Hyperspectral Superresolution Reconstruction via Decomposition of Low-Rank and Sparse Tensor

Huajing Wu, Kefei Zhang, Suqin Wu, Minghao Zhang, and Shuangshuang Shi

Abstract—Hyperspectral superresolution reconstruction technique obtains a high-resolution hyperspectral image (HR-HSI) by fusing both a low-resolution hyperspectral image (LR-HSI) and a high-resolution multispectral image. Existing methods of hyperspectral superresolution reconstruction are mostly concentrated on the global low-rank property while spatial and spectral information in individual regions is not considered. To address this issue, the decomposition of low-rank and sparse tensor proposed in this study. First, HR-HSI was decomposed into low-rank and sparse components. The former was further separated into spatial and spectral domains according to the spectral low-rank property, and the latter was used to compensate for information loss caused by the low-rank property. Then, a nonlocal constraint of adaptive manifold extracting structural details by the manifold structure was designed to enforce nonlocal self-similarity of the spatial domain. In order to ensure the same spatial structure of different bands and reduce the false individual regions in the sparse component, a surface-aware regularization combined with group sparsity was utilized. Finally, HR-HSI was constructed by the alternating direction method of multipliers. Experiment results on three datasets show that the proposed method outperforms five common existing methods by means of both visual and quantitative evaluations. It is concluded that the new method by taking into account the low-rank and sparse properties can improve the result of the reconstruction.

Index Terms—Adaptive manifold, hyperspectral superresolution reconstruction, nonlocal self-similarity.

I. INTRODUCTION

HYPERSPECTRAL imaging is a new means of observations and it can capture multiple narrow-band spectral images from the same scene [1]. With the increase in the spectral bands, the energy captured by each spectral band decreases, resulting in a reduction in the spatial resolution of the images from these spectral bands. The low spatial resolution of hyperspectral images (HSIs) makes it difficult to meet the needs of applications, including the accuracy of classification [2], [3], detection of changes [4], [5], and detection of objects [6], [7].

In order to improve the spatial resolution of the HSIs, the hyperspectral superresolution reconstruction has recently been introduced. Low-resolution HSIs (LR-HSIs) and high-resolution multispectral images (HR-MSIs) have a low spatial but high spectral resolution and high spatial but low spectral resolution, respectively. High-resolution HSIs (HR-HSIs) with high spatial and high spectral resolutions can be obtained by fusing LR-HSIs and HR-MSIs [8] in the hyperspectral superresolution reconstruction.

Pan-sharpening, as an early technique for a hyperspectral superresolution reconstruction, can be divided into four methods: 1) component substitution (CS), 2) multiresolution analysis (MRA), 3) Bayesian, and 4) variational methods. The CS method replaces hyperspectral spatial components from an LR-HSI with multispectral spatial components from an HR-MSI, and these components are obtained by specific algorithms such as intensity hue saturation (IHS) [9], [10], principal component analysis (PCA) [11], [12], [13], and Gram Schmidt (GS) spectral sharpening [14]. Different from the CS, the MRA saves all the contents of the LR-HSI and adds more information obtained from the HR-MSI through spatial frequency filtering. The MRA extracts the spatial details of the HR-MSI through wavelet transformation [15], [16], Laplacian pyramid [17], [18], and nonseparable transforms [19]. Then, these details are injected into the LR-HSI based on the pixel, window, and region of the LR-HSI. However, the CS and MRA methods only focus on spatial details, which often causes serious spectral distortion of the obtained HR-HSI. To maintain spectral information, Irmak and Wei regarded pan-sharpening as an appropriate method for reconstructing HR-HSIs and established a spectral image degradation model [20], [21]. Based on the spectral image degradation model, the Bayesian method [20], [21], [22] is under the assumption that the image noise follows a Gaussian distribution. The posterior distribution of the Bayesian model is used to simulate the hyperspectral superresolution reconstruction. Due to the incompleteness of spatial-spectral information of the observation images (the LR-HSI and HR-MSI), the reconstruction is regarded as an underdetermined inverse problem, and the potential HR-HSI requires a-priori constraints. The variational method resolves the illness of the hyperspectral superresolution reconstruction by adding total variational regularities [23], [24], [25] into the Bayesian method. Although the above methods can improve spatial resolution, the spectral correlation is significantly weakened, resulting in a spectral distortion of the reconstructed HR-HSI. The spectral unmixing, as a technique addressing spectral correlation, takes...
the spectra of mixed pixels as a linear combination of spectra from pure pixels [26]. Further research [27], [28], [29], [30] decomposed the HR-HSI into spectral basis and coefficients by using the spatial superresolution and spectral unmixing. Finally, the results of the spatial superresolution and spectral unmixing were obtained with interactive feedback. If only spectral correlation is used to decompose an unknown HR-HSI, it is difficult to obtain information of the spatial and spectral structure. To further explore this information, the factorization method has been investigated for reconstructing an HR-HSI. The unknown HR-HSI is regarded as a matrix, which can be factored into low-dimensional subspace and coefficients. The research [31], [32] used subspace-based regularization as the a-priori constraints to alternately update the low-dimensional subspace and coefficients. Inspired by the matrix factorization, Dian et al. [33] proposed a tensor factorization method. The unknown HR-HSI was regarded as a 3-dimensional tensor, which was decomposed into four factors using Tucker decomposition: a core tensor multiplied (n-mode products) by dictionaries (factor matrices) of three matrices along each mode. The research [34], [35], [36] utilized coupled sparse, nonlinear, and low-rank tensors to obtain the above four factors, respectively. To utilize the self-similarity of the unknown HR-HSI effectively, a-priori low-rank constraint was used to reconstruct the hyperspectral superresolution [37], [38], [39]. The research [40] designed structured sparse low-rank representation to extract spatial and spectral representation coefficients, constraining the unknown HR-HSI.

As deep learning rises, research [41], [42] used a convolutional neural network to construct deep prior for the reconstruction of HR-HSIs. Dian et al. [43] proposed a deep HSI sharpening method, which directly learned the image priors via deep convolutional neural network-based residual learning. Well-trained CNN denoiser is added as a regularization to the reconstruction of HR-HSIs [44]. These methods achieved relatively significantly improved performance.

The aforementioned approaches attempt to characterize the spatial structure of an unknown HR-HSI by decomposition. However, since the scenes of HSIs are often complex, it is difficult to characterize an unknown HR-HSI by a single global attribute such as low-rankness. Specifically, there are many individual regions in HSIs, which occupy a very small area and are different from their surrounding spectral properties. However, the existing methods can hardly characterize these small regions well, which leads to a loss of regional spatial and spectral information in the hyperspectral superresolution reconstruction. Therefore, the key point of the hyperspectral superresolution reconstruction is to accurately characterize the spatial and spectral information of the HR-HSI in either global or individual special regions.

To characterize the above individual special regions, in this study, a new method named decomposition of low-rank and sparse tensor (DLST) is proposed, and its implementation procedure is as follows.

1) The unknown HR-HSI was divided into two components: 1) a low-rank and 2) sparse ones. The low-rank component for the global attribute of the HR-HSI was further separated into two parts: 1) the spectral and 2) spatial parts by the spectral low-rank property. The sparse component extracted spectral information in individual special regions.

2) A nonlocal constraint of adaptive manifold was designed for the global region, which used manifolds for grouping and adaptive weight determination, and deeply balanced the local low-rank property and nonlocal self-similarity of spatial part.

3) Surface-aware regularization combined with group sparsity (GS) was used to constrain the sparse component, which enhanced the same spatial structure of different bands and reduced the false individual regions.

4) An iterative optimization method was used to recursively update both the low-rank and sparse components to reconstruct the HR-HSI.

The remainder of this article is organized as follows. Section II introduces the model of hyperspectral superresolution reconstruction and tensor train rank (TTR). The proposed approach is introduced in Section III. In Section IV, experimental results on three datasets are presented. Conclusions are given in Section V.

II. PRELIMINARIES AND NOTATIONS

In this section, the model of the hyperspectral superresolution reconstruction and TTR are introduced. The model of hyperspectral superresolution reconstruction building on an objective function transforms the reconstruction into an optimization problem by exploring the relationship among the HR-HSI, HR-MSI, and LR-HSI of the same scene. TTR is used to measure the low-rank property of the HR-HSI in the reconstruction of the HR-HSI.

A. Problem Model

An HR-HSI, the target of hyperspectral superresolution reconstruction, is expected to carry complete spatial and spectral information. Compared with the HR-HSI, the HR-MSI of the same scene lacks complete spectral information. The spectral information of the HR-MSI can be regarded as a linear combination of the complete spectral information of the HR-HSI to simulate the reduction of spectral resolution from the HR-HSI to the HR-MSI [45]. Let \( \mathcal{X} \in \mathbb{R}^{H \times M \times N} \) be the HR-HSI with \( M \times N \) spatial resolution and \( H \) spectral bands. Let \( \mathcal{Y}_h \in \mathbb{R}^{h \times M \times N} \) be the HR-MSI with \( M \times N \) spatial resolution and \( h \) spectral bands. The relationship of \( \mathcal{X} \) and \( \mathcal{Y}_h \) is often modeled as

\[
\mathcal{Y}_{h(3)} = H \mathcal{X}_{(3)}
\]

where \( \mathcal{X}_{(3)} \in \mathbb{R}^{H \times M \times N} \) and \( \mathcal{Y}_{h(3)} \in \mathbb{R}^{h \times M \times N} \) are the three-mode unfold of \( \mathcal{X} \) and \( \mathcal{Y}_h \), respectively; \( H \in \mathbb{R}^{h \times H} \) is such a matrix whose columns contain the spectral response of the multispectral sensor (\( H > h \)).

