Adaptive Nonnegative Sparse Representation for Hyperspectral Image Super-Resolution

Xuesong Li ⓘ, Youqiang Zhang ⓘ, *Member, IEEE*, Zixian Ge ⓘ, Guo Cao ⓘ, Hao Shi ⓘ, and Peng Fu

Abstract—As the hyperspectral images (HSIs) usually have a low spatial resolution, HSI super-resolution has recently attracted more and more attention to enhance the spatial resolution of HSIs. A common method is to fuse the low-resolution (LR) HSI with a multispectral image (MSI) whose spatial resolution is higher than the HSI. In this article, we proposed a novel adaptive nonnegative sparse representation-based model to fuse an HSI and its corresponding MSI. First, basing the linear spectral unmixing, the nonnegative structured sparse representation model estimates the sparse codes of the desired high-resolution HSI from both the LR-HSI and the MSI. Then, the adaptive sparse representation can balance the relationship between the sparsity and collaboration by generating a suitable coefficient. Finally, in order to obtain more accurate results, we alternately optimize the spectral basis and coefficients rather than keeping the spectral basis fixed. The alternating direction method of multipliers is applied to solve the proposed optimization problem. The experimental results on both ground-based HSIs and real remote sensing HSIs show the superiority of our proposed approach to some other state-of-the-art HSI super-resolution methods.

Index Terms—Adaptive sparse representation (ASR), hyperspectral image (HSI), spectral basis updating, super-resolution reconstruction.

I. INTRODUCTION

HYPERSPECTRAL (HS) imaging has attracted wide attention in recent years since it can simultaneously obtain images of the same scenario across plenty of different successive wavelengths at the same time [1]–[3]. Because hyperspectral image (HSI) has rich spectral information, it has been widely used in many fields, such as target detection [4], environmental monitoring [5], military [6], and remote sensing [7]. However, since there is a limited amount of incident energy in optical remote sensing systems, the imaging systems have to compromise between the spectral resolution and spatial resolution [8]. For example, HSIs captured by HYPXIM usually have more than

Xuesong Li, Zixian Ge, Guo Cao, Hao Shi, and Peng Fu are with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: cedar_xuesong@163.com; zxge727@foxmail.com; caoguo@njust.edu.cn; hao1227@njust.edu.cn;).

Youqiang Zhang is with the Jiangsu Key Laboratory of Broadband Wireless Communication and Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China, and also with the School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: zhangyouqiang@foxmail.com).
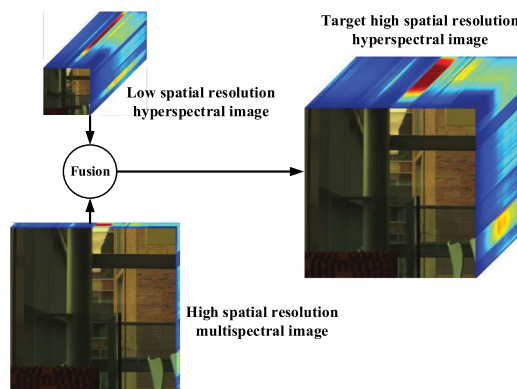
Fig. 1. HSI super-resolution problem.

one hundred spectral bands but only a decametric spatial resolution. Compared with HS imaging sensors, the multispectral (MS) imaging sensors can provide multispectral images (MSIs) with much higher spatial resolution but with a limited number of spectral bands. For example, the PLEIADES can provide MSIs with a spatial resolution of 70 cm but with only three or four spectral bands. In order to enhance the spatial resolution of HSIs, researchers have made much effort. A popular approach to reconstruct the high spatial resolution HSI (HR-HSI) is to fuse the high spatial resolution MSI (HR-MSI) with the low spatial resolution HSI (LR-HSI) [9], [10]. This approach is called HSI–MSI fusion or HSI super-resolution.

HSI super-resolution problem aims to reconstruct an HR-HSI by fusing the spectral information of an LR-HSI and the spatial information of an HR-MSI, as illustrated in Fig. 1. Note that the LR-HSI and the HR-MSI should be the same scene. The target HSI should not only have a good visual effect but also ensure the authenticity of each pixel.

A large number of studies have been done on HSI super-resolution. A special situation of HSI super-resolution is pan-sharpening, which fuses an LR-HSI with its corresponding panchromatic (PAN) image [9], [11]. A variety of pansharpening methods have been proposed over the past two decades. Generally, these methods can be categorized into two classes, i.e., transform-based methods [12]–[14] and variational methods [15]–[17]. However, because the PAN images have little spectral resolution, there are usually considerable spectral distortions in the HR-HSIs reconstructed by these pansharpening methods.

As the MSIs contain more spectral information than the PAN images, in recent work, HSI–MSI fusion, which can be seen

as the extension of pansharpening, has drawn more attention. Yokoya *et al.* [10] present a comparative review of several HSI–MSI fusion techniques. Typically, the HSI–MSI fusion methods can be divided into four categories: component substitution (CS), Bayesian, deep learning, and sparse representation.

In the CS-based approaches, a basic idea is to substitute one component of the HSI with the high-resolution (HR) image. For example, the intensity-hue-saturation (IHS) [18], [19] method replaces the intensity component in the IHS domain of the LR image with the PAN image. The principal component analysis [20] method uses the HR image to replace the first principal component of the LR-HSI. However, the CS-based approaches usually result in spectral distortions in the obtained HR-HSI.

The Bayesian-based approaches introduce the appropriate prior distribution of the HR-MSI, such as naive Gaussian [21], [22] and sparsity promoting prior [23], [24] to achieve the accurate estimation. The variational methods can be regarded as a special case of the Bayesian one. The target images are estimated by minimizing the objective function, which is structured by the posterior probability density of the fused image. Among these methods, HS super-resolution [25] uses the vector-total-variation-based regularization in the objective function. Zhang *et al.* [26], [27] introduced a method that works in the wavelet domain and later published an expectation–maximization algorithm to maximize the posterior distribution.

Since the deep learning has been demonstrated to be very effective in object detection [28]–[30], classification [31]–[33], and natural image super-resolution [34]–[36], many researchers have introduced deep learning into HSI super-resolution. Li *et al.* [37] proposed to learn an end-to-end spectral difference mapping between the LR-HSI and HR-HSI through a deep spectral difference convolution neural network. Yuan *et al.* [38] proposed a multiscale and multidepth convolution neural network to achieve the HR-HSI. In order to take the advantage of the spectral correlation and exploit the HR-MSI, Yang *et al.* [39] presented a convolution neural network with two branches. With the two branches convolution neural network, the spectrum features of each pixel and its corresponding spatial neighborhood are extracted from the LR-HSI and the HR-MSI, respectively. Dian *et al.* [40] proposed to learn the spectral prior of HSI via deep residual convolutional neural networks. In addition to the convolution neural network, a stacking sparse denoising autoencoder-based deep neural network is proposed by Huang *et al.* [41] for pansharpening. Although the deep learning based methods obtained great reconstruction results, these kinds of methods need large amounts of training samples to estimate the parameters.

In the past years, the sparse representation has been widely used in remote sensing applications [42]. The sparse representation-based HSI super-resolution methods usually represent the targeted HR-HSI image by the product of a spectral basis matrix and a coefficient matrix, where the spectral basis and coefficient matrices can be extracted from the LR-HSI and the HR-MSI. Besides, some matrix factorization and unmixing-based methods can also be regarded as the sparse representation-based method because the source images are decomposed into spectral bases and coefficients. Actually, the sparse representation-based methods are usually combined with matrix factorization and spectral unmixing. Based on the unsupervised spectral unmixing, Yokoya *et al.* [43] proposed a coupled nonnegative matrix factorization (CNMF) approach to estimate the HSI endmember matrix and the HR abundance matrix. However, the nonnegative matrix factorization is usually not unique. So, Yokoya *et al.* [43] cannot always obtain satisfactory results. Huang *et al.* [44] used the *k*-singular value decomposition (K-SVD) algorithm [45] to learn the spectral basis and proposed a sparse prior-based matrix factorization method to fuse the remote sensing MSI at different spatial and spectral resolution. Zhang *et al.* [46] used the group spectral embedding and low-rank factorization to fuse the LR-HSI and HR-MSI. Lanaras *et al.* [47] proposed to jointly solve the spectral unmixing problems for both input images. However, only using a spectral dictionary is insufficient for preserving spatial information, and vice-versa. To address this problem, an HSI–MSI fusion method termed optimized twin dictionaries (OTD) using optimized twin dictionaries was proposed by Han *et al.* [48]. Since the pixelwise sparse representation neglects the similarity among neighbor pixels, Akhtar *et al.* [49] proposed to utilize the similarities among the spectral pixels in the same local patch and obtain the coefficients with a generalization of simultaneous orthogonal matching pursuit (G-SOMP$_+$) algorithm for each local patch. Later, Akhtar *et al.* [50] proposed a Bayesian dictionary learning and Bayesian sparse coding approach for HSI super-resolution and achieved improved performance. Note that, the structures of MSI are usually very complex, and thus, a fixed local window may still contain different variations. Combined with superpixel segmentation methods, Fang *et al.* [51] proposed a superpixel-based sparse representation (SSR) model, which ensured that the shape and size of each superpixel can adaptively adjust according to the spatial structures of MSI, and therefore, the spatial structures of spectral pixels in each superpixel are similar for HSI super-resolution. Furthermore, Dong *et al.* [52] proposed a nonnegative structured sparse representation (NSSR) method, which exploited a clustering-based structured sparse coding approach to ensure the spatial correlation among the obtained sparse coefficients.

Sparse representation-based approaches are indeed effective for HSI super-resolution and achieve great reconstruction results. However, the existing methods usually use an l_1-norm to constrain the representation coefficients, and thus, only the sparsity is taken into consideration. Sometimes the constraint of l_1-norm is not reasonable because there is not only the sparsity but also some correlation among the representation coefficients. Another extreme case, which only uses an l_2-norm to constrain the representation coefficients, is to consider only the correlated information. Therefore, a more reasonable choice is to take the sparsity and correlation simultaneously into consideration. Inspired by the trace least absolute shrinkage and selection operator (LASSO) [53], [54], we propose a novel spatial–spectral adaptive nonnegative sparse representation (ANSR) method for HSI super-resolution by fusing the LR-HSI and the corresponding HR-MSI. The proposed method integrates sparsity and correlation effectively as a regularization term in the model and can produce more suitable coefficients adaptively with the constraint

between l_1-norm and l_2-norm. Specifically, the estimation of HR-HSI is formulated as a joint estimation of spectral basis and sparse coefficients from the LR-HSI and HR-MSI with the prior knowledge of spatial–spectral sparsity and spectral unmixing. According to the spectral mixture model [55], the spectral basis and sparse coefficients need to be nonnegative, and the sparse coefficients often meet the sum-to-one constraint. Besides, based on trace LASSO, we utilize the adaptive sparse representation (ASR) to balance the sparsity and correlation and can obtain more precise sparse coefficients. Furthermore, we also design an alternative optimization algorithm to update the spectral basis and sparse coefficients, which is more flexible and accurate than keeping the spectral basis fixed. Meanwhile, the alternating direction method of multipliers (ADMM) is adopted to solve both the updating of the spectral basis and coefficients.

The main contributions of this article can be summarized as follows.

1) We introduce the ASR, which can obtain more precise sparse coefficients by balancing the sparsity and correlation of the coefficients, into the HSI super-resolution model.

2) Instead of keeping the spectral basis fixed, we alternately optimize the spectral basis and sparse coefficients.

3) We design two specific ADMM methods to update the spectral basis and sparse coefficients, respectively.

4) Experimental results on both ground-based HSIs and real remote sensing HSIs show that our ANSR method performs better than some other state-of-the-art HSI super-resolution methods.

The remainder of this article is organized as follows. We briefly introduce the spectral dictionary learning method in [52] and the ASR in Section II. In Section III, we first formulate the problem of HSI super-resolution and then describe the details of the proposed ANSR method for HSI super-resolution. Extensive experiments and comparisons are shown in Section IV. Finally, Section V concludes this article.

II. RELATED WORK

In this section, we introduce the spectral dictionary learning method in [52] and the ASR, which are used in our method.

A. Spectral Dictionary Learning

We denote the LR-HSI as $\boldsymbol{X} \in \mathbb{R}^{B \times n}$, where B represents the spectral dimension and n represents the number of pixels. As each pixel in the LR-HSI $\boldsymbol{X}$ can be written as the linear combination of a small number of spectral pixels, we can express $\boldsymbol{X}$ as the product of a spectral dictionary $\boldsymbol{D}$ and a coefficient matrix $\boldsymbol{B}$. The formulation is as follows:

$$\boldsymbol{X} = \boldsymbol{DB} + \boldsymbol{V} \qquad (1)$$

where $\boldsymbol{V}$ denotes the approximation error matrix, which is assumed to be additive Gaussian.

