Tensor Completion via Nonlocal Low-Rank Regularization

Ting Xie, *Student Member, IEEE*, Shutao Li, *Senior Member, IEEE*, Leyuan Fang, *Senior Member, IEEE*, and Licheng Liu, *Member, IEEE*

***Abstract*—Tensor completion (TC), aiming to recover original high-order data from its degraded observations, has recently drawn much attention in hyperspectral images (HSIs) domain. Generally, the widely used TC methods formulate the rank minimization problem with a convex trace norm penalty, which shrinks all singular values equally, and may generate a much biased solution. Besides, these TC methods assume the whole high-order data is of low-rank, which may fail to recover the detail information in high-order data with diverse and complex structures. In this paper, a novel nonlocal low-rank regularization-based TC (NLRR-TC) method is proposed for HSIs, which includes two main steps. In the first step, an initial completion result is generated by the proposed low-rank regularization-based TC (LRR-TC) model, which combines the logarithm of the determinant with the tensor trace norm. This model can more effectively approximate the tensor rank, since the logarithm function values can be adaptively tuned for each input. In the second step, the nonlocal spatial-spectral similarity is integrated into the LRR-TC model, to obtain the final completion result. Specifically, the initial completion result is first divided into groups of nonlocal similar cubes (each group forms a 3-D tensor), and then the LRR-TC is applied to each group. Since similar cubes within each group contain similar structures, each 3-D tensor should have low-rank property, and thus further improves the completion result. Experimental results demonstrate that the proposed NLRR-TC method outperforms state-of-the-art HSIs completion techniques.**

***Index Terms*—Hyperspectral image (HSI), low-rank approximation, nonlocal strategy, tensor completion (TC).**

I. Introduction

GIVEN an observed incomplete image, the objective of image completion is to fill its missing part so that a plausible outcome is obtained. To achieve this objective, various methods have been designed, including the partial differential equations [1], total variation [2], and belief propagation [3]. Since these methods do not fully capture the global correlations hidden in data, their completion results are usually not so satisfying. Recently, the rank minimization has demonstrated to be a powerful global constraint [4]–[7], which has an extensive range of applications in computer vision [8], [9] and signal processing [10].

In the matrix case, since direct rank minimization is generally a nondeterministic polynomial hard problem [11], some literatures [12], [13] have adopted the trace norm as a convex proxy to approximate the matrix rank, and the researches in [14] and [15] have shown that the matrix rank minimization problem can be solved using a convex relaxation technique under certain linear constraints. Despite the theoretical soundness, the trace norm penalty over-penalizes large singular values, and may make the solution deviate from the original solution [16], [17]. To make a closer approximation to the matrix rank, some nonconvex researches, e.g., log-determinate function [18]–[20], weighted nuclear norm [21], tractable Schatten norms [16], [17], and iterative reweighting algorithm [22], [23], have been proposed. However, the log-determinate function in researches [18]–[20] are quite limited in application and they are designed for specific application. In addition, the matrix can only well address binary-factor variability of data, which has the limitation for extracting information from a multidimensional perspective.

Instead of matrices, high-order tensors have attracted much attention [24]–[27], since they can accurately reveal more complicated intrinsic structures underlying multidimensional data. In [28]–[30], low-rank tensor completion (LR-TC) has been applied to hyperspectral image (HSI) completion problem. In general, the current development of LR-TC methods can be roughly classified into two categories: 1) CANDECOMP/PARAFAC (CP) decomposition-based [31] and 2) Tucker decomposition-based [32].

For the CP decomposition-based LR-TC approaches [28], [33], [34], a tensor is first factorized into a sum of component rank-1 factors, and then a rank minimization constraint is imposed upon the factors. In specific, Liu *et al.* [28] have proposed a factor matrix rank minimization method for TC. In [33], the CP model has been formulated as a weighted least squares method. Although these approaches require less storage space, the CP decomposition-based LR-TC methods require a predefined rank. In fact, determining or even bounding the CP rank of an arbitrary tensor remains a difficult task [35]–[37], especially for the uncomplete data, and few researches have

Manuscript received December 16, 2017; revised April 1, 2018; accepted April 4, 2018. This work was supported by the National Natural Science Fund of China for Distinguished Young Scholars under Grant 61325007, and International Cooperation and Exchanges under Grant 61520106001. This paper was recommended by Associate Editor S. Cruces. *(Corresponding author: Shutao Li.)*

The authors are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: xting@hnu.edu.cn; shutao_li@hnu.edu.cn; fangleyuan@gmail.com; lichenghnu@gmail.com).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the author. The material is 3.58 MB in size.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2018.2825598

been presented to solve this problem. Zhao *et al.* [38], [39] have introduced the fully Bayesian CP factorization (FBCP) method to determine the CP rank by Bayesian inference. However, these methods are based on hierarchical probabilistic framework, which generally shows slow convergence and has complex computation.

For the Tucker decomposition-based LR-TC approaches [37], [40]–[42], researchers first transform the tensor into matrices, and then solve the matrix rank minimization problem. Tensor trace norm (TTN) [40], [41] has been introduced to approximate the tensor rank, which is defined as the weighted sum of all the trace norm of matrices unfolded along each mode. Meanwhile, many algorithms have been proposed to solve the TTN minimization problem by using the alternative direction multiplier method (ADMM) [40], Douglas–Rachford splitting technique [41], and a fixed point iterative method [42], etc. In addition, Mu *et al.* [37] have proposed a square deal model to approximate the tensor rank. In a word, studies in [37] and [40]–[42] formulate the tensor rank minimization problem with a convex trace norm minimization. Nevertheless, nonconvex researches [29], [43], [44] have shown that they outperform those based on conventional trace norm methods in tensor recovery. Moreover, the LR-TC methods mentioned above directly perform the low-rank constraint on the whole data to obtain the final completion result. By this way, the detail information may not be well recovered in high-order data with diverse and complex structures, especially in low sampling rate (SR).

To tackle the problem of the trace norm penalty shrinks each singular value equally in tensor rank approximation, and better recover the detail information in high-order data with diverse and complex structures, a novel nonlocal low-rank regularization-based TC (NLRR-TC) method is proposed for HSI completion, which consists of two steps. In the first step, the low-rank regularization-based TC (LRR-TC) model is introduced into the whole degraded tensor data, providing an initial completion result. The LRR-TC combines the logarithm of the determinant and the TTN, which can treat each singular value adaptively. In the second step, nonlocal spatial-spectral similarity is integrated with the LRR-TC model, to obtain the final completion result. More specifically, similar cubes in the initial completion result are first clustered into 3-D tensors, and then the LRR-TC is applied to each grouped tensor. Since similar cubes within each group contain similar structures, each 3-D tensor should have low-rank property. Applying the LRR-TC to each grouped tensor, the diverse and complex structural information can be better recovered. The contributions of this paper are summarized as follows.

1) We propose a novel nonconvex LRR-TC to effectively approximate the tensor rank. The logarithm of the trace norm can be skillfully transferred into the weighted trace norm, whose minimization can be efficiently solved.
2) We propose a nonlocal TC strategy, which can effectively utilize the nonlocal spatial-spectral similarity within HSI.
3) A tensor low-rank regularization (TLRR) algorithm based on ADMM is introduced to efficiently solve the LRR-TC, and the convergence of this algorithm is also analyzed. Experiments on HSIs completion show that the proposed method achieves the state-of-the-art performance among various quality assessments.

The rest of this paper is structured as follows. Sections II and III briefly overview preliminaries and related work. Section IV introduces the proposed NLRR-TC method for HSI completion. Experimental results and analysis are reported in Section V. Finally, the conclusions and future work are given in Section VI.

II. Notations and Preliminaries

In the rest of this paper, scalars, vectors, matrices, and tensors are denoted by nonbold lowercase letters (e.g., a), bold lowercase letters (e.g., $\mathbf{a}$), uppercase letters (e.g., A), and Euler script letters (e.g., $\mathcal{A}$), respectively. In addition, the trace norm of a matrix A is denoted as $\|A\|_*$, which is defined as the sum of the singular values of a matrix A, i.e., $\|A\|_* = \sum_i \sigma_i(A)$.

An N-dimensional tensor is denoted as $\mathcal{X} \in R^{I_1 \times I_2 \times \cdots \times I_N}$. Elements of $\mathcal{X}$ are denoted as $x_{i_1 \cdots i_k \cdots i_N}$, where $1 \le i_k \le I_k$ and $1 \le k \le N$, while the order of a tensor is the number of dimensions, also known as ways or modes. A fiber of a tensor $\mathcal{X}$ is defined by fixing all indices of $\mathcal{X}$ but one, i.e., a mode-n fibers of $\mathcal{X}$ are all vectors $\mathbf{x}_{i_1 \cdots i_{n-1} \, : \, i_{n+1} \cdots i_N}$ obtained by varying index i_n while keeping the others fixed, and a slice of $\mathcal{X}$ is defined by fixing all indices but two. The mode-n unfolding of tensor $\mathcal{X}$ is denoted by $X_{(n)} = \text{unfold}_{(n)}(\mathcal{X}) \in R^{I_n \times \prod_{j \ne n} I_j}$, which is obtained by arranging the mode-n fibers of $\mathcal{X}$ as its column vectors. The inverse operation "fold" is defined as $\text{fold}_n(X_{(n)}) = \mathcal{X}$.