Compared with the HR-HSI, the LR-HSI of the same scene has a lower spatial resolution. Every spectral band image of the LR-HSI can be regarded as spatial degradation from the corresponding spectral band image of the HR-HSI to simulate the reduction of spatial resolution from the HR-HSI to LR-HSI. Let \( \mathcal{Y}_H \in \mathbb{R}^{H \times m \times n} \) be the LR-HSI with \( m \times n \) spatial resolution and \( H \) spectral bands; then, the relationship of (for the HR-HSI)
and (for the LR-HSI) is often modeled as
\[
Y_{H(3)} = X_{(3)} M
\]
(2)
where \(Y_{H(3)} \in \mathbb{R}^{H \times m \times n}\) is the three-mode unfold of \(Y_H\); \(M \in \mathbb{R}^{M \times N \times m \times n}\) is the spatial degradation operator, which is often assumed to be composed of a cyclic convolution operator and a down-sampling matrix \((M > m \text{ and } N > n)\).

As a result, \(X\) can be obtained by solving the following objective function:
\[
\min_X \left\| Y_{h(3)} - HX_{(3)} \right\|_F^2 + \left\| Y_{H(3)} - X_{(3)} M \right\|_F^2
\]
(3)
where \(\left\| \cdot \right\|_F^2\) denotes the Frobenius norm, which is used to measure the similarity between matrix \(Y_{h(3)}\) and matrix \(HX_{(3)}\).

Hyperspectral superresolution reconstruction is transformed into the optimization problem of equation (3), which is severely ill-posed and its solution may not be unique. To ensure a stable solution of \(X\), various constraints can be added, and these constraints can be divided into implicit and explicit constraints. The implicit constraints are usually obscure, e.g., the matrix factorization [46] and Trucker factorization [37] of tensor \(X\) in the spectral direction are tight constraints on the factorization form of \(X\) by using the spectral low-rank property and tensor low-rank property of \(X\), respectively. The explicit constraints are used to regularize the hyperspectral superresolution reconstruction. The regularizations include spectral unmixing [30], sparseness [32], non-local spatial similarities [35], spatial smoothness, and low-rankness [46] of \(X\). These constraints restrain the ill-posed solution of \(X\), but they may also introduce errors to \(X\). Hence, there is still a lack of more appropriate constraints for solving \(X\).

B. Tensor Train Rank

For an \(N\)-dimensional tensor \(A \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}\), its tensor train (TT) decomposition [47] is composed of \(N\) 3-dimensional core factors \(G^1, G^2, \ldots, G^N\) where \(G^n \in \mathbb{R}^{I_{r_1} \times I_{r_2} \times \cdots \times I_{r_n}}\) for \(n = 1, 2, \ldots, N\) with \(r_1 = r_{N+1} = 0\). Each element in \(A\) is calculated as
\[
A_{i_1, i_2, \ldots, i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_{N-1}=1}^{R_{N-1}} g^1_{r_1, i_2} \otimes \cdots \otimes g^N_{r_{N-1}, r_N, i_N}
\]
(4)
where \(\otimes\) is the outer product; \([R_1, \ldots, R_N]\) is the TT rank of \(A\). However, the TT decomposition for a tensor is not unique, i.e., different TT decompositions often lead to different TT ranks. In practice, the main focus is often on the lower bound of the TT rank \(R_n \geq \text{rank}(A_{(n)})\), where \(A_{(n)}\) is the \(n\)-mode unfold of \(A\). Therefore, the TT rank of \(A\) can be defined as
\[
\text{rank}_{TT}(A) = \frac{1}{N} \sum_{n=1}^{N} \text{rank}(A_{(n)})
\]
(5)
where \(\text{rank}(A_{(n)})\) equates to the number of the positive singular values of \(A_{(n)}\), and it is nonconvex and thus difficult to optimize. To solve this problem, the convex nuclear norm is usually introduced to replace \(\text{rank}(A_{(n)})\), and \(\text{rank}(A_{(n)})\) is equal to the sum of the singular values of \(A_{(n)}\); thus
\[
\text{rank}(A_{(n)}) = \| A_{(n)} \|_2 = \sum \sigma_i(A_{(n)}) + \varepsilon
\]
(6)
where \(\sigma_i(A_{(n)})\) is the \(i\)th singular value of \(A_{(n)}\). The log sum of the singular values is used to shrink larger singular values slightly and shrink smaller singular values much [48]. The TTR is defined as
\[
\text{rank}_{TT}(A) = \frac{1}{N} \sum_{n=1}^{N} \sum_i \log (\sigma_i(A_{(n)}) + \varepsilon).
\]
(7)

III. Method

The architecture for the DLST method proposed in this study for the hyperspectral superresolution reconstruction is shown in Fig. 1. The DLST was designed to decompose the HR-HSI, see Fig. 1(a). To maintain the spatial information of HR-MSI, a nonlocal constraint of adaptive manifold via TTR was used to effectively maintain the local low-rank property and nonlocal self-similarity of the HR-HSI, see Fig. 1(b). To reduce the solution space of the sparse component, surface-aware regularization combined with GS was introduced, see Fig. 1(c). Finally, the reconstruction was transformed into an alternative optimization problem of multiple variables, see Fig. 1(d).

A. Decomposition of Low-Rank and Sparse Tensor

Recently some new models of tensor decomposition are proposed for HSI denoising [49], fusion [34], [39], and tensor completion task [50], [51], [52]. The global correlation across spectrum and nonlocal self-similarity are utilized by nonlocal low-rank regularized tensor decomposition [49]. Li et al. [34] and Dian et al. [39] utilized coupled sparse tensor factorization (CSTF) and low TTR representation to characterize HSIs, respectively. Xue et al. [50] encoded the structured sparsity of a tensor by the multiple-layer representation. Xue et al. [51] proposed a parametric tensor sparsity, which encoded the sparsity for a general tensor by Laplacian scale mixture modeling. Xue et al. [52] presented an enhanced sparsity prior model using both local and global sparsity information in HSIs. These models describe an HR-HSI using combinations of multiple attributes, and it is difficult to characterize HR-HSIs by a single attribute. Hence, DLST considering low-rank and sparse properties is constructed.

HR-HSIs are spectral low-rank, and an HR-HSI can be divided into pure pixel images, each of which has corresponding spatial coefficients [53]. Assuming that the number of the pure pixels in the HR-HSI is \(K\), then
\[
X = \sum_{k=1}^{K} W^k \circ h_k
\]
(8)
where \(W^k \in \mathbb{R}^{M \times N}\) and \(h_k \in \mathbb{R}^{1 \times H}\) are the spatial coefficients and pure pixel matrices, respectively; \(h_k\) is an expression of the spectrum of the \(k\)th pure pixel (\(k = 1, 2, \ldots, K\)).

Theoretically, an HR-HSI can be expressed by (8). However, \(K (K < M \text{ and } K \ll N)\), as spectral rank, is set to a small
value to constrain the HR-HSI. This causes the pure pixels in the individual regions covering small areas to be ignored as nonmain components. These pure pixels in the individual regions are just in accordance with sparsity because they cover very small areas and may even contain only a few pixels in the HR-HSI. Therefore, the sparse component is used to represent these individual regions and supplement the loss of information caused by spectral low-rank constraints. Euclidean distance (ED), as a measure of the similarity between $\mathcal{X}$ (the expected HR-HSI) and $\sum_{k=1}^{K} W_k \circ h_k$ [the HR-HSI expressed by (8)], was obtained by setting different values for $K$ shown in Fig. 2(b). As we can see from Fig. 2(a) that in the red box, the cars park sparsely beside the green belt, and the spatial and spectral features of the cars can hardly be expressed when $K$ is set to a small value as shown in Fig. 2(b); thus, the individual regions need to be emphasized by the sparse component. An HR-HSI can be divided into two components: 1) a low-rank component $\tilde{X}$ and 2) a sparse component $\mathcal{S}$, and thus, it can be written as

$$\mathcal{X} = \tilde{X} + \mathcal{S}$$

where $\tilde{X}$ is composed of $K$ pure pixels, and it can be expressed by pure pixel matrices and their corresponding spatial coefficients: $\tilde{X} = \sum_{k=1}^{K} W_k \circ h_k$. The HR-HSI $\mathcal{X}$ can be written as

$$\mathcal{X} = \sum_{k=1}^{K} W_k \circ h_k + \mathcal{S}.$$ (10)

