J. Nonlinear Var. Anal. 7 (2023), No. 4, pp. 527-548
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.7.2023.4.05

MULTITEMPORAL IMAGE CLOUD REMOVAL USING GROUP SPARSITY AND NONCONVEX LOW-RANK APPROXIMATION

ZHIHUI TU[1], LEI GUO[1], HUAN PAN[1,*], JIAN LU[1,2,*], CHEN XU[1,3], YURU ZOU[1]

[1]*Shenzhen Key Laboratory of Advanced Machine Learning and Applications, College of Mathematics and Statistics, Shenzhen University, Shenzhen 518060, China*
[2]*Pazhou Lab, Guangzhou, China*
[3]*National Center for Applied Mathematics Shenzhen (NCAMS), Shenzhen 518055, China*

Abstract. Remote sensing (RS) images are widely used in environmental monitoring, urban planning, and land surface classification. However, RS images are often polluted by cloud, which leads to the loss of some important information of RS images and hinders the development of relevant applications. The existing spatial-spectral total-variational regularization can only promote the spatial-spectral continuity of cloud component, but cannot maintain the shared group sparse mode of spatial difference images of different spectral bands. To solve this problem, we add the weighted $\ell_{2,1}$-norm to constrain the cloud component differential images, and use the nonconvex regularization term, namely the weighted nuclear norm, to replace the traditional nuclear norm, which solves the problem that the original nuclear norm violates the larger singular value. In summary, we propose a multitemporal image cloud removal model based on the weighted nuclear norm which is the nonconvex low-rank approximation and the group sparsity regularization (WNGS), where the group sparsity regularization and the weighted nuclear norm promote each other. The resulting problems are solved using the alternating direction method of multipliers. Numerical experiments both simulated and real multitemporal images demonstrate that the proposed method is superior to other advanced cloud-removal methods in different cloud-removal scenarios.

Keywords. Multitemporal image cloud removel; Group sparsity regularization; Weighted nuclear norm.

1. INTRODUCTION

Multitemporal images contain rich spatial, spectral, and temporal information, and are widely used in many fields, such as unmixing [1, 2, 3, 4, 5, 6], classification [7, 8, 9], target detection [10], fusion [11, 12, 13, 14, 15] and so on. However, in the process of acquisition, multitemporal RS images are inevitably damaged by various factors, such as environmental impact and transmission errors, which lead to the impact of cloud, cloud shadow or noise on multitemporal RS images. The damaged images seriously affect the subsequent application, so the study of multitemporal RS image restoration is of great significance.

For multitemporal RS image restoration, there are matrix-based methods, tensor based methods, and deep learning methods. In the matrix-based image restoration method, the early idea

*Corresponding author.
E-mail addresses: panhuan@szu.edu.cn (H. Pan), jianlu@szu.edu.cn (J. Lu).
Received February 18, 2022; Accepted March 7, 2023.

is to convert multidimensional images into matrices, and use the rank function of matrix to describe the low-rank property of images [16, 17, 18, 19, 20, 21]. Zhang et al. [22] proposed the low-rank matrix recovery (LRMR) model, which uses the nuclear norm to make convex substitution of the rank function to describe the low-rank property of the image. In order to make full use of the low-dimensional structure of the image and retain the image edge information more effectively, He et al. [23] proposed an image restoration model based on the total variational regularization low-rank matrix decomposition, which uses the space-spectral total variational regularization constraint on the cloud/shadow component at each time point. The total variational regularization in this model can effectively preserve the edge information and improve the segmentation smoothness. At the same time, low-rank matrix decomposition can make the model make full use of the low-dimensional structure of the image. However, the low-rank matrix decomposition method is sensitive to outliers and the low-rank based method only explores the correlation between spectral bands, ignoring the spatial correlation of local adjacent pixels. Therefore, He et al. [24] proposed an image restoration model based on local low-rank matrix restoration and global spatial spectral total variation, which combined local and global spatial correlation. However, this model only made use of local or non-local relationships in spatial dimensions, and could not reconstruct large-scale cloud/shadow polluted areas. On the basis of using nuclear norm to carry out convex replacement of rank function, in order to deal with different singular values adaptively and better promote low-rank characteristics, Gu et al. [25] proposed weighted nuclear norm for image restoration, which obtained good restoration effect. However, these methods do not take into account the temporal properties of multitemporal RS data.