There are two widely used definitions for the tensor rank. One is the Tucker rank [32] characterized by unfolded matrices along each mode, i.e.,

$$\text{rank}_{\text{tc}}(\mathcal{X}) = \big(\text{rank}\big(X_{(1)}\big), \ldots, \text{rank}\big(X_{(n)}\big), \ldots, \text{rank}\big(X_{(N)}\big)\big) \tag{1}$$

where $X_{(n)} \in R^{I_n \times \prod_{j \ne n} I_j}$ is the mode-n unfolding of $\mathcal{X}$. The other is CP rank [45], which is defined as the minimum number of rank-1 decomposition, formulated as

$$\text{rank}_{\text{cp}}(\mathcal{X}) = \min\left\{ r | \mathcal{X} = \sum\nolimits_{i=1}^{r} \boldsymbol{a}_1^{(i)} \circ \cdots \circ \boldsymbol{a}_j^{(i)} \circ \cdots \circ \boldsymbol{a}_N^{(i)} \right\} \tag{2}$$

where $\circ$ represents the vector outer product and $\boldsymbol{a}_j^{(i)} \in R^{I_j}, j = 1, \ldots, N$.

The inner product of two tensors $\mathcal{X}, \mathcal{Y} \in R^{I_1 \times \cdots \times I_N}$ is calculated by sum of the product of their entries, i.e., $\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1=1}^{n_1} \cdots \sum_{i_N=1}^{n_N} x_{i_1 \cdots i_N} y_{i_1 \cdots i_N}$. The Frobenius norm of a tensor $\mathcal{X}$ is defined as $\|\mathcal{X}\|_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle}$ and the L_1 norm is defined as $\|\mathcal{X}\|_1 = \sum_{i_1=1}^{n_1} \cdots \sum_{i_N=1}^{n_N} |x_{i_1 \cdots i_N}|$. More details about tensors can be found in [45].

III. Related Work

The general model of TC is to minimize the tensor rank function under limited sample constraints, i.e.,

$$\begin{aligned} &\min_{\mathcal{X}} : \text{rank}(\mathcal{X}) \\ &\text{s.t.,} \quad \mathcal{X}_\Omega = \mathcal{T}_\Omega \end{aligned} \tag{3}$$

where Ω is an index set of the same size as $\mathcal{T}$, and zeros in Ω indicate the missing entries in $\mathcal{T}$. This constraint reveals that the value of the estimated tensor $\mathcal{X}$ is the same as $\mathcal{T}$ in the observed entries. To address the tensor rank minimization problem (3), there are roughly two typical LR-TC methods: CP decomposition-based [31], and Tucker decomposition-based [32], respectively. Details of these two methods are presented below.

A. CP Decomposition-Based Tensor Completion

Given a general tensor $\mathcal{X} \in R^{I_1 \times I_2 \times \cdots \times I_N}$ with tensor rank r, the CP decomposition can be expressed as the following form:

$$\mathcal{X} = X_1 \circ X_2 \cdots \circ X_N = \sum_{i=1}^{r} \boldsymbol{x}_1^{(i)} \circ \boldsymbol{x}_2^{(i)} \circ \cdots \circ \boldsymbol{x}_N^{(i)} \tag{4}$$

where $\circ$ represents the vector outer product and $X_n = [\boldsymbol{x}_n^{(1)}, \boldsymbol{x}_n^{(2)} \cdots, \boldsymbol{x}_n^{(r)}] \in R^{(I_n \times r)}, n = 1, \ldots, N$ represents the factor matrix of $\mathcal{X}$. Liu *et al.* [28] have attempted to complete the low-rank tensor $\mathcal{X}$ by solving the following optimization problem:

$$\min_{\mathcal{X}, X_n} \sum_{n=1}^{N} \alpha_n \|X_n\|_* + \frac{\lambda}{2} \|\mathcal{X} - X_1 \circ X_2 \cdots \circ X_N\|_F^2 \quad \text{s.t.,} \ \ \mathcal{X}_\Omega = \mathcal{T}_\Omega \tag{5}$$

where $X_n \in R^{(I_n \times R)}, n = 1, \ldots, N$, the coefficients $\{\alpha_n\}$ are prespecified constants, and R is the upper bound of tensor rank r. Suppose the elements of $\mathcal{A}$ in the set Ω are sampled, work in [34] has proposed a parallel method to seek the components $\boldsymbol{x}_n^{(i)}$ in (4) for a tensor $\mathcal{X}$ of rank r, by solving the optimization problem

$$\min_{\boldsymbol{x}_n^{(i)}} \left\| P_\Omega \left(\mathcal{A} - \sum_{i=1}^{r} \boldsymbol{x}_1^{(i)} \circ \boldsymbol{x}_2^{(i)} \circ \cdots \circ \boldsymbol{x}_N^{(i)} \right) \right\|_2^2 + \lambda \sum_{i=1}^{r} \sum_{n=1}^{N} \left\| \boldsymbol{x}_n^{(i)} \right\|_2^2 \tag{6}$$

where λ is the regularization parameter. In general, determining or even bounding the CP rank r of an arbitrary tensor remains a difficult problem [35]–[37], especially for the uncomplete tensor.

B. Tucker Decomposition-Based Tensor Completion

Liu *et al.* [40] have introduced TC based on the minimization of Tucker rank defined in (1). More specifically, they use the nuclear norm as the convex relaxation for the matrix rank as follows:

$$\min_{\mathcal{X}} \sum_{i=1}^{N} w_i \left\| X_{(i)} \right\|_*, \quad \text{s.t.,} \ \ \mathcal{X}_\Omega = \mathcal{T}_\Omega \tag{7}$$

where the coefficients $\{w_i\}$ are constants meeting $w_i \geq 0$. Moreover, three efficient algorithms have been proposed in [40] to solve problem (7). In addition, Mu *et al.* [37] have proposed the square deal model to recover $\mathcal{X} \in R^{I_1 \times I_2 \times \cdots \times I_N}$. The square deal is the nuclear norm of a squarer matrix $X_{[j]}$ and its objective function is

$$\min_{\mathcal{X}} \left\| X_{[j]} \right\|_*, \quad \text{s.t.,} \ \ \mathcal{X}_\Omega = \mathcal{T}_\Omega \tag{8}$$

where $X_{[j]}$ is defined as

$$X_{[j]} = \text{reshape}\left(X_{(1)}, \prod_{i \leq j} I_i, \prod_{i > j} I_i \right) \tag{9}$$

and $j \in \{1, 2, \ldots, N\}$ aims to make $\prod_{i \leq j} I_i$ as close to $\prod_{i > j} I_i$ as possible. When the involved tensor order is no more than three, this square deal model may not deliver good performance [28]. The studies in [37] and [40] formulate the tensor rank minimization problem with a convex trace norm minimization. Nevertheless, the trace norm penalty makes all singular values shrink equally, and a solution deviated significantly from the truth may be generated [16], [17]. In addition, these methods assume the whole tensor data is of low-rank. However, it only approximates to low-rank for the high-order data with diverse and complex structures [40], [46]. In such cases, directly performing the low-rank constraint on the whole data to get the final completion result may not well recover the detail information, especially in low SR.

IV. Proposed NLRR-TC Approach

In this section, the NLRR-TC is presented for HSI completion, which can adaptively shrink each singular value of unfolded matrices, and effectively exploit the nonlocal spatial self-similarity within HSI. It consists of two main steps: 1) LRR-TC-based initial completion and 2) LRR-TC-based final completion. In the first step, the LRR-TC is introduced to the whole degraded tensor data to provide an initial completion result. The LRR-TC model combines the logarithm of the determinant with the TTN to approximate the tensor rank. Since the logarithm function values can be adaptively tuned for each input, the integration of the logarithm of the determinant with the TTN can effectively approximate the tensor rank. In the second step, the nonlocal spatial-spectral similarity is integrated with the LRR-TC model, to obtain the final completion result. More specifically, according to the high spatial-spectral similarity within HSI, the initial completion result is first divided into groups of nonlocal similar cubes, and each group forms a 3-D tensor, then the LRR-TC is applied to each formed 3-D tensor. Since similar cubes within each group contain similar structures, each 3-D tensor should have low-rank property. By applying the LRR-TC to each grouped tensor, one can better recover the diverse and complex structural information.

