional multispectral images.

The SSF describes the degradation of spectral information in color images. It can be measured with a monochromator and a spectrophotometer in laboratory [33], or be solved with observed data for convenience. Proper constraints can be used to reduce the feasible solution space of SSF. For example, the curve characteristics of SSF like smoothness or non-negativity can impose shape constraints. The rank orders of RGB image values can impose half-space constraints [34]. Following this strategy, SSF can be solved using standard color checker image through linear regression with unknown illuminations [35]. Using the chromaticity invariance of fluorescence, the work [36] estimates SSF under an arbitrary illumination. By cross-calibration methods proposed in [8] and [37], SSF is solved with the assistance of multispectral image.

III. Proposed Multispectral Image Super-Resolution Algorithm

In this section, we first introduce the flowchart of the SRIF algorithm and the mathematical model of image degradation. Then we provide the solutions to iCRF and SSF coupled in RGB image. Finally, we present the reconstruction process for HR multispectral image with wTV regularization.

In this work, we suppose the LR multispectral image $\widetilde{\mathbf{Y}}$ is of dimension $m \times n \times L$, where $m \times n$ denotes the spatial resolution and L denotes the number of spectral bands. The HR RGB image $\widetilde{\mathbf{Z}} \in \mathbb{R}^{M \times N \times 3}$ is of spatial resolution $M \times N$. Denoting the scale factor of resolution improvement with d, the spatial dimensions are related by $M = m \times d$ and $N = n \times d$. The objective of this work is to compute the HR multispectral image $\widetilde{\mathbf{X}} \in \mathbb{R}^{M \times N \times L}$ by fusing $\widetilde{\mathbf{Y}}$ and $\widetilde{\mathbf{Z}}$.

A. Flowchart

Fig. 2 shows the flowchart of our SRIF algorithm. An LR multispectral image $\widetilde{\mathbf{Y}}$ and an HR RGB image $\widetilde{\mathbf{Z}}$ are required at first. The radiometric calibration is then performed to solve iCRF $g(\cdot)$ and SSF $\mathbf{R}$ according to Algorithm 1. During image reconstruction, the edge weight matrices $\mathbf{W}_x$ and $\mathbf{W}_y$ are computed from the linearized RGB image $\overline{\mathbf{Z}}$, which are used in wTV to preserve edge structure. Then the basis $\mathbf{\Psi}$ of subspace is generated by applying PCA to $\widetilde{\mathbf{Y}}$, and the corresponding coefficient matrix $\mathbf{C}$ for all pixels is computed using Algorithm 2. The full HR multispectral image cube $\widetilde{\mathbf{X}}$ is finally obtained by combining the basis $\mathbf{\Psi}$ and coefficient matrix $\mathbf{C}$.

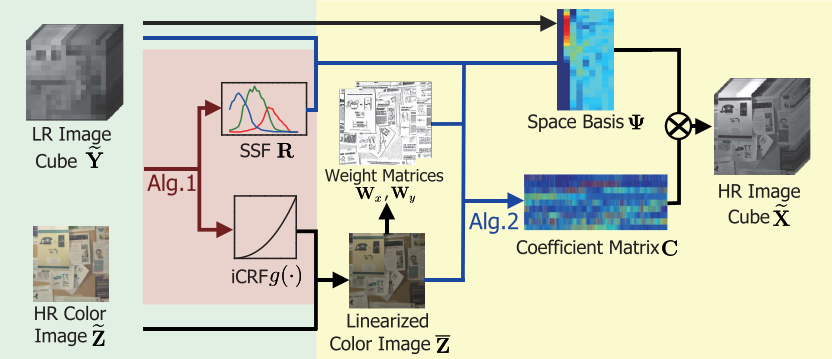


Fig. 2. Flowchart of our SRIF algorithm. Algorithm 1 computes the SSF $\mathbf{R}$ and iCRF $g(\cdot)$. The linearized RGB image $\overline{\mathbf{Z}}$ is used to compute edge weight matrices $\mathbf{W}_x$ and $\mathbf{W}_y$. Low-dimensional space basis $\mathbf{\Psi}$ is extracted from image $\mathbf{Y}$. Algorithm 2 computes the coefficient matrix $\mathbf{C}$ and reconstructs the HR multispectral image.

B. Observation Model

By indexing pixels in lexicographic order, the LR multispectral image cube $\widetilde{\mathbf{Y}}$ is represented by matrix $\mathbf{Y} \in \mathbb{R}^{L\times mn}$. The row vectors of $\mathbf{Y}$ are actually the vectorized band images. Similarly, the acquired HR color image cube $\widetilde{\mathbf{Z}}$ is represented by matrix $\mathbf{Z} \in \mathbb{R}^{3\times MN}$. Let $\mathbf{X} \in \mathbb{R}^{L\times MN}$ be the target HR image matrix, the spatial degradation model with respect to $\mathbf{Y}$ can be constructed as

$$\mathbf{Y} = \mathbf{XBS}, \tag{1}$$

where matrix $\mathbf{B} \in \mathbb{R}^{MN\times MN}$ is a spatial blurring matrix under circular boundary conditions. It denotes the PSF of multispectral sensor in the spatial domain of $\mathbf{X}$. The PSF can be modeled by Gaussian kernel [4], [8], [9]. Matrix $\mathbf{S} \in \mathbb{R}^{MN\times mn}$ is a decimation matrix, which accounts for a uniform downsampling of image with scale factor d.

The spectral degradation model with respect to $\mathbf{Z}$ can be formulated as

$$\mathbf{Z} = f(\mathbf{RX}). \tag{2}$$

where matrix $\mathbf{R} \in \mathbb{R}^{3\times L}$ denotes the SSF and holds in its rows the spectral sensitivities of RGB camera. Function $f(\cdot) : \mathbb{R}(0,1) \rightarrow \mathbb{R}(0,1)$ represents the CRF. To obtain the relationship between the linearized RGB image $\overline{\mathbf{Z}}$ and multispectral image $\mathbf{X}$, model (2) is converted to

$$\overline{\mathbf{Z}} := g(\mathbf{Z}) = \mathbf{RX}, \tag{3}$$

where $g(\cdot) := f^{-1}(\cdot)$ denotes the iCRF, which is monotonous and continuous.

C. Computing iCRF and SSF

By combing (1) and (3) with simple operation, we have

$$\mathbf{RY} = g(\mathbf{Z})\mathbf{BS}. \tag{4}$$

The embedded variables iCRF $g(\cdot)$ and SSF $\mathbf{R}$ can be solved alternatively by fitting this model to the acquired image data. As will be detailed in Section IV-A, the initial estimate of iCRF $g_0(\cdot)$ can be quite accurate. This guarantees that the following computation of iCRF and SSF in each iteration is reasonable.

Algorithm 1 Computing iCRF $g(\cdot)$ and SSF $\mathbf{R}$

Input: LR spectral matrix $\mathbf{Y} \in \mathbb{R}^{L\times mn}$, HR RGB matrix $\mathbf{Z} \in \mathbb{R}^{3\times MN}$, basis matrix $\mathbf{\Omega} \in \mathbb{R}^{100\times K_\Omega}$.
Output: $g(\cdot)$ and $\mathbf{R}$.
Initialize $g_0(\cdot)$;
for $j = 1$ **to** 5 **do**
  Compute $\mathbf{R}$ using (5);
  Compute $g(\cdot)$ using (6);
end

Given a fixed $g(\cdot)$, $\mathbf{R}$ can be solved as

$$\mathbf{R} = \arg\min_{\mathbf{R}} \ \|\mathbf{RY} - g(\mathbf{Z})\mathbf{BS}\|_F^2 \quad \text{s.t.} \ \left\|\nabla^2\mathbf{R}\right\|_F^2 < \lambda, \tag{5}$$

where $\|.\|_F$ denotes the element-wise Frobenius norm, ∇^2 denotes the second order derivative operator. The constant λ controls the smoothness of the estimated iCRF. It is set to 0.5 in this work. We note that the computation in (5) is similar to the HySure algorithm [8] except that HySure adopts a first order derivative constraint.

Given a fixed $\mathbf{R}$, $g(\cdot)$ can be solved as

$$g = \arg\min_{g} \ \|g(\mathbf{Z})\mathbf{BS} - \mathbf{RY}\|_F^2 \quad \text{s.t.} \ \nabla g > 0, \quad g \in \mathcal{R}(\mathbf{\Omega}), \tag{6}$$

where ∇ denotes the first order derivative operator and $\nabla g > 0$ guarantees the monotonicity of $g(\cdot)$. $\mathcal{R}(\mathbf{\Omega})$ denotes the linear vector space with respect to basis $\mathbf{\Omega} \in \mathbb{R}^{100\times K_\Omega}$. The basis $\mathbf{\Omega}$ is obtained by applying PCA to the iCRF curves in DoRF [29], with each curve sampled at 100 points. In this work, the number K_Ω of basis vectors is set to 9, and problem (6) is solved by the nonlinear programming operator fmincon in MATLAB®. Note that the solution g is in the form of column vector with its elements sampled from the corresponding curve, and the mapping effect of $g(\cdot)$ is achieved by the one-dimensional interpolation operation.

To summary, Algorithm 1 lists the computation of $g(\cdot)$ and $\mathbf{R}$ based on the discussion above. It is observed that the model (4) can fit the acquired data well after a few number (e.g., 5) of iterations.

D. Edge-Preserving Regularizer

It is essential to preserve the image edges during the super-resolution procedure. In this work, we introduce an edge-preserving regularizer, which is in the form of wTV, to incorporate the edge structure information of the RGB image into the reconstructed multispectral image. It is based on the reasonable assumption that the edge structure in RGB image is in good condition.

