



# Multispectral Image Super-Resolution Using Structure-Guided RGB Image Fusion

Zhi-Wei Pan and Hui-Liang Shen<sup>(✉)</sup>

College of Information Science and Electronic Engineering, Zhejiang University,  
Hangzhou 310027, China  
{pankdda, shenh1}@zju.edu.cn

**Abstract.** 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 by fusing the low-resolution multispectral image and the high-resolution RGB image. The fusion is formulated as an optimization problem according to the linear image degradation models. Meanwhile, the fusion is guided by the edge structure of RGB image via the directional total variation regularizer. Then the fusion problem is solved by the alternating direction method of multipliers algorithm through iteration. The subproblems in each iterative step is simple and can be solved in closed-form. The effectiveness of the proposed algorithm is evaluated on both public datasets and our image set. Experimental results validate that the algorithm outperforms the state-of-the-arts in terms of both reconstruction accuracy and computational efficiency.

**Keywords:** Multispectral imaging · Super-resolution  
Directional total variation · Image reconstruction · Image fusion

## 1 Introduction

Multispectral imaging has been widely applied in various application fields, including biomedicine [1], remote sensing [2], color reproduction [3], and etc. Multispectral imaging can achieve high spectral resolution, but lacks spatial information when compared with general RGB cameras. The objective of this work is to reconstruct a high-resolution (HR) multispectral image by fusing a low-resolution (LR) multispectral image and an HR RGB image of the same scene.

---

This work was supported 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.

Student as first author.

The fusion of multispectral and RGB image can be conveniently formulated in the Bayesian inference framework. The work [4] estimates the signal-dependent noise statistics to generate the conditional probability distribution of acquired images, and makes the reconstruction robust to noise corruption. Extracting auxiliary information in Bayesian framework requires additional calculations and influences the reconstruction efficiency to some degree.

Matrix factorization has been widely employed in image fusion. As spectral bands are highly correlated, principal component analysis (PCA) is used in [5] to decompose the image data. By adopting the coupled nonnegative matrix factorization criterion, the spectral unmixing principle is employed in [6] to unmix the hyperspectral and multispectral image in a coupled fashion. Meanwhile, tensor factorization has the potential to fully exploit the inherent spatial-spectral structures during image fusion. The work [7] 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.

Regularization techniques can be employed to produce a reasonable approximate solution when the fusion problem is ill-posed. The HySure algorithm [8] uses vector total variation as an edge-preserving regularizer to promote a piecewise-smooth solution. The NSSR algorithm [9] uses a clustering-based regularizer to exploit the spatial correlations among local and nonlocal similar pixels. The regularization problem is usually solved through iteration. To decrease the computational complexity, the R-FUSE algorithm [10] derives a robust and efficient solution to the regularized image fusion problem based on a generalized Sylvester equation. In addition, the work [11] explores the properties of decimation matrix and derives an analytical solution for the  $\ell_2$  norm regularized super-resolution problem.

Deep learning presents new solutions for the multispectral image super-resolution. The work [12] learns a mapping function between LR and HR images by training a deep neural network with the modified sparse denoising autoencoder. PanNet [13] 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.

Inspired by the above works, this paper proposes a super-resolution algorithm to reconstruct the target HR multispectral data via structure-guided RGB image fusion. In the algorithm, the spatial and spectral degradation models are used to fit the acquired image data. An edge-preserving regularizer, which is in the form of directional total variation (dTV) [14], is used to guide the image reconstruction. It is based on the reasonable assumption that the spectral images and RGB image share not only the edge location but also the edge direction. To avoid the singularity induced by spectral dependence, the reconstruction is performed on a subspace of the LR multispectral image. The fusion problem is finally solved by the alternating direction method of multipliers (ADMM) algorithm [15] through iteration. The solutions of subproblems are in closed-form and can be accelerated in frequency domain.

The main contributions of this paper include: (1) The image fusion accuracy is improved by guiding the recovered edge structure in accordance to that of RGB image, and (2) The image fusion efficiency is improved by solving the subproblems in closed-form and accelerating the solutions in frequency domain. These makes the proposed algorithm more suitable for practical applications.

## 2 Problem Formulation

The acquired LR multispectral image is denoted as  $\tilde{\mathbf{Y}} \in \mathbb{R}^{m \times n \times L}$ , where  $m \times n$  is the spatial resolution and  $L$  is the number of spectral bands. The acquired HR RGB image  $\tilde{\mathbf{Z}} \in \mathbb{R}^{M \times N \times 3}$  has the 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 goal of super-resolution is to estimate the HR multispectral image  $\tilde{\mathbf{X}} \in \mathbb{R}^{M \times N \times L}$  by fusing  $\tilde{\mathbf{Y}}$  and  $\tilde{\mathbf{Z}}$ .