The proposed NLRR-TC method is demonstrated in Fig. 1, and will be described in the following three sections: 1) the LRR-TC model; 2) the TLRR algorithm to solve LRR-TC; and 3) nonlocal clustering for TC.

A. LRR-TC Model

In general, recent models [40], [41] utilize the TTN as a convex relaxation of the tensor rank, and the tensor $\mathcal{X}$ can be recovered by solving the following optimization problem:

$$\mathcal{X} = \underset{\mathcal{X}}{\text{argmin}} \sum_{i=1}^{N} w_i \left\| X_{(i)} \right\|_* \quad \text{s.t.,} \ \ \mathcal{X}_\Omega = \mathcal{T}_\Omega. \tag{10}$$

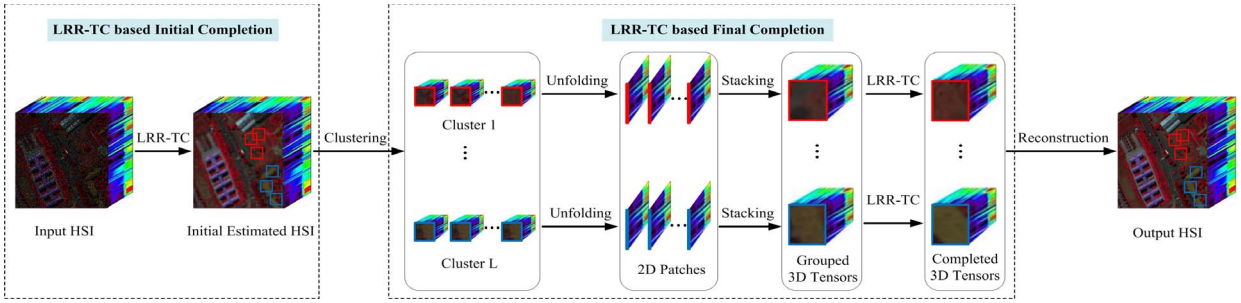


Fig. 1. Framework of the proposed NLRR-TC method.

However, TTN formulates the tensor rank minimization problem with a convex trace norm penalty, which may yield a biased solution. To make a closer approximation to the tensor rank, the LRR-TC model is introduced that combines the logarithm of the determinant with the TTN, and is defined as

$$\begin{aligned} &\min_{\mathcal{X}} : \sum_{i=1}^{N} \alpha_i R(X_{(i)}, \varepsilon) \\ &\text{s.t.,} \quad \mathcal{X}_\Omega = \mathcal{T}_\Omega \end{aligned} \tag{11}$$

where $X_{(i)}$ is the mode-i unfolding of $\mathcal{X}$ and the coefficients $\{\alpha_i\}$ are constants meeting minimal value of $\{\alpha_i\}$ is 1. $\mathrm{R}(X, \varepsilon) = \sum_j \log(\sigma_j(X) + \varepsilon)$, $\sigma_j(X)$ represents the jth singular value of matrix X, and ε is a small positive number.

Let $g(\sigma, \varepsilon) = \log(\sigma + \varepsilon)$. The function value becomes gentler as the singular value gets larger, and each singular value can be treated adaptively. More specifically, $g(\sigma, \varepsilon)$ can be approximated using the first-order Taylor expansion as $g(\sigma, \varepsilon) = g(\sigma^k, \varepsilon) + \langle \nabla g(\sigma^k, \varepsilon), \sigma - \sigma^k \rangle$, where σ^k is the solution obtained in the kth iteration. Constants in the first-order Taylor expansion can be ignored since they do not affect the minimization problem. The solution of $\min_\sigma g(\sigma, \varepsilon)$ can be approximated by that of $\min_\sigma(\sigma/(\sigma^k + \varepsilon))$. Therefore, the solution of $\min \mathrm{R}(X, \varepsilon)$ can be approximated by the solution of $\min \sum_j (\sigma_j(X)/(\sigma_j^k(X) + \varepsilon))$. As can be seen, for the larger singular value input, the logarithm function will shrink the singular value less, and thus more meaningful structural information can be preserved. By contrast, for the smaller one, the logarithm function will shrink the singular value more. Hence, the LRR-TC model can effectively approximate the tensor rank, leading to a better TC result.

Let $\mathrm{R}(X)$ represent $\mathrm{R}(X, \varepsilon)$. Therefore, instead of solving (3), our aim is then changed to solve the following minimization for TC:

$$\begin{aligned} &\min_{\mathcal{X}} : \sum_{i=1}^{N} \alpha_i R\big(X_{(i)}\big) \\ &\text{s.t.,} \quad \mathcal{X}_\Omega = \mathcal{T}_\Omega \end{aligned} \tag{12}$$

In the next section, a TLRR algorithm is introduced to efficiently solve the objective function (12).

B. TLRR Algorithm to Solve LRR-TC Model

Recently, it has been shown in [28] and [47] that the ADMM [48] is very efficient to handle large scale problems and solve optimization problems with multiple objective items. In this paper, we present the TLRR algorithm for solving the proposed model (12) by using the ADMM technique. First, we introduce N auxiliary tensors $\mathcal{M}_i (1 \le i \le N)$ and obtain the following equivalent formulation:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{M}_1, \ldots, \mathcal{M}_N} \quad & \sum_{i=1}^{N} \alpha_i R\big(M_{i(i)}\big) \\ \text{s.t.,} \quad & \mathcal{X}_\Omega = \mathcal{T}_\Omega, \mathcal{X} = \mathcal{M}_i, \ 1 \le i \le N \end{aligned} \tag{13}$$

where $M_{i(i)} = \mathrm{unfold}_i(\mathcal{M}_i)$. Then, the ADMM is applied to solve (13). The advantage of ADMM is that we can divide (13) into two subproblems that both have optimal solutions. By applying the ADMM to (13), we obtain

$$\begin{aligned} &L_\mu(\mathcal{X}, \mathcal{M}_1, \ldots, \mathcal{M}_N, y_1, \ldots, y_N) \\ &\quad = \sum_{i=1}^{N} \alpha_i \mathrm{R}\big(M_{i(i)}\big) + \big\langle \mathcal{X} - \mathcal{M}_i, y_i \big\rangle + \frac{\mu}{2} \|\mathcal{X} - \mathcal{M}_i\|_F^2 \end{aligned} \tag{14}$$

where μ is a positive scalar, and $y_i (1 \le i \le N)$ are the Lagrange multipliers. The problem (14) can be solved with the following iterations:

$$\begin{aligned} &\left\{\mathcal{M}_1^{k+1}, \ldots, \mathcal{M}_i^{k+1}, \mathcal{M}_N^{k+1}\right\} \\ &\quad = \mathrm{argmin}_{\mathcal{M}_1, \ldots, \mathcal{M}_i, \ldots, \mathcal{M}_N} : \alpha_i R\big(M_{i(i)}\big) \\ &\qquad + \frac{\mu^k}{2} \left\| \mathcal{X}^k + \frac{1}{\mu^k} y_i^k - \mathcal{M}_i \right\|_F^2 \\ &\mathcal{X}^{k+1} = \mathrm{argmin}_{\mathcal{X}} \sum_{i=1}^{N} \left\| \mathcal{X} + \frac{1}{\mu^k} y_i^k - \mathcal{M}_i^{k+1} \right\|_F^2 \\ &y_i^{k+1} = y_i^k - \mu^k \left(\mathcal{M}_i^{k+1} - \mathcal{X}^{k+1} \right) \\ &\mu^{k+1} = t * \mu^k \end{aligned} \tag{15}$$

where $t > 1$, denoted as the step length. With $\mathcal{M}_j^{k+1} (j \ne i)$ and other parameters fixed, $\mathcal{M}_i^{k+1}$ can be updated by solving the following problem:

$$\min_{\mathcal{M}_i} a_i R\big(M_{i(i)}\big) + \frac{1}{2} \left\| \mathcal{X}^k + \frac{1}{\mu^k} y_i^k - \mathcal{M}_i \right\|_F^2 \tag{16}$$

Algorithm 1 TLRR

Input: Degraded tensor $\mathcal{T}$, index set Ω, Lagrange multipliers $y_i (i = 1, \ldots, N)$, positive scale μ, and constant t.
Output: Completed tensor $\mathcal{X}$.
Initialize: Number of iterations $k = 0$, $\mathcal{X}_\Omega = \mathcal{T}_\Omega$, $\mathcal{X}_{\overline{\Omega}} = 0$, and $y_i = 0 (i = 1, \ldots, N)$.
while not converged **do**
(a) Compute $\mathcal{M}_i^{k+1}$, $i = 1, \ldots, N$, via Eq. (18).
(b) Compute $\mathcal{X}^{k+1}$ via Eq. (20).
(c) $y_i^{k+1} = y_i^k - \mu^k(\mathcal{M}_i^{k+1} - \mathcal{X}^{k+1})$, $i = 1, \ldots, N$.
(d) $\mu^{k+1} = t\mu^k$.
end while

where $a_i = (\alpha_i/\mu^k)$. By using the first-order Taylor expansion to approximate the $\mathrm{R}(M_{i(i)})$, (16) can be transform into

$$\min_{M_{i(i)}} a_i \sum\nolimits_{I_i} \frac{\sigma_{I_i}\left(M_{i(i)}\right)}{\sigma_{I_i}\left(M_{i(i)}^k\right) + \varepsilon} + \frac{1}{2}\left\| X_{(i)}^k + \frac{1}{\mu^k} Y_{i(i)}^k - M_{i(i)} \right\|_F^2. \tag{17}$$