In (1), both $\boldsymbol{D}$ and $\boldsymbol{B}$ are unknown. Generally, there are infinite possible decompositions of (1) and a unique decomposition cannot be determined. Fortunately, with the help of the sparsity assumption, we can solve $\boldsymbol{D}$ and $\boldsymbol{B}$ using the sparse nonnegative

matrix decomposition. Therefore, the spectral dictionary $\boldsymbol{D}$ can be estimated by solving the following sparse nonnegative matrix decomposition problem:

$$(\boldsymbol{D}, \ \boldsymbol{B}) = \arg \min_{\boldsymbol{D}, \boldsymbol{B}} \frac{1}{2} \|\boldsymbol{X} - \boldsymbol{DB}\|_F^2 + \lambda \|\boldsymbol{B}\|_1$$

$$\text{s.t. } \boldsymbol{B} \geq 0, \ \boldsymbol{D} \geq 0 \qquad (2)$$

Because that the sparse coefficient matrix $\boldsymbol{B}$ and the spectral dictionary $\boldsymbol{D}$ are constrained to be nonnegative, existing dictionary algorithms (e.g., K-SVD algorithm and online dictionary learning algorithm) are all invalid. In order to solve the above sparse nonnegative matrix decomposition problem, a computationally efficient nonnegative dictionary learning algorithm, which solves (2) by updating $\boldsymbol{D}$ and $\boldsymbol{B}$ alternately, is proposed in [52].

With $\boldsymbol{D}$ fixed, the subproblem with respect to $\boldsymbol{B}$ becomes

$$\boldsymbol{B} = \arg \min_{\boldsymbol{B}} \frac{1}{2} \|\boldsymbol{X} - \boldsymbol{DB}\|_F^2 + \lambda \|\boldsymbol{B}\|_1, \text{ s.t. } \boldsymbol{B} \geq 0 \quad (3)$$

which can be efficiently solved by the ADMM technique. To apply ADMM, we introduce $\boldsymbol{S} = \boldsymbol{B}$, and (3) can be reformulated as the following augmented Lagrangian function:

$$L(\boldsymbol{B}, \ \boldsymbol{S}, \ \boldsymbol{U}) = \frac{1}{2} \|\boldsymbol{X} - \boldsymbol{DB}\|_F^2 + \lambda \|\boldsymbol{B}\|_1 + \mu \|\boldsymbol{S} - \boldsymbol{B} + \frac{\boldsymbol{U}}{2\mu}\|_F^2$$

$$\text{s.t. } \boldsymbol{B} \geq 0 \qquad (4)$$

where $\boldsymbol{U}$ is the Lagrangian multiplier ($\mu \geq 0$). Then, we solve $\boldsymbol{B}, \boldsymbol{S}$, and $\boldsymbol{U}$ alternately until convergence.

With $\boldsymbol{B}$ fixed, the subproblem with respect to $\boldsymbol{D}$ becomes

$$\boldsymbol{D} = \arg \min_{\boldsymbol{D}} \|\boldsymbol{X} - \boldsymbol{DB}\|_F^2, \text{ s.t. } \boldsymbol{D} \geq 0. \qquad (5)$$

Similar to the online dictionary learning method, (5) is solved by using block coordinate descent. During each iteration, one column of $\boldsymbol{D}$ is updated while keeping the others fixed under the nonnegative constraint.

More information about the spectral dictionary learning method can be found in [52].

B. Adaptive Sparse Representation

As we all know, the goal of sparse representation is to encode a signal vector as a linear combination of a few dictionary atoms. Suppose that $\boldsymbol{x} \in \mathbb{R}^m$ is an input signal vector and $\boldsymbol{D_s} \in \mathbb{R}^{m \times K}$ ($m \ll K$) is a dictionary, with l_0-norm as the regularization term, the sparse representation model of $\boldsymbol{x}$ takes the form

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{x} - \boldsymbol{D_s \alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_0 \qquad (6)$$

where $\boldsymbol{\alpha}$ is the sparse coefficient of $\boldsymbol{x}$ and λ is a regularization parameter.

However, the l_0-minimization problem is NP-hard. Usually, the l_1-norm, a reasonably convex surrogate of l_0-norm, is chosen to replace the l_0-norm. Then, the sparse optimization model takes the form

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{x} - \boldsymbol{D_s \alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1. \qquad (7)$$

The result of (7) can be solved by quadratic programming techniques, including basis pursuit [56], LASSO [57], etc.

Although the l_1-regularization can make full use of the sparse information of signals, it completely ignores the correlated information of signals. Timofte *et al.* [58] proposed the collaborative representation that replaces the l_1-norm of sparse representation model by l_2-norm. The collaborative representation model takes the form as

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{x} - \boldsymbol{D_s}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_2^2. \tag{8}$$

In contrast, the collaborative representation model only takes the correlation into consideration and completely ignores the sparsity. Actually, the best choice is to balance the sparsity and correlation and make a compromise between l_1-norm and l_2-norm. In order to overcome this problem, Zhao *et al.* [59] proposed a natural image super-resolution method with the property of trace LASSO [60], [61]

$$\|\boldsymbol{\alpha}\|_2 \leq \|\boldsymbol{D_s}\mathrm{Diag}\left(\boldsymbol{\alpha}\right)\|_* \leq \|\boldsymbol{\alpha}\|_1 \tag{9}$$

where $\|\cdot\|_*$ represents the kernel norm, which computes the sum of the singular values of a matrix, and $\mathrm{Diag}(\boldsymbol{\alpha})$ obtains a diagonal matrix whose diagonal elements are the corresponding values in vector $\boldsymbol{\alpha}$. When the columns of basis $\boldsymbol{D_s}$ are almost uncorrelated, $\|\boldsymbol{D_s}\mathrm{Diag}(\boldsymbol{\alpha})\|_*$ will be close to $\|\boldsymbol{\alpha}\|_1$. Conversely, $\|\boldsymbol{D_s}\mathrm{Diag}(\boldsymbol{\alpha})\|_*$ will be close to $\|\boldsymbol{\alpha}\|_2$. In practice, the column vectors of a basis are neither too correlated nor too independent. So, the trace LASSO can make a compromise between l_1-norm and l_2-norm adaptively. We call it ASR. Then, the ASR model that can find a more suitable sparse coefficient takes the form

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{x} - \boldsymbol{D_s}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{D_s}\mathrm{Diag}\left(\boldsymbol{\alpha}\right)\|_*. \tag{10}$$

Based on the convexity of the model, this problem can be solved by some efficient methods. Among them, the alternation direction method of multipliers (ADMM) [62], [63] is a widely used method to find an approximate optimal solution.

III. Proposed ANSR Method

In this section, we first provide a general introduction of the HSI super-resolution problem, including the linear spectral mixture model. Then, we introduce our super-resolution model in detail. Finally, we brief readers on the alternating optimization method thoroughly, including the optimization of coefficients and spectral basis.

In this article, the bold lowercase letters stand for the vectors and the bold uppercase letters stand for the matrices. The plain lowercase letters stand for the scalars.

A. Problem Formulation

The HSI super-resolution aims to recover an HR-HSI $\boldsymbol{Z} \in \mathbb{R}^{B \times N}$ from an LR-HSI $\boldsymbol{X} \in \mathbb{R}^{B \times n}$ and an HR-MSI $\boldsymbol{Y} \in \mathbb{R}^{b \times N}$ of the same scene, where $N = W \times H$ and $n = w \times h \, (w \ll W, h \ll H)$ denote the number of pixels in the HR-HSI $\boldsymbol{Z}$ and LR-HSI $\boldsymbol{X}$, respectively. B and $b \, (b \ll B)$ indicate the spectral dimensions of $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively.

In the linear spectral mixture model, each spectral vector $\boldsymbol{z_i} \in \mathbb{R}^B$ of the target image $\boldsymbol{Z}$ can be represented by a linear
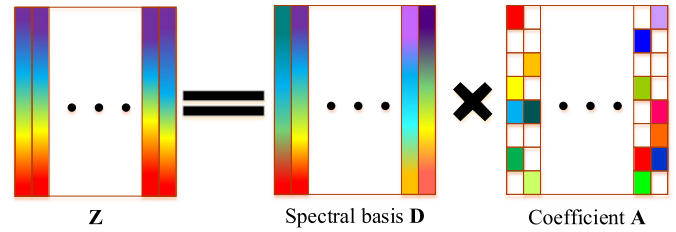

Fig. 2. Linear spectral mixture model.

combination of several spectral signatures [43], as shown in Fig. 2. Mathematically, we have

$$\boldsymbol{z_i} = \boldsymbol{D}\boldsymbol{\alpha_i} \tag{11}$$

where $\boldsymbol{D} \in \mathbb{R}_+^{B \times K}$ is the spectral basis with K atoms, and $\boldsymbol{\alpha_i} \in \mathbb{R}^K$ represents the corresponding coefficient. Each column of $\boldsymbol{D}$ denotes a spectral vector of the underlying material in the scene. Considering pixels of the whole HSI, (11) can be rewritten as

$$\boldsymbol{Z} = \boldsymbol{D}\boldsymbol{A} \tag{12}$$

where $\boldsymbol{Z} = [\boldsymbol{z_1}, \boldsymbol{z_2}, \ldots, \boldsymbol{z_N}]$ and $\boldsymbol{A} = [\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \ldots, \boldsymbol{\alpha_N}] \in \mathbb{R}^{K \times N}$.

Furthermore, both $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be regarded as linear combinations of the target HSI $\boldsymbol{Z}$. The LR-HSI $\boldsymbol{X}$ can be formulated as the linear spatial degradation of $\boldsymbol{Z}$

$$\boldsymbol{X} = \boldsymbol{Z}\boldsymbol{H} \tag{13}$$

where $\boldsymbol{H} \in \mathbb{R}^{N \times n}$ represents the spatial dimensionality degradation operator, including blurring and downsampling.

The HR-MSI $\boldsymbol{Y}$ can be formulated as the linear spectral degradation of $\boldsymbol{Z}$

$$\boldsymbol{Y} = \boldsymbol{P}\boldsymbol{Z} \tag{14}$$

where $\boldsymbol{P} \in \mathbb{R}^{b \times B}$ represents the spectral dimensionality downsampling matrix, which is the spectral response of the MS sensor.

By combining the linear mixture model (12) and the forward models (13) and (14), we have

$$\boldsymbol{X} = \boldsymbol{Z}\boldsymbol{H} = \boldsymbol{D}\boldsymbol{A}\boldsymbol{H} = \boldsymbol{D}\boldsymbol{B} \tag{15}$$

$$\boldsymbol{Y} = \boldsymbol{P}\boldsymbol{Z} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{A} \tag{16}$$

where $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{H} \in \mathbb{R}^{K \times n}$ is a coefficient matrix with each column of $\boldsymbol{B}$ being a sparse vector.

According to the linear spectral mixture model (12), the HSI super-resolution problem can be transformed into the estimation of spectral basis $\boldsymbol{D}$ and representation coefficients $\boldsymbol{A}$. And, $\boldsymbol{D}$ and $\boldsymbol{A}$ can be well estimated from the LR-HSI $\boldsymbol{X}$ and HR-MSI $\boldsymbol{Y}$ with (15) and (16), which will be elaborated in Section III-B.

B. Establishment of Our Model

As mentioned above, the HSI super-resolution problem can be transformed into the estimation of spectral basis $\boldsymbol{D}$ and representation coefficients $\boldsymbol{A}$. According to (15) and (16), we can jointly estimate the spectral basis $\boldsymbol{D}$ and coefficients $\boldsymbol{A}$ from both the LR-HSI and HR-MSI. In this way, the HSI super-resolution

problem can be written as

$$\min_{\boldsymbol{D},\,\boldsymbol{A}} \|\boldsymbol{Y} - \boldsymbol{P}\boldsymbol{D}\boldsymbol{A}\|_F^2 + \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{A}\boldsymbol{H}\|_F^2. \tag{17}$$

Obviously, the above optimization problem is ill-posed, and the solutions of $\boldsymbol{D}$ and $\boldsymbol{A}$ are not unique. Therefore, we need some prior knowledge to constrain the solution space. Some common and effective priors include sparsity prior, nonlocal spatial similarities, and nonnegative prior.