Tensors are higher-order generalizations of matrices, and matrices are second-order tensors. Compared with matrices, tensors can more effectively consider the spatial, spectral and temporal prior information of multi-temporal RS images. Ji et al. [26] used tensor ranks based on tensor singular value decomposition and group sparsity functions along the spectral dimension to characterize the global properties of image components and obtained good results. Lin et al. [27] used coupled tensor decomposition to explore the relationship between multi-temporal image abundances in the same scene, effectively removing thick clouds in images. The tensor decomposition method used in the above model aims to decompose a higher-order tensor into a series of low-dimensional factors, and tensor decomposition has a powerful ability to capture the global correlation of the tensor. However, unlike the rank of a matrix, the rank of a tensor is not unique.

In recent years, due to the rise of deep learning, image restoration methods based on deep learning have been proposed and widely applied, and image cloud removal tasks are gradually associated with deep learning methods [28, 29, 30]. Image cloud removal method based on deep learning can effectively remove clouds and achieve cloud removal effect, but there are some problems such as long training time, insufficient cloud removal effect, and color distortion. Recently, methods based on generative adversarial networks (GAN) have become dominant and have gone beyond the general convolutional neural network (CNN) -based approach to improve the most advanced performance [31, 32]. The high compatibility and capacity of the GAN model reduces the burden of dedicating network design to specific applications, but at the cost of larger and deeper network and training problems.

To solve the above problems, we combine the nonconvex low-rank approximation and group sparse regularization in the proposed model. First of all, the existing image restoration or image cloud removal models only use the weighted nuclear norm [25] or group sparse regularization terms [33]. The weighted nuclear norm is used as the nonconvex approximation of the rank function to describe the low rank of the image, and the group sparse regularization term to describe the shared group sparse pattern of the image along the spectral dimension. Both of these methods achieve good results in the image restoration task. The main idea of our approach is to add group sparse regularization to the difference cloud/shadow components while using the weighted nuclear norm as a nonconvex substitute for the image rank function. We expect that our model can adaptively assign weights to different singular values on the basis of the original nuclear norm, enhance the representation ability of the original nuclear norm, and solve the problem of giving the same weight to different singular values and violating the prior knowledge. At the same time, we do not directly perform group sparse regularization on the target image, but perform group sparse regularization on the difference image of cloud/shadow component, which describes the prior knowledge that the spatial difference image of the original ignored cloud/shadow has sparsity along the spectral dimension. In our proposed model, group sparse regularization and weighted nuclear norm promote each other, which can effectively improve the model's cloud removal ability. In addition, we use the alternating direction method of multipliers (ADMM) to solve the proposed cloud removal model, and investigate the algorithm's complexity, convergence and parameter settings. Extensive experiments on simulated and real datasets demonstrate that our proposed cloud removal method for multitemporal RS images is an effective method. To sum up, the contribution of this paper is as follows:

- Using the prior information of the real image to characterize the various parts of the cloud image and using the $\ell_{2,1}$-norm to represent the group sparsity along the spectral dimension which can not only consider the sparsity of the cloud/shadow component along the column direction, but also the sparsity of the spatial difference image of cloud/shadow along the spectral dimension.
- By using the weighted nuclear norm as the nonconvex approximation of the rank function, we can adaptively assign different weights to different singular values, improve the flexibility of the traditional nuclear norm minimization problem and enhance the representation ability of the original nuclear norm. At the same time, on the basis of the $\ell_{2,1}$-norm increases the adaptive weighting vector, increase the space difference image of cloud and cloud shadow group of sparse.
- We develop an efficient algorithm based on the ADMM to solve the proposed model. Simulations and real experiments are performed on datasets acquired by Sentinel-2 and Landsat-8. Experimental results demonstrate the effectiveness and robustness of the method.

The rest of this paper is organized as follows. In Section 2, we propose the weighted nuclear norm and group sparsity regularization model (WNGS) for multitemporal image cloud removal, and develop an ADMM algorithm to solve the proposed model. Experimental results are demonstrated in Section 3. Section 4 discusses the details of parameter selection and convergence behavior. Section 5 concludes the paper.