The wTV regularizer has been used for reconstructing three-dimensional medical images [38], where the weighting effect at each pixel location is isotopic. Considering that in most cases the horizontal and vertical components of image gradient are different, we employ a novel anisotropic weighting scheme and modify the edge-preserving wTV regularizer as

$$\begin{aligned}\mathrm{wTV}(\mathbf{XD}_x, \mathbf{XD}_y) &= \|\mathbf{W}_x \odot (\mathbf{XD}_x)\|_1 + \left\|\mathbf{W}_y \odot (\mathbf{XD}_y)\right\|_1 \\ &= \sum_{i=1}^{L}\sum_{j=1}^{MN} w_j^x |(\mathbf{XD}_x)_{ij}| + w_j^y |(\mathbf{XD}_y)_{ij}|,\end{aligned} \tag{7}$$

where $|\cdot|$, $\odot$, and $\|\cdot\|_1$ denote the absolute operation, Hadamard product, and element-wise ℓ_1 norm, respectively. Symbol $(\cdot)_{ij}$ denotes the element in the ith row and jth column of a matrix. Matrices $\mathbf{D}_x$ and $\mathbf{D}_y \in \mathbb{R}^{MN\times MN}$ represent the first-order horizontal and vertical gradient matrices under circular boundary conditions. Matrices $\mathbf{W}_x$ and $\mathbf{W}_y$ are the horizontal and vertical weighting matrices consisting of spatially varying edge-aware weights w_j^x and w_j^y, respectively. It is known that for natural images the distribution of gradients can be approximately modeled as Laplacian with zero mean [39], [40]. Motivated by this, we define the weights as

$$w_j^x = \exp\left(-\frac{|\mathbf{f}(\overline{\mathbf{Z}}\mathbf{D}_x)|_j}{b_x}\right), \quad w_j^y = \exp\left(-\frac{|\mathbf{f}(\overline{\mathbf{Z}}\mathbf{D}_y)|_j}{b_y}\right), \tag{8}$$

where $\mathbf{f}(\cdot)$ represents the grayscale conversion function that integrates image gradient information across the visible spectrum. Scale parameters b_x and b_y are estimated by the mean absolute value of the corresponding gradient components at every pixel location [41]. Through the regulating effect of w_j^x and w_j^y, the noise of the reconstructed image will be suppressed in flat area indexed by the edge structure of RGB image, and the real edges will be preserved.

It is of our interest to compare the effect of weights defined in (8) and the weight in [38] which is of the form

$$w_j = \frac{\eta}{\sqrt{\left(|\mathbf{f}(\overline{\mathbf{Z}}\mathbf{D}_x)|_j + |\mathbf{f}(\overline{\mathbf{Z}}\mathbf{D}_y)|_j\right)^2 + \eta^2}}. \tag{9}$$

The parameter η controls the extent that an edge can be considered as a structure. A limitation of this definition is that, even when the constant η is fine tuned, it still cannot cope with different image scenes. In comparison, the weights definition (8) describes the edge structure by a statistical model and does not involve any manual-setting parameters.

Fig. 3 demonstrates the effect of edge structure preservation by the R-FUSE [9] algorithm and our SRIF algorithm with the wTV in [38] and our wTV. Here we set $\eta = 0.01$ in (9) that is appropriate for common RGB images. It is observed that the R-FUSE [9] algorithm, which is based on dictionary learning and sparse representation, fails to recover entire edge structures because of serious image noise. Our SRIF algorithm with the wTV in [38] recovers the edge well when using a bright RGB image, but fails to maintain this edge preserving effect when the RGB image is dark. In comparison, our SRIF algorithm using the novel wTV keeps the edge structures well for both bright and dark scenes.

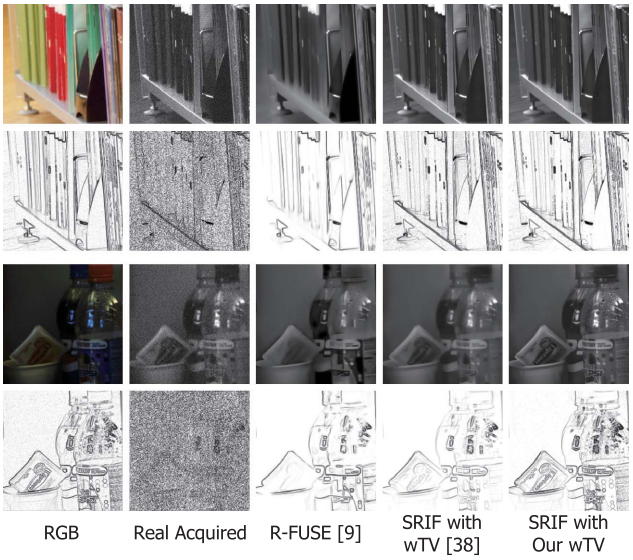


Fig. 3. Demonstration of edge structure preserving effect by our SRIF algorithm. From left to right: two HR RGB image regions and their edge structures, real acquired band images at band 420 nm, reconstructed band images using R-FUSE [9], reconstructed band images using SRIF with the wTV from [38], reconstructed band images using SRIF with our wTV. The spatial resolution is improved by 16×.

E. Super-Resolution Reconstruction

With the above treatment, the super-resolution reconstruction can be formulated as an optimization problem

$$\begin{aligned}\mathbf{X} = \arg\min_{\mathbf{X}} \ \frac{1}{2}\|\mathbf{Y} - \mathbf{XBS}\|_F^2 + \frac{\beta}{2}\left\|\overline{\mathbf{Z}} - \mathbf{RX}\right\|_F^2 \\ + \gamma\,\mathrm{wTV}(\mathbf{XD}_x, \mathbf{XD}_y),\end{aligned} \tag{10}$$

where β and λ are weighting and regularization parameters, respectively. We note this formulation is common in the field of hyperspectral image super-resolution [4], [8]. Our special treatments are the RGB linearization by CRF estimation and the novel wTV form for edge preservation in image reconstruction.

As color spectrum is intrinsically of low dimension, the HR multispectral matrix $\mathbf{X}$ is solved in a subspace with dimension less than L. This treatment can improve both computation efficiency and solution stability. The space basis $\boldsymbol{\Psi} \in \mathbb{R}^{L\times K_\Psi}$ ($K_\Psi \le L$), is obtained beforehand by applying PCA on the LR multispectral matrix $\mathbf{Y}$. Then problem (10) is converted to solving the corresponding coefficient matrix $\mathbf{C} \in \mathbb{R}^{K_\Psi \times MN}$ as

$$\begin{aligned}\mathbf{C} = \arg\min_{\mathbf{C}} \ \frac{1}{2}\|\mathbf{Y} - \boldsymbol{\Psi}\boldsymbol{C}\boldsymbol{B}\boldsymbol{S}\|_F^2 + \frac{\beta}{2}\left\|\overline{\mathbf{Z}} - \mathbf{R}\boldsymbol{\Psi}\mathbf{C}\right\|_F^2 \\ + \gamma\,\mathrm{wTV}(\boldsymbol{\Psi}\mathbf{C}\mathbf{D}_x, \boldsymbol{\Psi}\mathbf{C}\mathbf{D}_y).\end{aligned} \tag{11}$$

Considering that the wTV regularizer (7) is nonquadratic and nonsmooth, the ADMM algorithm [5] is employed to solve the problem (11) through the variable splitting technique. By introducing 3 auxiliary variables, the original optimization problem becomes

$$\begin{aligned}\min_{\mathbf{V_1},\mathbf{V}_2,\mathbf{V}_3\mathbf{C}} \ & \frac{1}{2}\left\|\mathbf{Y}-\mathbf{\Psi}\mathbf{C}\mathbf{B}\mathbf{S}\right\|_F^2+\frac{\beta}{2}\left\|\overline{\mathbf{Z}}-\mathbf{R}\mathbf{\Psi}\mathbf{V}_1\right\|_F^2 \\ & +\gamma\,\mathrm{wTV}(\mathbf{V}_2,\mathbf{V}_3) \\ \text{s.t. } & \mathbf{V}_1=\mathbf{C},\quad \mathbf{V}_2=\mathbf{\Psi}\mathbf{C}\mathbf{D}_x,\ \mathbf{V}_3=\mathbf{\Psi}\mathbf{C}\mathbf{D}_y. \end{aligned} \tag{12}$$

The auxiliary variable $\mathbf{V}_1$ helps avoid singularity, and the auxiliary variables $\mathbf{V}_2$ and $\mathbf{V}_3$ help generate closed-form solutions associated with the regularizer. The problem (12) has the following augmented Lagrangian

$$\begin{aligned}&\min\ \mathcal{L}_\rho(\mathbf{C},\mathbf{V}_1,\mathbf{V}_2,\mathbf{V}_3,\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3)\\ &=\frac{1}{2}\left\|\mathbf{Y}-\mathbf{\Psi}\mathbf{C}\mathbf{B}\mathbf{S}\right\|_F^2+\frac{\beta}{2}\left\|\overline{\mathbf{Z}}-\mathbf{R}\mathbf{\Psi}\mathbf{V}_1\right\|_F^2\\ &\quad+\gamma\,\mathrm{wTV}(\mathbf{V}_2,\mathbf{V}_3)+\frac{\rho}{2}\left\|\mathbf{C}-\mathbf{V}_1-\mathbf{A}_1\right\|_F^2\\ &\quad+\frac{\rho}{2}\left\|\mathbf{\Psi}\mathbf{C}\mathbf{D}_x-\mathbf{V}_2-\mathbf{A}_2\right\|_F^2+\frac{\rho}{2}\left\|\mathbf{\Psi}\mathbf{C}\mathbf{D}_y-\mathbf{V}_3-\mathbf{A}_3\right\|_F^2,\end{aligned} \tag{13}$$

where matrices $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ represent three scaled dual variables, and ρ denotes the penalty parameter.

The variables in (13) are solved through iteration. The subproblem of coefficient matrix $\mathbf{C}$ can be minimized in frequency domain, which uses the properties of convolution and decimation matrices. This implementation will be detailed in Section IV-B.

The auxiliary variable $\mathbf{V}_1$ is associated with an unconstrained least squares problem and has the closed-form solution

$$\begin{aligned}\mathbf{V}_1^{j+1}=(\beta(\mathbf{R}\mathbf{\Psi})^{\mathsf{T}}(\mathbf{R}\mathbf{\Psi})+\rho\mathbf{I})^{-1}(\beta(\mathbf{R}\mathbf{\Psi})^{\mathsf{T}}\overline{\mathbf{Z}}\\ +\rho(\mathbf{C}^{j+1}-\mathbf{A}_1^j)),\end{aligned} \tag{14}$$

where $(\cdot)^{\mathsf{T}}$ represents matrix transpose and $\mathbf{I}$ represents the unit matrix with proper dimensions.

By using soft shrinkage operator, the minimization problems involving $\mathbf{V}_2$ and $\mathbf{V}_3$ have the analytical solutions

$$\begin{aligned}\mathbf{V}_2^{j+1}=&\ \mathsf{sgn}(\mathbf{\Psi}\mathbf{C}^{j+1}\mathbf{D}_x-\mathbf{A}_2^j)\odot\mathsf{max}\{|\mathbf{\Psi}\mathbf{C}^{j+1}\mathbf{D}_x-\mathbf{A}_2^j|\\ &-\frac{\gamma}{\rho}\mathbf{W}_x,\mathbf{0}\}\end{aligned} \tag{15}$$

$$\begin{aligned}\mathbf{V}_3^{j+1}=&\ \mathsf{sgn}(\mathbf{\Psi}\mathbf{C}^{j+1}\mathbf{D}_y-\mathbf{A}_3^j)\odot\mathsf{max}\{|\mathbf{\Psi}\mathbf{C}^{j+1}\mathbf{D}_y-\mathbf{A}_3^j|\\ &-\frac{\gamma}{\rho}\mathbf{W}_y,\mathbf{0}\}.\end{aligned} \tag{16}$$

where $\mathsf{sgn}(\cdot)$ and $\mathsf{max}\{\cdot,\cdot\}$ denote sign and maximum functions, respectively.