The sparsity prior is known to be a very effective method to deal with the HSI super-resolution problem. With the sparsity constraint, we assume that each spectral pixel in the target HSI can be represented as a linear combination of a few distinct atoms of spectral basis. Then, the HSI super-resolution problem can be written as

$$\min_{\boldsymbol{D},\,\boldsymbol{A}} \|\boldsymbol{Y} - \boldsymbol{P}\boldsymbol{D}\boldsymbol{A}\|_F^2 + \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{A}\boldsymbol{H}\|_F^2 + \eta\|\boldsymbol{A}\|_1 \tag{18}$$

where $\|\cdot\|_1$ stands for the sum of the absolute values of all elements in a vector and η is a regularization parameter.

However, in (18), the sparse coefficients of each spectral pixel are estimated independently. It is generally known that a pixel of a typical HSI usually has a strong spatial correlation with its similar neighbors. In order to take advantage of the local and nonlocal similarities, we assume that a spectral pixel $\boldsymbol{z_i}$ in the target HSI can be approximately represented as a linear combination of the pixels, which are similar to it. Combined with this nonlocal spatial similarity prior, (18) can be improved to

$$\min_{\boldsymbol{D},\,\boldsymbol{A}} \|\boldsymbol{Y} - \boldsymbol{P}\boldsymbol{D}\boldsymbol{A}\|_F^2 + \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{A}\boldsymbol{H}\|_F^2 + \eta_2\|\boldsymbol{A}\|_1$$
$$+ \eta_1 \sum_{q=1}^{Q} \sum_{i \in S_q} \|\boldsymbol{D}\boldsymbol{\alpha_i} - \boldsymbol{\mu_q}\|_2^2 \tag{19}$$

where $\boldsymbol{\mu_q}$ represents the qth cluster center, which can be seen as the linear combinations of pixels that are similar to the reconstructed spectral pixel $\boldsymbol{z_i}$. The centre $\boldsymbol{\mu_q}$ of the qth cluster can be computed as

$$\boldsymbol{\mu_q} = \sum_{i \in S_q} \boldsymbol{\omega_i} \left(\boldsymbol{D}\boldsymbol{\alpha_i}\right) \tag{20}$$

where $\boldsymbol{\omega_i}$ denotes the weighting coefficients based on the similarity of the target HSI pixels. Since the HSI is unknown, we use the HR-MSI that has the same spatial information as the target HSI to computed $\boldsymbol{\omega_i}$. And the weighting coefficient $\boldsymbol{\omega_i}$ is computed as

$$\boldsymbol{\omega_i} = \frac{1}{c} \exp\left(\frac{-\|\boldsymbol{y_i} - \boldsymbol{y_q}\|_2^2}{h}\right) \tag{21}$$

where c represents the normalization constant, and $\boldsymbol{y_i}$ and $\boldsymbol{y_q}$ represent two pixels of HR-MSI, respectively. In practice, the vector $\boldsymbol{\alpha_i}$ is not known. We cannot achieve $\boldsymbol{\mu_q}$ directly using the (20). To overcome this difficulty, we iteratively estimating $\boldsymbol{\mu_q}$ from the current update of $\boldsymbol{\alpha_i}$. With the estimated $\boldsymbol{\mu_q}$ and taking the whole image into consideration, we can rewrite (19)

as

$$\min_{\boldsymbol{D},\,\boldsymbol{A}} \|\boldsymbol{Y} - \boldsymbol{P}\boldsymbol{D}\boldsymbol{A}\|_F^2 + \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{A}\boldsymbol{H}\|_F^2 + \eta_2\|\boldsymbol{A}\|_1$$
$$+ \eta_1 \|\boldsymbol{D}\boldsymbol{A} - \boldsymbol{U}\|_F^2 \tag{22}$$

where $\boldsymbol{U} = [\boldsymbol{\mu_1}, \boldsymbol{\mu_2}, \ldots, \boldsymbol{\mu_N}]$.

Besides, considering the physical characteristics of HSIs, the pixels of an HSI should be nonnegative. With this nonnegative prior, we can improve (22) to

$$\min_{\boldsymbol{D},\,\boldsymbol{A}} \|\boldsymbol{Y} - \boldsymbol{P}\boldsymbol{D}\boldsymbol{A}\|_F^2 + \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{A}\boldsymbol{H}\|_F^2 + \eta_2\|\boldsymbol{A}\|_1$$
$$+ \eta_1 \|\boldsymbol{D}\boldsymbol{A} - \boldsymbol{U}\|_F^2, \text{ s.t. } \boldsymbol{A} \geq 0,\ 0 \leq \boldsymbol{D} \leq 1. \tag{23}$$

Furthermore, in order to balance the sparsity and correlation, we propose to use trace LASSO instead of the l_1-norm to constrain the coefficients in this article. The trace LASSO can make a compromise between l_1-norm and l_2-norm adaptively, and we call it an ASR. Finally, the HSI super-resolution problem can be written as

$$\min_{\boldsymbol{D},\,\boldsymbol{A}} \|\boldsymbol{Y} - \boldsymbol{P}\boldsymbol{D}\boldsymbol{A}\|_F^2 + \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{A}\boldsymbol{H}\|_F^2 + \eta_1\|\boldsymbol{D}\boldsymbol{A} - \boldsymbol{U}\|_F^2$$
$$+ \eta_2 \sum_{i=1}^{N} \|\boldsymbol{P}\boldsymbol{D}\mathrm{Diag}\left(\boldsymbol{\alpha_i}\right)\|_*$$
$$\text{s.t. } \boldsymbol{A} \geq 0,\ 0 \leq \boldsymbol{D} \leq 1 \tag{24}$$

where $\boldsymbol{\alpha_i}$ represents the ith column of the coefficient matrix $\boldsymbol{A}$.

Once we have solved $\boldsymbol{D}$ and $\boldsymbol{A}$, the target HR-HSI can be obtained by multiplying $\boldsymbol{D}$ by $\boldsymbol{A}$.

C. Alternating Optimization of the Fusion Problem

It is obvious that (24) is highly nonconvex. However, the problem (24) is convex with respect to $\boldsymbol{D}$ and $\boldsymbol{A}$, respectively. Therefore, we propose to alternately optimize the $\boldsymbol{D}$ and $\boldsymbol{A}$, respectively, with the other one fixed. First, we initialize the spectral basis $\boldsymbol{D}$ using the spectral dictionary learning method in [52]. Then, the $\boldsymbol{D}$ and $\boldsymbol{A}$ are updated alternatively via (24). Specifically, we update $\boldsymbol{A}$ with $\boldsymbol{D}$ fixed, and then we update $\boldsymbol{D}$ with $\boldsymbol{A}$ fixed. These two steps are iterated until they converge. Finally, the target HR-HSI can be obtained through (12). The overall algorithm for the HSI super-resolution problem is summarized in Algorithm 1. In order to show the operation process of our method more intuitively, the flowchart of the proposed HSI super-resolution method is illustrated in Fig. 3.

D. Optimization of the Coefficients With the Spectral Basis Fixed

In this procedure, we fix the spectral basis $\boldsymbol{D}$. Then, the updating of coefficient matrix $\boldsymbol{A}$ can be written as

$$\min_{\boldsymbol{A}} \|\boldsymbol{Y} - \boldsymbol{P}\boldsymbol{D}\boldsymbol{A}\|_F^2 + \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{A}\boldsymbol{H}\|_F^2 + \eta_1\|\boldsymbol{D}\boldsymbol{A} - \boldsymbol{U}\|_F^2$$
$$+ \eta_2 \sum_{i=1}^{N} \|\boldsymbol{P}\boldsymbol{D}\mathrm{Diag}\left(\boldsymbol{\alpha_i}\right)\|_*, \text{ s.t. } \boldsymbol{A} \geq 0 \tag{25}$$
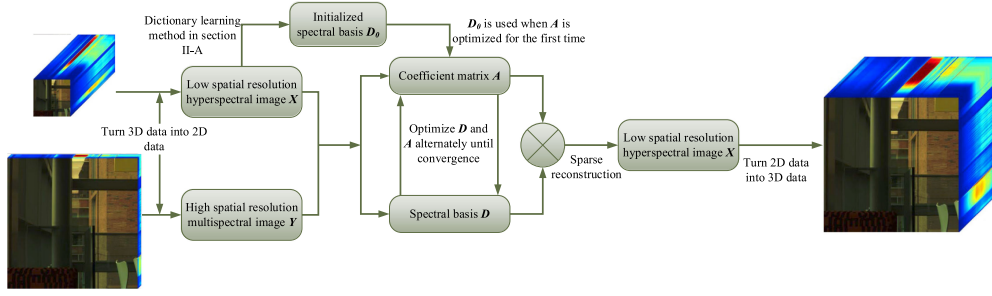
Fig. 3.   Flowchart of the proposed HSI super-resolution method.

Algorithm 1 : ANSR-Based HSI Super-Resolution.

1: **Input:** LR-HSI X; HR-MSI Y; spatial degradation operator
  H; spectral transform matrix P; and regularization parameters η_1 and η_2.

2: **Initialize** the spectral basis D.

3: **While** not converge **do**

4: Update coefficient matrix A with D fixed.

5: Update spectral basis D with A fixed.

6: **End while**

7: Compute the desired HR-HSI Z via (12).

8: **Output:** HR-HSI Z.

Obviously, the optimization problem (25) is convex and can be efficiently solved by ADMM, which can decompose the complex optimization problem into several easily solved subproblems. In specific, we introduce $S = A$, $Q_i = PD\text{Diag}(\alpha_i)$, and $Z = DA$ and can obtain the following augmented Lagrangian function:

$$L\left(A,\ S,\ Q, Z,\ V_1,\ V_2,\ V_3\right)$$
$$= \|Y - PDS\|_F^2 + \|X - ZH\|_F^2$$
$$+ \eta_1 \|DS - U\|_F^2 + \eta_2 \sum_{i=1}^{N} \|Q_i\|_* + \mu \|DS - Z + \frac{V_1}{2\mu}\|_F^2$$
$$+ \mu \|S - A + \frac{V_2}{2\mu}\|_F^2 + \mu \sum_{i=1}^{N} \|Q_i - PD\text{Diag}(\alpha_i) + \frac{V_3^{(i)}}{2\mu}\|_F^2$$

$$\text{s.t. } A \geq 0 \qquad (26)$$

where $V_1, V_2,$ and V_3 are the Lagrangian multipliers ($\mu > 0$). Minimizing the augmented Lagrangian function (26) leads to the following iterations:

$$S^{(t+1)} = \arg\min_{S} L\left(A^{(t)}, S, Q^{(t)}, Z^{(t)}, V_1^{(t)}, V_2^{(t)}, V_3^{(t)}\right)$$
$$Z^{(t+1)} = \arg\min_{S} L\left(A^{(t)}, S^{(t)}, Q^{(t)}, Z, V_1^{(t)}, V_2^{(t)}, V_3^{(t)}\right)$$
$$Q^{(t+1)} = \arg\min_{S} L\left(A^{(t)}, S^{(t)}, Q, Z^{(t)}, V_1^{(t)}, V_2^{(t)}, V_3^{(t)}\right)$$
$$A^{(t+1)} = \arg\min_{S} L\left(A, S^{(t)}, Q^{(t)}, Z^{(t)}, V_1^{(t)}, V_2^{(t)}, V_3^{(t)}\right).$$
$$(27)$$

Meanwhile, the Lagrangian multipliers are updated by

$$V_1^{(t+1)} = V_1^{(t)} + \mu \left(DS^{(t+1)} - Z^{(t+1)}\right)$$
$$V_2^{(t+1)} = V_2^{(t)} + \mu \left(S^{(t+1)} - A^{(t+1)}\right)$$
$$V_3^{(i)(t+1)} = V_3^{(i)(t)} + \mu \left(Q_i^{(t+1)} - PD\text{Diag}\left(\alpha_i^{(t+1)}\right)\right).$$
$$(28)$$

All the subproblems in (27) can be solved analytically, i.e.,

$$S = \left[(PD)^T (PD) + (\eta_1 + \mu) D^T D + \mu I\right]^{-1}$$
$$\left[(PD)^T Y + \eta_1 D^T U + \mu D^T \left(Z - \frac{V_1}{2\mu}\right)\right.$$
$$\left. + \mu \left(A - \frac{V_2}{2\mu}\right)\right]$$
$$Z = \left[XH^T + \mu \left(DS + \frac{V_1}{2\mu}\right)\right]\left(HH^T + \mu I\right)^{-1}$$

as to the solutions of Q and A, we need to solve them pixel by pixel

$$Q_i = \mathcal{J}_{\frac{\eta_2}{2\mu}}\left[PD\text{Diag}(\alpha_i) - \frac{V_3^{(i)}}{2\mu}\right]$$

$$\alpha_i = \left\{\left[2\mu I + 2\mu\text{Diag}\left(\text{Diag}\left((PD)^T (PD)\right)\right)\right]^{-1}\right.$$
$$\left[2\mu s_i + V_2^i + \text{Diag}\left((PD)^T V_3^{(i)}\right)\right.$$
$$\left.\left. + 2\mu\text{Diag}\left((PD)^T Q_i\right)\right]\right\}_+$$
$$(29)$$

where $\mathcal{J}.(\cdot)$ represents the singular value soft-thresholding operator [64].