2. Proposed Weighted Nuclear Norm and Group Sparsity Regularization Model (WNGS)

2.1. **Notations.** In this section, we introduce the expressions and basic terminology used by our proposed model. A matrix is represented by X. For a given matrix $X \in R^{n_1 \times n_2}$, $X(k,l)$, $X_{k,l}$, and x_{kl} are the (k,l)-th element of the matrix, $\sigma_i(X)$ is the i-th singular value of X, $\|X\|_1$ is ℓ_1-norm of X, $\|X\|_1 = \max_l \sum_{k=1}^{n_1} |a_{kl}|$ is the maximum value of the sum of the absolute values of all matrix column vectors, $\|X\|_F$ is Frobenius norm of the matrix X, $\|X\|_F = (\sum_{k=1}^{n_1}\sum_{l=1}^{n_2} |a_{kl}|^2)^{\frac{1}{2}}$ is the sum of squares matrix elements, $\|X\|_*$ is the nuclear norm of matrix X, where $\|X\|_* = \Sigma_i \sigma_i(X)$ is the accumulation of all the singular of the matrix X, $\|X\|_{w,*}$ is the weighted nuclear norm of matrix X, and $\|X\|_{w,*} = \Sigma_i w_i \sigma_i(X)$, w_i is a non-negative weight assigned to $\sigma_i(X)$. The weight vector will enhance the representation capability of the original nuclear norm. $\|X\|_{2,1}$ is $\ell_{2,1}$-norm of the matrix X, where $\|X\|_{2,1} = \sum_{k=1}^{n_1} (\sum_{l=1}^{n_2} X_{k,l}^2)^{\frac{1}{2}} = \sum_{k=1}^{n_1} \|X_{k,:}\|_2$.

2.2. **WNGS for Multitemporal Image Cloud Removal.** In the process of image formation, recording, processing, and transmission, due to the imperfection of imaging system, recording equipment, transmission medium and processing method, the image quality will decrease, which is called image degradation. The degradation models mainly include nonlinear degradation, fuzzy degradation, motion degradation, and random noise degradation. We use the degradation process of thick, thin clouds and cloud shadows as an additive model, and approximate the cloud/shadow effect to additive noise. The image degradation model we describe is as follows:

$$Y = X + S + N,$$

where $Y \in R^{n_1 \times n_2}$, $X \in R^{n_1 \times n_2}$, $S \in R^{n_1 \times n_2}$, $N \in R^{n_1 \times n_2}$, Y is the image matrix contaminated by cloud/shadow, X is the clean image matrix without cloud, S is the cloud/shadow component matrix, and N is the residual component matrix.

We expect to propose a multitemporal image removal model. The goal is to separate clean cloudless images and cloud/shadow images from images contaminated by clouds and cloud shadows. However, when we solve the problem only through the above degradation model, we will find that the solution obtained by this method cannot fully satisfy the three conditions of existence, uniqueness and stability, that is, it is a strongly ill-posed problem. An effective way to solve this discomfort is to introduce prior information about the image in the image processing process. So we propose a preliminary cloud removal model as follows:

$$\begin{aligned} &\min_{X,S,N} \phi_X(X) + \phi_S(S) + \phi_N(N), \\ &s.t.\ Y = X + S + N, \end{aligned}$$

where $\phi_X(X)$, $\phi_S(S)$, and $\phi_N(N)$ represent the corresponding constraints of X, S, and N based on prior knowledge. The specific constraints are introduced in the following text.

(1) **The nonconvex approximation of the rank function of the clean image** X. Observed images Y and clean images X are multitemporal images with and without clouds, respectively. Multitemporal images refer to multi-time images of the same scene. The difference between images at different time points is that the light wavelengths of each band are different, and the time nodes of imaging are different. This special source of multitemporal images makes it

highly correlated in the time dimension. Meanwhile, like other natural images, there is a lot of redundant information in multitemporal images, that is, each spectral feature can be represented by a linear combination of a few basic spectral features. Therefore, combining the two features mentioned above, we believe that multitemporal images have a strong correlation in the spectral-temporal dimension. This prior prompts us to propose a solution model based on the low rank of clean image X in the model of removing clouds and cloud shadows.

In the research of image processing, such as image denoising and image filling, researchers usually use rank minimization to study the low rank of image X. However, direct rank minimization is an NP-hard and difficult problem to solve. Usually, the constraint problem is relaxed by substituting the minimum nuclear norm. This is a convex relaxation method for minimizing the matrix, which is called nuclear norm minimization (NNM). Nuclear norm minimization can realize the inherent low rank attribute of the image. The expression describing the nuclear norm minimization problem is as follows:

$$\hat{X} = \text{prox}_{\lambda\|\cdot\|_*}(Y) = \arg\min_X \|X - Y\|_F^2 + \lambda\|X\|_*,$$

where the nuclear norm of matrix X is the sum of the singular values of matrix X:

$$\|X\|_* = \sum_i \sigma_i(X).$$

The closed form solution of the above nuclear norm proximal problem can be obtained by soft threshold operation on the singular value of the observation matrix. The expression is as follows:

$$\hat{X} = US_{\frac{\lambda}{2}}(\Sigma)V^T,$$

where

$$S_{\frac{\lambda}{2}}(\Sigma)_{ii} = \max(\Sigma_{ii} - \frac{\lambda}{2}, 0).$$