### 2.1 Observation Model

By indexing pixels in lexicographic order, the image cubes  $\tilde{\mathbf{Y}}$ ,  $\tilde{\mathbf{Z}}$  and  $\tilde{\mathbf{X}}$  can be represented by matrices  $\mathbf{Y} \in \mathbb{R}^{L \times mn}$ ,  $\mathbf{Z} \in \mathbb{R}^{3 \times MN}$  and  $\mathbf{X} \in \mathbb{R}^{L \times MN}$  respectively. The row vectors of these matrices are actually the vectorized band images. With this treatment, the spatial degradation model can be constructed as

$$\mathbf{Y} = \mathbf{X}\mathbf{B}\mathbf{S}, \quad (1)$$

where matrix  $\mathbf{B} \in \mathbb{R}^{MN \times MN}$  is a spatial blurring matrix representing the point spread function (PSF) of multispectral sensor in the spatial domain of  $\mathbf{X}$ . It is assumed under circular boundary conditions. Matrix  $\mathbf{S} \in \mathbb{R}^{MN \times mn}$  accounts for a uniform downsampling of image with scale factor  $d$ .

The spectral degradation model can be formulated as

$$\mathbf{Z} = \mathbf{R}\mathbf{X}, \quad (2)$$

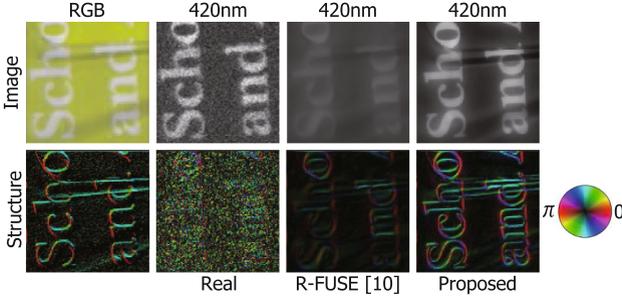
where matrix  $\mathbf{R} \in \mathbb{R}^{3 \times L}$  denotes the spectral sensitivity function (SSF) and holds in its rows the spectral responses of RGB camera.

### 2.2 Edge-Preserving Regularizer

A regularizer, which is in the form of dTV [14], is used to preserve both the location and direction of image edges during the super-resolution procedure. It is based on a priori knowledge that the RGB image and spectral images are likely to show very similar edge structures.

The edge-preserving dTV regularizer is formulated as

$$\begin{aligned} \text{dTV}(\mathbf{X}\mathbf{D}_x, \mathbf{X}\mathbf{D}_y) &= \|\mathbf{X}\mathbf{D}_x - [\mathbf{G}_x \odot (\mathbf{X}\mathbf{D}_x) + \mathbf{G}_y \odot (\mathbf{X}\mathbf{D}_y)] \odot \mathbf{G}_x\|_1 \\ &\quad + \|\mathbf{X}\mathbf{D}_y - [\mathbf{G}_x \odot (\mathbf{X}\mathbf{D}_x) + \mathbf{G}_y \odot (\mathbf{X}\mathbf{D}_y)] \odot \mathbf{G}_y\|_1, \end{aligned} \quad (3)$$



**Fig. 1.** Demonstration of edge structure preserving effect by the proposed algorithm. From left to right: An HR image region and its edge structure, real band image at band 420 nm, reconstructed band image using R-FUSE [10], reconstructed band image using the proposed algorithm. The spatial resolution is improved by  $16\times$ .

where  $\odot$  and  $\|\cdot\|_1$  denote the Hadamard product and element-wise  $\ell_1$  norm respectively. Matrices  $\mathbf{D}_x$  and  $\mathbf{D}_y \in \mathbb{R}^{MN \times MN}$  represent the first-order horizontal and vertical derivative matrices under circular boundary conditions. Matrix  $\mathbf{G}_x$  and  $\mathbf{G}_y$  denote the normalized horizontal and vertical gradient components of RGB image  $\mathbf{Z}$ , which can be computed in advance as

$$\mathbf{G}_* = \frac{\mathbf{f}(\mathbf{Z}\mathbf{D}_*)}{\sqrt{\mathbf{f}(\mathbf{Z}\mathbf{D}_x) \odot \mathbf{f}(\mathbf{Z}\mathbf{D}_x) + \mathbf{f}(\mathbf{Z}\mathbf{D}_y) \odot \mathbf{f}(\mathbf{Z}\mathbf{D}_y) + \eta^2}}, \quad * := x, y$$

where  $\cdot/\cdot$  and  $\sqrt{\cdot}$  are element-wise division and square root operators. Grayscale conversion function  $\mathbf{f}(\cdot)$  integrates image gradient information across the visible spectrum. Constant  $\eta$  adjusts the relative magnitude of edges and is set to 0.01 in this work. Through the regulating effect of Eq. (3), the component of reconstructed gradient that is orthogonal to the one from RGB image in the same edge location will be penalized. Thus the reconstructed image  $\mathbf{X}$  tends to share the same edge direction with RGB image  $\mathbf{Z}$ . Meanwhile, the noise of the reconstructed image will be suppressed in flat area since Eq. (3) reduces to total variation there. Figure 1 shows that the proposed algorithm keeps the edge structure of reconstructed band image in consistent with the one of RGB image, and also suppresses the band image noise. In comparison, the R-FUSE [10] algorithm, which is based on dictionary learning and sparse representation, fails to recover the edge structure.