The update of variables and multipliers is alternately iterated until convergence. The overall algorithm for updating coefficient matrix A is summarized in Algorithm 2.

Algorithm 2 : Update $\boldsymbol{A}$ With $\boldsymbol{D}$ Fixed.

1: **Input:** LR-HSI $\boldsymbol{X}$; HR-MSI $\boldsymbol{Y}$; spectral basis $\boldsymbol{D}$; spatial degradation operator $\boldsymbol{H}$; spectral transform matrix $\boldsymbol{P}$; and regularization parameters η_1 and η_2.

2: **Initialization:** $\boldsymbol{A} = 0$; $\boldsymbol{Q} = 0$; $\boldsymbol{Z} = 0$; $\boldsymbol{V}_1 = 0$; $\boldsymbol{V}_2 = 0$; $\boldsymbol{V}_3 = 0$; $\boldsymbol{U} = 0$.

3: **While** not converge **do**

4: Update variables $\boldsymbol{S}$, $\boldsymbol{Z}$, $\boldsymbol{Q}$, and $\boldsymbol{A}$ by using (29).

5: Update Lagrangian multipliers $\boldsymbol{V}_1$, $\boldsymbol{V}_2$, and $\boldsymbol{V}_3$ by using (28).

6: Update μ: $\mu = \rho\mu \, (\rho > 1)$.

7: Update $\boldsymbol{U}$ by using (20).

8: **End while**

9: **Output:** coefficient matrix $\boldsymbol{A}$.

E. Optimization of the Spectral Basis With the Coefficients Fixed

In this procedure, we fix the coefficient matrix $\boldsymbol{A}$. Then, the updating of spectral basis $\boldsymbol{D}$ can be written as

$$\min_{\boldsymbol{D}} \|\boldsymbol{Y} - \boldsymbol{PDA}\|_F^2 + \|\boldsymbol{X} - \boldsymbol{DAH}\|_F^2, \text{ s.t. } 0 \leq \boldsymbol{D} \leq 1. \tag{30}$$

As the nonlocal spatial similarity prior is mainly reflected by the coefficient matrix $\boldsymbol{A}$, the constraint term $\eta_1 \|\boldsymbol{DA} - \boldsymbol{U}\|_F^2$ can be excluded. Similarly, the ADMM can also be used to solve problem (30). More specifically, we introduce $\boldsymbol{W} = \boldsymbol{D}$, and obtain the following augmented Lagrangian function:

$$L\left(\boldsymbol{D}, \boldsymbol{W}, \boldsymbol{V}_4\right) = \|\boldsymbol{Y} - \boldsymbol{PDA}\|_F^2 + \|\boldsymbol{X} - \boldsymbol{DAH}\|_F^2$$
$$+ \mu \left\|\boldsymbol{W} - \boldsymbol{D} + \frac{\boldsymbol{V}_4}{2\mu}\right\|_F^2$$
$$\text{s.t. } 0 \leq \boldsymbol{W} \leq 1 \tag{31}$$

where $\boldsymbol{V}_4$ is the Lagrangian multiplier. Minimizing the augmented Lagrangian function (31) leads to the following iterations:

$$\boldsymbol{D}^{(t+1)} = \arg\min_{\boldsymbol{D}} L\left(\boldsymbol{D}, \boldsymbol{W}^{(t)}, \boldsymbol{V}_4^{(t)}\right)$$
$$\boldsymbol{W}^{(t+1)} = \arg\min_{\boldsymbol{W}} L\left(\boldsymbol{D}^{(t)}, \boldsymbol{W}, \boldsymbol{V}_4^{(t)}\right). \tag{32}$$

Meanwhile, the Lagrangian multiplier is updated by

$$\boldsymbol{V}_4^{(t+1)} = \boldsymbol{V}_4^{(t)} + \mu\left(\boldsymbol{W}^{(t+1)} - \boldsymbol{D}^{(t+1)}\right). \tag{33}$$

The two subproblems in (32) can be easily solved analytically. For the updating of $\boldsymbol{D}$, with the auxiliary variable $\boldsymbol{W}$ and the Lagrangian multiplier $\boldsymbol{V}_4$ fixed, we can acquire the following equation:

$$\boldsymbol{D}^{(t+1)} \boldsymbol{H}_1 + \boldsymbol{H}_2 \, \boldsymbol{D}^{(t+1)} = \boldsymbol{H}_3 \tag{34}$$

where

$$\boldsymbol{H}_1 = \left[(\boldsymbol{AH})(\boldsymbol{AH})^T + \mu\boldsymbol{I}\right]\left(\boldsymbol{AA}^T\right)^{-1}$$
$$\boldsymbol{H}_2 = \boldsymbol{P}^T \boldsymbol{P}$$
$$\boldsymbol{H}_3 = \left[\boldsymbol{X}(\boldsymbol{AH})^T + \boldsymbol{P}^T\boldsymbol{Y}\boldsymbol{A}^T + \mu\left(\boldsymbol{W}^{(t)} + \frac{\boldsymbol{V}_4^{(t)}}{2\mu}\right)\right]\left(\boldsymbol{AA}^T\right)^{-1}. \tag{35}$$

Then, vectorizing $\boldsymbol{D}^{(t+1)}$ and $\boldsymbol{H}_3$ of (34), we can acquire the following equation:

$$\left(\boldsymbol{H}_1^T \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{H}_2\right) \text{vec}\left(\boldsymbol{D}^{(t+1)}\right) = \text{vec}\left(\boldsymbol{H}_3\right) \tag{36}$$

where $\otimes$ represents the Kronecker product, and $\text{vec}(\cdot)$ is the vectorization operation. Therefore, $\boldsymbol{D}^{(t+1)}$ can be computed as

$$\text{vec}\left(\boldsymbol{D}^{(t+1)}\right) = \left(\boldsymbol{H}_1^T \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{H}_2\right)^{-1} \text{vec}\left(\boldsymbol{H}_3\right). \tag{37}$$

For the updating of $\boldsymbol{W}$, the solution of $\boldsymbol{W}^{(t+1)}$ can be analytically obtained by

$$\boldsymbol{W}^{(t+1)} = \min\left(\max\left(\boldsymbol{D}^{(t+1)} - \frac{\boldsymbol{V}_4^{(t)}}{2\mu}, 0\right), 1\right). \tag{38}$$

The update of variables and multiplier is alternately iterated until convergence. The overall algorithm for updating spectral basis $\boldsymbol{D}$ is summarized in Algorithm 3.

Algorithm 3: Update $\boldsymbol{D}$ With $\boldsymbol{A}$ Fixed.

1: **Input:** LR-HSI $\boldsymbol{X}$; HR-MSI $\boldsymbol{Y}$; coefficient matrix $\boldsymbol{A}$; spatial degradation operator $\boldsymbol{H}$; and spectral transform matrix $\boldsymbol{P}$.

2: **Initialization:** $\boldsymbol{D} = 0$; $\boldsymbol{W} = 0$; and $\boldsymbol{V}_4 = 0$;

3: **While** not converge **do**

4: Update variables $\boldsymbol{D}$ and $\boldsymbol{W}$ by using (37) and (38), respectively.

5: Update Lagrangian multipliers $\boldsymbol{V}_4$ by using (33).

6: Update μ: $\mu = \rho\mu \, (\rho > 1)$.

7: **End while**

8: **Output:** spectral basis $\boldsymbol{D}$.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, to evaluate the performance of our proposed HSI super-resolution method, we conduct ample experiments on both ground-based HSIs datasets and real remote sensing HSI. To objectively evaluate the quality of the reconstructed HSIs, we adopt four objective evaluation indices, which are peak signal to noise ratio (PSNR), root-mean-square error (RMSE), relative dimensionless global error in synthesis (ERGAS), and spectral angle mapper (SAM), in our experiments.

A. Experimental Datasets

In our experiments, we use two categories of images to show the effectiveness of our method. For the ground-based HSIs, we use a public HSI dataset, which is named Columbia Computer Vision Laboratory (CAVE) [65]. The CAVE dataset includes 32 HSIs of everyday objects, which are captured by generalized
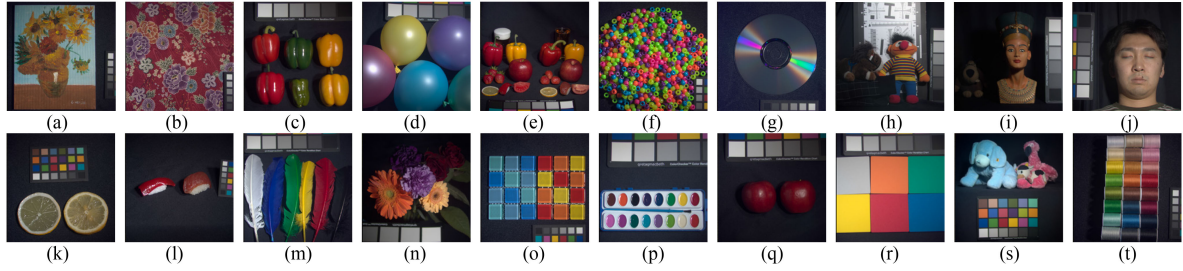
Fig. 4.   Total of 20 representative testing images from the CAVE datasets. (a) Oil_painting. (b) Cloth. (c) Fake_and_real_peppers. (d) Balloons. (e) Fake_and_real_food. (f) Beads. (g) CD. (h) Chart_and_stuffed_toy. (i) Egyptain_statue. (j) Face. (k) Fake_and_real_lemon_slices. (l) Fake_and_real_sushi. (m) Feathers. (n) Flowers. (o) Glass_tiles. (p) Paints. (q) Real_and_fake_apples. (r) Sponges. (s) Stuffed_toys. (t) Thread_spools.
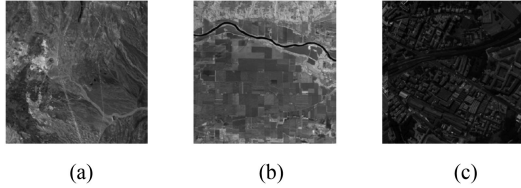

Fig. 5.   Three popular remote sensing HSIs. (a) Cuprite Mine Nevada. (b) Indian Pines. (c) Pavia Center.

TABLE I

AVERAGE RESULTS OF THE TEST METHODS FOR DIFFERENT SCALING FACTORS ON THE GROUND-BASED HSIs

Scale	Method	GSOMP+	SNNMF	CNMF	NSSR	Proposed ANSR
s=8	PSNR	32.8298	43.0446	39.9982	45.0708	**45.5406**
	RMSE	6.3952	2.0674	2.8039	1.5709	**1.4982**
	SAM	12.9536	3.4603	4.8674	2.9698	**2.7925**
	ERGAS	3.3588	1.0727	1.4279	0.8398	**0.8096**
s=16	PSNR	32.6574	40.5578	39.1698	43.3402	**43.7289**
	RMSE	6.3989	2.6583	3.1324	1.9442	**1.8639**
	SAM	12.7897	4.7690	5.3377	3.5393	**3.3019**
	ERGAS	1.6677	0.6971	0.8078	0.5151	**0.4996**
s=32	PSNR	32.9479	38.5262	36.4959	41.2033	**41.9801**
	RMSE	6.4140	3.3121	4.3493	2.5062	**2.2990**
	SAM	12.9236	6.3776	6.8718	4.5235	**4.0085**
	ERGAS	0.8456	0.4475	0.5779	0.3294	**0.3020**

TABLE II

PSNR, RMSE, SAM, AND ERGAS RESULTS OF THE TEST METHODS FOR DIFFERENT SCALING FACTORS ON CUPRITE MINE NEVADA

Scale	Method	GSOMP+	CNMF	NSSR	SSR	OTD	Proposed ANSR
s=8	PSNR	40.9012	40.0068	42.0377	42.3001	44.2961	**44.7354**
	RMSE	2.2987	2.5480	2.0168	1.9567	1.5077	**1.4780**
	SAM	1.0766	1.4100	0.8222	0.9028	0.7872	**0.7178**
	ERGAS	0.2720	0.3641	0.2330	0.2309	0.2059	**0.1745**
s=16	PSNR	25.3635	39.5162	41.8116	42.1340	43.3441	**44.2738**
	RMSE	13.7521	2.6961	2.0700	1.9945	1.6932	**1.5590**
	SAM	5.7583	1.4704	0.8460	0.9134	0.8220	**0.7483**
	ERGAS	0.8133	0.1913	0.1193	0.1177	0.1008	**0.0917**
s=32	PSNR	27.5796	39.3012	40.9364	41.9813	42.8931	**43.9323**
	RMSE	10.6552	2.7636	2.2894	2.0299	1.8661	**1.6215**
	SAM	4.6252	1.5211	0.9023	0.9152	0.8593	**0.7788**
	ERGAS	0.3374	0.0977	0.0658	0.0598	0.0502	**0.0476**