Then the scaled dual variables are updated by

$$\begin{aligned}\mathbf{A}_1^{j+1}&=\mathbf{A}_1^j-(\mathbf{C}^{j+1}-\mathbf{V}_1^{j+1}),\\ \mathbf{A}_2^{j+1}&=\mathbf{A}_2^j-(\mathbf{\Psi}\mathbf{C}^{j+1}\mathbf{D}_x-\mathbf{V}_2^{j+1}),\\ \mathbf{A}_3^{j+1}&=\mathbf{A}_3^j-(\mathbf{\Psi}\mathbf{C}^{j+1}\mathbf{D}_y-\mathbf{V}_3^{j+1}).\end{aligned} \tag{17}$$

At the end of iteration, the target HR image $\mathbf{X}$ is recovered by combing basis $\mathbf{\Psi}$ and the corresponding coefficient $\mathbf{C}$.

Algorithm 2 Reconstruct $\mathbf{X}$ Using ADMM

Input: LR multispectral matrix $\mathbf{Y}\in\mathbb{R}^{L\times mn}$, HR RGB matrix $\overline{\mathbf{Z}}\in\mathbb{R}^{3\times MN}$, SSF $\mathbf{R}\in\mathbb{R}^{3\times L}$.
Output: HR multispectral matrix $\mathbf{X}$.
Compute weighting matrices $\mathbf{W}_x$ and $\mathbf{W}_y$ from $\overline{\mathbf{Z}}$;
Train the subspace basis $\mathbf{\Psi}$ from $\mathbf{Y}$;
for $j=1$ **to** 20 **do**
  Compute $\mathbf{C}^j$ according to Section IV-B;
  Compute $\mathbf{V}_1^j$ using (14);
  Compute $\mathbf{V}_2^j$ and $\mathbf{V}_3^j$ using (15);
  Update $\mathbf{A}_1^j$, $\mathbf{A}_2^j$, and $\mathbf{A}_3^j$ using (16);
end
Compute $\mathbf{X}=\mathbf{\Psi}\mathbf{C}$.

Our study reveals that 20 iterations are sufficient to obtain a satisfactory HR image. Algorithm 2 lists the procedure of this reconstruction.

Compared with Hysure [8], Algorithm 2 introduces different auxiliary variables to generate closed-form solutions with numerical stability. Also, Algorithm 2 handles the subproblem associated with coefficient matrix in a different way.

IV. Algorithm Implementation Details

In this section, we introduce the algorithm implementation details including the estimation of initial iCRF $g_0(\cdot)$ and the efficient solution of coefficient $\mathbf{C}$.

A. Estimating Initial iCRF

There is an intrinsic linear relationship among the pixels in multispectral image. However, it is destroyed in the RGB image rendered by nonlinear CRF. The estimation of initial iCRF can be treated as a process of recovering this intrinsic linear relationship from the rendered RGB image.

The intrinsic linear relationship reveals that each pixel spectrum $\mathbf{y}_i\in\mathbb{R}^L$ in $\mathbf{Y}$ can be composed by a collection of intrinsic basis spectra $\mathbf{\Phi}\in\mathbb{R}^{L\times K_\Phi}$ and the corresponding coefficient vector $\boldsymbol{\xi}_i\in\mathbb{R}^{K_\Phi}$ as

$$\mathbf{y}_i=\mathbf{\Phi}\boldsymbol{\xi}_i. \tag{18}$$

Each intrinsic spectrum $\boldsymbol{\phi}_k^y\in\mathbb{R}^L$ in $\mathbf{\Phi}$ is extracted from $\mathbf{Y}$ at location $\mathcal{I}(k)$, i.e., $\boldsymbol{\phi}_k^y=\mathbf{Y}_{:,\mathcal{I}(k)}$. The selection of intrinsic spectra is carried out in a iterative manner. An intrinsic spectrum to be selected should be as different as possible from those already selected. The number K_Φ of intrinsic spectra should be larger than L, and is set to 70 in this work. The coefficient vector $\boldsymbol{\xi}_i$ is solved as

$$\boldsymbol{\xi}_i=\arg\min_{\boldsymbol{\xi}_i}\left\|\boldsymbol{\xi}_i\right\|_1\quad\text{s.t. }\left\|\mathbf{y}_i-\mathbf{\Phi}\boldsymbol{\xi}_i\right\|_F^2<\varepsilon, \tag{19}$$

where the element-wise ℓ_1 norm promotes the sparsity of $\boldsymbol{\xi}_i$. The constant ε denotes the representation tolerance, and is set to 10^{-5} in this work.

Concatenating (18) for all pixel spectra leads to a self-representation of $\mathbf{Y}$,

$$\mathbf{Y}=\mathbf{Y}\mathbf{N}, \tag{20}$$

where the sparse matrix $\mathbf{N}$ is only constituted by the elements of coefficient vectors, i.e., $(\mathbf{N})_{\mathcal{I}(k),i} = (\boldsymbol{\xi}_i)_k$. With this treatment, the linear relationship existing in $\mathbf{Z}$ can be derived by combining (4) and (20), i.e., $g(\mathbf{Z})\mathbf{BS} = g(\mathbf{Z})\mathbf{BSN}$. Based on this equation, the initial CRF $g_0(\cdot)$ can be estimated without knowing SSF,

$$g_0 = \arg\min_{g_0} \frac{\|g_0(\mathbf{Z})\mathbf{BS} - g_0(\mathbf{Z})\mathbf{BSN}\|_p}{\|g_0(\mathbf{Z})\mathbf{BS} + g_0(\mathbf{Z})\mathbf{BSN}\|_p}, \quad \text{s.t. } \nabla g_0 > 0, \quad g_0 \in \mathcal{R}(\boldsymbol{\Omega}). \tag{21}$$

where $\|.\|_p$ denotes the element-wise p-norm. Our analysis indicates that the estimates of g_0 are close when using $p = 1$ or $p = 2$. In this work we let $p = 1$ considering that potentially the solution g_0 is insensitive to large errors when adopting the ℓ_1 norm. The denominator in (21) empirically prevents $g_0(\cdot)$ from being a zero function.

B. Solving Coefficient Matrix

Although a fast analytical solution of coefficient matrix $\mathbf{C}$ can be derived in terms of solving a Sylvester function [22], this work further reduces the amount of computation by rearranging the subproblem as

$$\begin{aligned}\mathbf{C}^{j+1} &= \arg\min_{\mathbf{C}} \ \widehat{\mathcal{L}}_\rho(\mathbf{C}) \\ &= \arg\min_{\mathbf{C}} \ \frac{1}{2}\left\|\widehat{\mathbf{Y}} - \mathbf{CBS}\right\|_F^2 + \frac{\rho}{2}\left\|\mathbf{C} - \mathbf{V}_1^j - \mathbf{A}_1^j\right\|_F^2 \\ &\quad + \frac{\rho}{2}\left\|\mathbf{CD}_x - \widehat{\mathbf{V}}_2^j - \widehat{\mathbf{A}}_2^j\right\|_F^2 + \frac{\rho}{2}\left\|\mathbf{CD}_y - \widehat{\mathbf{V}}_3^j - \widehat{\mathbf{A}}_3^j\right\|_F^2,\end{aligned} \tag{22}$$

where $\widehat{\mathbf{Y}}$ represents the projection coefficients of $\mathbf{Y}$ on the subspace $\mathcal{R}(\boldsymbol{\Psi})$, i.e., $\widehat{\mathbf{Y}} = (\boldsymbol{\Psi}^\mathsf{T}\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^\mathsf{T}\mathbf{Y}$, and matrices $\widehat{\mathbf{V}}_\mathbf{2}$, $\widehat{\mathbf{V}}_\mathbf{3}$, $\widehat{\mathbf{A}}_\mathbf{2}$ and $\widehat{\mathbf{A}}_\mathbf{3}$ are obtained in the same way. This rearrangement is based on the assumption that little loss of information is incurred during the projection.

Setting the derivative $\nabla_\mathbf{C}\widehat{\mathcal{L}}_\rho(\mathbf{C}) = \mathbf{0}$ leads to the matrix equation

$$\begin{aligned}&\mathbf{CBSS}^\mathsf{H}\mathbf{B}^\mathsf{H} + \rho\mathbf{CD}_x\mathbf{D}_x^\mathsf{H} + \rho\mathbf{CD}_y\mathbf{D}_y^\mathsf{H} + \rho\mathbf{C} \\ &\quad = \widehat{\mathbf{Y}}\mathbf{S}^\mathsf{H}\mathbf{B}^\mathsf{H} + \rho(\mathbf{V}_1^j + \mathbf{A}_1^j) + \rho(\widehat{\mathbf{V}}_2^j + \widehat{\mathbf{A}}_2^j)\mathbf{D}_x^\mathsf{H} \\ &\qquad + \rho(\widehat{\mathbf{V}}_3^j + \widehat{\mathbf{A}}_3^j)\mathbf{D}_y^\mathsf{H},\end{aligned} \tag{23}$$

where $(\cdot)^\mathsf{H}$ denotes conjugate transpose. Convolution matrices $\mathbf{B}$, $\mathbf{D}_x$ and $\mathbf{D}_y$ can be diagonalized by Fourier matrix $\mathbf{F} \in \mathbb{R}^{MN\times MN}$ [42], i.e., $\mathbf{B} = \mathbf{F}\boldsymbol{\Lambda}_B\mathbf{F}^\mathsf{H}$, $\mathbf{D}_x = \mathbf{F}\boldsymbol{\Lambda}_x\mathbf{F}^\mathsf{H}$ and $\mathbf{D}_y = \mathbf{F}\boldsymbol{\Lambda}_y\mathbf{F}^\mathsf{H}$. With this treatment, (23) is consolidated as

$$\mathbf{C}^{j+1} = \mathbf{Q}\mathbf{K}^{-1}\mathbf{F}^\mathsf{H}, \tag{24}$$

where

$$\begin{aligned}\mathbf{Q} &= \widehat{\mathbf{Y}}\mathbf{S}^\mathsf{H}\mathbf{F}\boldsymbol{\Lambda}_B^\mathsf{H} + \rho(\mathbf{V}_1^j + \mathbf{A}_1^j)\mathbf{F} + \rho(\mathbf{V}_2^j + \mathbf{A}_2^j)\mathbf{F}\boldsymbol{\Lambda}_x^\mathsf{H} \\ &\quad + \rho(\mathbf{V}_3^j + \mathbf{A}_3^j)\mathbf{F}\boldsymbol{\Lambda}_y^\mathsf{H},\end{aligned} \tag{25}$$

$$\mathbf{K}^{-1} = \left(\boldsymbol{\Lambda}_B\mathbf{F}^\mathsf{H}\mathbf{SS}^\mathsf{H}\mathbf{F}\boldsymbol{\Lambda}_B^\mathsf{H} + \rho\boldsymbol{\Lambda}_x^2 + \rho\boldsymbol{\Lambda}_y^2 + \rho\mathbf{I}\right)^{-1}. \tag{26}$$