### 2.3 Optimization Problem

The target HR multispectral image  $\mathbf{X}$  usually lives in a linear subspace, i.e.,

$$\mathbf{X} = \Psi\mathbf{C}, \quad (4)$$

where matrix  $\Psi \in \mathbb{R}^{L \times K_\Psi}$  is the subspace basis that can be obtained in advance by applying PCA on the LR multispectral image  $\mathbf{Y}$ , and the dimension  $K_\Psi$  is

set to 10 in this work. Matrix  $\mathbf{C} \in \mathbb{R}^{K_\Psi \times MN}$  is the corresponding projection coefficients of  $\mathbf{X}$ .

In this case, based on degradation models with the proposed regularizer, the reconstruction problem can be converted to the problem of estimating the unknown coefficient matrix  $\mathbf{C}$  from the following optimization equation

$$\mathbf{C} = \arg \min_{\mathbf{C}} \frac{1}{2} \|\mathbf{Y} - \Psi \mathbf{C} \mathbf{B} \mathbf{S}\|_F^2 + \frac{\beta}{2} \|\mathbf{Z} - \mathbf{R} \Psi \mathbf{C}\|_F^2 + \gamma \text{dTV}(\Psi \mathbf{C} \mathbf{D}_x, \Psi \mathbf{C} \mathbf{D}_y), \quad (5)$$

where  $\beta$  and  $\lambda$  are weighting and regularization parameters, respectively, and  $\|\cdot\|_F$  denotes the Forbenious norm.

### 3 Optimization Method

Due to the nature of dTV regularizer, which is nonquadratic and nonsmooth, the ADMM algorithm [15] is employed to solve problem (5) through the variable splitting technique. Each subproblem can be efficiently solved.

#### 3.1 ADMM for Problem (5)

By introducing 5 auxiliary variables, the original problem (5) is reformulated as

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathbf{Y} - \Psi \mathbf{C} \mathbf{B} \mathbf{S}\|_F^2 + \frac{\beta}{2} \|\mathbf{Z} - \mathbf{R} \Psi \mathbf{V}_1\|_F^2 + \gamma \{\|\mathbf{V}_2\|_1 + \|\mathbf{V}_3\|_1\}_{\text{dTV}} \\ \text{s.t.} \quad & \mathbf{V}_1 = \mathbf{C}, \\ & \mathbf{V}_2 = \mathbf{V}_x - (\mathbf{G}_x \odot \mathbf{V}_x + \mathbf{G}_y \odot \mathbf{V}_y) \odot \mathbf{G}_x, \quad \mathbf{V}_x = \Psi \mathbf{C} \mathbf{D}_x, \\ & \mathbf{V}_3 = \mathbf{V}_y - (\mathbf{G}_x \odot \mathbf{V}_x + \mathbf{G}_y \odot \mathbf{V}_y) \odot \mathbf{G}_y, \quad \mathbf{V}_y = \Psi \mathbf{C} \mathbf{D}_y. \end{aligned} \quad (6)$$

The auxiliary variable  $\mathbf{V}_1$  helps bypass singularity. The auxiliary variables  $\mathbf{V}_2$  and  $\mathbf{V}_3$  help generate closed-form solutions associated with the dTV regularizer. The auxiliary variables  $\mathbf{V}_x$  and  $\mathbf{V}_y$  help compute the coefficient matrix  $\mathbf{C}$  in frequency domain. Problem (6) has the following augmented Lagrangian

$$\begin{aligned} \min \quad & \mathcal{L}_\rho(\mathbf{C}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_x, \mathbf{V}_y, \mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_x, \mathbf{A}_y) \\ = \quad & \frac{1}{2} \|\mathbf{Y} - \Psi \mathbf{C} \mathbf{B} \mathbf{S}\|_F^2 + \frac{\beta}{2} \|\mathbf{Z} - \mathbf{R} \Psi \mathbf{V}_1\|_F^2 + \frac{\rho}{2} \|\mathbf{C} - \mathbf{V}_1 - \mathbf{A}_1\|_F^2 \\ & + \gamma \|\mathbf{V}_2\|_1 + \frac{\rho}{2} \|\mathbf{V}_x - (\mathbf{G}_x \odot \mathbf{V}_x + \mathbf{G}_y \odot \mathbf{V}_y) \odot \mathbf{G}_x - \mathbf{V}_2 - \mathbf{A}_2\|_F^2 \\ & + \gamma \|\mathbf{V}_3\|_1 + \frac{\rho}{2} \|\mathbf{V}_y - (\mathbf{G}_x \odot \mathbf{V}_x + \mathbf{G}_y \odot \mathbf{V}_y) \odot \mathbf{G}_y - \mathbf{V}_3 - \mathbf{A}_3\|_F^2 \\ & + \frac{\rho}{2} \|\Psi \mathbf{C} \mathbf{D}_x - \mathbf{V}_x - \mathbf{A}_x\|_F^2 + \frac{\rho}{2} \|\Psi \mathbf{C} \mathbf{D}_y - \mathbf{V}_y - \mathbf{A}_y\|_F^2, \end{aligned} \quad (7)$$

where matrices  $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_x, \mathbf{A}_y$  represent five scaled dual variables, and  $\rho$  denotes the penalty parameter.