TABLE III

PSNR, RMSE, SAM, AND ERGAS RESULTS OF THE TEST METHODS FOR DIFFERENT SCALING FACTORS ON INDIAN PINES

Scale	Method	GSOMP+	CNMF	NSSR	SSR	OTD	Proposed ANSR
s=8	PSNR	55.2181	59.2645	57.0355	56.2136	61.0195	**61.5516**
	RMSE	0.4422	0.2766	0.3587	0.3943	0.2573	**0.2130**
	SAM	0.7063	0.7573	0.5293	0.5218	0.4007	**0.3718**
	ERGAS	0.7950	0.1899	0.1397	0.1641	0.1139	**0.0817**
s=16	PSNR	42.3580	58.4174	57.1458	55.3233	59.9324	**60.3075**
	RMSE	1.9437	0.3060	0.3542	0.4369	0.2977	**0.2461**
	SAM	3.2670	0.8246	0.5441	0.5742	0.4903	**0.4153**
	ERGAS	0.4658	0.1206	0.0713	0.0969	0.0510	**0.0476**
s=32	PSNR	39.8384	58.1921	54.8004	53.5612	56.7631	**59.5190**
	RMSE	2.5979	0.3140	0.4640	0.5352	0.4229	**0.2695**
	SAM	3.8740	0.8758	0.6780	0.7284	0.5726	**0.4540**
	ERGAS	0.3349	0.0624	0.0480	0.0616	0.0401	**0.0265**

assorted pixel camera with high quality. The spatial size of each HSI in CAVE is 512×512. And each HSI has 31 spectral bands ranging from 400 to 700 nm at an interval of 10 nm. Because some images in the CAVE dataset are similar, we select 20 representative HSIs of CAVE as our experimental data, which are shown in Fig. 4. The selected HSIs are served as ground truth images and used to generate the LR-HSIs and HR-MSIs.

For the real remote sensing HSI, we use three popular remote sensing HSIs: Cuprite Mine Nevada, Indian Pines, and Pavia Center, which are adopted in [51]. The three HSIs are shown in Fig. 5. The wavelength of the Cuprite Mine Nevada image ranges from 400 to 2500 nm at an interval of 10 nm and the spatial resolution of the Cuprite mine Nevada image is 20 m. We crop the top left region of size 512×512 as the ground truth after abandoning the bands with water absorptions and low SNR. The final size of the ground truth image in our experiment is $512 \times 512 \times 200$. The Indian Pines image is captured by the airborne visible and infrared imaging spectrometer over the northwestern

Indiana. The wavelength of Indian Pines ranges from 400 to 2500 nm with an interval of 10 nm. We crop the bottom right part of size 512×512 and remove the water absorption bands (104–108, 150–163, and 220). The final size of the ground truth image in our experiment is $512 \times 512 \times 200$. Pavia Center image is taken by the reflective optics system imaging spectrometer over the Center of Pavia area. The spectrum of Pavia Center ranges from 430 to 860 nm at an interval of 4 nm. After abandoning the nosiest bands, we crop the bottom right region with the size of $512 \times 512 \times 102$, which is used as the ground truth image in our experiment.
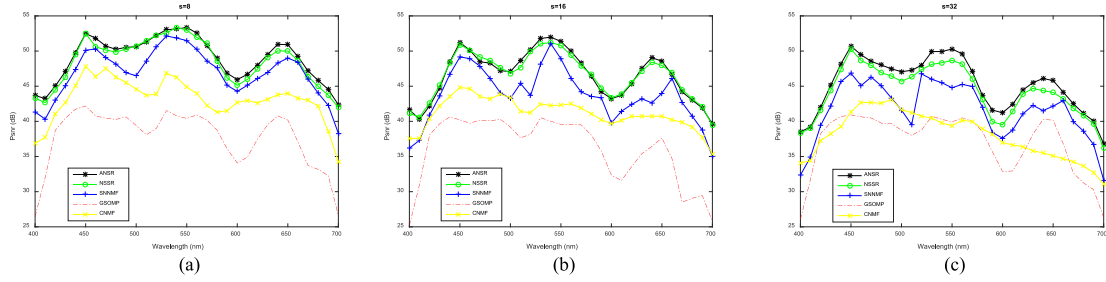
Fig. 6. PSNR curves of all the wavelengths of spectral bands over the testing image "fake_and_real_food." (a) PNSR curves with scaling factor $s = 8$. (b) PNSR curves with scaling factor $s = 16$. (c) PNSR curves with scaling factor $s = 32$.

TABLE IV
PSNR, RMSE, SAM, AND ERGAS RESULTS OF THE TEST METHODS FOR DIFFERENT SCALING FACTORS ON PAVIA CENTER

Scale	Method	GSOMP$_+$	CNMF	NSSR	SSR	OTD	Proposed ANSR
s=8	PSNR	38.6672	37.9549	35.4237	39.5669	39.6273	**40.7791**
	RMSE	2.9729	3.2269	4.3187	2.6804	2.6618	**2.3312**
	SAM	2.0827	3.7851	2.5027	1.8620	1.7339	**1.6992**
	ERGAS	0.6541	1.2273	0.8165	0.5860	0.5709	**0.4893**
s=16	PSNR	38.3515	37.8615	35.1459	37.0282	39.2168	**39.4593**
	RMSE	3.0829	3.2619	4.4591	3.5903	2.7906	**2.7138**
	SAM	2.2853	3.9106	2.6195	2.6088	1.9488	**1.8959**
	ERGAS	0.3479	0.6376	0.4184	0.3986	0.3006	**0.2744**
s=32	PSNR	31.4349	37.5873	34.6226	36.3915	37.1246	**38.5534**
	RMSE	6.8359	3.3664	4.7360	3.8634	3.5507	**3.0121**
	SAM	4.6033	4.4077	2.8697	2.9245	2.3945	**2.0906**
	ERGAS	0.3592	0.3311	0.2252	0.2124	0.1895	**0.1511**

B. Evaluation Indices

In this article, we use four indices to evaluate the reconstruction quality. The first index is PSNR, which is defined as the average PSNR value of all spectral bands. The formulation is as follows:

$$\text{PSNR}\left(\hat{Z}, Z\right) = \frac{1}{S} \sum_{i=1}^{S} \text{PSNR}\left(\hat{Z}_i, Z_i\right) \tag{39}$$

where Z_i and $\hat{Z}_i$ denote the ith band of the ground truth HSI Z and the estimated HSI $\hat{Z}$, respectively. S represents the number of spectral bands. PSNR index measures the similarities between the two images. The larger the PSNR, the better the reconstruction result.

The second index is RMSE, which is defined as the average RMSE of all spectral bands, i.e.,

$$\text{RMSE}\left(\hat{Z}, Z\right) = \frac{1}{S} \sum_{i=1}^{S} \text{RMSE}\left(\hat{Z}_i, Z_i\right). \tag{40}$$

The smaller the RMSE, the better the reconstruction result. The third index is ERGAS, whose formulation is

$$\text{ERGAS}\left(\hat{Z}, Z\right) = \frac{100}{c} \sqrt{\frac{1}{S} \sum_{i=1}^{S} \frac{\text{MSE}\left(\hat{Z}_i, Z_i\right)}{\mu_{\hat{Z}_i}^2}} \tag{41}$$

where c represents the spatial downsampling factor, and $\mu_{\hat{Z}_i}$ is the mean value of $\hat{Z}_i$. The smaller the ERGAS, the better the reconstruction result.

The fourth index is SAM, which is defined as

$$\text{SAM}\left(\hat{Z}, Z\right) = \frac{1}{N} \sum_{j=1}^{N} \cos^{-1} \frac{\hat{z}_j^T z_j}{\left\|\hat{z}_j\right\|_2 \left\|z_j\right\|_2} \tag{42}$$

where z_j and $\hat{z}_j$ denote the jth pixel of the ground truth HSI Z and the estimated HSI $\hat{Z}$, respectively. N represents the number of pixels. SAM measures the spectral quality of reconstructed HSI. The smaller the SAM, the better the reconstruction result.

C. Experimental Settings for the Comparison Methods

For the sake of fairness, we describe the experimental settings in this section. The LR-HSI X and the HR-MSI Y are obtained according to the same settings for all the comparison methods.

For the ground-based HSIs, as in [52], the ground truth HR-HSI Z is downsampled by averaging the disjoint $s \times s$ blocks to simulate the LR-HSI X, where s denotes the scaling factor ($s = 8, 16,$ and 32). The HR-MSI Y is generated directly downsampling the spectral dimension of Z using the spectral transform matrix P, which is derived from the response of a Nikon D700 camera. [1]

For the real remote sensing HSIs, as the operations in [51], the ground truth HR-HSI Z is downsampled s times ($s = 8, 16,$ and 32) to obtain the LR-HSI X. Specifically, each pixel $x_i \in X$ is generated by averaging pixels in a $s \times s$ window of HR-HSI Z centering on location i. For the HR-MSI Y, we directly select several bands from the ground truth HR-HSI Z. The Landsat7-like reflectance spectral response filter is used as the spectral transform matrix P. That is, for Cuprite Mine and Indian Pine images, the bands whose center wavelengths are 480, 560, 660, 830, 1650, and 2220 nm will be selected. For the Pavia Center image, we choose 480, 560, 660, and 830 nm (corresponding to the blue, green, red, and near-infrared channel, respectively) of the HR-HSI Z to simulate the HR-MSI Y.

D. Experimental Results of Our Method

In this section, we will show the experimental results of our HSI super-resolution method compared with some typical

---

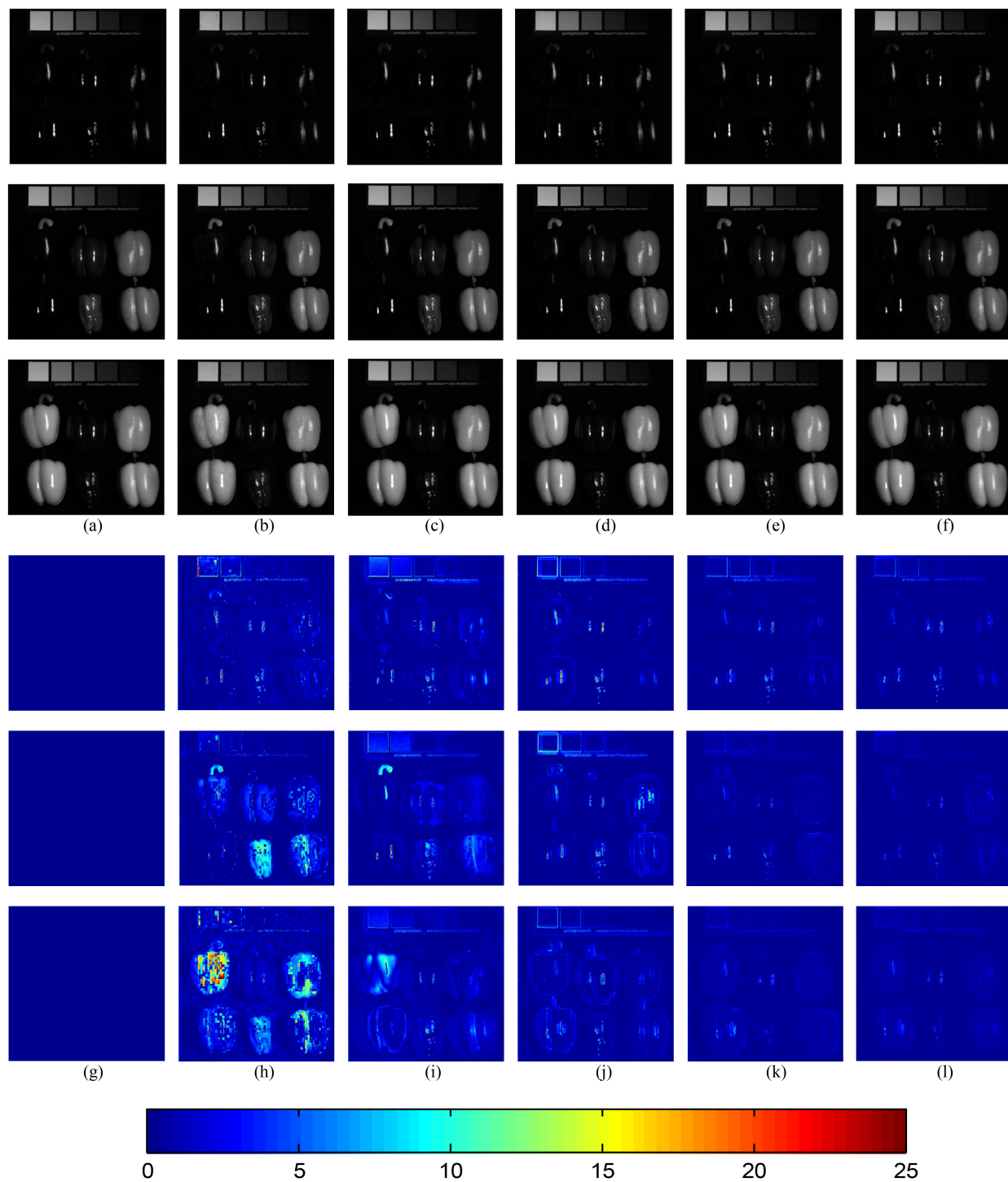[1]Online. [Available]: https://www.maxmax.com/spectral_respinse.htm