This minimization problem (17) can be solved in a closed-form [21], and $\mathcal{M}_i^{k+1}$ can be updated by the following:

$$\mathcal{M}_i^{k+1} = \mathrm{fold}_i\left(V_1 \Sigma_{a_i} V_2^T\right) \tag{18}$$

where $\Sigma_{a_i} = \mathrm{diag}(S_{a_i,\varepsilon}(\sigma_1), \ldots, S_{a_i,\varepsilon}(\sigma_{I_i}))$ and $V_1 \mathrm{diag}(\sigma_1, \ldots, \sigma_{I_i}) V_2^T$ is the SVD of $\mathrm{unfold}_i(\mathcal{X}^k + (1/\mu^k) y_i^k)$. Here, $S_{b,\varepsilon}(\cdot)$ is the singular value thresholding operator, which is defined as

$$S_{b,\varepsilon}(x) = \begin{cases} 0 & \text{if } r_2 \le 0 \\ \mathrm{sign}(x)\left(\frac{r_1 + \sqrt{r_2}}{2}\right) & \text{if } r_2 > 0 \end{cases} \tag{19}$$

with that $r_1 = |x| - \varepsilon$, and $r_2 = (r_1)^2 - 4(b - \varepsilon|x|)$. Similarly, for fixed $\mathcal{M}_i^{k+1}$s, y_i^k and μ^k, $\mathcal{X}^{k+1}$ admits a closed-form solution

$$\mathcal{X}^{k+1} = \mathcal{P}_\Omega(\mathcal{T}) + \mathcal{P}_{\overline{\Omega}}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\mathcal{M}_i^{k+1} - \frac{1}{\mu^k} y_i^k\right)\right) \tag{20}$$

where $\overline{\Omega}$ is the complements of Ω, and $\mathcal{P}_\Omega$ is the projection operator of Ω, formulated as

$$\mathcal{P}_\Omega(\mathcal{T}) = \begin{cases} \mathcal{T}_{i_1, i_2, \ldots, i_N} & \text{if } (i_1, i_2, \ldots, i_N) \in \Omega \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

Such process is iterated until convergence, and the TLRR algorithm to solve the proposed LRR-TC is summarized in Algorithm 1.

Theorem 1: In general, the arbitrary sequence $\{\mathcal{X}^k\} \in R^{I_1 \times I_2 \times \cdots \times I_N}$ generated by the TLRR algorithm satisfies

$$\lim_{k\to\infty} \left\| \mathcal{X}^{k+1} - \mathcal{X}^k \right\|_F = 0$$
$$\lim_{k\to\infty} \left\| \mathcal{M}_i^k - \mathcal{X}^k \right\|_F = 0, \forall i = 1, 2, \ldots, N. \tag{22}$$

The detailed proof of Theorem 1 is given in the supplementary file. It is worth noting that the proof of Theorem 1 relies on $\mu > 0$ and the unboundedness of the positive scalar μ^k.

C. Nonlocal Clustering for Tensor Completion

Considering an initial completion HSI $\mathcal{X} \in R^{I_1 \times I_2 \times I_3}$, we first partition it into overlapped cubes with the spatial size of $W \times W$ and full spectrum size. These cubes are denoted as $\{C_{mn}\} \in R^{W \times W \times I_3}$, where $m = ((I_1 - W/W - p) + 1)$, $n = ((I_2 - W/W - p) + 1)$, and p is the size of overlaps. A group of 2-D patches $\{P_{mn}\} \in R^{W^2 \times I_3}$ is constructed to represent $\mathcal{X}$, by reshaping each band of a cube into a 1-D column vector. All cubes can be reformulated as a group of 2-D full band patches $\varphi = \{Y_i \in R^{W^2 \times I_3}\}_{i=1}^{I}$, where $I = m \times n$. The efficient K-means++ method [49] is employed to obtain clusters $\mathcal{X}^l \in R^{W^2 \times I_3 \times l^n}$, $l = 1, 2, \ldots, L$, where L is the number of clusters and l^n is the number of nonlocal similar 2-D full band patches in the lth cluster. Clusters of the index set Ω cubes are also built with the same spatial structure, denoted as $\Omega^l \in R^{W^2 \times I_3 \times l^n}$, $l = 1, 2, \ldots, L$, where L is the number of clusters and the spatial locations of Ω^l are the same as $\mathcal{X}^l$. In this way, both the spatial and spectral correlations are well preserved in such representation along with its spectral and nonlocal-similar-cube-number modes.

After nonlocal clustering of the similar cubes, the LRR-TC is applied to each grouped 3-D tensor $\mathcal{X}^l \in R^{W^2 \times I_3 \times l^n}$, $l = 1, 2, \ldots, L$. With the high similarity in the grouped cubes, the formed 3-D tensors $\mathcal{X}^l \in R^{W^2 \times I_3 \times l^n}$, $l = 1, 2, \ldots, L$, should have low-rank property and can be recovered by solving the following optimization problem:

$$\min_{\hat{\mathcal{X}}^l} : \sum_{i=1}^{3} \alpha_i R\left(\hat{X}_{(i)}^l\right)$$
$$\text{s.t.,} \quad \left(\hat{\mathcal{X}}^l\right)_{\Omega^l} = \left(\mathcal{X}^l\right)_{\Omega^l}, l = 1, 2, \ldots, L. \tag{23}$$

where Ω^l is an index set of the same size as $\mathcal{X}^l$. All reconstructed 3-D tensors $\hat{\mathcal{X}}^l$, $l = 1, 2, \ldots, L$ can be returned to the original position to reconstruct the completed HSI $\hat{\mathcal{X}}$. The proposed NLRR-TC is summarized in Algorithm 2, which mainly consists of two steps. In step 1, the LRR-TC is introduced into the whole degraded tensor to provide an initial completion result. In step 2, the nonlocal clustering process is first conducted to group the 3-D tensors, and then, the LRR-TC is applied to each grouped tensor.

V. Experimental Results

A. Experimental Setup

In this section, the effectiveness of the proposed NLRR-TC method in HSIs completion is evaluated. Meanwhile, to verify the effectiveness of the proposed LRR-TC model that combines the logarithm of the determinant with the TTN to approximate the tensor rank, LRR-TC-based initial completion result in NLRR-TC method, denoted as LRR-TC, is also included for comparison. Both natural hyperspectral data and remote sensing hyperspectral data are used in our experiments. In addition, we uniformly generate the index set Ω at Gaussian random distribution, and define the SR as $\mathrm{SR} = (|\Omega|/(I_1 \times I_2 \times I_3))$, where $|\Omega|$ represents the number of observation elements in the index set Ω. All experiments

TABLE I
AVERAGE QUANTITATIVE RESULTS OF THE TEST METHODS WITH DIFFERENT SRS ON THE FOUR NATURAL HSIS. THE SRS ARE 10%, 20%, AND 40%, RESPECTIVELY. THE BEST RESULT IN EACH CASE IS HIGHLIGHTED IN BOLD

Method	10%				20%				40%			
	PSNR	SSIM	FSIM	SAM	PSNR	SSIM	FSIM	SAM	PSNR	SSIM	FSIM	SAM
HaLRTC [40]	25.872	0.711	0.832	0.248	30.476	0.846	0.907	0.111	35.101	0.911	0.949	0.089
Square Deal [37]	31.854	0.892	0.932	0.094	34.439	0.932	0.956	0.078	39.023	0.970	0.981	0.053
t-SVD [50]	32.119	0.854	0.924	0.147	36.663	0.931	0.963	0.100	42.256	0.975	0.987	0.056
TMac [30]	29.536	0.769	0.871	0.155	31.107	0.812	0.891	0.127	32.146	0.840	0.907	0.111
FBCP [38]	34.122	0.878	0.932	0.116	36.054	0.914	0.951	0.099	37.041	0.927	0.958	0.090
TNCP [28]	32.728	0.849	0.917	0.142	34.285	0.879	0.935	0.113	35.916	0.909	0.953	0.099
LRR-TC	39.354	0.949	0.973	0.070	43.291	0.980	0.990	0.043	47.643	0.992	0.997	0.025
NLRR-TC	**43.593**	**0.987**	**0.993**	**0.032**	**46.860**	**0.992**	**0.996**	**0.024**	**50.132**	**0.995**	**0.998**	**0.018**

Algorithm 2 NLRR-TC

Input: Degraded HSI $\mathcal{T}$, index set Ω, the number of clusters L, and number of iterations K.
Output: Final completed HSI $\hat{\mathcal{X}}$.