Equation (10) can be rewritten as the following unfolded form:

$$X_{(3)} = \sum_{k=1}^{K} h_k^T \tilde{W}^k + S_{(3)}$$ (11)
where $\tilde{W}^k = \text{vec}(W^k)^T$. According to Section II-A, combined with (1) and (2), HR-MSI and LR-HSI can be rewritten as

$$Y_{H(3)} = \left( \sum_{k=1}^{K} h_k^T \tilde{W}^k + S_{(3)} \right) M$$

$$= \sum_{k=1}^{K} h_k^T \tilde{W}^k M + S_{(3)} M$$

$$Y_{h(3)} = H \left( \sum_{k=1}^{K} h_k^T \tilde{W}^k + S_{(3)} \right)$$

$$= \sum_{k=1}^{K} H h_k^T \tilde{W}^k + HS_{(3)}. \quad (12)$$

Equation (3) can be rewritten as

$$f(W^k, h_k, S) = \left\| Y_{h(3)} - \sum_{k=1}^{K} \sum_{i=1}^{G} W_{g(i)}^k \nabla W_{g(i)}^k M \right\|_F^2$$

$$+ \left\| Y_{H(3)} - \sum_{k=1}^{K} h_k^T \tilde{W}^k M - S_{(3)} M \right\|_F^2. \quad (13)$$

The reconstruction problem expressed by (13) is now transformed into solving $W^k, h_k, S$. Some constraints are needed to narrow down the solution space for $W^k$ and $S$ in the ill-posed problem. $h_k$ can be implicitly constrained by accurately estimating of $W^k$ and $S$. In this study, a nonlocal constraint of adaptive manifold and the surface-aware regularization combined with GS was adopted to constrain $W^k$ and $S$, respectively.

B. Nonlocal Constraint of Adaptive Manifold via TTR

Nonlocal self-similarity and local low-rank property over space are two important characteristics for HR-HSIs [45]. In order to enhance nonlocal self-similarity, the spatial coefficient $W^k$ is divided into many nonlocal patches and the most similar nonlocal patches are searched in the whole image. In this process, the similarity measurement of the patches based on image intensity is rough. Patches in Fig. 3(c) are divided into two groups by the image intensity [see Fig. 3(a)]. Textures in patches are ignored. In order to solve this problem, the manifold is introduced as the basis of the similarity measurement of the patches. Patches in Fig. 3(d) are divided into four groups by the manifold of the image [see Fig. 3(b)]. We can see that the manifold structure emphasizes texture details better than intensity. Therefore, the grouping of the patches based on manifold structures is better than that based on the image intensity because the former focuses more on texture.

The spatial coefficients $W^k$ are divided into patches, and those patches that have similar manifold structures in the HR-HSI are grouped together, then the patches in the same group constitute a 3-dimensional tensor $W_g^k$. All patches in Fig. 1 are divided into $G$ groups. The TTR of the 3-dimensional tensor $W_g^k$ is used to make a unified tradeoff between the local low-rank property and the nonlocal self-similarity of the manifold structure for $W^k$, which is written as

$$\Psi(W^k) = \sum_{g=1}^{G} \text{Rank}_T(W_g^k) \text{s.t. } W^k = \{ W_1^k, \ldots, W_g^k, \ldots, W_G^k \}. \quad (14)$$

Additionally, as the HR-HSI is unknown, its patches cannot be directly grouped. To solve this problem, the grouping of the patches that are in the HR-MSI was conducted since the spatial information is mainly contained in the HR-MSI. In the TTR, $\text{Rank}_T(W_g^k)$ and $\text{Rank}(W_g^k)$ reflect the local low-rank property of $W^k$, and $\text{Rank}(W_g^k)$ reflects the nonlocal self-similarity of the manifold structure for $W^k$. Their weights need to be determined. The weight of $\text{Rank}_T(W_g^k)$ is expected to be larger for the group with notable manifold structure. The weights of $\text{Rank}_T(W_g^k)$ and $\text{Rank}(W_g^k)$ are expected to be larger for the group with insignificant manifold structure. In this study, the weight was determined by the sparse norm of the manifold vector gradient since the sparse norm can measure the manifold strength. Thus, the nonlocal constraint of the adaptive manifold via TTR can be formulated as

$$\Psi(W^k) = \sum_{g=1}^{G} \sum_{i=1}^{3} \left\| \nabla W_{g(i)}^k \right\|_2 \text{Rank}_T(W_g^k) \text{s.t. } W^k = \{ W_1^k, \ldots, W_g^k, \ldots, W_G^k \}. \quad (15)$$

For the manifold of an image, more attention is paid to the texture details rather than the intensity of the image, and also more attention is paid to the nontexture region rather than the gradient of the image. The use of the adaptive weights based on the image’s manifold further magnifies this advantage. Therefore, the nonlocal constraint of the adaptive manifold via TTR can be used as an effective constraint on the spatial information of HR-HSIs in both simple and complex environments.

C. Surface-Aware Regularization Combined With GS

The sparse attribute of the sparse components $S$ is mainly reflected in the spatial domain. Standard sparsity is used to minimize the nonzero elements of $S$ when each element in $S$ is independent. Obviously, this ignores the fact that different bands of $S$ have the same spatial structure. In order to further preserve the potential relationship of different bands, the GS is used to ensure nonzero values in the same spatial locations of different
bands. The GS not only ensures that different bands of S are sparse but also preserve the potential relationship of different bands. The group sparse regularity of S in the spatial domain is expressed by the L1,2-norm

$$\varphi(S) = \sum_{i=1}^{MN} \left( \sum_{h=1}^{H} S(3(i) + h)^2 \right)^{1/2}. \quad (16)$$

The GS ensures that the sparsity of sparse components preserves individual regions, but it may also obtain false individual regions. In order to eliminate the false individual regions, surface-aware regularization is increased to reduce the false individual regions and disadvantageous structures and retain individual regions.

The sparse components S can be divided into H forward slices: $S = \{S^{(1)}, \ldots, S^{(h)}, \ldots, S^{(H)}\}$. In geometric frameworks for image processing, intensity images are considered as surfaces in the spatial-feature space. A image is thereby a 2-dimensional surface in the 3-dimensional space [54], and the surface of $S^{(h)}$ is

$$\Phi(S^{(h)}) = \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{1 + |\nabla_{i,j} S^{(h)}|^2} \quad \text{where} \quad |\nabla_{i,j} S^{(h)}|^2 = |\nabla_i S^{(h)}|^2 + |\nabla_j S^{(h)}|^2 \quad \text{with} \quad |\nabla_i S^{(h)}|^2,$$

and $|\nabla_j S^{(h)}|^2$ represent the difference matrices in the horizontal and vertical directions, respectively, and they are calculated via

$$\nabla_i S^{(h)} = \begin{cases} S^{(h)}_{i+1,j} - S^{(h)}_{i,j}, & 1 \leq j \leq N - 1 \\ S^{(h)}_{i,N+1} - S^{(h)}_{i,N}, & j = N \end{cases}$$

and

$$\nabla_j S^{(h)} = \begin{cases} S^{(h)}_{i+1,j} - S^{(h)}_{i,j}, & 1 \leq i \leq M - 1 \\ S^{(h)}_{M+1,j} - S^{(h)}_{M,j}, & i = M \end{cases}. \quad (18)$$

Surface-aware regularization of S is the sum of the surface-aware regularization of all forward slices, which can be written as

$$\Phi(S) = \sum_{h=1}^{H} \Phi(S^{(h)}). \quad (19)$$

The surface-aware regularization combined with GS can be defined as

$$\lambda \Phi(S) + \gamma \varphi(S) \quad (20)$$

where $\lambda$ and $\gamma$ are the weights of the surface-aware regularization and GS, respectively.

D. Alternating Optimization

Obviously, it is an inverse problem to estimate the HR-HSI from HR-MSI and LR-HSI. We incorporated the nonlocal constraint of adaptive manifold via TTR and surface-aware regularization combined with GS into (15) to transform the problem well-posed into

$$g(W^k, h_k, S) = f(W^k, h_k, S) + \lambda \Phi(S)$$

$$+ \gamma \varphi(S) + \mu \sum_{k=1}^{K} \Psi(W^k). \quad (21)$$

Although the variables in (21) are coupled, they can be solved alternatively, i.e., one variable is solved by fixing the others in the corresponding minimization problem. Problem (21) can be broken into three subproblems, in which all variables are iteratively updated with the others fixed. The proximal alternating optimization (PAO) scheme [55], [56] is used to solve (21), which can be guaranteed to converge to a critical point under certain conditions and can be written as

$$W^k = \text{argmin}_{W^k} g(W^k, h_k, S) + \beta \|W^k - W^k_{\text{pre}}\|_F^2 \quad (22)$$

$$h_k = \text{argmin}_{h_k} g(W^k, h_k, S) + \beta \|h_k - h_{k_{\text{pre}}}\|_F^2 \quad (23)$$

$$S = \text{argmin}_{S} g(W^k, h_k, S) + \beta \|S - S_{\text{pre}}\|_F^2. \quad (24)$$

The details for the optimizations of $W^k$, $h_k$, and $S$ are as follows.