However, the nuclear norm minimization problem has some limitations. In the traditional nuclear norm minimization problem, all singular values are treated equally, but the prior knowledge that we usually have is that larger singular values should be reduced slightly and given greater weights, and smaller singular values should be reduced greatly, that is, given smaller weights. Weighted nuclear norm minimization (WNNM) can adaptively assign weights to different singular values, where the weight vector can enhance the representation ability of the original nuclear norm, which can improve the flexibility of traditional nuclear norm minimization. So in our model, we use the weighted nuclear norm as a nonconvex approximation of the rank function of clean images, $\phi_X(X) = \lambda_1\|X\|_{w,*}$, λ_1 is the regularization coefficient of $\phi_X(X)$. The weighted nuclear norm of the matrix is as follows:

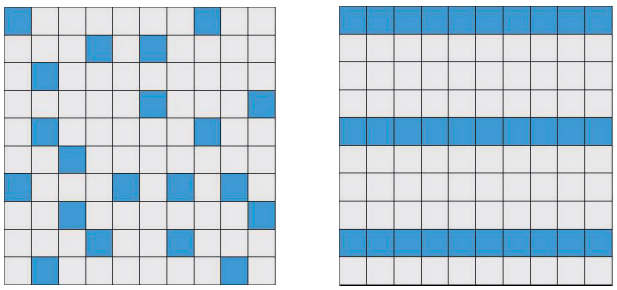
$$\|X\|_{w,*} = \Sigma_i w_i \sigma_i(X).$$

That is, based on the sum of singular values of matrix X, weights suitable for different singular values are added. A closed form optimal solution of weighted nulcear norm proximal (WNNP) problem can be obtained by weighted singular value soft threshold operation:

$$\text{prox}_{\lambda\|\cdot\|_{w,*}}(Y) = US_{\frac{\lambda w}{2}}(\Sigma)V^T,$$

where

$$S_{\frac{\lambda w}{2}}(\Sigma)_{ii} = \max(\Sigma_{ii} - \frac{\lambda w_i}{2}, 0).$$



(a) (b)

FIGURE 1. Description of group sparsity. White squares represent zero elements, and blue squares represent non-zero elements. (a) Elements are randomly distributed. (b) Elements show group distribution.

In fact, the nuclear norm minimization problem is a special case of the weighted nuclear norm minimization problem. When we set the weights corresponding to each singular value the same, the nuclear norm minimization problem is equivalent to the weighted nuclear norm minimization problem.

(2) **Using the $\ell_{2,1}$-norm characterizes the sparsity.** The cloud/shadow component is characterized by the fact that the density and position of the cloud/shadow component are usually different at different times, which indicates that the cloud/shadow component has little correlation in the time dimension. This is contrary to the strong correlation of multitemporal images in time dimension. However, when the proportion of cloud/shadow in the whole multitemporal image is low, it is feasible to use sparse prior to describe the cloud/shadow component. In order to realize the sparse prior, the ℓ_0-norm is the best choice. However, due to the nonconvexity of the ℓ_0-norm and its NP-hard problem, the ℓ_1-norm is used as the convex proxy of the ℓ_0-norm. When the ℓ_1-norm is used to represent the sparse priors of the cloud/cloud shadow components, the correlation between bands is ignored. However, when we assume that the imaging scenes of different bands are the same, the piecewise smooth structure of different bands is also the same, which indicates that the spatial difference image of cloud/shadow is also sparse along the spectral dimension. Therefore, we use the $\ell_{2,1}$-norm to characterize the group sparsity of cloud/shadow components in spectral difference images. Fig. 1 visualizes sparsity and group sparsity, where gray squares represent zero elements and blue squares represent nonzero elements. By observing Fig. 1 (a), it can be seen that the sparsity element distribution is random, and the group sparsity makes the element distribution be a group distribution (see Fig. 1 (b)).