Step 1. LRR-TC based initial completion
Estimate an initial HSI $\mathcal{X}$ via LRR-TC.
Step 2. LRR-TC based final completion
(a) Group similar cubes of the $\mathcal{X}$ into L tensors $\mathcal{X}^l, l = 1, 2, \ldots, L$, via K-means++ [49].
(b) The cubes of the index set Ω are also grouped into L tensors $\Omega^l, l = 1, 2, \ldots, L$, using the spatial location of the clusters obtained in (a).
For l=1:L **do**
(c) Solving Eq. (23) via LRR-TC.
end for
Return $\{\hat{\mathcal{X}}^l\}, l = 1, 2, \ldots, L$, to the original place to reconstruct the final completed HSI $\hat{\mathcal{X}}$.

were performed by a PC (Intel I7-6850K 3.6 GHz, 64 GB memory).

Here, two widely known data sets, CAVE Database [51][1] and Natural Scenes 2002 [52],[2] are used for test the performance of the proposed method. Both of CAVE database and Natural Scenes 2002 are widely used in HSI recovery literature [53]. Two images from CAVE Database (*Jelleybeans* and *Watercolors*) and two images from Natural Scenes 2002 (*Scenes 4* and *Scenes 8*) are used in our experiments for natural HSIs completion, and the test images are subimages (31 spectral bands, the size of each band is 300×300) of the original images.

Two remote sensing hyperspectral data sets, i.e., the "Pavia University" and "AVIRIS Cuprite," are also used for testing the completion performance of the proposed method. The Pavia University data set contains 103 spectral bands after the seriously noise-corrupted bands being discarded, with spatial size of 610×340, and it is widely used in HSI processing literatures [54], [55]. The test image is the subimage (103 spectral bands, the size of each band is 200×200) of the original image. The AVIRIS Cuprite data set is of spatial size of 250×191. This image is initially composed of 224 bands, which has been reduced to 188 bands (selected the same as [56]) after removing the water vapor absorption bands. Four evaluation measures and relative square error (RSE) are used to objectively assess the completion results in our experiments. The RSE is defined as $\text{RSE} = \|\mathcal{X} - \mathcal{T}\|_F / \|\mathcal{T}\|_F$, where $\mathcal{X}$ and $\mathcal{T}$ are the recovered and ground truth data, respectively. The four evaluation measures are peak signal-to-noise ratio (PSNR), structure similarity (SSIM), feature similarity (FSIM), and spectral angle mapper (SAM). Good completion results correspond to larger values in PSNR, SSIM, and FSIM, while smaller values in RSE and SAM.

Before the experiments, the values of the test data sets are normalized to the interval [0, 1], and all experiment results are the average of five trials. The stopping criterion for the proposed method is measured by the relative change of the two estimated tensors, i.e., $\|\mathcal{X}^{k+1} - \mathcal{X}^k\|_F / \|\mathcal{X}^{k+1}\|_F < \text{tol}$. The completion results of the proposed NLRR-TC methods are quantitatively and visually compared with six state-of-the-art TC algorithms, i.e., HaLRTC [40], Square Deal [37], tensor-SVD (t-SVD) [50], TMac [30], FBCP [38], and TNCP [28]. All parameters involved in the compared algorithms are set empirically to achieve the best performance in terms of objective metrics.

B. Parameter Settings

There are two important parameters in the proposed NLRR-TC method, i.e., the spatial patch size ($W \times W$) of cubes and the maximum iteration K, which are discussed in the supplementary file to this paper for space limitation.

The value of the spatial patch size $W \times W$ is set to 5×5 and the maximum iteration $K = 300$. To balance the computational complexity and recovery performance, overlapping patch size is set to be 2×2 ($p = 2$). The step length t used to determine the rate of convergence is set to be 1.1, and the number of clusters L =10, tol = 10^{-4} are set empirically in experiments. For better TC performance, the coefficients $\{\alpha_i\}$ are tuned separately for each case. Since the LRR-TC model is used twice in the NLRR-TC method, let $v^{(0)}$ and

[1] http://www.cs.columbia.edu/CAVE/database/multispectral/
[2] http://personalpages.manchester.ac.uk/staff/d.h.foster/

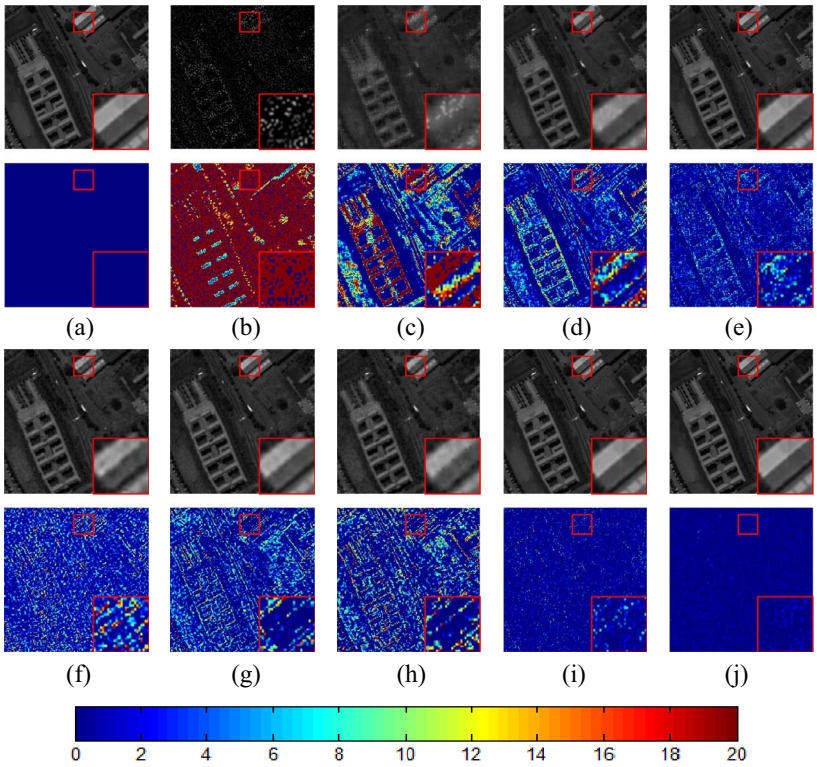


Fig. 2. Reconstructed images of the Pavia University data set at the 68th band with 20% SR. The first and third rows show the recovered images for the 68th band. The second and fourth rows show the corresponding error images. (a) Original image. (b) Degraded image 80% randomly entries from the original image is removed and fill the value "0" on them. (c)–(j) Reconstructed images by HaLRTC [40], Square Deal [37], t-SVD [50], TMac [30], FBCP [38], TNCP [28], LRR-TC, and NLRR-TC, respectively.

$v^{(1)}$ denote the values of v in the first and second time, respectively. Larger sparseness of the unfolded matrices corresponds to larger weights. Based on the spatial resolution and spectral resolution characteristics of natural and remote sensing HSIs, we set $\{\alpha_i^{(0)}\} = \{1, 1, 25\}$, $\{\alpha_i^{(1)}\} = \{1, 1.5, 1.2\}$ in natural HSIs completion experiments, and $\{\alpha_i^{(0)}\} = \{1, 1, 2.25\}$, $\{\alpha_i^{(1)}\} = \{1, 6, 5\}$ in remote sensing HSIs completion experiments. Positive scale μ is also set separately for each case. We empirically set $\mu^{(0)} = (1/120)$, $\mu^{(1)} = (1/160)$ in natural HSIs completion experiments, and $\mu^{(0)} = (1/100)$, $\mu^{(1)} = (1/160)$ in remote sensing HSIs completion experiments.

C. Experimental Results on Natural HSIs

The average quantitative results of each method over four natural HSIs under different SRs are tabulated in Tables I and II. It can be seen that, in all cases, the proposed LRR-TC and NLRR-TC methods outperform other methods in terms of all the evaluation indices. Even at very low SR (10%), our proposed NLRR-TC method can obtain good recovery (PSNR = 43.593 dB). More specifically, the proposed NLRR-TC method outperforms the best compared method (i.e., t-SVD) by 9.849 dB on PSNR, 0.071 on SSIM, 0.038 on FSIM, 0.077 on SAM, and 5.19e-2 on RSE on average results of four natural HSIs. In addition, compared with the LRR-TC, the NLRR-TC can more accurately reconstruct the HSI. Since there is spatial self-similarity within natural HSI, grouping those nonlocal similar cubes offers contribution for its better completion. For space limitation, we leave the visual comparisons of the natural HSIs completion results in the supplementary file.