1) Optimization of $W^k$: With fixed $h_k$ and $S$, the optimization of $W^k$ in (22) can be written as

$$\text{argmin}_{W^k} \|Y^*_h^{(3)} - H h^{(3)} T \hat{W}^k\|_F^2 + \|Y^*_h^{(3)} - h^{(3)} T \hat{W}^k M\|_F^2$$

$$+ \mu \Psi(W^k) + \beta \|\hat{W}^k - \hat{W}^k_{\text{pre}}\|_F^2 \quad (25)$$

where $Y^*_h^{(3)} = Y_{h^{(3)}} - \sum_{k=1, k \neq k}^{K} H h^{(3)} T \hat{W}^k + H S_{(3)}$ and $Y^*_h^{(3)} = Y_{h^{(3)}} - \sum_{k=1, k \neq k}^{K} h^{(3)} T \hat{W}^k M + S_{(3)} M$. Problem (25) is an unconstrained optimization problem and is solved by the ADMM [57]. By introducing the variable $U^k = W^k$, the following augmented Lagrangian function is acquired:

$$\text{argmin}_{W^k, U^k} \|Y^*_h^{(3)} - H h^{(3)} T \hat{W}^k\|_F^2 + \|Y^*_h^{(3)} - h^{(3)} T \hat{W}^k M\|_F^2$$

$$+ \beta \|\hat{W}^k - \hat{W}^k_{\text{pre}}\|_F^2 + \rho \|U^k - W^k - \frac{\rho}{2} \|_F^2 + \mu \Psi(U^k). \quad (26)$$

The augmented Lagrangian function (26) can be minimized by iteratively solving the following subproblems.

a) Update $W^k$: With $U^k$ fixed, the subproblem to update $W^k$ in (26) can be written as

$$W^k = \text{argmin}_{W^k} \|Y^*_h^{(3)} - H h^{(3)} T \hat{W}^k\|_F^2$$

$$+ \|Y^*_h^{(3)} - h^{(3)} T \hat{W}^k M\|_F^2$$

$$+ \beta \|\hat{W}^k - \hat{W}^k_{\text{pre}}\|_F^2 + \rho \|U^k - W^k - \frac{\rho}{2} \|_F^2 \quad (27)$$

Problem (27) is quadratic and its unique solution amounts to compute the general Sylvester matrix equation:

$$h_k h^{(3)} T \hat{W}^k M T + h_k H^T H h^{(3)} T \hat{W}^k = h_k Y^*_h^{(3)} M T$$

$$+ h_k H^T Y^*_h^{(3)} + \beta \hat{W}^k_{\text{pre}} + \rho \left( \hat{U}^k - \frac{\rho}{2} \right). \quad (28)$$
Algorithm 1: Optimization of $W^k$.

**Initialize** $W^k, h_k, S$.

**While** not converged do

1. **Update** $W^k$: With fixed $W^k$, the subproblem to update $U^k$ in (28) can be written as
   \[
   U^k = \arg \min_{U^k} \rho \left\| U^k - W^k - \frac{O}{2\rho} \right\|_F^2 + \mu \Psi(U^k)
   \] (29)
   where $U^k = \{U_1^k, \ldots, U_g^k, \ldots, U_{i+1}^k\}$ with $U_g^k = V_g^k$. Problem (29) can be rewritten as
   \[
   \rho \left\| U_g^k - W_g^k - \frac{\gamma}{2\rho} \right\|_F^2 + \mu \sum_{i=1}^3 \gamma_i \text{Rank}(U_{g(i)}^k)
   \] (30)
   where $\gamma_i = \|\nabla W_{g(i)}^k\|_2$, which is obtained using the last updated $W^k$ to simplify the process of calculation. Problem (30) is a low-rank recovery problem, thus it has the following solution:
   \[
   U_{g(i)}^k = U_i \hat{\Sigma}^i V_i
   \] (31)
   where $U_i$, $\hat{\Sigma}^i$, and $V_i$ are acquired by the SVD of $W_{g(i)}^k - \frac{\gamma}{2\rho}$.

2. **Update** $h_k$: With $W^k$ and $S$ fixed, the optimization of $h_k$ in (23) can be written as
   \[
   \arg \min_{h_k} \left\| Y_{h(3)} - H h_k^T \hat{W}^k \right\|_F^2 + \left\| Y_{h(3)} - h_k^T \tilde{W}^k M \right\|_F^2 + \beta \left\| h_k - h_{\text{pre}} \right\|_F^2.
   \] (32)
   Problem (32) is quadratic and its unique solution amounts to compute the general Sylvester matrix equation:
   \[
   \hat{W}^k M (\hat{W}^k M)^T h_k + \tilde{W}^k (\tilde{W}^k)^T h_k H^T H + \beta h_k
   \]
   \[
   = \hat{W}^k M (Y_{h(3)}^*)^T + \tilde{W}^k (Y_{h(3)}^*)^T H + \beta h_{\text{pre}}.
   \] (33)

3) **Optimization of $S$**: With $W^k$ and $h_k$ fixed, the optimization of $S$ in (24) can be written as
   \[
   \arg \min_{S} \left\| \hat{Y}_{h(3)} - H S(3) \right\|_F^2 + \left\| \hat{Y}_{H(3)} - S(3) M \right\|_F^2 + \lambda \Phi(S) + \gamma \phi(S) + \beta \left\| S - S_{\text{pre}} \right\|_F^2
   \] (34)
   where $\hat{Y}_{h(3)} = Y_{h(3)} - \sum_{i=1}^K H h_i^T \hat{W}^k$ and $\hat{Y}_{H(3)} = Y_{H(3)} - \sum_{i=1}^K h_i^T \tilde{W}^k M$. To solve (34) efficiently, we introduce two auxiliary tensor $\mathcal{V}$ and $\mathcal{L}$ to substitute the $S$ in the terms $\Phi(S)$ and $\phi(S)$, respectively. Then, the minimization problem (34) can be reformulated as the following constrained optimization problem:
   \[
   \arg \min_{\mathcal{S}, \mathcal{L}, \mathcal{V}} \left\| \hat{Y}_{h(3)} - H S(3) \right\|_F^2 + \left\| \hat{Y}_{H(3)} - S(3) M \right\|_F^2 + \lambda \Phi(\mathcal{V}) + \gamma \phi(\mathcal{L}) + \beta \left\| \mathcal{S} - S_{\text{pre}} \right\|_F^2 + \alpha \left\| \mathcal{V} - S \right\|_F^2 + v \left\| \mathcal{L} - S \right\|_F^2.
   \] (35)
   The problem (35) can be minimized by iteratively solving the following subproblems.

a) **Update $\mathcal{S}$**: With $\mathcal{V}$ and $\mathcal{L}$ fixed, the subproblem to update $\mathcal{S}$ in (35) can be written as
   \[
   \arg \min_{\mathcal{S}} \left\| \hat{Y}_{h(3)} - H S(3) \right\|_F^2 + \left\| \hat{Y}_{H(3)} - S(3) M \right\|_F^2 + \lambda \Phi(\mathcal{V})
   \]
   \[
   + \gamma \phi(\mathcal{L}) + \beta \left\| \mathcal{S} - S_{\text{pre}} \right\|_F^2 + \alpha \left\| \mathcal{V} - S \right\|_F^2 + v \left\| \mathcal{L} - S \right\|_F^2.
   \] (36)
   Problem (36) is quadratic and its unique solution amounts to compute the general Sylvester matrix equation:
   \[
   S(3) M M^T + H^T H S(3) + (\beta + \alpha + v) S(3)
   \]
   \[
   = \hat{Y}_{H(3)} M^T + H^T \hat{Y}_{h(3)} + \beta S_{\text{pre}}(3) + \alpha \mathcal{V}(3) + v \mathcal{L}(3).
   \] (37)
   Finally, CG is used to solve (37) due to its symmetric positive definite system matrix, simple calculation steps and fast convergence.

b) **Update $\mathcal{L}$**: With $\mathcal{V}$ and $\mathcal{S}$ fixed, the subproblem to update $\mathcal{L}$ in (35) can be written as
   \[
   \sum_{h=1}^{MN} \left( \sum_{h=1}^{H} L(3) h \right)^{1/2}
   \]
   \[
   + v \left\| \mathcal{L} - S \right\|_F^2.
   \] (38)
   The subproblem can be solved by half-quadratic optimization strategy [58]
   \[
   \mathcal{L}_{[i,j,:]} = \begin{cases}
   \frac{\| S_{[i,j,:]} \|_2^2}{\| S_{[i,j,:]} \|_2} S_{[i,j,:]}, & \| S_{[i,j,:]} \|_2 > \frac{\| S_{[i,j,:]} \|_2}{\alpha}, \\
   0, & \| S_{[i,j,:]} \|_2 \leq \frac{\| S_{[i,j,:]} \|_2}{\alpha}.
   \end{cases}
   \] (39)

c) **Update $\mathcal{V}$**: With $\mathcal{L}$ and $\mathcal{S}$ fixed, the subproblem to update $\mathcal{V}$ in (35) can be written as
   \[
   \lambda \sum_{h=1}^{H} \sum_{i=1}^{M} \sqrt{1 + \| \nabla_{i,j} \mathcal{V}(h) \|_2^2} + \alpha \left\| \mathcal{V} - S \right\|_F^2.
   \] (40)

Finally, conjugate gradient (CG) [54] is used to solve (33). The symmetric positive definite system matrix, simple calculation steps, and fast convergence prompted the CG to be selected.
Algorithm 2: DLST.