The variables in (7) are solved through iteration. The subproblem of coefficient matrix  $\mathbf{C}^{j+1}$  can be fast minimized in frequency domain, which will be detailed in Subsect. 3.2.

The auxiliary variable  $\mathbf{V}_1$  has the following closed-form solution of an unconstrained least squares problem

$$\mathbf{V}_1^{j+1} = (\beta(\mathbf{R}\Psi)^H(\mathbf{R}\Psi) + \rho\mathbf{I})^{-1} \left( \beta(\mathbf{R}\Psi)^H\mathbf{Z} + \rho(\mathbf{C}^{j+1} - \mathbf{A}_1^j) \right), \quad (8)$$

where  $(\cdot)^H$  denotes matrix conjugate 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

$$\begin{aligned} \mathbf{V}_2^{j+1} &= \text{shrink} \left\{ \left[ \mathbf{V}_x^j - (\mathbf{G}_x \odot \mathbf{V}_x^j + \mathbf{G}_y \odot \mathbf{V}_y^j) \odot \mathbf{G}_x \right] - \mathbf{A}_2^j, \gamma/\rho \right\}, \\ \mathbf{V}_3^{j+1} &= \text{shrink} \left\{ \left[ \mathbf{V}_y^j - (\mathbf{G}_x \odot \mathbf{V}_x^j + \mathbf{G}_y \odot \mathbf{V}_y^j) \odot \mathbf{G}_y \right] - \mathbf{A}_3^j, \gamma/\rho \right\}, \end{aligned} \quad (9)$$

where  $\text{shrink}\{y, \kappa\} := \text{sgn}(y) \cdot \max(|y| - \kappa, 0)$ , with the sign and maximum functions denoted by  $\text{sgn}(\cdot)$  and  $\max(\cdot, \cdot)$  respectively.

Under the definitions of Hadamard product and Forbenious norm, every matrix element of  $\mathbf{V}_x^{j+1}$  and  $\mathbf{V}_y^{j+1}$  can be solved independently by minimizing a simple quadratic function. The solution details are omitted for the sake of simplicity.

Then the scaled dual variables are updated according to the ADMM iterative framework [15]. At the end of iteration, the target HR image  $\mathbf{X}$  is recovered as  $\mathbf{X} = \Psi\mathbf{C}$ . Algorithm 1 lists the procedure of this reconstruction. For any  $\beta > 0$ ,  $\gamma > 0$ , and  $\rho > 0$ , Algorithm 1 will converge to a solution of (5) as its ADMM steps are all closed, proper, and convex [15]. Our study reveals that 20 iterations are enough to obtain a satisfactory HR image.

---

**Algorithm 1.** Reconstruct  $\mathbf{X}$  using ADMM

---

**Input:** LR multispectral matrix  $\mathbf{Y} \in \mathbb{R}^{L \times mn}$ , HR RGB matrix  $\mathbf{Z} \in \mathbb{R}^{3 \times MN}$ ,  
SSF  $\mathbf{R} \in \mathbb{R}^{3 \times L}$ .

**Output:** HR multispectral matrix  $\mathbf{X}$ .

Compute gradient matrices  $\mathbf{G}_x$  and  $\mathbf{G}_y$  from  $\mathbf{Z}$ ;

Train the subspace basis  $\Psi$  from  $\mathbf{Y}$ ;

**for**  $j = 1$  **to** 20 **do**

Compute  $\mathbf{C}^j$  according to Section 3.2;  
 Compute  $\mathbf{V}_1^j$  using (8);  
 Compute  $\mathbf{V}_2^j$  and  $\mathbf{V}_3^j$  using (9);  
 Compute  $\mathbf{V}_x^j$  and  $\mathbf{V}_y^j$ ;  
 Update  $\mathbf{A}_1^j, \mathbf{A}_2^j, \mathbf{A}_3^j, \mathbf{A}_4^j$ , and  $\mathbf{A}_5^j$ ;

**end**

Compute  $\mathbf{X} = \Psi\mathbf{C}$ .