Fig. 7. Reconstructed results of image "fake_and_real_peppers" at 480, 550, and 640 nm with scaling factor $s = 16$. From top to bottom, the first three rows show the reconstructed images of different methods at 480, 550, and 640 nm, respectively; the last three rows show the errors of different methods at 480, 550, and 640 nm, respectively. (a) Original images. (b) Results of G-SOMP$_+$. (c) Results of CNMF. (d) Results of SNNMF. (e) Results of NSSR. (f) Results of our ANSR. (g) Errors of the original images. (h) Errors of G-SOMP$_+$ results. (i) Errors of CNMF results. (j) Errors of SNNMF results. (k) Errors of NSSR results. (l) Errors of ANSR results.

existing HSI super-resolution methods, including G-SOMP$_+$ [49], SNNMF [66], CNMF [43], NSSR [52], SSR [51], and OTD [48]. Note that we do not compare our method with SSR in the experiments of the ground-based HSIs. We will explain it in the experimental analysis of real remote sensing HSI.

Similarly, we do not compare our method with OTD in the experiments of the ground-based HSIs. It is because that Han *et al*. [48] only conducts experiments with real remote sensing HSIs. For the sake of fairness, we only compared our method with OTD on the real remote sensing HSIs. In recent years,
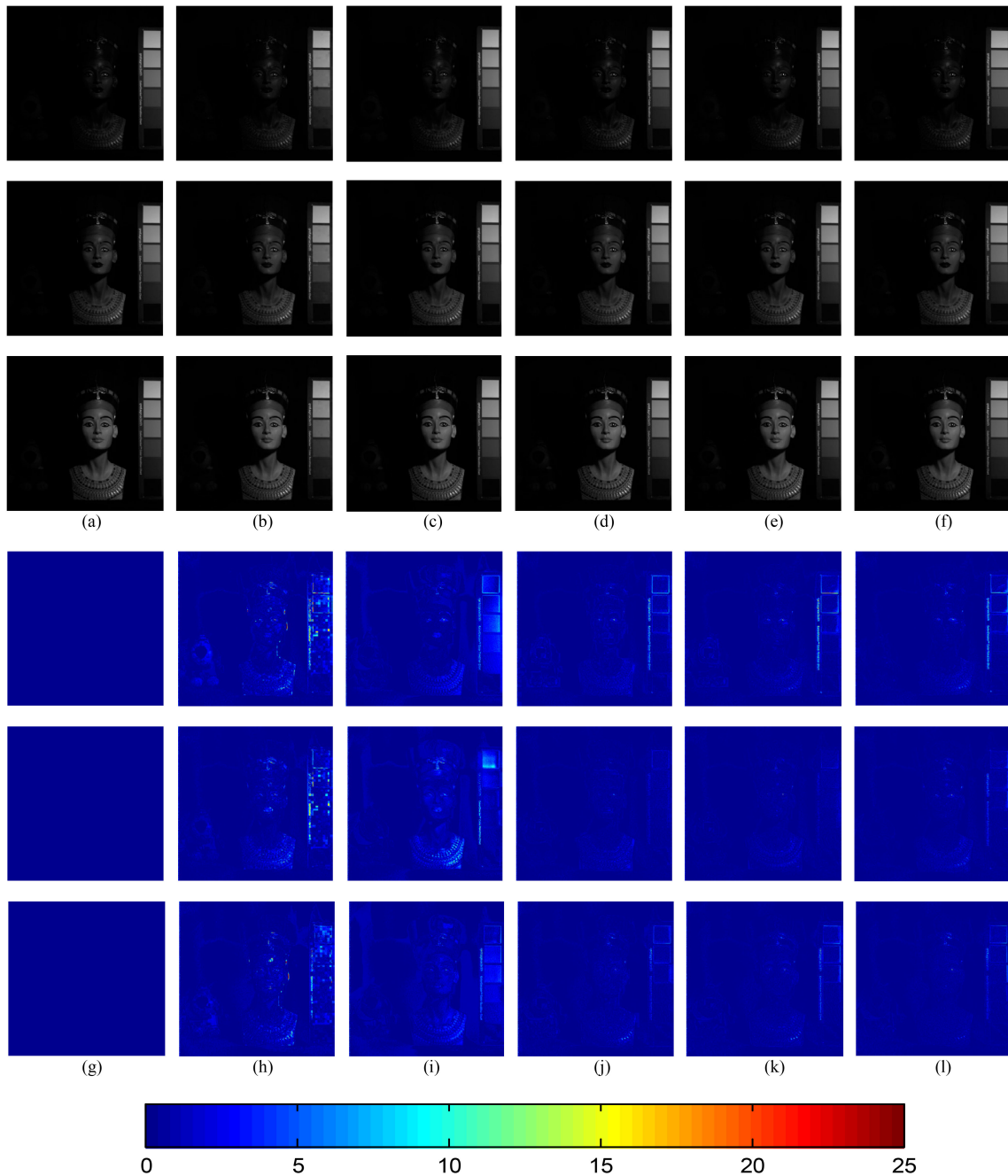
Fig. 8. Reconstructed results of image "Egyptian_statue" at 480, 550, and 640 nm with scaling factor $s = 16$. From top to bottom, the first three rows show the reconstructed images of different methods at 480, 550, and 640 nm, respectively; the last three rows show the errors of different methods at 480 nm, 550 nm, and 640 nm, respectively. (a) Original images. (b) Results of G-SOMP$_+$. (c) Results of CNMF. (d) Results of SNNMF. (e) Results of NSSR. (f) Results of our ANSR. (g) Errors of the original images. (h) Errors of G-SOMP$_+$ results. (i) Errors of CNMF results. (j) Errors of SNNMF results. (k) Errors of NSSR results. (l) Errors of ANSR results.

deep learning based HSI super-resolution methods have shown good results, but this kind of method needs a large number of training samples. Considering that our method do not need any training sample, we do not compare our approach with the deep learning based methods. In our experiments, we assume that the spatial degradation operator H and the spectral transform matrix P are known. In the real-world situation, both the spatial degradation operator H and the spectral transform matrix P can be estimated from the LR-HSI and the HR-MSI. Some major

parameters include regularization parameters η_1 and η_2, and the number of atoms of spectral basis K. The selection of these parameters will be discussed in Section IV-E.

For the ground-based HSIs, the comparison of results is given in Table I. The best results are bolded for clarity. As presented in Table I, our ANSR method achieves the best results among all the compared methods, and the NSSR is the second-best method, although it was proposed in 2016. According to Table I, the large average PSNR gains of our method over the
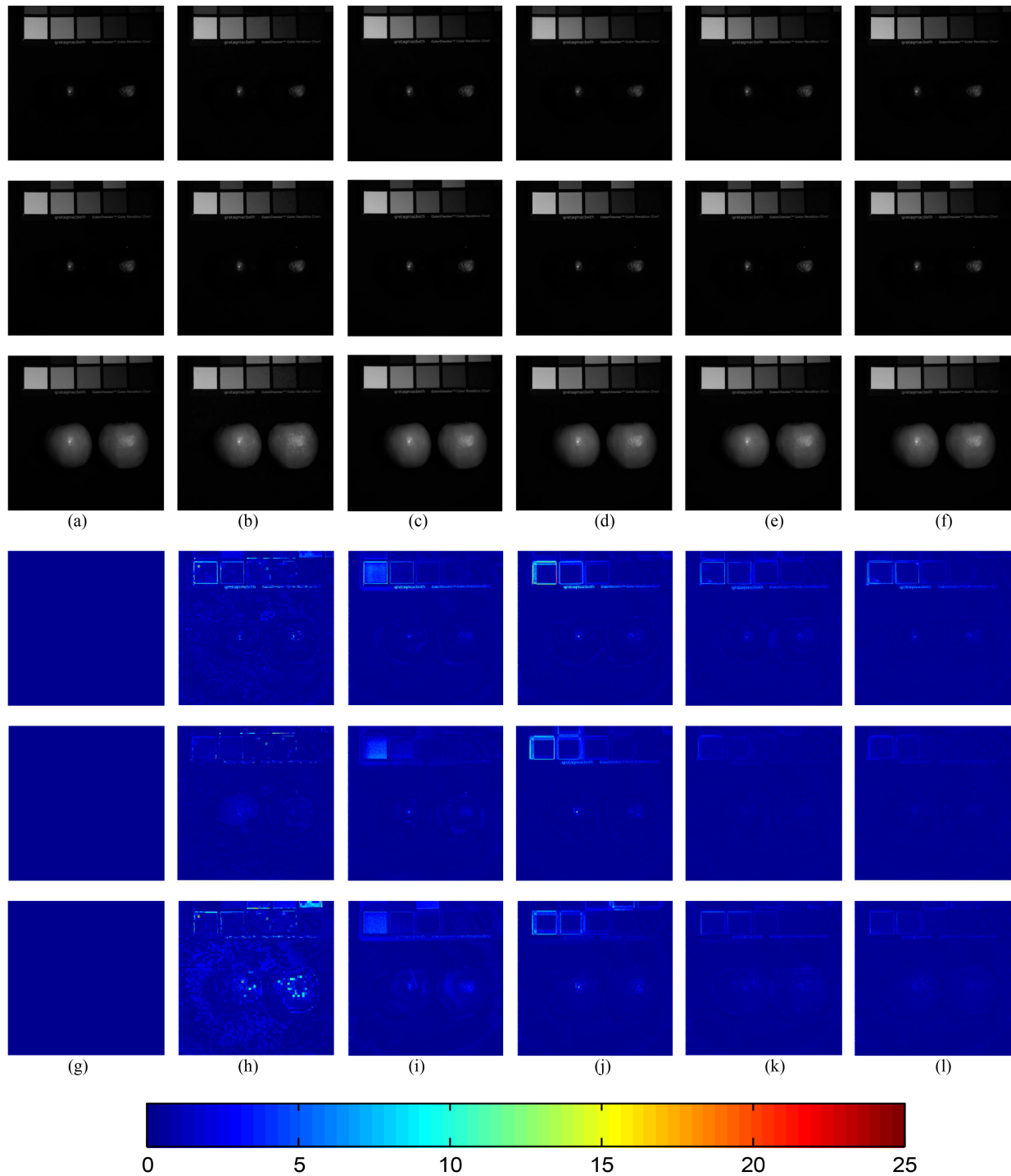
Fig. 9. Reconstructed results of image "real_and_fake_apples" at 480, 550, and 640 nm with scaling factor $s = 16$. From top to bottom, the first three rows show the reconstructed images of different methods at 480, 550, and 640 nm, respectively; the last three rows show the errors of different methods at 480, 550, and 640 nm, respectively. (a) Original images. (b) Results of G-SOMP$_+$. (c) Results of CNMF. (d) Results of SNNMF. (e) Results of NSSR. (f) Results of our ANSR. (g) Errors of the original images. (h) Errors of G-SOMP$_+$ results. (i) Errors of CNMF results. (j) Errors of SNNMF results. (k) Errors of NSSR results. (l) Errors of ANSR results.