Initialize $W^k, h_k, S$.

While not converged do
  Step1: update $W$
  for $k = 1:K$
    update $W^k$ by Algorithm 1
  end for
  Step2: update $h_k$ by solving (33) with CG
  $h_{k,pre} = h_k$
  Step3: update $S$ by solving (37) with CG
  $S_{pre} = S$
  Step4: update $L$ via (39)
  Step5: update $\lambda$
  for $h = 1:H$
    update $\lambda(h)$ via (43)
  end for

end while

Output estimated $\lambda'$ for the HR-HSI via (12)

The subproblem of update $\lambda$ can be reformulated to solve the forward slice $\lambda(h)$

$$
\min_{\lambda(h)} \max_{Z} \left< \nabla \lambda(h), Z \right> + \lambda \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{1 - |Z_{i,j}|^2}^2 + \alpha \left< \lambda(h) - S^{(h)} \right>^2_F \quad (41)
$$

Due to its simplicity and stability, we apply the primal-dual algorithm [59], [60]. $\nabla_{\lambda} \lambda(h)$ is the matrix of $\nabla_{ij} \lambda(h)$. Then, the subproblem (41) can be easily rewritten in terms of a saddle point problem

$$
\min_{\lambda(h)} \max_{Z} \left< \nabla \lambda(h), Z \right> + \lambda \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{1 - |Z_{i,j}|^2}^2 + \alpha \left< \lambda(h) - S^{(h)} \right>^2_F \quad (42)
$$

The following iteration equations are used to solve the primal-dual problem (41):

$$
\begin{align*}
\hat{Z}^{k+1} &= Z^{k} + \sigma \left( \nabla \lambda(h)^{k} - \lambda \frac{Z^{k}}{\sqrt{1 - |Z^{k}|^2}} \right) \\
Z^{k+1} &= \frac{\hat{Z}^{k+1}}{\max(|Z^{k+1}|, 1)} \\
\lambda^{(h)}^{k+1} &= \lambda^{(h)}^{k} - \tau \left( \nabla^T Z^{k+1} + \alpha \left( \lambda^{(h)}^{k} - S^{(h)} \right) \right)
\end{align*}
$$

Algorithm 2: Process for the optimization of the proposed DLST method.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. Experiment

1) Datasets: Three datasets were used to test the performance of the proposed method. The first dataset (Pavia University) was acquired from the reflective optics system imaging spectrometer ROSIS sensor during a flight campaign at Pavia University, Northern Italy. The image is of size $610 \times 340 \times 115$ with a 1.3-m spatial resolution and the spectral coverage ranging from 0.43 $\mu$m to 0.86 $\mu$m. The second dataset (CAVE) was acquired by a Cooled CCD camera (Apogee Alta U260) from 32 scenes. The images are of size $512 \times 512 \times 31$ with the spectral coverage ranging from 0.4 $\mu$m to 0.7 $\mu$m with a 0.01-nm interval. The third dataset (Chikusei) was taken by a Headwall Hyperspec-VNIR-C imaging sensor over the agricultural and urban areas in Chikusei, Ibaraki, Japan. The image is of size $2517 \times 2335 \times 128$ with a 2.5-m spatial resolution and the spectral coverage ranging from 0.363 $\mu$m to 1.018 $\mu$m. We generate the LR-HSI via down-sampling the reference by averaging over disjoint $s \times \alpha$ blocks, where $s$ is the scaling factor, and simulate HR-MSI via down-sampling the reference along the spectral dimension using the spectral response matrix $H$ derived from a Nikon D700 camera.

2) Datasets Quantitative Metrics: To evaluate the performance of the hyperspectral superresolution reconstruction from all methods, in this study, the following six image quality indices were adopted: 1) peak signal-to-noise ratio (PSNR), 2) universal image quality index (UIQI), 3) root-mean-squared error (RMSE) [61], 4) structured similarity (SSIM) [62], 5) spectral angle mapper (SAM), and 6) erreur relative globale adimensionnelle de synthèse (ERGAS) [63]. The PSNR, RMSE, and UIQI were for evaluating the similarity between the estimated HR-HSI and the reference. The SSIM, SAM, and ERGAS measure the structural consistency and the spectral distortion between the reconstructed HR-HSI and the reference. The larger the PSNR, SSIM, and UIQI values, the better the reconstructed results, and the smaller the RMSE, SAM, and ERGAS values, the less the error in the reconstruction and less the distortion in the spectral.

B. Selection of Parameters

For the DLST method, the two key parameters $K$ and $G$ need to be set. $K$ is the number of pure pixels and $G$ is the number of groups. The former affects the decomposition process of the low-rank component and the latter affects the grouping process of the manifold structure. To set appropriate values for these two parameters, the Stuffed Toy (a close-range HSI in the CAVE) and Pavia University of remote sensing aerial perspective were selected for reconstruction. The aforementioned PSNR was used to evaluate the performance of reconstructed images. The PSNR curves of the reconstructed image were obtained by setting different values for the above two parameters, as shown in Fig. 4.

Fig. 4(a) shows the PSNR curves of the reconstructed Stuffed Toy and of Pavia University as a function of $K$. When $K$ changes...
from 1 to 6, the PSNR of CAVE increases, but when $K$ is greater than 8, the PSNR starts to decrease. Therefore, $K = 7$ was adopted for CAVE. The PSNR of Pavia University rises rapidly with the increase of $K$ from 1 to 11, and then, it changes little. As a result, $K$ of Pavia University was set to 11. For those scenes with strong spectral similarity, such as small scenes captured by CAVE, which capture much smaller scenes and contain much fewer objects, a smaller value was set to $K$. Due to strong spectral similarity in small scenes from the CAVE, smaller scenes are captured, fewer objects are contained and a smaller value is set to $K$. Due to complex spectra in the scene from the Pavia University, which is remote sensed and contains many different objects, a larger value is set to $K$.

C. Convergence Analysis

In order to analyze the convergence of the method, DLST was used to reconstruct HR-HSIs of the Pavia University and the Stuffed Toy. The RMSE curves of the reconstructed HR-HSIs were obtained by setting the different number of iterations, as shown in Fig. 5.

Fig. 5(a) and (b) shows the RMSE curve of the reconstructed Pavia University and Stuffed Toy as a function of the number of iterations, respectively. The RMSE of Pavia University and Stuffed Toy decreases with the rise of the number of iterations in the range from 1 to 32, and then, it converges gradually at 1.5971 and 1.557, respectively. It shows that This method has good convergence in the reconstruction of HR-HSI. The number of iterations from DLST was set to 32.

D. Ablation Test

In order to verify the effectiveness of DLST, three ablation tests about DLST, nonlocal constraint of adaptive manifold via TTR and surface-aware regularization combined with GS were carried out.

1) Decomposition Test: In order to verify the promotion of the DLST to hyperspectral superresolution reconstruction, DLST without sparse component (SC) and DLST were used to reconstruct HR-HSIs of the Pavia University and the Stuffed Toy. $K$, as an important parameter of decomposition, is set to different values to control decomposition. The PSNR curves of the reconstructed image were obtained by setting different values for $K$ on the two methods, as shown in Fig. 6.

Fig. 6(a) and (b) shows the PSNR curves of the reconstructed Pavia University and Stuffed Toy as a function of $K$, respectively. The PSNRs of Pavia University by DLST without SC and DLST rise with the rise of $K$ in the range from 1 to 15 and from 1 to 11, and then, they keep at 42.91 and 44.9, respectively. The PSNRs of CAVE by DLST without SC and DLST rise with the rise of $K$ in the range from 1 to 7 and from 1 to 12, and then, they keep at 42.1 and 44.45, respectively. It shows that DLST can reconstruct HSIs more effectively than DLST without SC. The sparse component in the DLST can compensate for the information loss brought by the constraint of low rankness for the reconstruction of the HR-HSI.

2) Nonlocal Constraint of Adaptive Manifold Test: In order to verify the promotion of nonlocal constraint of adaptive manifold via TTR to hyperspectral superresolution reconstruction, DLST without nonlocal self-similarity (NS), which replaces nonlocal self-similarity via TTR with matrix low-rankness via nuclear norm, DLST without adaptive manifold (AM), which replaces the manifold with the image intensity, and DLST were used to reconstruct HR-HSIs of the Pavia University and Stuffed Toy.

Table I shows the means of the PSNR and RMSE values of the HR-HSIs constructed by the three methods for the Pavia University and the Stuffed Toy. Compared with DLST without
NS and DLST without AM, the PSNR and RMSE values of HR-HSIs reconstructed by DLST are the highest and lowest, respectively. Hence, the combination of adaptive manifold and nonlocal self-similarity can constrain spatial information more effectively.