TABLE II
AVERAGE QUANTITATIVE RESULTS ON RSE OF THE TEST METHODS WITH DIFFERENT SRS ON THE FOUR NATURAL HSIS. THE SRS ARE 10%, 20%, AND 40%, RESPECTIVELY. THE BEST RESULT IN EACH CASE IS HIGHLIGHTED IN BOLD

Method	10%	20%	40%
HaLRTC [40]	2.26e-1	1.33e-1	6.55e-2
Square Deal [37]	1.09e-1	8.14e-2	4.85e-2
t-SVD [50]	1.13e-1	6.84e-2	3.62e-2
TMac [30]	1.53e-1	1.22e-1	1.01e-1
FBCP [38]	8.54e-2	6.80e-2	6.05e-2
TNCP [28]	1.13e-1	8.38e-2	6.94e-2
LRR-TC	4.76e-2	2.97e-2	1.79e-2
NLRR-TC	**2.85e-2**	**1.96e-2**	**1.37e-2**

D. Experimental Results on Remote Sensing HSIs

The average quantitative results of each method on the Pavia University and AVIRIS Cuprite data set under different SRs are reported in Tables III and IV. The results show that our LRR-TC and NLRR-TC methods outperform other compared methods in terms of five objective metrics. Specifically, the proposed NLRR-TC method outperforms the best compared method (i.e., t-SVD) by 9.837 dB on

TABLE III
QUANTITATIVE RESULTS OF THE TEST METHODS WITH DIFFERENT SRS ON THE PAVIA UNIVERSITY AND AVIRIS CUPRITE DATA SET. THE SRS ARE 10%, 20%, AND 40%, RESPECTIVELY. THE BEST RESULT IN EACH CASE IS HIGHLIGHTED IN BOLD

Image	Method	10%				20%				40%			
		PSNR	SSIM	FSIM	SAM	PSNR	SSIM	FSIM	SAM	PSNR	SSIM	FSIM	SAM
Pavia University	HaLRTC [40]	21.366	0.519	0.695	0.232	24.612	0.682	0.806	0.174	29.569	0.866	0.914	0.115
	Square Deal [37]	29.937	0.870	0.914	0.096	31.828	0.912	0.943	0.084	35.542	0.959	0.974	0.063
	t-SVD [50]	31.807	0.872	0.935	0.121	36.813	0.943	0.973	0.081	42.736	0.977	0.990	0.049
	TMac [30]	20.928	0.403	0.702	0.441	30.172	0.821	0.898	0.150	33.533	0.908	0.945	0.086
	FBCP [38]	31.594	0.863	0.920	0.103	32.840	0.891	0.937	0.093	33.076	0.896	0.939	0.091
	TNCP [28]	29.141	0.790	0.879	0.131	30.080	0.824	0.899	0.112	31.521	0.869	0.928	0.109
	LRR-TC	35.327	0.933	0.967	0.076	40.223	0.972	0.989	0.049	49.980	0.992	0.997	0.024
	NLRR-TC	**44.208**	**0.984**	**0.995**	**0.035**	**46.596**	**0.989**	**0.996**	**0.029**	**50.121**	**0.994**	**0.998**	**0.021**
AVIRIS Cuprite	HaLRTC [40]	29.252	0.749	0.825	0.041	32.620	0.840	0.895	0.032	37.368	0.932	0.958	0.023
	Square Deal [37]	31.672	0.824	0.873	0.034	36.121	0.917	0.943	0.024	42.639	0.977	0.986	0.014
	t-SVD [50]	41.029	0.962	0.980	0.022	45.236	0.981	0.990	0.015	49.391	0.990	0.996	0.010
	TMac [30]	27.981	0.650	0.805	0.121	37.316	0.918	0.955	0.027	41.599	0.963	0.980	0.015
	FBCP [38]	36.845	0.915	0.949	0.026	39.245	0.946	0.969	0.021	40.352	0.956	0.975	0.020
	TNCP [28]	27.906	0.647	0.802	0.123	37.651	0.921	0.956	0.026	41.598	0.963	0.979	0.015
	LRR-TC	46.011	0.981	0.992	0.014	51.388	0.990	0.996	0.010	56.557	0.995	0.998	0.007
	NLRR-TC	**52.879**	**0.992**	**0.997**	**0.009**	**55.112**	**0.994**	**0.998**	**0.008**	**57.117**	**0.996**	**0.999**	**0.006**

PSNR, 0.037 on SSIM, 0.020 on FSIM, 0.032 on SAM, and 3.24e-2 on RSE on average results of the Pavia University and AVIRIS Cuprite data set. In addition, compared with the LRR-TC, the NLRR-TC can more accurately reconstruct the HSI, especially in low SR (10%). The average gain of NLRR-TC over LRR-TC on the Pavia University and AVIRIS Cuprite data set under different SRs are up to 4.430 dB on PSNR, 0.014 on SSIM, 0.007 on FSIM, 0.012 on SAM, and 1.25e-2 on RSE. This is due to the fact that HSIs have good property of nonlocal spatial self-similarity. By clustering nonlocal similar cubes, the formed 3-D tensor rank is low, which can contribute significantly to its better reconstruction.

The visual comparison results of the 68th band of the Pavia University image with 20% SR can be found in Fig. 2. It can be seen that the LRR-TC and NLRR-TC methods get better reconstruction results compared with other methods (see the corresponding error results). By exploiting the nonlocal self-similarity within HSI, the NLRR-TC method performs much better than the LRR-TC. From the zoomed red rectangular regions, it can be observed that the compared methods generate more or less artifacts across the recovered images. Specifically, textures are better recovered by the proposed LRR-TC and NLRR-TC methods. It can demonstrate that the LRR-TC model is effectively approximate the tensor rank. In addition, different from the compared TC methods that treat the estimated tensor data as a whole, nonlocal TC strategy is proposed in the step two in our NLRR-TC method, which can further improve the reconstruction capability. For space limitation, we leave the visual comparison results of the AVIRIS Cuprite data set in the supplementary file.

TABLE IV
QUANTITATIVE RESULTS ON RSE OF THE TEST METHODS WITH DIFFERENT SRS ON THE PAVIA UNIVERSITY AND AVIRIS CUPRITE DATA SET. THE SRS ARE 10%, 20%, AND 40%, RESPECTIVELY. THE BEST RESULT IN EACH CASE IS HIGHLIGHTED IN BOLD

Image	Method	10%	20%	40%
Pavia University	HaLRTC [40]	3.76e-1	2.59e-1	1.47e-1
	Square Deal [37]	1.41e-1	1.13e-1	7.41e-2
	t-SVD [50]	1.26e-1	7.67e-2	4.34e-2
	TMac [30]	4.75e-1	1.55e-1	9.63e-2
	FBCP [38]	1.16e-1	1.01e-1	9.77e-2
	TNCP [28]	1.55e-1	1.39e-1	1.18e-1
	LRR-TC	7.75e-2	4.65e-2	2.18e-2
	NLRR-TC	**3.37e-2**	**2.68e-2**	**1.94e-2**
AVIRIS Cuprite	HaLRTC [40]	8.75e-2	5.91e-2	3.40e-2
	Square Deal [37]	6.59e-2	3.93e-2	1.89e-2
	t-SVD [50]	2.47e-2	1.61e-2	1.08e-2
	TMac [30]	1.79e-1	4.12e-2	2.14e-2
	FBCP [38]	3.59e-2	2.76e-2	2.45e-2
	TNCP [28]	3.57e-1	3.28e-2	2.75e-2
	LRR-TC	1.46e-2	1.03e-2	7.50e-3
	NLRR-TC	**9.22e-3**	**7.85e-3**	**6.41e-3**

E. Comparison Results of Nonlocal Completion

The initial completion tensor already has a satisfactory result and we only divide it into ten clusters, thus a good clustering result could be obtained. To further explain it, we conduct an experiment on the Pavia University data set to

TABLE V
QUANTITATIVE RESULTS OF THE TEST METHODS INTEGRATED WITH THE SECOND STEP WITH DIFFERENT SRS ON THE FOUR NATURAL HSIS DATA SET. THE SRS ARE 10%, 20%, AND 40%, RESPECTIVELY. THE BEST RESULT IN EACH CASE IS HIGHLIGHTED IN BOLD

Method	10%				20%				40%			
	PSNR	SSIM	FSIM	SAM	PSNR	SSIM	FSIM	SAM	PSNR	SSIM	FSIM	SAM
N-HaLRTC	28.532	0.849	0.908	0.135	35.448	0.950	0.967	0.071	42.418	0.987	0.992	0.035
N-Square Deal	35.558	0.958	0.968	0.060	40.548	0.984	0.989	0.037	46.921	0.995	0.997	0.020
N-t-SVD	31.226	0.868	0.918	0.111	36.307	0.951	0.969	0.071	44.436	0.990	0.994	0.033
N-TMac	29.783	0.841	0.912	0.203	34.033	0.902	0.945	0.154	36.451	0.931	0.962	0.108
N-FBCP	34.965	0.898	0.944	0.105	36.995	0.927	0.961	0.090	39.105	0.949	0.974	0.074
N-TNCP	41.372	0.982	0.988	0.039	45.721	0.991	0.995	0.027	48.606	0.995	0.998	0.021
NLRR-TC	**43.593**	**0.987**	**0.993**	**0.032**	**46.860**	**0.992**	**0.996**	**0.024**	**50.132**	**0.995**	**0.998**	**0.018**

compare the cluster results on the original data and the initial completion tensor. The clustering result on original data is regarded as the reference data. The mean clustering errors of 2.2%, 1.0%, and 0.6% are respectively achieved, on the initial completion tensors under SRs 10%, 20%, and 40%. Therefore, the cluster results on the initial completion tensors are effective. It is also valid to use the cluster results of the initial completion tensors for nonlocal completion to further improve the completion results.