second-best method with $s = 8$, $s = 16$, and $s = 32$ are 0.4698, 0.3887, and 0.7768 dB, respectively. The effects of G-SOMP$_+$ are much worse than the other compared methods. The reason is that the G-SOMP$_+$ method does not make use of the prior knowledge of the spatial degradation operator H, which is usually unknown and needs to be estimated in practical applications. Fig. 6 shows the PSNR curves of the wavelengths of the spectral bands over the testing image "fake_and_real_food" in the CAVE dataset for the compared methods. It can be seen from Fig. 6 that the proposed ANSR
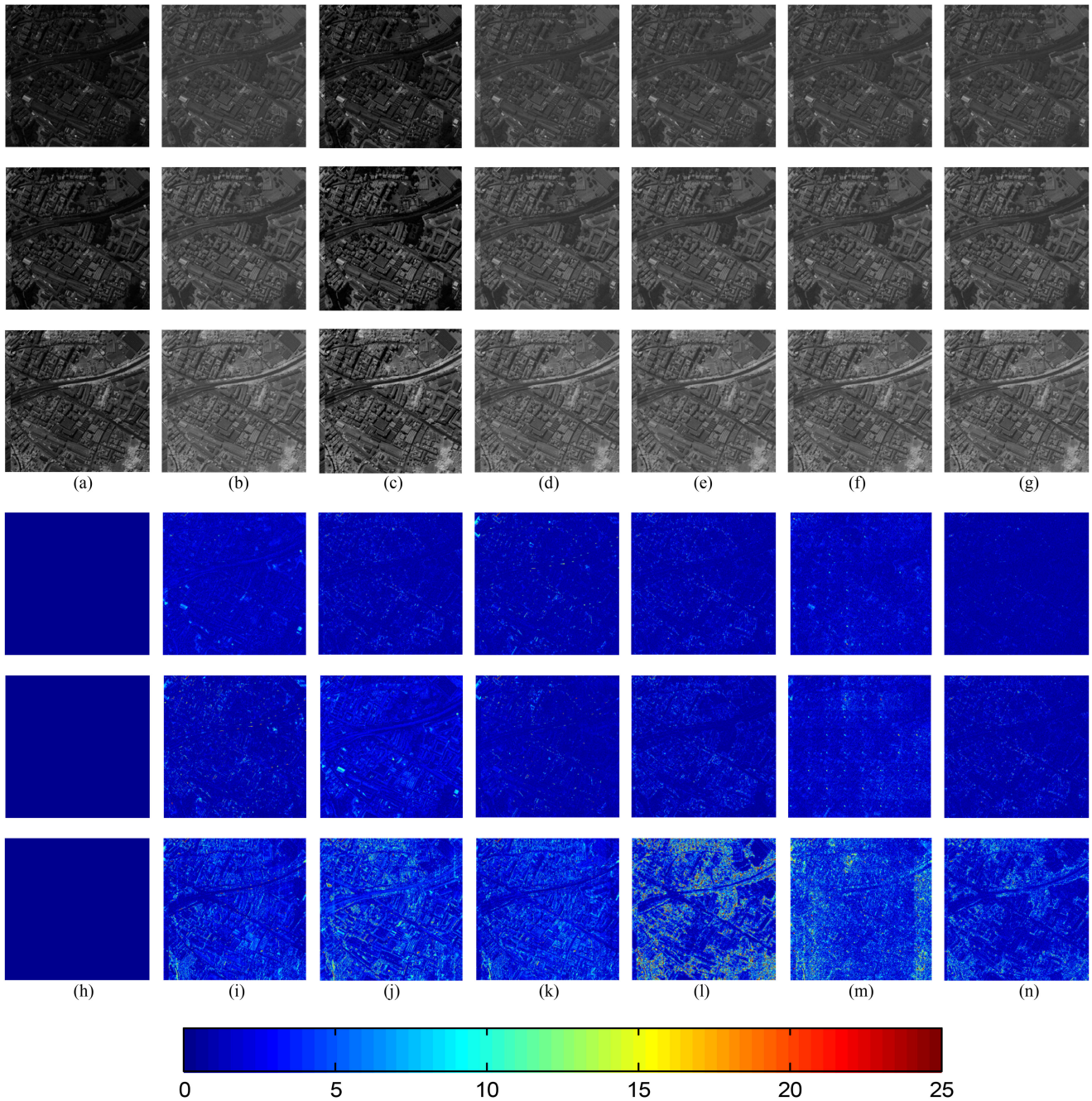
Fig. 10. Reconstructed results of image "Pavia Center" at the 25th, 55th, and 85th bands with scaling factor $s = 8$. From top to bottom, the first three rows show the reconstructed images of different methods at 25th, 55th, and 85th bands, respectively; the last three rows show the errors of different methods at 25th, 55th, and 85th bands, respectively. (a) Original images. (b) Results of G-SOMP$_+$. (c) Results of CNMF. (d) Results of SSR. (e) Results of NSSR. (f) Results of OTD. (g) Results of our ANSR. (h) Errors of the original images. (i) Errors of G-SOMP$_+$ results. (j) Errors of CNMF results. (k) Errors of SSR results. (l) Errors of NSSR results. (m) Errors of OTD. (n) Errors of ANSR results.

method outperforms other compared methods at each wavelength for all the scaling factors. Meanwhile, in order to further demonstrate the effect of our method, some comparisons of the vision effects can be seen in Figs. 7–9. Figs. 7–9 show the reconstructed HR-HSI at different wavelengths of test images "fake_and_real_peppers," "egyptian_statue," and "real_and_fake_apples," respectively. It can be seen from

Figs. 7–9 that our ANSR method has the best visual results and achieves the minimum reconstruction errors.

For the real remote sensing HSI, the comparison of results is given in Tables II–IV. The best results are bolded for clarity. In this part, we compare our ANSR method with SSR and OTD, which are not compared in the experiments of the ground-based HSIs. It is because that the SSR method needs to cluster the
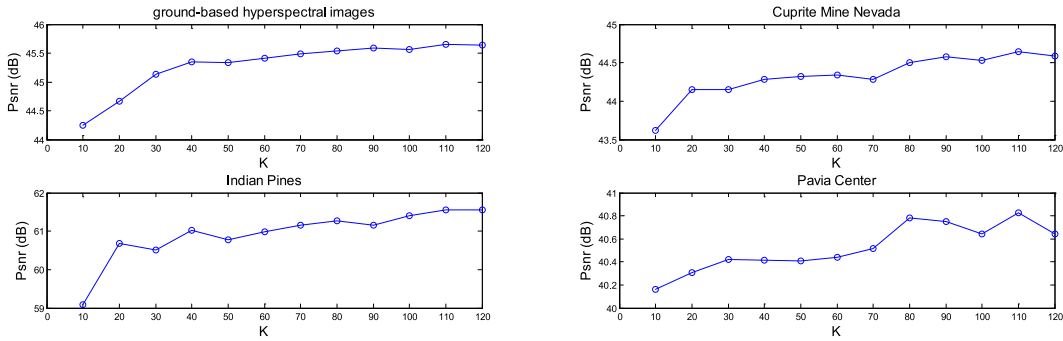
Fig. 11.   PSNR curves of the ground-based HSIs, Cuprite Mine Nevada, Indian Pines, and Pavia Center as functions of the number of atoms K.
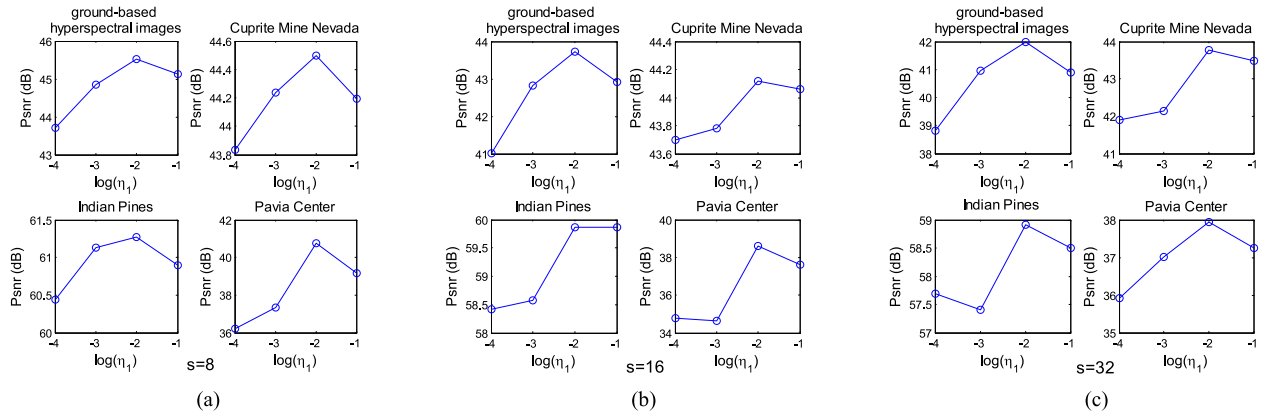

Fig. 12.   PSNRs of the ground-based HSIs, Cuprite Mine Nevada, Indian Pines, and Pavia Center as functions of $\log \eta_1$ with different scaling factors. (a) PNSR curves with scaling factor $s = 8$. (b) PNSR curves with scaling factor $s = 16$. (c) PNSR curves with scaling factor $s = 32$.
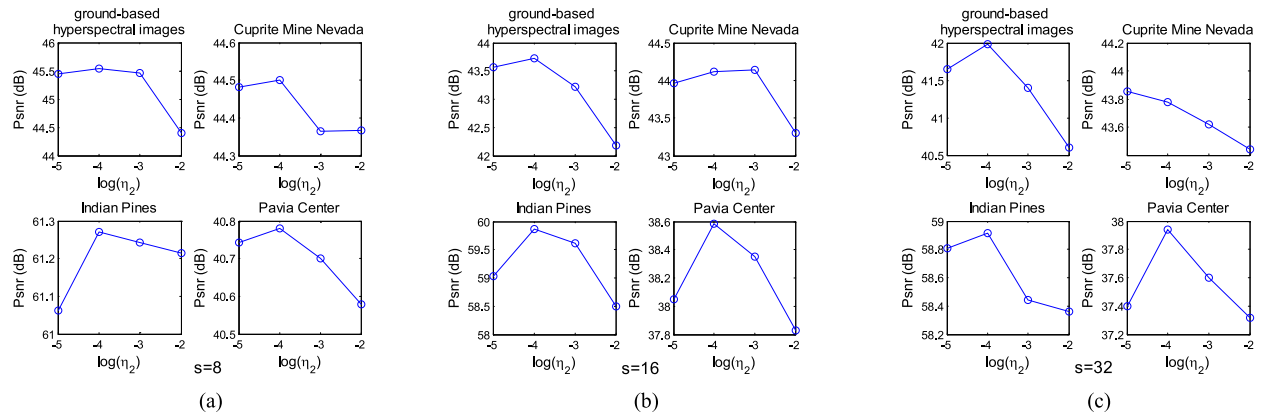

Fig. 13.   PSNRs of the ground-based HSIs, Cuprite Mine Nevada, Indian Pines, and Pavia Center as functions of $\log \eta_2$ with different scaling factors. (a) PNSR curves with scaling factor $s = 8$. (b) PNSR curves with scaling factor $s = 16$. (c) PNSR curves with scaling factor $s = 32$.

HR-MSI into superpixels, whose size and shape are adaptively adjusted according to the local structures. However, the images in the CAVE dataset are always very simple, and there is not much information in them. Therefore, the SSR method will fall into an endless loop because some superpixels only contain invalid information, when the number of superpixels is too large (e.g., 6000, which is used in real remote sensing HSI). But if we reduce the number of superpixels, the results of the SSR method become very poor because it loses its advantage. As presented in Tables II–IV, our ANSR method performs best among all the compared methods and the OTD is the second-best method. According to Table II, the large PSNR gains of our method over the second-best method on the "Cuprite Mine Nevada" with $s = 8$, $s = 16$, and $s = 32$ are 0.4393 dB, 0.9297 dB, and

1.0392 dB, respectively. According to Table III, the large PSNR gains of our method over the second-best method on the "Indian Pines" with $s = 8$, $s = 16$, and $s = 32$ are 0.5321 dB, 0.3751 dB, and 2.7559 dB, respectively. According to Table IV, the large PSNR gains of our method over the second-best method on the "Pavia Center" with $s = 8$, $s = 16$, and $s = 32$ are 1.1518 dB, 0.2425 dB, and 1.4288 dB, respectively. Also, we can find that the SSR method performs better than the NSSR method on "Cuprite Mine Nevada" and "Pavia Center," but the opposite is true on the "Indian Pines." It is because that the "Indian Pines" image is relatively simple and contains less surface features information. In Fig. 10, we show the reconstructed HR-HSI at different wavelengths of test images "Pavia Center." It can be seen from Fig. 10 that the proposed ANSR method has the best visual results and achieves the minimum reconstruction errors.

According to the above experimental discussion, we can find that the larger the scaling factor, the more obvious the advantage of our method. In our experiments, the PSNR values of our method have the most obvious advantage over the second-best method when scaling factor $s = 32$.