3) Surface-Aware Regularization Combined With GS Test: In order to verify the promotion of surface-aware regularization combined with GS to hyperspectral superresolution reconstruction, DLST without GS, DLST without surface-aware regularization (SA), and DLST were used to reconstruct HR-HSIs of the Pavia University and the Stuffed Toy. The PSNR curves of the reconstructed image were obtained by setting different values for \( K \) on the three methods, as shown in Fig. 7.

Fig. 7(a) and (b) shows the PSNR curves of the reconstructed Pavia University and Stuffed Toy as a function of \( K \) from DLST without GS, DLST without SA, and DLST.

In Fig. 7(a) and (b), the large red boxes indicate that DLST is closer to the truth than all the other methods (see the large red box). The PSNR values of the reconstructed HR-HSIs for the Pavia University and Stuffed Toy, as two HSIs, are from indoor close and outdoor remote sensing scenes, respectively. The spatial reconstructions of the two regions (cars and cluttered background) in the reconstructed HR-HSIs are shown in Figs. 8 and 9, respectively. Again, the large red box is a zoom-in of the small red box for clearly observing the small region. The two special areas in the small red boxes of the upper two rows and the bottom two rows, respectively, are the cars parking sparsely at the edge of the green belt on the 30th band and the duplicate T-shaped building on the 80th band, respectively. Due to the small area of the cars and the large difference between the cars and their surrounding regions in the spatial domain, their spatial structure is difficult to be reconstructed. The duplicate T-shaped building can be used as an investigation area due to its clear spatial low-rank structure. From subregion (see the large red box) comparison results of the error images, the HR-HSI constructed by DLST is closer to the ground truth in both the cars and duplicate T-shaped building regions than all the other methods. Therefore, the DLST method retains more spatial information of the cars and duplicate T-shaped building than the other five methods.

In Fig. 9, the reconstructed images/corresponding error images are shown in the first/second and third/fourth rows for the Stuffed Toy on the 5th and 26th bands, respectively. The small and large red boxes in each subfigure are similar to that in Fig. 5. The two special areas in the small red boxes of the upper two rows and the bottom two rows, respectively, are the cluttered background on the 5th band and the striped sweater on the 26th band, respectively. The cluttered background has complex and irregular texture while the striped sweater has a simple low-rank texture. From (see large red box) comparisons of the error images from different methods, the HR-HSI recovered by DLST is closer to the ground truth in both complex and simple structured texture regions than all the other methods. This result suggests that the DLST method reserves more detailed characteristics than the other five methods.

For quantitative comparison, Fig. 10 shows the mean of the PSNR values of all bands in the reconstructed HR-HSIs resulting from each method. Fig. 10(a) and (b) shows the mean of the PSNR values of the reconstructed HR-HSIs for the Pavia University and Stuffed Toy, respectively. The results in both subfigures indicate that the proposed DLST method outperforms the other methods on all bands.

2) Spectral Reconstruction: The spatial reconstructions of the six methods were compared in both visual and quantitative aspects. For visual comparison, the reconstructed images and corresponding error images of the above two HSIs at the two bands that had the largest and smallest spectral reflections are shown in Figs. 8 and 9, respectively. For quantitative comparison, the mean of the PSNR values of all bands in the reconstructed HR-HSIs of the two HSIs is shown in Fig. 10.

In Fig. 8, the reconstructed images/corresponding error images are shown in the first/second and third/fourth rows for Pavia University on the 30th and 80th bands, respectively. The last column (g) is the ground truth. In each subfigure, the large red box is a zoom-in of the small red box for clearly observing the small region. The two special areas in the small red boxes of the upper two rows and the bottom two rows, respectively, are the cars parking sparsely at the edge of the green belt on the 30th band and the duplicate T-shaped building on the 80th band, respectively. Due to the small area of the cars and the large difference between the cars and their surrounding regions in the spatial domain, their spatial structure is difficult to be reconstructed. The duplicate T-shaped building can be used as an investigation area due to its clear spatial low-rank structure. From subregion (see the large red box) comparison results of the error images, the HR-HSI constructed by DLST is closer to the ground truth in both the cars and duplicate T-shaped building regions than all the other methods. Therefore, the DLST method retains more spatial information of the cars and duplicate T-shaped building than the other five methods.

In Fig. 9, the reconstructed images/corresponding error images are shown in the first/second and third/fourth rows for the Stuffed Toy on the 5th and 26th bands, respectively. The small and large red boxes in each subfigure are similar to that in Fig. 5. The two special areas in the small red boxes of the upper two rows and the bottom two rows, respectively, are the cluttered background on the 5th band and the striped sweater on the 26th band, respectively. The cluttered background has complex and irregular texture while the striped sweater has a simple low-rank texture. From (see large red box) comparisons of the error images from different methods, the HR-HSI recovered by DLST is closer to the ground truth in both complex and simple structured texture regions than all the other methods. This result suggests that the DLST method reserves more detailed characteristics than the other five methods.
Fig. 8. Comparison between the reconstructed images/error images in the first/second and third/fourth rows for the dataset of Pavia University at the 30th and 80th bands, respectively resulting from methods (a) Hysure, (b) NLSTF, (c) LTMR, (d) LTTR, (e) CSTF, (f) DLST, and (g) the ground truth.

Fig. 9. Comparison between the reconstructed images/error images in the first/second and third/fourth rows for the Stuffed Toy dataset on the 5th and 26th bands, respectively resulting from methods (a) Hysure, (b) NLSTF, (c) LTMR, (d) LTTR, (e) CSTF, (f) DLST, and (g) the ground truth.

The ED and cosine similarity (CS) were used to measure the similarity between the spectral reflection curves constructed by all six methods and the true spectral reflection curve, and the results are listed in Table II. The smaller the ED and CS values, the more similar the spectral reflection curves are. Among all the six results, the DLST result is the smallest, meaning that DLST is the best performer.

3) Total Quality Evaluation: Table III shows the means of the PSNR, RMSE, UIQI, SSIM, SAM, and ERGAS values of the HR HSIs constructed by all the six methods for the Pavia University, Chikusei, and the two HSIs (Flowers and Lemons in the CAVE). One can see that DLST is the best performer in terms of all the six PQIs. The performance of the NLSTF method is the closest to that of DLST for scenes with high-texture
Fig. 10. Mean of the PSNR values of all bands in the reconstructed HR-HSIs resulting from six methods and the datasets of (a) Pavia University and (b) Stuffed Toy.

Fig. 11. Spectral reflection curves of all bands for the reconstructed HR-HSIs from two regions resulting from six methods for the datasets of (a) Pavia University and (b) Stuffed Toy.

TABLE II
SIMILARITY BETWEEN THE SPECTRAL REFLECTION CURVES CONSTRUCTED BY EACH OF THE SIX METHODS AND THE TRUTH

| Dataset        | PQIs | Hysure | NLSTF | LTTR | CSTF | DLST |
|----------------|------|--------|-------|------|------|------|
| Pavia University | ED   | 0.25  | 0.16  | 0.43 | 0.42 | 0.15 | 0.12 |
|                | CS   | 0.21  | 0.17  | 0.28 | 0.26 | 0.11 | 0.10 |
| Stuffed Toy    | ED   | 0.38  | 0.15  | 0.21 | 0.26 | 0.29 | 0.16 |
|                | CS   | 0.47  | 0.13  | 0.16 | 0.18 | 0.25 | 0.14 |

information, e.g., Pavia University and Chikusei, and lower than that of DLST for scenes with low-texture information, e.g., the lemons and flowers. The mean of all the six PQIs from DLST is close to that of NLSTF for the Pavia University. The mean of the PSNR, SSIM, and UIQI values of DLST is larger than that of CSTF by 4, 0.01, and 0.004 for the Flowers, respectively. The mean of the RMSE, ERGAS, and SAM values of DLST is lower than that of NLSTF by 0.5, 0.5, and 2 for the Flowers, respectively. The results of the LTTR and LTMR methods are the closest to that of DLST for scenes with high-texture information, e.g., the Pavia University and Chikusei. The mean of the PSNR, SSIM, and UIQI values of DLST is larger than that of LTMR by 10, 0.05, and 0.04 for the Chikusei, respectively. The mean of the RMSE, ERGAS and SAM values of DLST is lower than that of LTMR by 4, 0.6, and 2 for the Chikusei, respectively.

The results of all the methods in Table III were further analyzed for their own characteristics. The performance of Hysure is poor for scenes with low texture information because the TV regularization hardly characterizes these scenes, e.g., the lemons. The CSTF method uses coupled sparse tensor factorization to exploit spatial-spectral correlations, leading to the better results than Hysure. The LTTR and LTMR methods explore the low-rank property based on the nonlocal spatial self-similarity, hardly maintaining texture details for scenes rich in complex texture, e.g., the Pavia University and Chikusei. The NLSTF method based on the NLSTF, lacking the a-priori constraints in homogeneous regions. Therefore, NLSTF leads to more errors in homogeneous regions than LTTR and LTMR for the lemons. In contrast, the DLST method has the highest accuracy in both spatial and spectral preservation for the three datasets.