Matrix $\mathbf{Q}$ in (25) can be fast solved just through upsampling, fast Fourier transform (FFT) and element-wise multiplication operations. As the inverse of large-scale matrix in (26) is difficult, $\mathbf{K}$ is simplified as

$$\mathbf{K} = \frac{1}{d^2}\boldsymbol{\Lambda}_B\mathbf{PP}^\mathsf{H}\boldsymbol{\Lambda}_B^\mathsf{H} + \boldsymbol{\Lambda}_K, \tag{27}$$

where the diagonal matrices in (26) are combined into $\boldsymbol{\Lambda}_K = \rho\boldsymbol{\Lambda}_x^2 + \rho\boldsymbol{\Lambda}_y^2 + \rho\mathbf{I}$, and the property of downsampling matrix $\mathbf{S}$, i.e.,

$$\mathbf{F}^\mathsf{H}\mathbf{SS}^\mathsf{H}\mathbf{F} = \mathbf{PP}^\mathsf{H}/d^2, \tag{28}$$

is utilized (see Appendix for proof). $\mathbf{P} \in \mathbb{R}^{MN\times mn}$ is a transform matrix with 0 and 1 elements. Right multiplying $\mathbf{P}$ and $\mathbf{P}^\mathsf{H}$ can be achieved by performing sub-block accumulating and image copying operations to the corresponding image. Using the Woodbury inversion lemma [21], the inverse of (27) can be computed as

$$\mathbf{K}^{-1} = \boldsymbol{\Lambda}_K^{-1} - \boldsymbol{\Lambda}_K^{-1}\boldsymbol{\Lambda}_B\mathbf{P}\left(d^2\mathbf{I} + \mathbf{P}^\mathsf{H}\boldsymbol{\Lambda}_B^\mathsf{H}\boldsymbol{\Lambda}_K^{-1}\boldsymbol{\Lambda}_B\mathbf{P}\right)^{-1} \times \mathbf{P}^\mathsf{H}\boldsymbol{\Lambda}_B^\mathsf{H}\boldsymbol{\Lambda}_K^{-1}, \tag{29}$$

where matrix $\mathbf{P}^\mathsf{H}\boldsymbol{\Lambda}_B^\mathsf{H}\boldsymbol{\Lambda}_K^{-1}\boldsymbol{\Lambda}_B\mathbf{P}$ is diagonal.

Inserting (29) into (24) yields the final solution

$$\mathbf{C}^{j+1} = \mathbf{Q}\boldsymbol{\Lambda}_K^{-1}\mathbf{F}^\mathsf{H} - \mathbf{Q}\boldsymbol{\Lambda}_K^{-1}\boldsymbol{\Lambda}_B\mathbf{P}\left(d^2\mathbf{I} + \mathbf{P}^\mathsf{H}\boldsymbol{\Lambda}_B^\mathsf{H}\boldsymbol{\Lambda}_K^{-1}\boldsymbol{\Lambda}_B\mathbf{P}\right)^{-1} \times \mathbf{P}^\mathsf{H}\boldsymbol{\Lambda}_B^\mathsf{H}\boldsymbol{\Lambda}_K^{-1}\mathbf{F}^\mathsf{H}. \tag{30}$$

Note that this solution procedure only contains FFT, element-wise multiplication, sub-block accumulating, and image copying operations.

V. Experiments

In the experiments, we evaluate the SRIF algorithm on both public datasets (Harvard [43], CAVE [44], and *Pavia University* [45]) and our acquired LR multispectral images.

The Harvard[1] dataset [43] has 77 multispectral images of indoor and outdoor scenes. The CAVE[2] dataset [44] contains 32 multispectral images of everyday objects. The images in these two datasets have 31 bands within the visible spectrum 400-720 nm,[3] which serve as ground truth to generate LR spectral images and HR RGB images. The *Pavia University*[4] dataset [45] contains a remotely sensed hyperspectral image consisting of 103 bands ranging from 430 nm to 860 nm. It is used for the extended experiment of fusing a LR hyperspectral image with an HR multispectral image.

The LR spectral images are generated by applying Gaussian blur and downsampling operations to the images in datasets. The size and standard deviation of Gaussian blur kernel are set to $(3d+1)\times(3d+1)$ and $3d/4$ empirically with respect to scale factor d = 4, 8, and 16, which makes the downsampled images clear and free of sawtooth artifacts. The HR RGB images are obtained from Harvard and CAVE datasets using the SSFs and CRFs provided in the CamSpec database [35] and DoRF database [29]. The 4-band HR multispectral image is obtained

[1] http://vision.seas.harvard.edu/hyperspec/download.html.

[2] http://www1.cs.columbia.edu/CAVE/databases/multispectral/.

[3] The spectrum is 420-720 nm in the Harvard dataset and 400-700 nm in the CAVE dataset.

[4] http://www.ehu.eus/ccwintco/uploads/e/ee/PaviaU.mat.

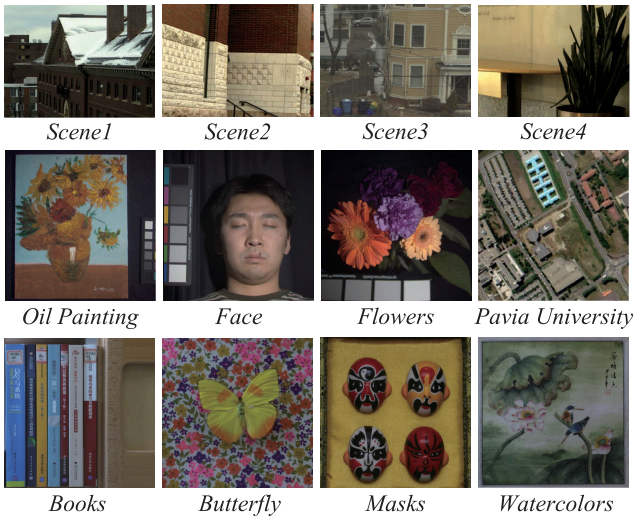


Fig. 4. Example images (displayed in RGB) used in the experiments. The 4 images in the first row are from the Harvard dataset, the 4 images in the middle row are from the CAVE and *Pavia University* datasets, and the 4 images in the last row are acquired in our laboratory.

by filtering *Pavia University* dataset with the LANDSAT-like spectral reflectance response [22].

In our real image dataset, the LR multispectral images with 31 bands are acquired by an imaging system consisting of a liquid crystal tunable filters (LCTF) and a CoolSnap monochrome camera. The HR RGB images are captured using a Canon 70D camera. The acquired multispectral and RGB images are aligned according to [46]. Fig. 4 shows 12 typical image samples (displayed in RGB) mentioned above.

Five objective quality metrics are used to evaluate the quality of reconstructed multispectral images, i.e., cross correlation (CC) [47], spectral angle mapper (SAM) [47], root mean squared error (RMSE) [47], relative dimensionless global error in synthesis (ERGAS) [47], and peak signal to noise ration (PSNR) [10]. The CC and SAM metrics measure the spatial quality and spectral quality respectively. The RMSE, ERGAS, and PSNR metrics measure the global quality.

In the following we discuss some issues on the SRIF algorithm, including parameter setting and the effect of reconstruction terms in (10). We also verify the reliability of iCRF and SSF estimation using simulation. Throughout the experiments, we compare the reconstruction accuracy of SRIF with those of six state-of-the-art ones, namely, SNNMF [7], HySure [8], R-FUSE [9], NSSR [4], CO-CNMF [10], and CSTF [11], whose source codes are publicly available online.[5,6,7,8,9,10] Finally we analyze the computational complexity of our algorithm.

A. Algorithm Settings

We evaluate the effect of three key parameters (weighting parameter β, regularization parameter γ, and penalty parameter ρ) on the reconstruction accuracy of SRIF.

[5]http://www.csse.uwa.edu.au/~ajmal/code/HSISuperRes.zip.
[6]https://github.com/alfaiate/HySure.
[7]https://github.com/qw245/BlindFuse.
[8]http://see.xidian.edu.cn/faculty/wsdong/Code_release/NSSR_HSI_SR.rar.
[9]http://www1.ee.nthu.edu.tw/cychi/sourcecode/CO-CNMF_demo.rar.
[10]https://sites.google.com/view/renweidian.

TABLE II

AVERAGE RMSE/SAM VALUES PRODUCED BY THE FULL OPTIMIZATION MODEL (10), AND THE MODEL WITHOUT ICRF (NONLINEAR) ESTIMATION OR WITHOUT USING EDGE-AWARE WEIGHTING IN WTV REGULARIZATION. THE QUANTITIES ARE COMPUTED ON ALL THE IMAGES IN HARVARD AND CAVE DATASETS WITH 16× RESOLUTION IMPROVEMENT

Conditions	Harvard	CAVE
Full model	1.46 / 3.72	2.64 / 4.92
iCRF: Nonliear $\rightarrow$ Linear	6.78 / 6.21	12.38 / 16.49
wTV: $w_j \rightarrow$ Const.	1.57 / 4.04	2.94 / 5.32

Fig. 5 plots the average RMSE values of all the reconstructed images in the Harvard and CAVE datasets with respect to these parameters. Parameter β is the weight of RGB data term in (10). We can see from Fig. 5(a) that the RMSE for both datasets drops when $\log_{10}\beta$ varies from -2 to 1, and keeps relatively stable when $\log_{10}\beta$ varies from 1 to 4. A larger β may overemphasize the importance of RGB data term. Based on this observation, β is set to 10 in this work.

Parameter γ regularizes the degree of edge preservation. A larger value of γ will increase the role of RGB edge guidance. However, if γ goes too large, the reconstructed edges will be over smoothed. As shown in Fig. 5 (b), the RMSE stays low for both datasets when $\log_{10}\gamma \in [-10, -8]$. Hence we set $\gamma = 10^{-8}$ in this work.