---

### 3.2 Solving Coefficient Matrix

By forcing the derivative of (5) w.r.t.  $\mathbf{C}$  to be zero, an efficient analytical solution can be derived in terms of solving the following Sylvester function

$$\mathbf{C}^{j+1}\mathbf{W}_1 + \mathbf{W}_2\mathbf{C}^{j+1} = \mathbf{W}_3, \quad (10)$$

where

$$\begin{aligned} \mathbf{W}_1 &= \mathbf{B}\mathbf{S}\mathbf{S}^H\mathbf{B}^H + \rho\mathbf{D}_x\mathbf{D}_x^H + \rho\mathbf{D}_y\mathbf{D}_y^H, \\ \mathbf{W}_2 &= \rho(\Psi^H\Psi)^{-1}, \end{aligned}$$

and

$$\begin{aligned} \mathbf{W}_3 &= (\Psi^H\Psi)^{-1}[\Psi^H\mathbf{Y}\mathbf{S}^H\mathbf{B}^H + \rho(\mathbf{V}_1^j + \mathbf{A}_1^j) + \rho\Psi^H(\mathbf{V}_x^j + \mathbf{A}_x^j)\mathbf{D}_x^H \\ &\quad + \rho\Psi^H(\mathbf{V}_y^j + \mathbf{A}_y^j)\mathbf{D}_y^H]. \end{aligned}$$

Using the decomposition  $\mathbf{W}_2 = \mathbf{Q}\Lambda\mathbf{Q}^{-1}$  and multiplying both sides of (10) by  $\mathbf{Q}^{-1}$  leads to

$$\overline{\mathbf{C}}\mathbf{W}_1 + \Lambda\overline{\mathbf{C}} = \overline{\mathbf{W}}_3,$$

where  $\overline{\mathbf{C}} = \mathbf{Q}^{-1}\mathbf{C}^{j+1}$  and  $\overline{\mathbf{W}}_3 = \mathbf{Q}^{-1}\mathbf{W}_3$ . Thus each row of  $\overline{\mathbf{C}}$  can be solved independently as

$$\overline{\mathbf{C}}_i = \overline{\mathbf{W}}_3(\mathbf{W}_1 + \lambda_i\mathbf{I})^{-1}, \quad 1 \leq i \leq K_\Psi, \quad (11)$$

where  $i$  denotes the row index, and  $\lambda_i$  denotes the  $i$ th eigenvalue of  $\mathbf{W}_2$ .

Utilizing the properties of convolution and decimation matrices, the solution (11) can be accelerated in frequency domain. Convolution matrices  $\mathbf{B}$ ,  $\mathbf{D}_x$  and  $\mathbf{D}_y$  can be diagonalized by Fourier matrix  $\mathbf{F} \in \mathbb{R}^{MN \times MN}$ , i.e.,  $\mathbf{B} = \mathbf{F}\Lambda_B\mathbf{F}^H$ ,  $\mathbf{D}_x = \mathbf{F}\Lambda_x\mathbf{F}^H$  and  $\mathbf{D}_y = \mathbf{F}\Lambda_y\mathbf{F}^H$ . Then when computing  $\overline{\mathbf{W}}_3$ , right multiplying with these matrices can be achieved through fast Fourier transform (FFT) and entry-wise multiplication operations. Meanwhile, right multiplying with  $\mathbf{S}^H$  is equivalent to the simple upsampling operation.

For further simplification, the matrix inverse in (11) is represented as

$$\mathbf{F} \left( \Lambda_B\mathbf{F}^H\mathbf{S}\mathbf{S}^H\mathbf{F}\Lambda_B^H + \rho\Lambda_x^2 + \rho\Lambda_y^2 + \lambda_i\mathbf{I} \right)^{-1} \mathbf{F}^H := \mathbf{F}\mathbf{K}^{-1}\mathbf{F}^H.$$

By translating the frequency properties of decimation matrix [10] into

$$\mathbf{F}^H\mathbf{S}\mathbf{S}^H\mathbf{F} = \mathbf{P}\mathbf{P}^H/d^2,$$

$\mathbf{K}$  can be consolidated as

$$\mathbf{K} = \frac{1}{d^2}\Lambda_B\mathbf{P}\mathbf{P}^H\Lambda_B^H + \Lambda_K,$$

where  $\Lambda_K = \rho\Lambda_x^2 + \rho\Lambda_y^2 + \lambda_i\mathbf{I}$  is a diagonal matrix,  $\mathbf{P} \in \mathbb{R}^{MN \times mn}$  is a transform matrix with 0 and 1 elements. Right multiplying with  $\mathbf{P}$  and  $\mathbf{P}^H$  can be

achieved by performing sub-block accumulating and image copying operations to the corresponding image. As the inverse of large-scale matrix is difficult, the Woodbury inversion lemma [11] is used to decompose  $\mathbf{K}^{-1}$  as

$$\mathbf{K}^{-1} = \Lambda_K^{-1} - \Lambda_K^{-1} \Lambda_B \mathbf{P} (d^2 \mathbf{I} + \mathbf{P}^H \Lambda_B^H \Lambda_K^{-1} \Lambda_B \mathbf{P})^{-1} \mathbf{P}^H \Lambda_B^H \Lambda_K^{-1}, \quad (12)$$

where matrix  $d^2 \mathbf{I} + \mathbf{P}^H \Lambda_B^H \Lambda_K^{-1} \Lambda_B \mathbf{P}$  is diagonal.