Then, we compare the NLRR-TC with other methods integrated with the second step to assess the nonlocal completion performance. For convenience, the six compared methods integrated with the second step are called as, N-HaLRTC, N-Square Deal, N-t-SVD, N-TMac, N-FBCP, and N-TNCP, respectively.

The average quantitative results of the test methods integrated with the second step over four natural HSIs under different SRs are reported in Tables V and VI. From these quantitative comparisons, it can be observed that the NLRR-TC outperforms all other competing methods with respective to all evaluation measures. Specifically, the NLRR-TC outperforms the best compared method (i.e., N-TNCP) by 1.629 dB on PSNR, 0.002 on SSIM, 0.002 on FSIM, 0.004 on SAM, and 4.20e-3 on RSE on average results of four natural HSIs. Furthermore, by quantitative comparisons Table I with Table V, and Table II with Table VI, it can be observed that by integrating the second step with the compared methods, the completion performance has been improved except for t-SVD. This is due to the fact that the t-SVD is mainly proposed for data with linear camera motion, which is not good at handle the clustered tensor formed in this paper. It is worth noting that the performance of the HaLRTC, Square Deal, and TNCP, has been significantly improved with nonlocal TC strategy. More comparison results can be found in the supplementary file.

F. Computational Complexity

Considering Algorithm 1 for an N-dimensional input tensor $\mathcal{X} \in R^{I_1 \times I_2 \times \cdots \times I_N}$, the main running time is consumed by performing SVD decomposition on the unfold$_i(\mathcal{X}^k + (1/\mu^k)y_i^k)$ with size of $I_i \times \prod_{j \neq i} I_j (1 \leq i \leq N)$ in each iteration. Therefore, the computational complexity of Algorithm 1 is $O(\sum_i I_i^2 \times \prod_{j \neq i} I_j)$.

TABLE VI
QUANTITATIVE RESULTS ON RSE OF THE TEST METHODS INTEGRATED WITH THE SECOND STEP WITH DIFFERENT SRS ON THE FOUR NATURAL HSIS. THE SRS ARE 10%, 20%, AND 40%, RESPECTIVELY. THE BEST RESULT IN EACH CASE IS HIGHLIGHTED IN BOLD

Method	10%	20%	40%
N-HaLRTC	1.73e-1	7.75e-2	3.34e-2
N-Square Deal	7.59e-2	4.29e-2	1.98e-2
N-t-SVD	1.18e-1	6.59e-2	2.69e-2
N-TMac	2.27e-2	1.48e-1	1.04e-1
N-FBCP	7.67e-2	6.11e-2	4.79e-2
N-TNCP	3.65e-2	2.21e-2	1.58e-2
NLRR-TC	**2.85e-2**	**1.96e-2**	**1.37e-2**

For Algorithm 2 with an input HSI $\mathcal{X} \in R^{I_1 \times I_2 \times I_3}$, the main running time is consumed by performing SVD decomposition and the clustering of similar cubes. In this paper, we use the K-means++ method for clustering. The computational complexity of the K-means++ algorithm is $O(Ls^2)$, where L is the number of clusters and $s = O(I_1 I_2)$ is number of samples. In addition, the size of the jth clustered tensor is $W^2 \times I_3 \times s_j (1 \leq j \leq L, s_1 + s_2 + \cdots + s_L = s)$, where W is the spatial size of cube and s_j is the number of nonlocal similar cubes in the jth cluster. The computation cost is not competitive for quite large s. However, completion on the L clustered tensors can be processed in parallel, each with relatively small computation complexity $O(W^2 \times I_3 \times s_j(W^2 + I_3 + s_j))$. Therefore, the overall computational complexity of Algorithm 2 is $O(\sum_n I_n^2 \times \prod_{m \neq n} I_m + W^2 \times I_3 \times s_j(W^2 + I_3 + s_j) + Ls^2)(1 \leq n \leq 3)$.

VI. CONCLUSION

In this paper, a novel NLRR-TC method is presented for HSIs completion, which consists of two main steps. In the first step, the LRR-TC is introduced into the whole degraded data to generate an initial completion result. The LRR-TC model that combines the logarithm of the determinant with the TTN, can effectively approximate the tensor rank. The TLRR algorithm is developed to effectively solve the tensor rank minimization problem, whose convergence is also analyzed. In the second

step, similar 3-D image cubes in the initial completion result are first clustered into 3-D tensors, which enable themselves to be better approximated by low-rank tensors, the LRR-TC is then applied to each 3-D tensor to further improve the completion result. Experimental results demonstrate that the proposed NLRR-TC method is capable of achieving better reconstruction than state-of-the-art HSIs completion approaches in tested HSIs containing diverse and complex structures.

In our future work, we will extend the method to other types of application, e.g., recommendation system, medical image, and video completion. In addition, the correlations in different channels will be utilized for multichannel image completion. The proposed model may be also used to address the fourth-order or higher-order tensors completion problem, and other kinds of robust tensor recovery problems.

ACKNOWLEDGMENT

The authors would like to thank the Editor-in-Chief, the Associate Editor, and the anonymous reviewers for their insightful comments and constructive suggestions, which have significantly improved this paper.

REFERENCES

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. ACM SIGGRAPH*, 2000, pp. 417–424.

[2] T. F. Chan and J. Shen, "Mathematical models for local nontexture inpaintings," *SIAM J. Appl. Math.*, vol. 62, no. 3, pp. 1019–1043, Feb. 2002.

[3] N. Komodakis and G. Tziritas, "Image completion using global optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, New York, NY, USA, 2006, pp. 442–452.

[4] X. Fang, Y. Xu, X. Li, Z. Lai, and W. K. Wong, "Robust semi-supervised subspace clustering via non-negative low-rank representation," *IEEE Trans. Cybern.*, vol. 46, no. 8, pp. 1828–1838, Aug. 2016.

[5] Y. Yang, W. Hu, Y. Xie, W. Zhang, and T. Zhang, "Temporal restricted visual tracking via reverse-low-rank sparse learning," *IEEE Trans. Cybern.*, vol. 47, no. 2, pp. 485–498, Feb. 2017.

[6] J. Chen and J. Yang, "Robust subspace segmentation via low-rank representation," *IEEE Trans. Cybern.*, vol. 44, no. 8, pp. 1432–1445, Aug. 2014.

[7] Y. Lu, Z. Lai, Y. Xu, X. Li, D. Zhang, and C. Yuan, "Low-rank preserving projections," *IEEE Trans. Cybern.*, vol. 46, no. 8, pp. 1900–1913, Aug. 2016.

[8] T. Korah and C. Rasmussen, "Spatiotemporal inpainting for recovering texture maps of occluded building facades," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2262–2271, Sep. 2007.

[9] S. Mao, L. Xiong, L. Jiao, T. Feng, and S.-K. Yeung, "A novel Riemannian metric based on Riemannian structure and scaling information for fixed low-rank matrix completion," *IEEE Trans. Cybern.*, vol. 47, no. 5, pp. 1299–1312, May 2017.

[10] Y. Ren, G. Li, J. Zhang, and W. Zhou, "Lazy collaborative filtering for data sets with missing values," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 1822–1834, Dec. 2013.

[11] M. Fazel, "Matrix rank minimization with applications," Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, Mar. 2002.

[12] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Mar. 2010.

[13] X. Lu, T. Gong, P. Yan, Y. Yuan, and X. Li, "Robust alternative minimization for matrix completion," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 3, pp. 939–949, Jun. 2012.

[14] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053–2080, May 2010.

[15] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SAIM Rev.*, vol. 52, no. 3, pp. 471–501, Aug. 2010.

[16] F. Shang, Y. Liu, and J. Cheng, "Scalable algorithms for tractable Schatten quasi-norm minimization," in *Proc. 30th AAAI Conf. Artif. Intell.*, 2016, pp. 2016–2022.