Parameter ρ penalizes the equality constraint residuals in (13). Fig. 5(c) shows that the RMSE for Harvard dataset reaches a minimum as $\log_{10}\rho = -4$ after finite reconstruction iterations. This value is also appropriate for the CAVE dataset. For this reason we set $\rho = 10^{-4}$.

We also evaluate the effects of iCRF estimation and wTV regularization respectively on the reconstruction accuracy of SRIF. Table II shows that the RMSE and SAM values become very large when not considering the iCRF estimation (i.e., assuming a linear iCRF). The accuracy of SRIF also degrades if not incorporating the edge-aware weighting in the wTV regularization. These observations validate the necessity of the related terms in the optimization model (10).

B. Accuracy of iCRF and SSF

Fig. 6(a) shows the RMSE values of 201 initial iCRFs obtained using our algorithm and the single-image based technique introduced in [31]. Thanks to the employment of multispectral image, the RMSE values of our algorithm are much lower than those of [31]. Fig. 6(b) demonstrates 4 initial iCRF curves estimated by the our algorithm. It is observed that the first 2 estimates (Nos. 15 and 72) with RMSE 0.0019 and 0.0016 are close to ground truths. The third estimate (No. 128), which corresponds to the maximum RMSE 0.031, only slightly deviates from ground truth. The last estimate (No. 156) is linear and its RMSE error is 0.0062. In general, our algorithm can provide satisfactory initial iCRFs.

An accurate initial iCRF benefits the estimation of iCRF and SSF. Fig. 7 shows the initial iCRFs and final iCRFs, as well as the estimated SSFs of the red, green, and blue

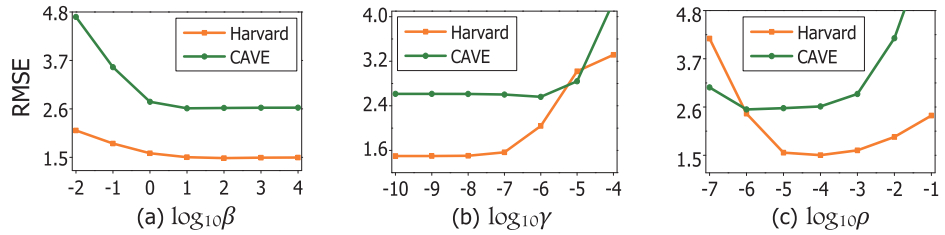


Fig. 5. The average RMSE values of all the reconstructed images in the Harvard and CAVE datasets with respect to parameters (a) $\log_{10}\beta$, (b) $\log_{10}\gamma$, and (c) $\log_{10}\rho$.

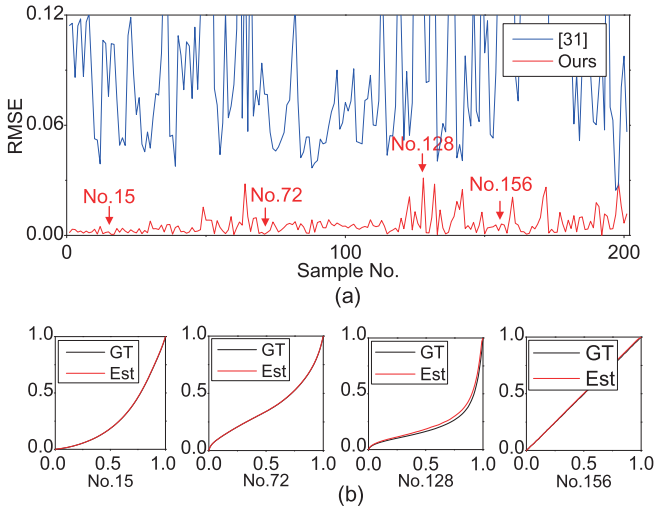


Fig. 6. Estimated initial iCRFs from *Scene1*. (a) RMSE values of 201 iCRFs from DoRF database produced by our algorithm and the technique [31]. Their mean RMSE values are 0.0059 and 0.1004, respectively. (b) Our 4 estimated iCRFs whose sample nos. are indicated in (a).

TABLE III

SAM AND PSNR VALUES OF RECONSTRUCTED HR MULTISPECTRAL IMAGES WHEN USING THE INITIAL ICRFS PRODUCED BY OUR ALGORITHM AND THE TECHNIQUE IN [31]. THE QUANTITIES ARE COMPUTED ON *Scene1* WITH 16× RESOLUTION IMPROVEMENT

	SAM		PSNR	
iCRF Estimation	[31]	Ours	[31]	Ours
SNNMF [7]	11.79	5.95	24.92	35.03
HySure [8]	6.90	5.25	23.11	36.44
R-FUSE [9]	4.64	3.88	24.00	37.90
NSSR [4]	6.34	3.12	23.82	39.56
CO-CNMF [10]	10.40	4.13	24.91	36.12
CSTF [11]	9.10	3.94	26.23	36.87
SRIF	3.89	2.97	27.03	40.41

channels. The RGB images are rendered using the SSF of Canon 60D camera. It is observed that the estimates of iCRF and SSF are quite close to ground truths.

It should be noted the accurate estimation of iCRF is essential to the ultimate HR image reconstruction. Table III shows that, when compared with the iCRF estimation technique [31], the super-resolution algorithms adopting our iCRF estimation algorithm produce lower SAM distortion and higher PSNR values. Hence, our iCRF estimation algorithm will be used for

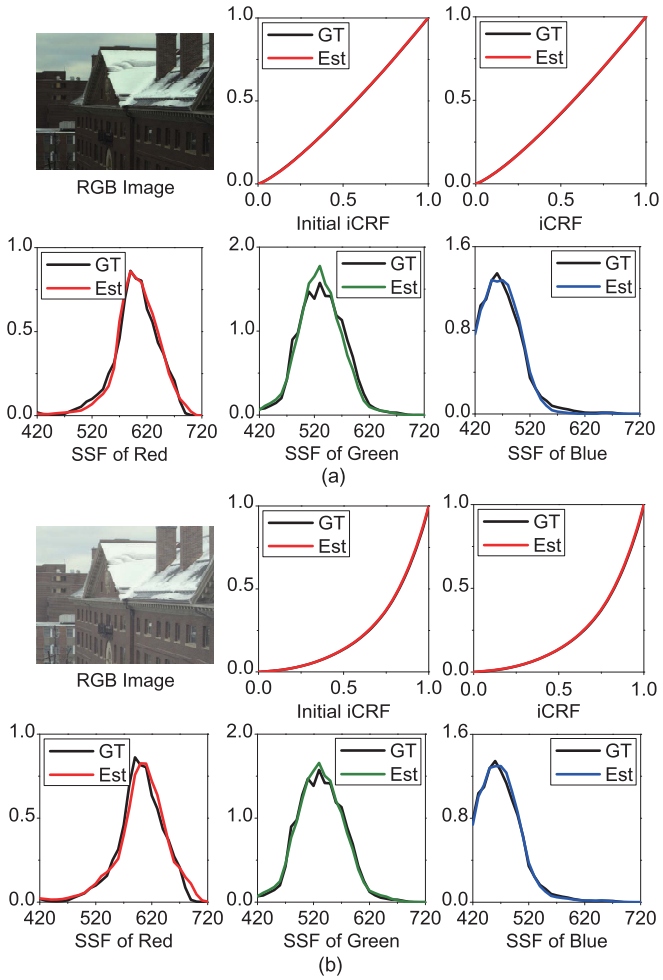


Fig. 7. Initial and final iCRFs, and estimated SSFs of the red, green, and blue channels.

both the SRIF algorithm and the competitors in the following experiments if not otherwise noted.

C. Image Reconstruction on Public Datasets

The HR multispectral images are reconstructed after the iCRF and SSF have been obtained. To fairly compare the reconstruction performance, all the super-resolution algorithms use the same and correct radiometric calibration. Fig. 8 shows the PSNR values at individual spectral bands of *Scene2* when the spatial resolution is improved with scale factor $d = 4$, 8, and 16. It is observed that our SRIF algorithm

TABLE IV

AVERAGE CC, SAM, RMSE, ERGAS, AND PSNR VALUES PRODUCED BY DIFFERENT ALGORITHMS ON THREE DATASETS. THE RESOLUTION IS IMPROVED WITH SCALE FACTORS $d = 8$ AND 16

	d = 8					d = 16				
	CC	SAM	RMSE	ERGAS	PSNR	CC	SAM	RMSE	ERGAS	PSNR
Harvard Dataset										
SNNMF [7]	0.979	5.14	1.91	1.35	38.94	0.970	8.10	2.64	0.99	36.35
HySure [8]	0.967	8.23	2.51	1.91	37.05	0.966	8.36	2.57	0.96	36.92
R-FUSE [9]	0.973	5.37	2.38	1.77	36.69	0.968	5.70	2.72	1.00	35.57
NSSR [4]	0.987	3.79	1.44	1.04	41.75	0.983	4.65	1.85	0.68	40.00
CO-CNMF [10]	0.980	5.23	1.84	1.36	38.98	0.977	5.60	2.07	0.76	37.97
CSTF [11]	0.980	4.95	1.94	1.30	38.82	0.978	5.37	2.06	0.71	38.30
SRIF	**0.987**	**3.53**	**1.35**	**0.96**	**42.49**	**0.986**	**3.72**	**1.46**	**0.52**	**41.95**
CAVE Dataset										
SNNMF [7]	0.995	5.66	2.63	1.28	42.64	0.992	15.49	4.01	0.97	40.42
HySure [8]	0.995	10.74	2.78	1.47	41.20	0.992	16.70	4.08	1.00	38.81
R-FUSE [9]	0.993	5.14	3.11	1.56	40.57	0.991	6.38	3.85	0.95	38.71
NSSR [4]	0.983	4.09	5.21	2.23	39.66	0.985	5.34	4.71	1.01	39.60
CO-CNMF [10]	0.987	8.94	4.70	2.04	38.48	0.988	10.42	4.73	1.08	37.99
CSTF [11]	0.992	7.16	3.04	1.47	41.08	0.990	8.03	3.48	0.82	40.22
SRIF	**0.996**	**3.90**	**2.12**	**1.08**	**44.17**	**0.994**	**4.92**	**2.64**	**0.66**	**42.60**
Pavia University Dataset										
SNNMF [7]	0.996	1.82	1.95	0.59	43.11	0.995	1.97	2.13	0.31	42.38
HySure [8]	0.996	1.97	1.95	0.56	43.70	0.996	2.02	2.01	0.28	43.44
R-FUSE [9]	0.995	2.10	2.20	0.71	42.00	0.994	2.31	2.42	0.39	41.17
NSSR [4]	0.993	2.47	2.69	0.76	41.41	0.994	2.46	2.60	0.35	41.75
CO-CNMF [10]	0.981	4.57	4.78	1.50	36.97	0.983	3.93	5.04	0.73	35.20
CSTF [11]	0.990	3.12	3.32	0.97	38.04	0.989	3.20	3.37	0.49	37.91
SRIF	**0.997**	**1.69**	**1.75**	**0.51**	**44.35**	**0.997**	**1.82**	**1.86**	**0.27**	**43.83**