Inserting (12) into (11) yields the final solution

$$\begin{aligned} \bar{\mathbf{C}}_i = \bar{\mathbf{W}}_3 \mathbf{F} \Lambda_K^{-1} \mathbf{F}^H - \bar{\mathbf{W}}_3 \mathbf{F} \Lambda_K^{-1} \Lambda_B \mathbf{P} (d^2 \mathbf{I} + \mathbf{P}^H \Lambda_B^H \Lambda_K^{-1} \Lambda_B \mathbf{P})^{-1} \\ \mathbf{P}^H \Lambda_B^H \Lambda_K^{-1} \mathbf{F}^H, \quad 1 \leq i \leq K_\Psi, \end{aligned} \quad (13)$$

and the coefficient matrix is computed as  $\mathbf{C}^{j+1} = \mathbf{Q} \bar{\mathbf{C}}$ . Noting that this solution procedure mainly contains the efficient FFT, entry-wise multiplication, sub-block accumulating, and image copying operations.

## 4 Experiments

Experiments are performed on both simulated and our acquired LR multispectral images. In the simulation, the LR multispectral images with 31 bands are generated by applying Gaussian blur and downsampling operations to the images in the Harvard scene dataset [16]<sup>1</sup> and CAVE object dataset [17]<sup>2</sup>. The HR RGB images are generated using the SSF of Canon 60D camera provided in the CamSpec database [18]. In our real image set, the LR multispectral images with 31 bands are acquired across the visible spectrum 400–720 nm by an imaging system consisting of a liquid crystal tunable filters 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 [19].

To evaluate the quality of reconstructed multispectral images, four objective quality metrics namely spectral angle mapper (SAM) [6], root mean squared error (RMSE) [6], relative dimensionless global error in synthesis (ERGAS) [6], and peak signal to noise ration (PSNR) [6] are used in our study. For comparison, three leading super-resolution methods namely HySure [8], R-FUSE [10], and NSSR [9] are also implemented under the same environment. Their source codes are publicly available online<sup>3,4,5</sup>.

### 4.1 Parameter Setting

We evaluate the effect of three key parameters (weighting parameter  $\beta$ , regularization parameter  $\gamma$ , and penalty parameter  $\rho$ ) on the reconstruction accuracy

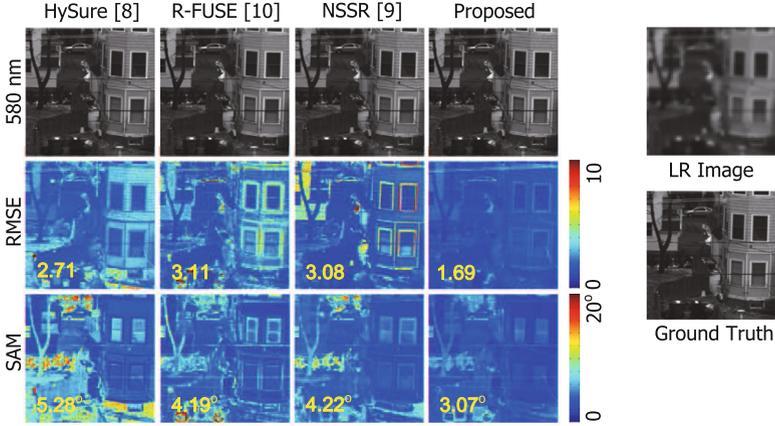
<sup>1</sup> <http://vision.seas.harvard.edu/hyperspec/download.html>.

<sup>2</sup> <http://www1.cs.columbia.edu/CAVE/databases/multispectral/>.

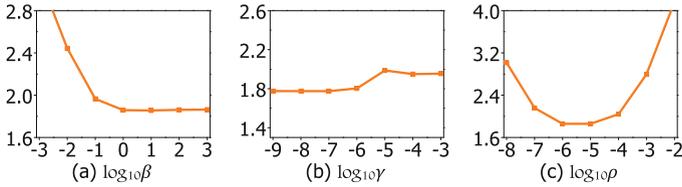
<sup>3</sup> <https://github.com/alfaiate/HySure>.

<sup>4</sup> <https://github.com/qw245/BlindFuse>.

<sup>5</sup> [http://see.xidian.edu.cn/faculty/wsdong/Code\\_release/NSSR\\_HSL\\_SR.rar](http://see.xidian.edu.cn/faculty/wsdong/Code_release/NSSR_HSL_SR.rar).



**Fig. 2.** Reconstruction results of *imgc4* with  $16\times$  spatial resolution improvement. The 1st row shows the reconstructed HR images at 580 nm using different algorithms. The LR image and ground truth image are listed on the right. The remaining rows illustrate the corresponding RMSE maps and SAM maps calculated across all the spectral bands.



**Fig. 3.** The average RMSE values of all the reconstructed images with respect to parameters (a)  $\log_{10}\beta$ , (b)  $\log_{10}\gamma$ , and (c)  $\log_{10}\rho$ .

of proposed algorithm. Figure 3 plots the average RMSE values of all the reconstructed images with respect to these parameters. In this work, we set  $\beta = 1$ ,  $\gamma = 10^{-6}$ , and  $\rho = 10^{-5}$  that result in small RMSE value. We note that setting the  $\beta$  value too large will overemphasize the importance of RGB data term, and setting the  $\gamma$  value too small will decrease the role of RGB edge guidance.