4) Runtime Analysis: To show the computation efficiency of the six methods, Table IV reports the average computational time in seconds of the compared methods on the Pavia University, CAVE, and Chikusei. All experiments are implemented at Mat- lab R2021b, and the computer has an Intel Core-i5-11400F CPU with 2.6-GHz and 16-GB random access memory. As can be seen from the table, the Hysure method needs the least computational time on the three datasets. Our method needs relatively more computation time on the three datasets, and the main computation burden of our method comes from solving (25), where we compute TTR for each cluster. The time complexity for solving (29) is mainly related to the size of patches, the number of patches in each group, and the number of spectral ranks.

5) Scaling Analysis: As the ratio of the HR-MSI to the LR-HSI in the spatial domain, the scale factor is an important factor affecting reconstruction. With the increase of the scale factor, the spectral and the spatial information contained in the LR-HSI decreases under the same HR-MSI. To compare the influence of the scale factor on the methods, the values of $2^1$, $2^2$, ..., $2^6$ were set for the scale factor to obtain the LR-HSIs. Then, the LR-HSIs and the same HR-MSI were used to construct the HR-HSIs using all the six methods. Fig. 12 shows the PSNR and RMSE results for evaluating the quality of the HR-HSIs. From Fig. 12(a) and (b), with the increase of the scale factor, the PSNR and RMSE values of the HR-HSIs constructed by the DLST method decrease and increase more slowly, respectively, than that of all the other methods. This implies that the proposed DLST method can better resist information reduction caused by the increase of the scale factor.
TABLE III
QUANTITATIVE METRICS OF THE HR-HSIS CONSTRUCTED BY SIX METHODS FOR THE PAVIA UNIVERSITY (TOP), CHIKUSEI (SECOND), AND THE TWO HSIS (FLOWERS AND LEMONS IN THE CAVE)

| HSI          | PQIs       | Hysure | NLSTF | LTRM | LLTR | CSTF | DLST  |
|--------------|------------|--------|-------|------|------|------|-------|
| Pavia University |           |        |       |      |      |      |       |
| PSNR         | 41.64      | 44.34  | 37.61 | 38.55| 43.53| 44.49|       |
| RMSE         | 2.27       | 1.52   | 4.71  | 4.28 | 1.93 | 1.48 |       |
| SSIM         | 0.9876     | 0.9927 | 0.9576| 0.9569| 0.9892| 0.9971|       |
| ERGAS        | 0.64       | 2.05   | 1.19  | 1.12 | 0.54 | 0.47 |       |
| UIQI         | 0.9922     | 0.9952 | 0.968 | 0.9711| 0.994 | 0.9963|       |
| SAM          | 2.29       | 1.94   | 4.39  | 3.96 | 1.92 | 1.55 |       |
| Chikusei     |           |        |       |      |      |      |       |
| PSNR         | 43.84      | 47.07  | 37.45 | 39.13| 46.32| 47.25|       |
| RMSE         | 2.01       | 1.60   | 5.59  | 4.42 | 1.64 | 1.58 |       |
| SSIM         | 0.9834     | 0.9962 | 0.9284| 0.9496| 0.9863| 0.9986|       |
| ERGAS        | 0.59       | 2.19   | 1.55  | 1.09 | 0.45 | 0.42 |       |
| UIQI         | 0.978      | 0.9921 | 0.9083| 0.947 | 0.9985| 0.999 |       |
| SAM          | 1.34       | 1.5417 | 2.85  | 2.28 | 1.11 | 1.09 |       |
| Flowers      |           |        |       |      |      |      |       |
| PSNR         | 42.01      | 44.89  | 47.88 | 47.3 | 43.12| 48.61|       |
| RMSE         | 2.24       | 1.53   | 1.07  | 1.14 | 2.09 | 1.01 |       |
| SSIM         | 0.9563     | 0.9843 | 0.9916| 0.9913| 0.9679| 0.9957|       |
| ERGAS        | 1.75       | 1.24   | 0.84  | 0.88 | 1.56 | 0.78 |       |
| UIQI         | 0.731      | 0.8165 | 0.8481| 0.8377| 0.711 | 0.8207|       |
| SAM          | 18.7       | 7.76   | 6.40  | 5.64 | 9.18 | 5.60 |       |
| Lemons       |           |        |       |      |      |      |       |
| PSNR         | 41.57      | 46.95  | 50.49 | 50.03| 43.81| 51.21|       |
| RMSE         | 2.57       | 1.08   | 0.81  | 0.86 | 1.68 | 0.81 |       |
| SSIM         | 0.928      | 0.9913 | 0.9955| 0.9946| 0.9767| 0.9958|       |
| ERGAS        | 1.74       | 2.97   | 0.56  | 0.6  | 1.14 | 0.54 |       |
| UIQI         | 0.6        | 0.7069 | 0.8548| 0.831 | 0.6876| 0.8644|       |
| SAM          | 15.67      | 3.81   | 2.84  | 3.03 | 5.56 | 2.74 |       |

TABLE IV
RUNTIME OF THE SIX METHODS

| Dataset     | Hysure | NLSTF | LTRM | LLTR | CSTF | DLST  |
|-------------|--------|-------|------|------|------|-------|
| Pavia University | 22     | 32    | 73   | 180  | 37   | 73    |
| CAVE        | 55     | 67    | 231  | 350  | 88   | 231   |
| Chikusei    | 95     | 109   | 367  | 450  | 112  | 327   |

V. CONCLUSION

In this study, to explore spatial and spectral information in individual regions of interest for hyperspectral superresolution reconstruction, a new method named DLST is proposed to divide an unknown HR-HSI into two components: 1) low-rank and 2) sparse components, focusing on the spectral low-rank property in the global region and the sparsity in the individual special regions, respectively. The low-rank component was constrained by the nonlocal constraint of adaptive manifold via TTR for balancing the local low-rank property and nonlocal self-similarity of manifold structure. Surface-aware regularization combined with GS was used to constrain the sparse component in the reconstruction. Experimental results using three datasets show that the new method outperforms the other five methods in spatial and spectral reconstruction. The DLST method can extract spatial information from HR-MSIs and restore spectral information in LR-HSIs.

REFERENCES

[1] W. Su and D. Sun, “Fourier transform infrared and raman and hyperspectral imaging techniques for quality determinations of powdery foods: A review,” Comprehensive Rev. Food Sci. Food Saf., vol. 17 no. 1, pp. 104–122, 2018.
[2] L. Mou, S. Saha, Y. Hua, F. Bovolo, L. Bruzzone, and X. Zhu, “Deep reinforcement learning for band selection in hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., vol. 60, 2022, Art. no. 5504414.
[3] Y. Xu, B. Du, and L. Zhang, “Self-attention context network: Addressing the threat of adversarial attacks for hyperspectral image classification,” IEEE Trans. Image Process., vol. 60, no. 5, pp. 8671–8685, 2021.
[4] A. A. Nielsen, “The regularized iteratively reweighted mad method for change detection in multi- and hyperspectral data,” IEEE Trans. Image Process., vol. 16, no. 2, pp. 463–478, Feb. 2007.
[5] X. Li, Z. Yuan, and Q. Wang, “Unsupervised deep noise modeling for hyperspectral image change detection,” Remote. Sens., vol. 11, 2019, Art. no. 258.
[6] Y. Zhang, W. Ke, B. Du, and X. Hu, “Independent encoding joint sparse representation and multitask learning for hyperspectral target detection,” IEEE Trans. Image Process., vol. 14, no. 11, pp. 8694–8696, Aug. 2017.
[7] Y. Shi, J. Li, Y. Li, and Q. Du, “Sensor-independent hyperspectral target detection with semisupervised domain adaptive few-shot learning,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 8, pp. 6894–6906, Aug. 2021.
[8] 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.
[9] M. Choi, “A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter,” IEEE Trans. Geosci. Remote Sens., vol. 44, no. 6, pp. 1672–1682, Jun. 2006.
[10] Q. Guo, M. Ehlers, Q. Wang, C. Pohl, S. Hornberg, and A. Li, “Ehlers pan-sharpening performance enhancement using HCS transform for n-band data sets,” Int. J. Remote Sens., vol. 38, pp. 4974–5002, 2017.
[11] W. Shi, D. Guo, and H. Zhang, “A reliable and adaptive spatiotemporal data fusion method for blending multi-spatiotemporal-resolution satellite images,” Remote Sens. Environ., vol. 268, Jan. 2022, Art. no. 112770.
[12] M. González-Audícana, J. L. Saleta, R. G. Catalán, and R. García, “Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition,” IEEE Trans. Geosci. Remote Sens., vol. 42, no. 6, pp. 1323–1335, May 2008.
[15] J.-L. Starck, M.-J. Fadili, and F. Murtagh, “The undecimated wavelet de-composition and its reconstruction,” IEEE Trans. Image Process., vol. 16, no. 2, pp. 297–309, Feb. 2007.

[16] J. Jayanth, T. A. Kumar, and S. Koliwad, “Fusion of multispectral and panchromatic data using regionally weighted principal component analysis and wavelet,” IEEE Trans. Geosci. Remote Sens., vol. 17, no. 4, pp. 651–655, Apr. 2020.