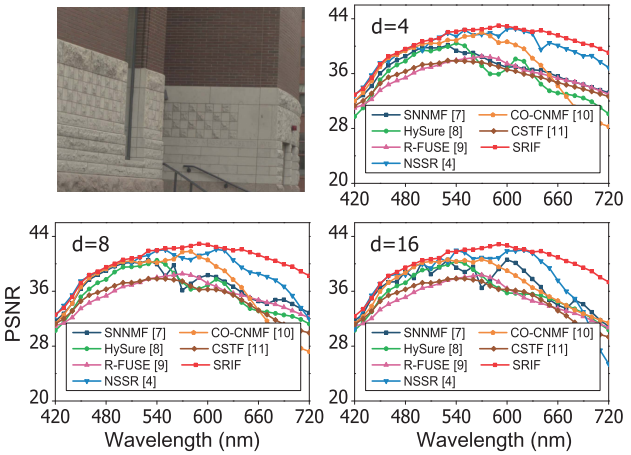


Fig. 8. PSNR values of individual spectral bands on *Scene2* when d = 4, 8, and 16.

produces the highest PSNR values in most cases when compared with the competing ones, and also produces relatively stable PSNR curves when the scale factor varies.

In the following, the resolution scale factor is set as d = 16 if not otherwise noted. Fig. 9(a) presents the fusion results of *Scene3* using different algorithms. The first row shows parts of the reconstructed band images at wavelength 580 nm. The second and third rows further show the corresponding RMSE and SAM maps computed from the spectra of all pixels. The average RMSE and SAM values are also shown with the corresponding images. The RMSE maps reveal that the images reconstructed by SNNMF [7] and R-FUSE [9] exhibit large spatial errors, while the SAM maps reveal that the spectral details are not well handled in the results produced by SNNMF [7] and HySure [8]. The CSTF algorithm [11] produces better image reconstruction, but still fails to recover the scene in the bottom area. In comparison, the HR images produced by the SRIF algorithm are of highest accuracy.

Fig. 9(b) shows the reconstruction results of the remotely sensed hyperspectral image *Pavia University* at the near-infrared wavelength 860 nm. It is observed that the CO-CNMF [10] and CSTF [11] algorithms are not good at handling this hyperspectral image, while the SNNMF [7] and HySure [8] algorithms perform relatively better. In comparison, our SRIF algorithm produces lowest spatial and spectral errors. This indicates that SRIF is capable of fusing remote-sensing LR hyperspectral image and HR multispectral image.

Table IV shows the average CC, SAM, RMSE, ERGAS, and PSNR values of all the reconstructed multispectral (or hyperspectral) images in the Harvard, CAVE, and *Pavia University* datasets. It is observed that our SRIF algorithm outperforms all the competitors when evaluated using these metrics. More specifically, Fig. 10(a) and Fig. 10(b) show the overall RMSE and SAM results on the 109 multispectral images in the CAVE

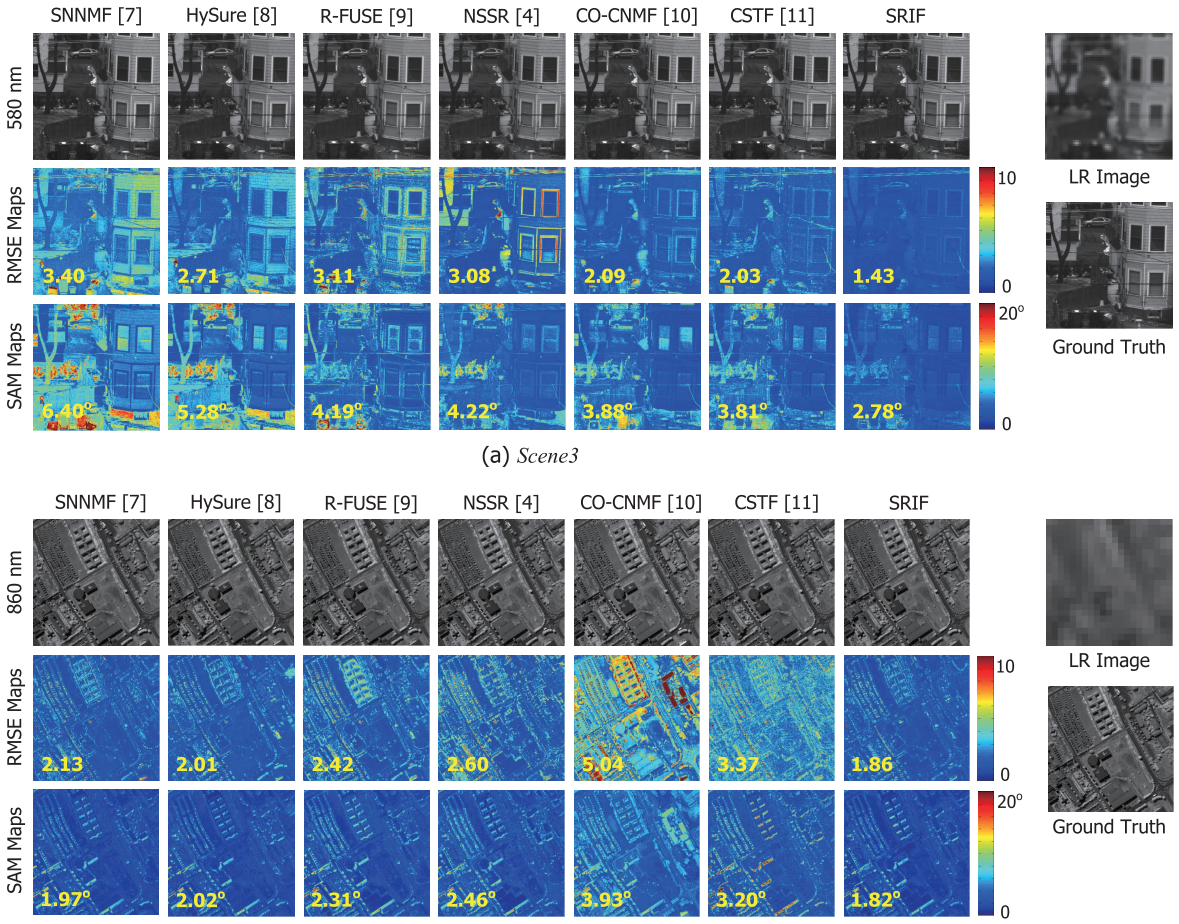


Fig. 9. Reconstruction results of (a) *Scene3* and (b) *Pavia University* with 16× spatial resolution improvement. The 1st and 4th rows show parts of the reconstructed band images of *Scene3* at 580 nm and parts of the reconstructed band images of *Pavia University* at 860 nm using different algorithms. Their LR images and ground truth images are listed on the right. The remaining rows illustrate the corresponding RMSE maps and SAM maps calculated across all the spectral bands.

and Harvard datasets. For clear demonstration, the image indexes are sorted in ascending order with respect to the metric values produced by the SRIF algorithm. It is observed that the RMSE and SAM values of SRIF are always of the lowest magnitude for individual images. This verifies that our algorithm performs better than the other six ones in terms of both spatial and spectral metrics.

We further evaluate the performance of the SRIF algorithm in case of noise corruption. In the experiment, two types of image noise are added to the LR multispectral image of the *Oil Painting* scene. The corresponding HR RGB image is kept clear as practically it is easy to guarantee enough exposure time during its imaging process. We first adopt the independent identically distributed (i.i.d) Gaussian noise to degrade the spectral images with the signal-to-noise ratio (SNR) of 20 dB and 25 dB, respectively. In this case, every pixel shares the same noise level. We also corrupt the spectral images with realistic signal-dependent noise. This noise model is derived from the assumption that the number of detected photons at each pixel location $\mathbf{p} \in \mathbb{Z}^2$ follows Poisson distribution. The sensor noise can be approximated by additive Gaussian noise with signal-dependent variance, which is linear to pixel intensity as $\sigma^2(\mathbf{p}) = a\mathrm{I}(\mathbf{p}) + b$ [48]. To generate spectral band images with different signal-dependent noise levels, we first scale the pixel intensity of *Oil Painting* such that the mean intensities of the resultant images are 0.3× and 1× as the mean intensity of the original image, and then add the noise by setting the linear coefficients $a = 0.04$ and $b = 0.13$ estimated in [48]. Due to the property of signal-dependent

TABLE V

RMSE/SAM RESULTS OF NOISY CASES. THE QUANTITIES ARE COMPUTED ON *Oil Painting* WITH 8× RESOLUTION IMPROVEMENT

	i.i.d Noise		Signal-dependent Noise	
	SNR=20dB	SNR=25dB	0.3×	1×
SNNMF [7]	4.11/9.72	4.13/8.90	3.42/7.67	3.11/6.65
HySure [8]	3.34/8.91	3.19/8.17	3.75/10.45	3.35/9.91
R-FUSE [9]	3.99/7.11	3.91/6.25	4.03/**6.79**	3.91/5.95
NSSR [4]	9.37/13.05	7.37/8.44	8.73/8.86	7.00/6.16
CO-CNMF [10]	4.63/8.35	5.11/7.88	6.49/8.30	4.98/7.67
CSTF [11]	3.79/7.22	3.50/6.70	3.82/7.05	3.54/6.67
SRIF	**3.09/7.10**	**2.83/5.88**	**3.27**/6.82	**2.89/5.77**

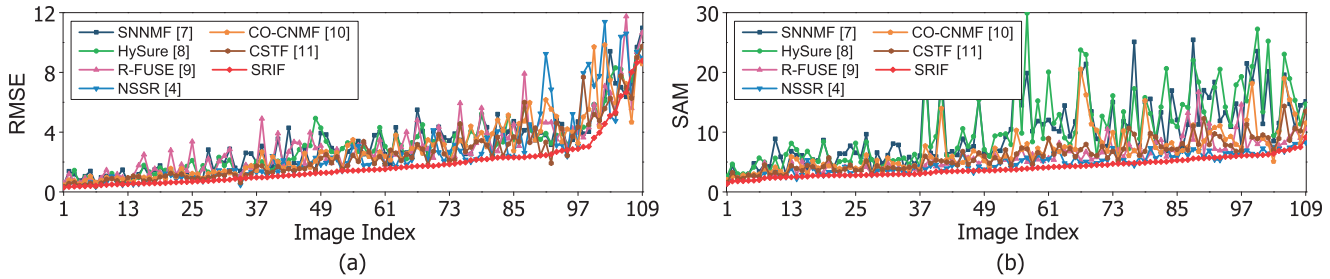


Fig. 10. RMSE (a) and SAM (b) values produced by different algorithms on all the images in the CAVE and Harvard datasets. The image indexes are sorted in ascending order with respect to the metric values produced by the SRIF algorithm to facilitate visual comparison.