[17] F. Shang, Y. Liu, and J. Cheng, "Tractable and scalable Schatten quasi-norm approximations for rank minimization," in *Proc. 19th Int. Conf. Artif. Intell. Stat.*, 2016, pp. 620–629.

[18] M. Fazel, H. Hindi, and S. P. Boyd, "Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices," in *Proc. Amer. Control Conf.*, Denver, CO, USA, 2003, pp. 2156–2162.

[19] C. Peng, Z. Kang, H. Li, and Q. Cheng, "Subspace clustering using log-determinant rank approximation," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2015, pp. 925–934.

[20] K. Mohan and M. Fazel, "Reweighted nuclear norm minimization with application to system identification," in *Proc. Amer. Control Conf.*, Baltimore, MD, USA, 2010, pp. 2953–2959.

[21] S. Gu *et al.*, "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, Jan. 2017.

[22] Y. Peng, J. Suo, Q. Dai, and W. Xu, "Reweighted low-rank matrix recovery and its application in image restoration," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2418–2430, Dec. 2014.

[23] L. Liu, L. Chen, C. L. P. Chen, Y. Y. Tang, and C. M. Pun, "Weighted joint sparse representation for removing mixed noise in image," *IEEE Trans. Cybern.*, vol. 47, no. 3, pp. 600–611, Mar. 2017.

[24] A. Barmpoutis, "Tensor body: Real-time reconstruction of the human body and avatar synthesis from RGB-D," *IEEE Trans. Cybern.*, vol. 43, no. 5, pp. 1347–1356, Oct. 2013.

[25] X. Li, S. Lin, S. Yan, and D. Xu, "Discriminant locally linear embedding with high-order tensor data," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 2, pp. 342–352, Apr. 2008.

[26] X. Geng, K. Smith-Miles, Z.-H. Zhou, and L. Wang, "Face image modeling by multilinear subspace analysis with missing values," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 41, no. 3, pp. 881–892, Jun. 2011.

[27] X. Cao, X. Wei, Y. Han, and D. Lin, "Robust face clustering via tensor decomposition," *IEEE Trans. Cybern.*, vol. 45, no. 11, pp. 2546–2557, Nov. 2015.

[28] Y. Liu, F. Shang, L. Jiao, J. Cheng, and H. Cheng, "Trace norm regularized CANDECOMP/PARAFAC decomposition with missing data," *IEEE Trans. Cybern.*, vol. 45, no. 11, pp. 2437–2448, Nov. 2015.

[29] Q. Zhao *et al.*, "A novel sparsity measure for tensor recovery," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 271–279.

[30] Y. Xu, R. Hao, W. Yin, and Z. Su, "Parallel matrix factorization for low-rank tensor completion," *Inverse Problems Imag.*, vol. 9, no. 2, pp. 601–624, Mar. 2015.

[31] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multi-modal factor analysis," *UCLA Working Paper Phonetics*, vol. 16, no. 1, pp. 1–84, 1970.

[32] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, Sep. 1966.

[33] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, "Scalable tensor factorizations with missing data," in *Proc. SIAM Int. Conf. Data Min.*, 2010, pp. 701–711.

[34] L. Karlsson, D. Kressner, and A. Uschmajew, "Parallel algorithms for tensor completion in the CP format," *Parallel Comput.*, vol. 57, pp. 222–234, Sep. 2016.

[35] E. S. Allman, P. D. Jarvis, J. A. Rhodes, and J. G. Sumner, "Tensor rank, invariants, inequalities, and applications," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 1014–1045, Jul. 2013.

[36] C. J. Hillar and L.-H. Lim, "Most tensor problems are NP-hard," *J. ACM*, vol. 60, no. 6, pp. 1–39, Nov. 2013.

[37] C. Mu, B. Huang, J. Wright, and D. Goldfarb, "Square deal: Lower bounds and improved relaxations for tensor recovery," in *Proc. Int. Conf. Mach. Learn.*, 2014, pp. 73–81.

[38] Q. Zhao, L. Zhang, and A. Cichocki, "Bayesian CP factorization of incomplete tensors with automatic rank determination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1751–1763, Sep. 2015.

[39] Q. Zhao, G. Zhou, L. Zhang, A. Cichocki, and S.-I. Amari, "Bayesian robust tensor factorization for incomplete multiway data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 4, pp. 736–748, Apr. 2016.

[40] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

[41] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-n-rank tensor recovery via convex optimization," *Inverse Problems*, vol. 27, no. 2, pp. 1–19, Jan. 2011.

[42] L. Yang, Z.-H. Huang, and X. Shi, "A fixed point iterative method for low n-rank tensor pursuit," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2952–2962, Jun. 2013.
[43] Y. Wu, H. Tan, Y. Li, F. Li, and H. He, "Robust tensor decomposition based on Cauchy distribution and its applications," *Neurocomputing*, vol. 223, no. 5, pp. 107–117, Feb. 2017.
[44] Y. Yang, Y. Feng, and J. A. K. Suykens, "Robust low-rank tensor recovery with regularized redescending M-estimator," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 9, pp. 1933–1946, Sep. 2016.
[45] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.
[46] Y. Liu, F. Shang, W. Fan, J. Cheng, and H. Cheng, "Generalized higher-order orthogonal iteration for tensor decomposition and completion," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 1763–1771.
[47] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3618–3632, Aug. 2014.
[48] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
[49] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proc. Annu. ACM-SIAM Symp. Discrete Algorithms*, 2007, pp. 1027–1035.
[50] Z. Zhang, G. Ely, S. Aeron, and N. Hao, "Novel methods for multilinear data completion and de-noising based on tensor-SVD," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 3842–3849.
[51] 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.
[52] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes," *J. Opt. Soc. Amer. A*, vol. 23, no. 10, pp. 2359–2372, Oct. 2006.
[53] 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.*, Honolulu, HI, USA, 2017, pp. 3862–3871.
[54] L. Fang, N. He, S. Li, P. Ghamisi, and J. A. Benediktsson, "Extinction profiles fusion for hyperspectral images classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1803–1815, Mar. 2018.
[55] L. Fang, C. Wang, S. Li, and J. A. Benediktsson, "Hyperspectral image classification via multiple-feature-based adaptive sparse representation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1646–1657, Jul. 2017.
[56] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan. 2014.

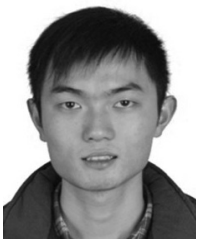
Ting Xie (S'17) received the B.S. degree in information and computational science and the M.S. degree in applied mathematics from the University of South China, Hengyang, China, in 2013 and 2016, respectively. She is currently pursuing the Ph.D. degree with the Laboratory of Vision and Image Processing, Hunan University, Changsha, China.

Her current research interests include image quality improvement, hyperspectral image processing, and low-rank representation.

Shutao Li (M'07–SM'15) received the B.S., M.S., and Ph.D. degrees from Hunan University, Changsha, China, in 1995, 1997, and 2001, respectively.

He was a Research Associate with the Department of Computer Science, Hong Kong University of Science and Technology, Hong Kong, in 2011. From 2002 to 2003, he was a Post-Doctoral Fellow with the Royal Holloway College, University of London, London, U.K., with Prof. J. Shawe-Taylor. In 2005, he was a Visiting Professor with the Department of Computer Science, Hong Kong University of Science and Technology. He joined the College of Electrical and Information Engineering, Hunan University in 2001, where he is currently a Full Professor. He has authored or co-authored over 200 refereed papers. His current research interests include compressive sensing, sparse representation, image processing, and pattern recognition.

Dr. Li was a recipient of two Second-Grade National Awards at the Science and Technology Progress of China in 2004 and 2006. He is currently an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, and a member of the Editorial Board of *Information Fusion* and *Sensing and Imaging*.



Leyuan Fang (S'10–M'14–SM'17) received the Ph.D. degree from the College of Electrical and Information Engineering, Hunan University, Changsha, China, in 2015.

From 2011 to 2012, he was a visiting Ph.D. student with the Department of Ophthalmology, Duke University, Durham, NC, USA, supported by the China Scholarship Council. Since 2017, he has been an Associate Professor with the College of Electrical and Information Engineering, Hunan University. His current research interests include sparse representation and multiresolution analysis in remote sensing and medical image processing.

Dr. Fang was a recipient of the Scholarship Award for Excellent Doctoral Student granted by the Chinese Ministry of Education in 2011.

Licheng Liu (M'16) received the B.S. degree in information and computational science from the China University of Geosciences, Wuhan, China, in 2010, the M.S. degree in applied mathematics from Hunan University, Changsha, China, in 2012, and the Ph.D. degree in software engineering from the University of Macau, Macau, China, in 2016.

He is currently an Assistant Professor with the College of Electrical and Information Engineering, Hunan University. His current research interests include image processing, sparse representation, pattern recognition, and machine learning.