## 4.2 Results on Simulated Images

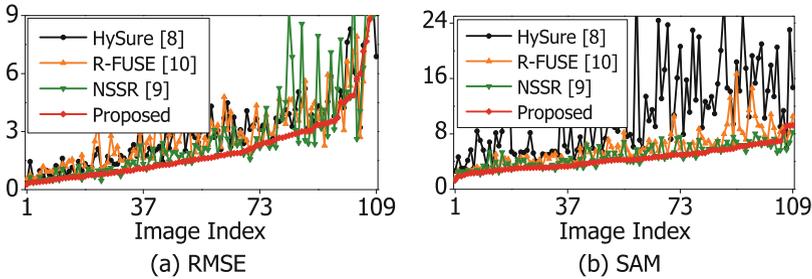
Figure 2 shows the reconstruction results of *imgc4* with  $16\times$  spatial resolution improvement, as well as the detailed RMSE maps and SAM maps. The average RMSE and SAM values are also listed for quantitative comparison. It is observed that the HySure [8] algorithm exhibits large spectral errors, and the R-FUSE [10] and NSSR [9] algorithms do not handle the spatial details well. In comparison, the proposed algorithm produces relatively accurate HR images.

Table 1 shows the average SAM, RMSE, ERGAS, and PSNR values of all the reconstructed multispectral images in Harvard and CAVE datasets. The spatial

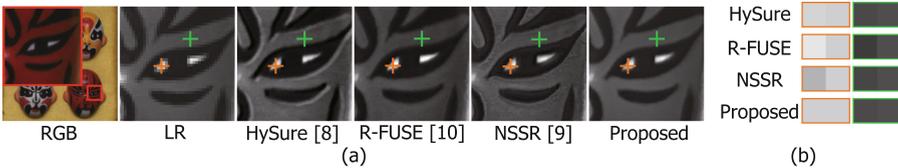
**Table 1.** Average SAM, RMSE, ERGAS, and PSNR values produced by different algorithms on two datasets. The resolution is improved with  $16\times$

	Harvard dataset				CAVE dataset			
	SAM	RMSE	ERGAS	PSNR	SAM	RMSE	ERGAS	PSNR
HySure [8]	8.36	2.57	0.96	36.92	16.70	4.08	1.00	38.81
R-FUSE [10]	5.70	2.72	1.00	35.57	6.38	3.85	0.95	38.71
NSSR [9]	4.65	1.85	0.68	40.00	5.34	4.71	1.01	39.60
Proposed	<b>4.06</b>	<b>1.69</b>	<b>0.57</b>	<b>40.56</b>	<b>5.24</b>	<b>3.42</b>	<b>0.75</b>	<b>40.97</b>

resolution is improved by 16 times. It is observed that the proposed algorithm outperforms all the competitors when evaluated using these metrics. Furthermore, Fig. 4 shows the overall reconstruction accuracy on the 109 multispectral images of the two datasets in terms of RMSE and SAM. For clear demonstration, the image indexes are sorted in ascending order with respect to the metric values produced by the proposed algorithm. It is observed that in most cases the proposed algorithm performs better than the competing methods when evaluated using either spatial or spectral metrics.



**Fig. 4.** (a) RMSE and (b) SAM values produced by different algorithms on all the stimulated data with scale factors  $d = 16$ .



**Fig. 5.** (a) Reconstruction results on real data *Masks* at band 590 nm with  $8\times$  spatial resolution improvement. (b) Marked pixels in reconstructed images compared with the ones in original LR image.

### 4.3 Results on Real Images

We also evaluate the performance of the proposed algorithm on real images acquired in our laboratory. The RGB image is linearized beforehand with the inverse camera response function estimated by [20]. The SSF is computed through linear regression with existing image data. Figure 5(a) shows the original HR RGB image and LR band image at 590 nm of *Masks*, as well as the corresponding reconstructed results with  $8\times$  spatial resolution improvement. Figure 5(b) shows the marked pixels in smooth regions. Each marked pixel in the reconstructed HR image is compared with the one in the original LR image, and it is desired that the intensity of the two pixels should be close. It is observed that the face edges produced by HySure and NSSR are not clear, and the intensity of eye produced by R-FUSE is too high. In comparison, the proposed algorithm performs well in handling these details.

### 4.4 Computational Complexity

The complexity of the proposed algorithm is dominated by the FFTs when computing coefficient matrix  $\mathbf{C}$ , and is of order  $\mathcal{O}(K_{\Psi}MN\log(MN))$  per ADMM iteration. Table 2 shows the running times of the HySure [8], R-FUSE [10], NSSR [9], and proposed algorithms for reconstructing an HR multispectral image with 31 spectral bands and  $1392 \times 1040$  spatial resolution. These algorithms are all implemented using MATLAB R2016a on a personal computer with 2.60 GHz CPU (Intel Xeon E5-2630) and 64 GB RAM. The proposed algorithm gains improvement in computational efficiency.