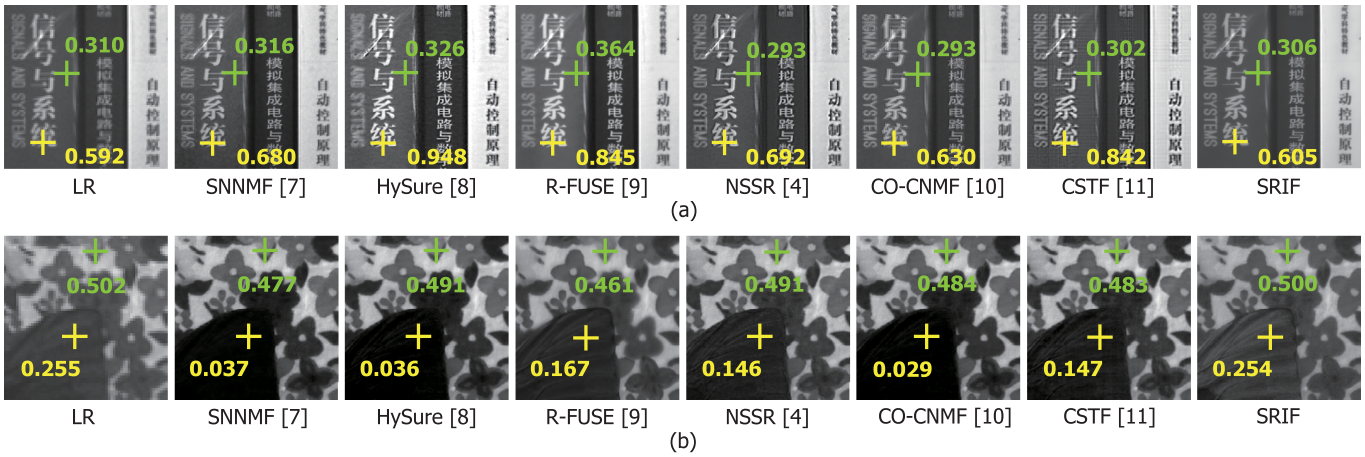


Fig. 11. Reconstruction results on real data. The intensities of the marked pixels are shown on the images. (a) *Books* at band 690 nm, (b) *Butterfly* at band 460 nm.

TABLE VI

PER-ITERATION AND TOTAL RUNNING TIMES (IN SECONDS) OF DIFFERENT ALGORITHMS FOR RECONSTRUCTING AN HR MULTISPECTRAL IMAGE WITH $L = 31$ BANDS AND 1392×1040 SPATIAL RESOLUTION. THE NUMBER OF ITERATIONS OF EACH ALGORITHM IS GIVEN IN THE PARENTHESES

	SNNMF [7]	HySure [8]	R-FUSE [9]	NSSR [4]	CO-CNMF [10]	CSTF [11]	SRIF
Per-iteration Running Time	375.9	6.9	258.3	33.3	44.0	9.9	**6.7**
Total Running Time (Iterations)	2365.5 (5)	1393.6 (200)	7533.2 (10)	1457.5 (26)	2938.3 (30)	213.1 (12)	**144.7** (20)

noise, the higher image intensities will result in relatively better image quality. Table V lists the RMSE and SAM values of the HR multispectral images reconstructed from the noisy LR images. It is observed that when adding the i.i.d noise, our SRIF algorithm produces lowest RMSE and SAM values. Meanwhile, we observe that R-FUSE [9], which is based on sparse representation, produces the lowest SAM value when dealing with large signal-dependent noise. However, its RMSE value is slightly unsatisfactory. Overall, our SRIF algorithm still outperforms other algorithms when considering both spatial and spectral reconstruction accuracy.

D. Image Reconstruction on Our Real Images

We also evaluate the SRIF algorithm on real images acquired in our laboratory. Fig. 11 shows the original LR band images of *Books* at 690 nm and *Butterfly* at 460 nm, as well as the corresponding reconstructed results with $8\times$ spatial resolution improvement. For a clear comparison, two pixels locating in smooth regions are marked and their corresponding intensities are shown. It is desired that the intensities of the marked pixels in the reconstructed HR images should be close to those in the original LR images. We observe from Fig. 11(a) that the shapes of English alphabets in the images produced by SNNMF, HySure, NSSR, and CO-CNMF are not clear and the intensity of Chinese characters produced by R-FUSE is too high. The image produced by CSTF exhibits undesired ring effect. Fig. 11(b) shows that the intensities of butterfly wings produced by the competing algorithms are relatively low. In comparison, our SRIF algorithm performs well in handling these details in both Fig. 11(a) and Fig. 11(b).

E. Complexity Analysis

Algorithm 1 solves the iCRF and SSF alternatively. In each iteration of minimizing (5) or (6), the most computationally intensive step is the multiplication of matrices $\mathbf{B}$ and $g(\mathbf{Z})$. Its complexity is $\mathcal{O}(MN\log(MN))$. Algorithm 2 reconstructs

TABLE VII
RMSE VALUES WITH RESPECT TO THE NUMBER OF ITERATIONS. THE QUANTITIES ARE COMPUTED ON *Pavia University* WITH 8× RESOLUTION IMPROVEMENT

Iterations	1	5	20	50	100	200
SNNMF [7]	2.05	2.19	2.49	2.52	2.51	2.71
HySure [8]	54.46	12.66	2.16	2.09	2.06	2.01
R-FUSE [9]	2.65	2.20	2.22	2.26	2.28	2.31
NSSR [4]	5.00	6.92	3.18	2.36	2.31	2.78
CO-CNMF [10]	5.49	3.59	4.01	10.44	13.14	11.25
CSTF [11]	3.63	3.36	3.30	3.30	3.31	3.30
SRIF	6.93	3.29	1.75	1.70	1.70	1.70

the HR multispectral image by solving several subproblems. The computational complexity of the prime variable $\mathbf{C}$ in frequency domain is of order $\mathcal{O}(LMN\log(MN))$. The computational complexity of the auxiliary variable $\mathbf{V}_1$ is of order $\mathcal{O}(K_\Psi MN)$. The computational complexities of the auxiliary variables $\mathbf{V}_2$ and $\mathbf{V}_3$ are both of order $\mathcal{O}(LK_\Psi MN)$, where the computation associated with gradient matrices can be achieved by subtracting the adjacent image rows or columns. In addition, an order of $\mathcal{O}(K_\Psi MN)$ is required in the computation of dual variable $\mathbf{A}_1$. An order of $\mathcal{O}(LMN)$ is required in the following computation of dual variables $\mathbf{A}_2$ and $\mathbf{A}_3$. Thus the complexity of Algorithm 2 is dominated by the FFTs and is of order $\mathcal{O}(LMN\log(MN))$ per iteration. We note that the HySure algorithm [8] has nearly the same computational complexity. However, HySure uses more auxiliary variables in ADMM, which increases the computation burden accordingly.

All the competing algorithms are implemented using MATLAB® and run on a personal computer with 2.60 GHz CPU (Intel Xeon E5-2630) and 64 GB RAM. The competing algorithms adopt their corresponding default iteration numbers in the computation. Table VI shows the per-iteration running times and total running times of these algorithms for reconstructing an HR multispectral image with $L = 31$ spectral bands and 1392×1040 spatial resolution. It is observed that our SRIF algorithm costs the least time in each iteration and also gains improvement in total computational efficiency when compared with the state-of-the-arts.

Table VII further lists the RMSE results with respect to the number of iterations on reconstructing the *Pavia University* image with 8× resolution improvement. It is observed that our SRIF algorithm always produces the lowest RMSE values when the number of iterations is larger than 20.

VI. CONCLUSIONS

This paper has introduced a multispectral image super-resolution algorithm, referred as SRIF, to improve the spatial resolution of multispectral image with an HR RGB image. The iCRF and SSF are first solved from the acquired data. Then the HR multispectral image is efficiently reconstructed, and the wTV operator is used to keep the recovered edge structure in accordance to the one of RGB image. Experimental results validate that our algorithm performs better than the state-of-the-arts in terms of both reconstruction accuracy and computational efficiency.

APPENDIX
PROOF OF (28)

Proof: The properties of Fourier transform matrix $\mathbf{F}$ and downsampling matrix $\mathbf{S}$ are first introduced. Matrix $\mathbf{F}$ can be decomposed as

$$\mathbf{F} = \mathbf{D}_N \otimes \mathbf{D}_M,$$

where $\mathbf{D}_N \in \mathbb{R}^{N\times N}$ and $\mathbf{D}_M \in \mathbb{R}^{M\times M}$ are one-dimensional discrete Fourier transform matrices, and $\otimes$ denotes Kronecker product. Right multiplying an image by $\mathbf{SS}^{\mathsf{H}}$ can be achieved through performing element-wise multiplication with a mask matrix, which has ones in the sampled position and zeros elsewhere. With this treatment, $\mathbf{SS}^{\mathsf{H}}$ can be expressed as

$$\mathbf{SS}^{\mathsf{H}} = \left(\mathbf{I}_{N/d} \otimes \mathbf{T}_d\right) \otimes \left(\mathbf{I}_{M/d} \otimes \mathbf{T}_d\right),$$

where $\mathbf{I}_{N/d}$ and $\mathbf{I}_{M/d}$ are identity matrices with dimensions N/d and M/d, respectively, and matrix $\mathbf{T}_d \in \mathbb{R}^{d\times d}$ has the form

$$\mathbf{T}_d = \begin{pmatrix} 1\ 0\ \ldots\ 0 \\ 0\ 0\ \ldots\ 0 \\ .\ .\ .\ . \\ .\ .\ .\ . \\ 0\ 0\ \ldots\ 0 \end{pmatrix}.$$