[17] T. Benzenati, Y. Kessentini, A. Kallel, and H. Hallabia, “Generalized Laplacian pyramid pan-sharpening gain injection prediction based on cnn,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 4, pp. 1806–1815, May 2014.

[18] J. Shen, Y. Zhao, W. Liao, and J. C.-W. Chan, “A pan-sharpening based on the non-subsampled contourlet transform: Application to worldview-2 imagery,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 7, no. 5, pp. 1117–1127, Sep. 2015.

[19] J. Xue, Y. Zhao, W. Liao, and J. C.-W. Chan, “Enhanced hyperspectral image fusion,” IEEE Trans. Cybern., vol. 53, no. 10, pp. 5174–5189, Jul. 2019.

[20] C. Lanaras, E. Baltsavias, and K. Schindler, “When Laplacian pyramid pan-sharpening gain injection prediction based on cnn,” IEEE Trans. Cybern., vol. 32, no. 7, pp. 3122–3135, Jul. 2021.

[21] P. Bianchi, W. Hachem, and F. Iutzeler, “A coordinate descent primal-dual algorithm and application to distributed asynchronous optimization,” IEEE Trans. Autom. Control, vol. 66, no. 5, pp. 2050–2065, May 2021.

[22] S. Bitterlich, R. I. Boˇk, and M. Simões, “A convex formulation for hyperspectral image super-resolution via subspace-based regularization,” Inf. Sci., vol. 405, 2022, Art. no. 113972.

[23] M. Simões, J. M. Bioucas-Dias, L. B. Almeida, and J. Chanussot, “Using non-local sparse tensor factorization for hyperspectral nonlinear unmixing,” IEEE Trans. Geosci. Remote Sens., vol. 55, no. 7, pp. 3823–3834, Jul. 2017.

[24] 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.

[25] R. Dian and S. Li, “Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization,” IEEE Trans. Image Process., vol. 28, no. 10, pp. 5135–5146, Oct. 2019.

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

[27] R. Dian, L. Fang, and S. Li, “Hyperspectral image super-resolution via non-local sparse tensor factorization,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 3862–3871.

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

[29] L. Gao, Z. Wang, L. Zhihuan, H. Yu, B. Zhang, and J. Chanussot, “Using low-rank representation of abundance maps and nonnegative tensor factorization for hyperspectral nonlinear unmixing,” IEEE Trans. Geosci. Remote Sens., vol. 60, 2022, Art. no. 5504017.

[30] F. Xiong, J. Zhou, and Y. tao Qian, “Hyperspectral restoration via l_0 gradient regularized low-rank tensor factorization,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 12, pp. 10410–10425, Dec. 2019.

[31] J. Xue, Y. Zhao, W. Liao, and J. C.-W. Chan, “Hyper-Laplacian regularized nonlocal low-rank matrix recovery for hyperspectral image compressive sensing reconstruction,” Inf. Sci., vol. 501, pp. 406–420, 2019.

[32] M. Simões, J. M. Bioucas-Dias, L. B. Almeida, and J. Chanussot, “Using low-rank representation of abundance maps and nonnegative tensor factorization for hyperspectral nonlinear unmixing,” IEEE Trans. Geosci. Remote Sens., vol. 55, no. 7, pp. 3823–3834, Jul. 2017.

[33] J. Xue, Y. Zhao, W. Liao, and J. C.-W. Chan, “Hyper-Laplacian regularized nonlocal low-rank matrix recovery for hyperspectral image compressive sensing reconstruction,” Inf. Sci., vol. 501, pp. 406–420, 2019.

[34] Y. Wang, X. Chen, Z. Han, and S. He, “Hyperspectral image super-resolution via nonlocal low-rank tensor approximation and total variation regularization,” Remote Sens., vol. 9, 2017, Art. no. 1286.

[35] R. Dian, S. Li, and L. Fang, “Learning a low tensor-train rank representation for hyperspectral image super-resolution,” IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 9, pp. 2672–2683, Sep. 2019.

[36] J. Xue, Y.-Q. Zhao, Y. Bu, W. Liao, J. C.-W. Chan, and W. Philips, “Spatial-spectral structured sparse low-rank representation for hyperspectral image super-resolution,” IEEE Trans. Image Process., vol. 30, pp. 3084–3097, 2021.

[37] J. Yang, L. Xiao, Y. Zhao, and J. C.-W. Chan, “Variational regularization network with attentive deep prior for hyperspectral–multispectral image fusion,” IEEE Trans. Geosci. Remote Sens., vol. 60, 2022, Art. no. 5508817.

[38] J. Yang, L. Xiao, Y. Zhao, and J. Chan, “Hybrid local and nonlocal 3-D attentive CNN for hyperspectral image super-resolution,” IEEE Geosci. Remote Sens. Lett., vol. 18, no. 7, pp. 1274–1278, Jul. 2021.

[39] 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.

[40] R. Dian, S. Li, and X. Kang, “Regularizing hyperspectral and multispectral image fusion by CNN denoiser,” IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 3, pp. 1124–1135, Mar. 2021.

[41] J. Jayanth, T. A. Kumar, and S. Koliwad, “Fusion of multispectral and panchromatic data using regionally weighted principal component analysis and wavelet,” IEEE Trans. Geosci. Remote Sens., vol. 17, no. 4, pp. 651–655, Apr. 2020.

[42] T. Benzenati, Y. Kessentini, A. Kallel, and H. Hallabia, “Generalized Laplacian pyramid pan-sharpening gain injection prediction based on cnn,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 4, pp. 651–655, Apr. 2020.

[43] M. C. El-Mezouar, K. Kpalma, N. Taleb, and J. Ronson, “A pan-sharpening based on the non-subsampled contourlet transform: Application to worldview-2 imagery,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 7, no. 5, pp. 1806–1815, May 2014.

[44] H. Irmak, G. B. Akar, and S. E. Yüksel, “A map-based approach for hyperspectral imagery super-resolution,” IEEE Trans. Image Process., vol. 27, no. 6, pp. 2942–2951, Jun. 2018.
[61] S. C. Satapathy, J. K. Mandal, S. K. Udga, and V. Bhateja, “Information systems design and intelligent applications,” in Advances in Intelligent Systems and Computing. Springer, 2019.
[62] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.
[63] P. E. Dennison, K. Halligan, and D. A. Roberts, “A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper,” Remote Sens. Environ., vol. 93, pp. 359–367, 2004.

Huajing Wu received the B.Sc. and M.S. degrees in surveying engineering from the Faculty of Land Resources Engineering, Kunming University of Science and Technology, Kunming, China, in 2017 and 2020, respectively. She is currently working toward the Ph.D. degree in geodesy and surveying engineering with the School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, China.

Her research interests include hyperspectral image processing and remote sensing data fusion.

Kefei Zhang received the Ph.D. degree in geodesy and surveying engineering from Curtin University, Bentley, WA, Australia, in 1998.

He is currently a Chair Professor and the Director of the Institute of Resources and Environment Technologies, University of Mining and Technology (CUMT), Beijing, China. He is also an honorary Professor with RMIT University, Melbourne, VIC, Australia, where he is the Founding Director of the Satellite Positioning for Atmosphere, Climate and Environment (SPACE) Research Centre. He has more than 30 years of research experience in geodesy, satellite positioning, and geospatial sciences in general. He has authored about 400 peer-reviewed publications in these fields and a co-inventor of 30 patents, and attracted in excess of 50 million dollars in funding from the Australian Research Council, the National Science Foundation of China, as well as funds from other national and international governments and industries. He is an Australian pioneer in cutting-edge technologies for smart tracking, GNSS atmospheric sounding for weather and climate, and space environment management, etc. His satellite-to-satellite tracking frontier research has led to more than 10-h weather forecast improvement and successful integration of the GPS radio occultation data into the Australian weather forecasting system in 2012. He currently leads a team of more than 40 researchers at CUMT. His current research interests primarily include algorithm development and innovative applications of satellite technologies for high-accuracy positioning, atmospheric studies, space situational awareness, and space resources utilization.

Prof. Zhang is a Fellow of The International Association of Geodesy (IAG).

Suqin Wu received the Ph.D. degree in geodesy and surveying engineering from RMIT University, Melbourne, VIC, Australia, in 2009.

She is currently a Professor with the China University of Mining and Technology, Beijing, China. Her research interests mainly include GNSS precise positioning and GNSS meteorology, GPS radio occultation, atmospheric error modeling for high accuracy GNSS applications, and precise orbit determination and prediction for space debris objects, etc.

Minghao Zhang received the B.Sc. degree in surveying engineering in 2020 from the China University of Mining and Technology, Xuzhou, China, where he is currently working toward the Ph.D. degree in geodesy and surveying engineering with the School of Environment Science and Spatial Informatics.

His research interests include Global Navigation Satellite System Meteorology (GNSS/MET).

Shuangshuang Shi received the M.S. degree in geodesy and surveying engineering from the School of Geodesy and Geomatics, Wuhan University, Wuhan, China, in 2015. She is currently working toward the Ph.D. degree in geodesy and surveying engineering with the School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, China.

Her research interests include global ionospheric modeling using multisource geodesy observations, and ionospheric anomaly analysis under abnormal conditions.