**Table 2.** Running times (in seconds) of different algorithms for reconstructing an HR multispectral image with 31 bands and  $1392 \times 1040$  spatial resolution. The numbers in parentheses are the speedup of the proposed algorithm over the corresponding competitors

HySure [8]	R-FUSE [10]	NSSR [9]	Proposed
1256.8 (7 $\times$ )	6758.5 (36 $\times$ )	998.8 (5 $\times$ )	185.7

## 5 Conclusions

This paper has proposed a super-resolution algorithm to improve the spatial resolution of multispectral image with an HR RGB image. The HR multispectral image is efficiently reconstructed according to the linear image degradation models, and the dTV operator is used to keep the recovered edge locations and directions in accordance with those of the RGB image. Experimental results validate that the proposed algorithm performs better than the state-of-the-arts in terms of both reconstruction accuracy and computational efficiency.

## References

1. Levenson, R.M., Mansfield, J.R.: Multispectral imaging in biology and medicine: slices of life. *Cytom. Part A* **69**(8), 748–758 (2006)
2. Shaw, G.A., Burke, H.H.K.: Spectral imaging for remote sensing. *Linc. Lab. J.* **14**(1), 3–28 (2003)
3. Berns, R.S.: Color-accurate image archives using spectral imaging. In: *Scientific Examination of Art: Modern Techniques in Conservation and Analysis*, pp. 105–119 (2005)
4. Pan, Z.W., Shen, H.L., Li, C., Chen, S.J., Xin, J.H.: Fast multispectral imaging by spatial pixel-binning and spectral unmixing. *IEEE Trans. Image Process.* **25**(8), 3612–3625 (2016)
5. Wei, Q., Bioucas-Dias, J., Dobigeon, N., Tourneret, J.Y.: Hyperspectral and multispectral image fusion based on a sparse representation. *IEEE Trans. Geosci. Remote. Sens.* **53**(7), 3658–3668 (2015)
6. Lin, C.H., Ma, F., Chi, C.Y., Hsieh, C.H.: A convex optimization-based coupled nonnegative matrix factorization algorithm for hyperspectral and multispectral data fusion. *IEEE Trans. Geosci. Remote. Sens.* **56**(3), 1652–1667 (2018)
7. Dian, R., Fang, L., Li, S.: Hyperspectral image super-resolution via non-local sparse tensor factorization. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5344–5353. IEEE (2017)
8. Simões, M., Bioucas-Dias, J., Almeida, L.B., Chanussot, J.: A convex formulation for hyperspectral image superresolution via subspace-based regularization. *IEEE Trans. Geosci. Remote. Sens.* **53**(6), 3373–3388 (2015)
9. Dong, W., et al.: Hyperspectral image super-resolution via non-negative structured sparse representation. *IEEE Trans. Image Process.* **25**(5), 2337–2352 (2016)
10. Wei, Q., Dobigeon, N., Tourneret, J.Y., Bioucas-Dias, J., Godsill, S.: R-FUSE: robust fast fusion of multiband images based on solving a Sylvester equation. *IEEE Signal Process. Lett.* **23**(11), 1632–1636 (2016)
11. Zhao, N., Wei, Q., Basarab, A., Kouamé, D., Tourneret, J.Y.: Single image super-resolution of medical ultrasound images using a fast algorithm. In: *IEEE 13th International Symposium on Biomedical Imaging*, pp. 473–476. IEEE (2016)
12. Huang, W., Xiao, L., Wei, Z., Liu, H., Tang, S.: A new pan-sharpening method with deep neural networks. *IEEE Geosci. Remote. Sens. Lett.* **12**(5), 1037–1041 (2015)
13. Yang, J., Fu, X., Hu, Y., Huang, Y., Ding, X., Paisley, J.: PanNet: a deep network architecture for pan-sharpening. In: *IEEE International Conference on Computer Vision*, pp. 1753–1761. IEEE (2017)
14. Ehrhardt, M.J., Betcke, M.M.: Multicontrast MRI reconstruction with structure-guided total variation. *SIAM J. Imaging Sci.* **9**(3), 1084–1106 (2016)
15. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends® Mach. Learn.* **3**(1), 1–122 (2011)
16. Chakrabarti, A., Zickler, T.: Statistics of real-world hyperspectral images. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 193–200. IEEE (2011)
17. Yasuma, F., Mitsunaga, T., Iso, D., Nayar, S.K.: Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE Trans. Image Process.* **19**(9), 2241–2253 (2010)

18. Jiang, J., Liu, D., Gu, J., Süsstrunk, S.: What is the space of spectral sensitivity functions for digital color cameras? In: IEEE Workshop on Applications of Computer Vision, pp. 168–179. IEEE (2013)
19. Chen, S.J., Shen, H.L., Li, C., Xin, J.H.: Normalized total gradient: a new measure for multispectral image registration. *IEEE Trans. Image Process.* **27**(3), 1297–1310 (2018)
20. Lee, J.Y., Matsushita, Y., Shi, B., Kweon, I.S., Ikeuchi, K.: Radiometric calibration by rank minimization. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(1), 144–156 (2013)