Then Equation (28) can be proved as

$$\begin{aligned}
&\mathbf{F}^{\mathsf{H}}\mathbf{SS}^{\mathsf{H}}\mathbf{F} \\
&= \mathbf{D}_N^{\mathsf{H}} \otimes \mathbf{D}_M^{\mathsf{H}} \left(\mathbf{I}_{N/d} \otimes \mathbf{T}_d\right) \otimes \left(\mathbf{I}_{M/d} \otimes \mathbf{T}_d\right) \mathbf{D}_N \otimes \mathbf{D}_M \\
&= \left(\mathbf{D}_N^{\mathsf{H}} \left(\mathbf{I}_{N/d} \otimes \mathbf{T}_d\right) \mathbf{D}_N\right) \otimes \left(\mathbf{D}_M^{\mathsf{H}} \left(\mathbf{I}_{M/d} \otimes \mathbf{T}_d\right) \mathbf{D}_M\right) \\
&= \left(\frac{1}{d}\mathbf{J}_{d\times d} \otimes \mathbf{I}_{N/d}\right) \otimes \left(\frac{1}{d}\mathbf{J}_{d\times d} \otimes \mathbf{I}_{M/d}\right) \\
&= \frac{1}{d^2}\left(\mathbf{J}_d\mathbf{J}_d^{\mathsf{H}}\right) \otimes \left(\mathbf{I}_{N/d}\mathbf{I}_{N/d}^{\mathsf{H}}\right) \otimes \left(\mathbf{J}_d\mathbf{J}_d^{\mathsf{H}}\right) \otimes \left(\mathbf{I}_{M/d}\mathbf{I}_{M/d}^{\mathsf{H}}\right) \\
&= \frac{1}{d^2}\left(\mathbf{J}_d \otimes \mathbf{I}_{N/d} \otimes \mathbf{J}_d \otimes \mathbf{I}_{M/d}\right)\left(\mathbf{J}_d^{\mathsf{H}} \otimes \mathbf{I}_{N/d}^{\mathsf{H}} \otimes \mathbf{J}_d^{\mathsf{H}} \otimes \mathbf{I}_{M/d}^{\mathsf{H}}\right) \\
&= \frac{1}{d^2}\mathbf{P}\mathbf{P}^{\mathsf{H}},
\end{aligned}$$

where the third equation holds according to the *Lemma 3* in [22]. Matrix $\mathbf{J}_{d\times d} \in \mathbb{R}^{d\times d}$ and vector $\mathbf{J}_d \in \mathbb{R}^d$ have all the elements being ones, and matrix $\mathbf{P}$ is defined as

$$\mathbf{P} = \mathbf{J}_d \otimes \mathbf{I}_{N/d} \otimes \mathbf{J}_d \otimes \mathbf{I}_{M/d},$$

which means that right multiplying an image by $\mathbf{P}$ can be achieved through accumulating $d \times d$ sub-blocks divided from the origin image into one small image. The corresponding adjoint operation, i.e., right multiplying an image by $\mathbf{P}^{\mathsf{H}}$, is achieved through copying the origin image $d \times d$ times to form a large image. ■

REFERENCES

[1] R. M. Levenson and J. R. Mansfield, "Multispectral imaging in biology and medicine: Slices of life," *Cytometry A*, vol. 69, no. 8, pp. 748–758, Aug. 2006.

[2] G. A. Shaw and H. K. Burke, "Spectral imaging for remote sensing," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 3–28, Nov. 2003.

[3] R. S. Berns, "Color-accurate image archives using spectral imaging," in *Scientific Examination of Art: Modern Techniques in Conservation and Analysis*. Washington, DC, USA: National Academies Press, Jan. 2005, pp. 105–119.

[4] W. Dong *et al.*, "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2337–2352, May 2016.

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

[6] G. Vivone *et al.*, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.

[7] E. Wycoff, T.-H. Chan, K. Jia, W.-K. Ma, and Y. Ma, "A non-negative sparse promoting algorithm for high resolution hyperspectral imaging," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 1409–1413.

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

[9] Q. Wei, N. Dobigeon, J.-Y. Tourneret, J. Bioucas-Dias, and S. Godsill, "R-FUSE: Robust fast fusion of multiband images based on solving a Sylvester equation," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1632–1636, Nov. 2016.

[10] C.-H. Lin, F. Ma, C.-Y. Chi, and C.-H. Hsieh, "A convex optimization-based coupled nonnegative matrix factorization algorithm for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1652–1667, Mar. 2018.

[11] S. Li, R. Dian, L. Fang, and J. M. Bioucas-Dias, "Fusing hyperspectral and multispectral images via coupled sparse tensor factorization," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4118–4130, Aug. 2018.

[12] M. T. Eismann and R. C. Hardie, "Application of the stochastic mixing model to hyperspectral resolution enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 9, pp. 1924–1933, Sep. 2004.

[13] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Bayesian fusion of hyperspectral and multispectral images," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2014, pp. 3176–3180.

[14] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, Jul. 2015.

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

[16] B. Huang, H. Song, H. Cui, J. Peng, and Z. Xu, "Spatial and spectral image fusion using sparse matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1693–1704, Mar. 2014.

[17] Y. Zhou, L. Feng, C. Hou, and S.-Y. Kung, "Hyperspectral and multispectral image fusion based on local low rank and coupled spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5997–6009, Oct. 2017.

[18] K. Zhang, M. Wang, and S. Yang, "Multispectral and hyperspectral image fusion based on group spectral embedding and low-rank factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1363–1371, Mar. 2017.

[19] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y.-W. Tai, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. IEEE CVPR*, Jun. 2011, pp. 2329–2336.

[20] R. Dian, L. Fang, and S. Li, "Hyperspectral image super-resolution via non-local sparse tensor factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3862–3871.

[21] N. Zhao, Q. Wei, A. Basarab, D. Kouamé, and J.-Y. Tourneret, "Single image super-resolution of medical ultrasound images using a fast algorithm," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 473–476.

[22] 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, Nov. 2015.

[23] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, "A new pan-sharpening method with deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1037–1041, May 2015.

[24] J. Yang, X. Fu, Y. Hu, Y. Huang, X. Ding, and J. Paisley, "PanNet: A deep network architecture for pan-sharpening," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1753–1761.

[25] R. Dian, S. Li, A. Guo, and L. Fang, "Deep hyperspectral image sharpening," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5345–5355, Nov. 2018.

[26] Y. Qu, H. Qi, and C. Kwan, "Unsupervised sparse Dirichlet-net for hyperspectral image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Apr. 2018, pp. 2511–2520.

[27] H. Farid, "Blind inverse gamma correction," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1428–1433, Oct. 2001.

[28] J.-Y. Lee, Y. Matsushita, B. Shi, I. S. Kweon, and K. Ikeuchi, "Radiometric calibration by rank minimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 144–156, Jan. 2013.

[29] M. D. Grossberg and S. K. Nayar, "What is the space of camera response functions?" in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2003, pp. 602–609.

[30] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1999, pp. 374–380.

[31] S. Lin, J. Gu, S. Yamazaki, and H.-Y. Shum, "Radiometric calibration from a single image," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, Jun./Jul. 2004, p. 2.

[32] B. Wilburn, H. Xu, and Y. Matsushita, "Radiometric calibration using temporal irradiance mixtures," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–7.

[33] J. Nakamura, *Image Sensors and Signal Processing for Digital Still Cameras*. Boca Raton, FL, USA: CRC Press, 2016.

[34] G. Finlayson, M. M. Darrodi, and M. Mackiewicz, "Rank-based camera spectral sensitivity estimation," *J. Opt. Soc. Amer. A, Opt. Image Sci. Vis.*, vol. 33, no. 4, pp. 589–599, Apr. 2016.

[35] J. Jiang, D. Liu, J. Gu, and S. Süsstrunk, "What is the space of spectral sensitivity functions for digital color cameras?" in *Proc. IEEE Workshop Appl. Comput. Vis. (WACV)*, Jan. 2013, pp. 168–179.

[36] S. Han, Y. Matsushita, I. Sato, T. Okabe, and Y. Sato, "Camera spectral sensitivity estimation from a single image under unknown illumination by using fluorescence," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 805–812.

[37] N. Yokoya, N. Mayumi, and A. Iwasaki, "Cross-calibration for data fusion of EO-1/hyperion and Terra/ASTER," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 419–426, Apr. 2013.

[38] M. J. Ehrhardt and M. M. Betcke, "Multicontrast MRI reconstruction with structure-guided total variation," *SIAM J. Imag. Sci.*, vol. 9, no. 3, pp. 1084–1106, Aug. 2016.

[39] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.

[40] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-laplacian priors," in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2009, pp. 1033–1041.

[41] S. Kotz, T. Kozubowski, and K. Podgorski, *The Laplace Distribution and Generalizations: A Revisit With Applications to Communications, Economics, Engineering, and Finance*. New York, NY, USA: Springer, 2012.

[42] P. C. Hansen, J. G. Nagy, and D. P. O'leary, *Deblurring Images: Matrices, Spectra, Filtering*. Philadelphia, PA, USA: SIAM, 2006.

[43] A. Chakrabarti and T. Zickler, "Statistics of real-world hyperspectral images," in *Proc. CVPR*, Jun. 2011, pp. 193–200.

[44] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2241–2253, Sep. 2010.

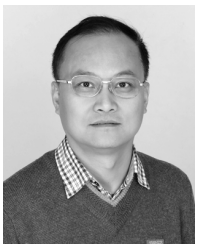
[45] F. Dell'Acqua, P. Gamba, A. Ferrari, J. A. Palmason, J. A. Benediktsson, and K. Árnason, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 322–326, Oct. 2004.

[46] S.-J. Chen, H.-L. Shen, C. Li, and J. H. Xin, "Normalized total gradient: A new measure for multispectral image registration," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1297–1310, Mar. 2018.

[47] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.

[48] Z. W. Pan, H. L. Shen, C. Li, S.-J. Chen, and J. H. Xin, "Fast multispectral imaging by spatial pixel-binning and spectral unmixing," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3612–3625, Aug. 2016.

Zhi-Wei Pan received the B.Eng. degree in information engineering from Zhejiang University in 2014, where he is currently pursuing the Ph.D. degree with the College of Information Science and Electronic Engineering. His research interests are multispectral imaging and image processing.



Hui-Liang Shen received the B.Eng. and Ph.D. degrees in electronic engineering from Zhejiang University, China, in 1996 and 2002, respectively. He was a Research Associate and a Research Fellow with The Hong Kong Polytechnic University from 2001 to 2005. He is currently a Full Professor with the College of Information Science and Electronic Engineering, Zhejiang University. His research interests include multispectral imaging, image processing, and 3D computer vision.