E. Parameters Selection in Our Method

In our HSI super-resolution method, there are three crucial parameters, i.e., regularization parameters η_1 and η_2, and the number of atoms of spectral basis K. To evaluate the sensitivity of the three parameters, we conduct plenty of experiments for different values of them.

First, we all know that the number of atoms of the basis is very important in a sparse coding-based method. We perform some experiments with different values of K. Fig. 11 plots the PSNR curves of the ground-based HSIs, Cuprite Mine Nevada, Indian Pines, and Pavia Center as functions of the number of atoms K. It can be seen from Fig. 11 that the four PSNR curves increase with slight fluctuation when the number of atoms K becomes larger from 10 to 120. But the change curves of the ground-based HSIs, Cuprite Mine Nevada, and Indian Pines tend to be flat when K is larger than 50. The change curve of Pavia Center tends to be flat when K is larger than 80. As the computational burden increases sharply with the increase of K, we decide to set $K = 80$.

Second, we test and verify the effect of η_1 on the reconstructed results. We make $\log \eta_1$ (log is base 10) range from -4 to -1. Fig. 12 plots the PSNRs of the ground-based HSIs, Cuprite Mine Nevada, Indian Pines, and Pavia Center as functions of $\log \eta_1$. We can see from Fig. 12 that all the PSNR curves increase as $\log \eta_1$ increases from -4 to -2, and then they decrease as $\log \eta_1$ increases. Therefore, we choose 1×10^{-2} as the optimal value of η_1.

Third, we test and verify the effect of η_2 on the reconstructed results. We make $\log \eta_2$ (log is base 10) range from -5 to -2. Fig. 13 plots the PSNRs of the ground-based HSIs, Cuprite Mine Nevada, Indian Pines, and Pavia Center as functions of $\log \eta_2$. As can be seen from Fig. 13 that as the $\log \eta_2$ increases from -5 to -4, all the PSNR values (except for the PSNR curve of Cuprite Mine Nevada with scaling factor $s = 32$, which always decreases as $\log \eta_2$ increases from -5 to -2) increase.

Then, the PSNRs decrease as $\log \eta_2$ increases. Thus, we set $\eta_2 = 1 \times 10^{-4}$ in our experiment.

V. CONCLUSION

In this article, we presented a novel sparse representation-based HSI super-resolution method, termed ANSR, to fuse an LR-HSI and its corresponding HR-MSI. In the base of the NSSR model, we introduce the ASR, which can balance the relationship between the sparsity and collaboration by generating a suitable coefficient, into our ASNR method. Also, we design an alternative optimization algorithm to optimize the spectral basis rather than keeping it fixed. ADMM method is applied to solve the proposed optimization problem. In order to show the performance of the proposed method, we conduct plenty of experiments. The experimental results on both ground-based HSI dataset and real remote sensing HSIs show the superiority of our proposed approach to some other state-of-the-art HSI super-resolution methods.

In further work, we aim to improve the method in several directions. We will focus on the estimation of the spatial degradation operator H and the spectral transform matrix P, and conduct experiments of blind fusion, which is more common in the real-world situation. Then, we will optimize our solution process to improve the computational efficiency.

REFERENCES

[1] Y. Yuan, X. Zheng, and X. Lu, "Discovering diverse subset for unsupervised hyperspectral band selection," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 51–64, Jan. 2017.

[2] L. Gao, Q. Du, B. Zhang, W. Yang, and Y. Wu, "A comparative study on linear regression-based noise estimation for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 488–498, Apr. 2013.

[3] W. Sun and Q. Du, "Graph-regularized fast and robust principal component analysis for hyperspectral band selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3185–3195, Jun. 2018.

[4] Y. Zhang, B. Du, L. Zhang, and T. Liu, "Joint sparse representation and multitask learning for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 894–906, Feb. 2017.

[5] D. Lorente, N. Aleixos, J. Gómez-Sanchis, S. Cubero, O. L. García-Navarrete, and J. Blasco, "Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment," *Food Bioprocess Technol.*, vol. 5, no. 4. pp. 1121–1142, 2012.

[6] J. Yang, Y.-Q. Zhao, J. C.-W. Chan, and S. G. Kong, "Coupled sparse denoising and unmixing with low-rank constraint for hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1818–1833, Mar. 2016.

[7] F. D. van der Meer *et al.*, "Multi- and hyperspectral geologic remote sensing: A review," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 14, no. 1, pp. 112–128, Feb. 2012.

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

[9] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.

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

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

[12] S. Li, J. T. Kwok, and Y. Wang, "Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images," *Inf. Fusion*, vol. 3, no. 1, pp. 17–23, 2002.

[13] F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, "Remote sensing image fusion using the curvelet transform," *Inf. Fusion*, vol. 8, no. 2, pp. 143–156, 2007.

[14] S. Yang, M. Wang, and L. Jiao, "Fusion of multispectral and panchromatic images based on support value transform and adaptive principal component analysis," *Inf. Fusion*, vol. 13, no. 3, pp. 177–184, 2012.

[15] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, Feb. 2011.

[16] W. Wang, L. Jiao, and S. Yang, "Fusion of multispectral and panchromatic images via sparse representation and local autoregressive model," *Inf. Fusion*, vol. 20, no. 1, pp. 73–87, 2014.

[17] S. Yang, K. Zhang, and M. Wang, "Learning low-rank decomposition for pan-sharpening with spatial-spectral offsets," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3647–3657, Aug. 2018.

[18] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, "The use of intensity hue saturation transformations for merging SPOT panchromatic and multispectral image data," *Photogramm. Eng. Remote Sens.*, vol. 56, no. 4, pp. 459–467, 1990.

[19] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, 2001.

[20] V. P. Shah, N. H. Younan, and R. L. King, "An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1323–1335, May 2008.

[21] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Bayesian fusion of multi-band images," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1117–1127, Sep. 2015.

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

[23] 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.

[24] Q. Wei, J. M. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Fusion of multispectral and hyperspectral images based on sparse representation," in *Proc. 22nd Eur. Signal Image Process. Conf.*, 2014, pp. 1577–1581.

[25] M. Simoes, 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.

[26] Y. Zhang, S. De Backer, and P. Scheunders, "Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3834–3843, Nov. 2009.

[27] Y. Zhang, A. Duijster, and P. Scheunders, "A Bayesian restoration approach for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3453–3462, Sep. 2012.

[28] J. Han, D. Zhang, G. Cheng, N. Liu, and D. Xu, "Advanced deep-learning techniques for salient and category-specific object detection: A survey," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 84–100, Jan. 2018.

[29] N. Liu and J. Han, "A deep spatial contextual long-term recurrent convolutional network for saliency detection," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3264–3274, Jul. 2018.

[30] G. Cheng, P. Zhou, and J. Han, "Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7405–7415, Dec. 2016.

[31] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, "When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2811–2821, May 2018.

[32] H. Wu and S. Prasad, "Semi-supervised deep learning using pseudo labels for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1259–1270, Mar. 2018.

[33] X. Lu, X. Zheng, and Y. Yuan, "Remote sensing scene classification by unsupervised representation learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5148–5157, Sep. 2017.

[34] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.

[35] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Fast and accurate image super-resolution with deep Laplacian pyramid networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 11, pp. 2599–2613, Nov. 2019.

[36] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 105–114.

[37] Y. Li, J. Hu, X. Zhao, W. Xie, and J. Li, "Hyperspectral image super-resolution using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 29–41, 2017.

[38] Q. Yuan, Y. Wei, X. Meng, H. Shen, and L. Zhang, "A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 978–989, Mar. 2018.

[39] J. Yang, Y. Zhao, and J. C.-W. Chan, "Hyperspectral and multispectral image fusion via deep two-branches convolutional neural network," *Remote Sens.*, vol. 10, no. 5, 2018, Art. no. 800.

[40] 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.

[41] 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.

[42] X. Lu, Y. Yuan, and X. Zheng, "Joint dictionary learning for multispectral change detection," *IEEE Trans. Cybern.*, vol. 47, no. 4, pp. 884–897, Apr. 2017.

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

[44] 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.

[45] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[46] 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.

[47] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 3586–3594.

[48] X. Han, J. Yu, J.-H. Xue, and W. Sun, "Hyperspectral and multispectral image fusion using optimized twin dictionaries," *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 4709–4720, Feb. 2020.

[49] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *European Conference on Computer Vision*, vol. 8695, no. 7. New York, NY, USA: Springer, 2014, pp. 63–78.

[50] N. Akhtar, F. Shafait, and A. Mian, "Bayesian sparse representation for hyperspectral image super resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3631–3640.

[51] L. Fang, H. Zhuo, and S. Li, "Super-resolution of hyperspectral image via superpixel-based sparse representation," *Neurocomputing*, vol. 273, pp. 171–177, 2018.

[52] 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.

[53] É. Grave, G. Obozinski, and F. Bach, "Trace Lasso: A trace norm regularization for correlated designs," in *Proc. 24th Int. Conf. Neural Inf. Process. Syst.*, 2011, pp. 2187–2195.

[54] X. Li, G. Cao, Y. Zhang, and B. Wang, "Single image super-resolution via adaptive sparse representation and low-rank constraint," *J. Vis. Commun. Image Represent.*, vol. 55, pp. 319–330, 2018.

[55] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[56] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[57] R. Tibshirani, "Regression shrinkage and selection via the lasso: A retrospective," *J. Roy. Statist. Soc., B, Statist. Methodol.*, vol. 73, no. 3, pp. 273–282, 2011.

[58] R. Timofte, V. De, and L. van Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 1920–1927.

[59] J. Zhao, H. Hu, and F. Cao, "Image super-resolution via adaptive sparse representation," *Knowl.-Based Syst.*, vol. 124, pp. 23–33, 2017.

[60] R. Angst, C. Zach, and M. Pollefeys, "The generalized trace-norm and its application to structure-from-motion problems," in *Proc. Int. Conf. Comput. Vis.*, 2011, pp. 2502–2509.

[61] J. Wang, C. Lu, M. Wang, P. Li, S. Yan, and X. Hu, "Robust face recognition via adaptive sparse representation," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2368–2378, Dec. 2014.

[62] Z. Lin, M. Chen, Y. Ma, and L. Wu, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices,", 2009 *arXiv:1009.5055* .

[63] J. Yang and X. Yuan, "Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization," *Math. Comput.*, vol. 82, no. 281, pp. 301–329, 2013.

[64] T. Zhang, B. Ghanem, S. Liu, C. Xu, and N. Ahuja, "Low-rank sparse coding for image classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 281–288.

[65] 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.

[66] 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.*, 2013, pp. 1409–1413.

Xuesong Li received the B.S. degree in computer science and technology in 2014 from the Nanjing University of Science and Technology (NUST), Nanjing, China, where he is currently working toward the Ph.D. degree in computer science and technology.

His research interests include image processing and sparse representation.

Youqiang Zhang (Member, IEEE) received the Ph.D. degree in computer science from the Nanjing University of Science and Technology, Nanjing, China, in 2020.

He joined the School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing. His research interests include machine learning, remote sensing image processing, and rough sets.

Dr. Zhang is currently an active Reviewer of several journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, IEEE ACCESS, PATTERN RECOGNITION, and the KNOWLEDGE-BASED SYSTEMS.

Zixian Ge received the M.S. degree in electronics and communications engineering from the Lanzhou University of Technology, Lanzhou, China, in 2019. He is currently working toward the Ph.D. degree in computer science and technology from the Nanjing University of Science and Technology, Nanjing, China.

His research interests include deep learning and remote sensing image processing.

Guo Cao received the Ph.D. degree in pattern recognition and intelligence system from Shanghai Jiao Tong University, Shanghai, China, in 2006.

Since 2007, he has been with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China, where he is currently a Full Professor. From January 2012 to 2013, he was a Visiting Scholar with the Department of Radiology, University of Chicago, Chicago, IL, USA, and from August 2016 to 2017, he was a Visiting Scholar with the Department of Geography, University of Florida, Gainesville, FL, USA, where he focused on change detection. His research interests include machine learning, remote sensing image processing, and biometrics.

Hao Shi received the B.S. degree in computer science and technology in 2019 from the Nanjing University of Science and Technology (NUST), Nanjing, China, where he is currently working toward the Ph.D. degree in computer science and technology.

His research interests include deep learning and remote sensing image processing.

Peng Fu received the B.E. degree in computer science and technology and the Ph.D. degree in pattern recognition and intelligence system from the Nanjing University of Science and Technology (NUST), Nanjing, China, in 2009 and 2015, respectively.

He is currently a Lecturer and a Postdoctoral Fellow with the School of Computer Science and Engineering, NUST. His research interests include image processing, noise estimation, and pattern recognition.