To improve the group sparsity of spatial difference images of cloud/shadow, we further utilize weighted group sparsity to improve its group sparsity. Therefore, in our proposed WNGS model,

$$\phi_S(S) = \lambda_2\|S\|_1 + \lambda_3 \sum_{i=1}^{t} \|W \odot DS_i\|_{2,1},$$

where $\odot$ denotes the Hadamard product. For $\ell_{2,1}$-norm, the expression of $\ell_{2,1}$-norm of matrix S is as follows:

$$\|S_i\|_{2,1} = \sum_{k=1}^{n_1} (\sum_{l=1}^{n_2} S_i^2(k,l))^{\frac{1}{2}}.$$

We perform a group sparse constraint on the difference of cloud/shadow images, which is expressed as $\|DS_i\|_{2,1}$, where the difference D of S contains two spatial dimensional differential operators D_x and D_y. The specific expressions of the three differential operators are as follows:

$$D_xS_i(k,l) = S_i(k,l+1) - S_i(k,l),$$
$$D_yS_i(k,l) = S_i(k+1,l) - S_i(k,l).$$

Like the weighted nuclear norm, adding an adaptive weight vector can improve the expression effect of the group sparse regularization term. The specific expression of the weighted group sparse regularization is as follows :

$$\|W \odot DS_i\|_{2,1} = \sum_{k=1}^{n_1} W_x(k,l)\|D_xS_i(k,l)\|_2 + \sum_{l=1}^{n_2} W_y(k,l)\|D_yS_i(k,l)\|_2.$$

For the residual component matrix N, we impose the Frobenius norm constraint on it, that is, $\phi_N(N) = \frac{1}{2}\|N\|_F^2$.

(3) **The proposed model.**

$$\begin{aligned} &\min_{X,S,N} \frac{1}{2}\|N\|_F^2 + \lambda_1\|X\|_{w,*} + \lambda_2\|S\|_1 + \lambda_3\sum_{i=1}^{t}\|W \odot DS_i\|_{2,1}, \\ &s.t.\ Y = X + S + N, \end{aligned} \tag{2.1}$$

where $X \in R^{n_1 \times n_2}$ is the clean cloudless image, $Y \in R^{n_1 \times n_2}$ is the observed cloud/shadow contaminated image, $S \in R^{n_1 \times n_2}$ is the cloud/shadow component, and $N \in R^{n_1 \times n_2}$ is the residual component matrix, λ_1, λ_2, and λ_3 are regularization coefficients, and W is weight matrix. Next, we propose the ADMM algorithm for solving the model, and give the complexity analysis of the algorithm.

2.3. **ADMM-WNGS Optimization Algorithm.** Based on our prior knowledge of natural images and cloud/shadow, we propose a multitemporal image cloud removal model based on group sparse priors and weighted nuclear norm.

After fully analyzing the underlying features and priors of multitemporal image and cloud, the weighted nuclear norm is used to represent the low-rank property of multitemporal image, and the ℓ_1-norm is used to represent the sparse prior of cloud/shadow. In addition, the piecewise smooth structure of multitemporal images in different bands should be the same, so group sparsity is used to represent the group sparsity structure of cloud/shadow in the spectral dimension.

To effectively solve model (2.1), we introduce auxiliary variables R and Q_i. The problem can be rewritten as:

$$\begin{aligned} &\min_{X,S,N,R,Q_i} \frac{1}{2}\|N\|_F^2 + \lambda_1\|X\|_{w,*} + \lambda_2\|R\|_1 + \lambda_3\sum_{i=1}^{t}\|W \odot Q_i\|_{2,1}, \\ &s.t.\ Y = X + S + N, R = S, Q_i = DS_i. \end{aligned} \tag{2.2}$$

Since ADMM strategy is an important method for solving convex optimization problems. It has the advantages of fast processing speed and good convergence performance. We adopt ADMM based on augmented Lagrangian multipliers (ALM) to solve the optimization problem. Next we demonstrate the augmented Lagrangian function.

$$\begin{aligned} L_\beta(X,S,N,R,Q_i,W_j,\Lambda_i) =& \frac{1}{2}\|N\|_F^2 + \lambda_1\|X\|_{w,*} + \lambda_2\|R\|_1 + \lambda_3\sum_{i=1}^{t}\|W\odot Q_i\|_{2,1} \\ &+ \langle W_1, Y-X-S-N\rangle + \frac{\beta}{2}\|Y-X-S-N\|_F^2 \\ &+ \langle W_2, R-S\rangle + \frac{\beta}{2}\|R-S\|_F^2 \\ &+ \sum_{i=1}^{t}(\langle \Lambda_i, Q_i - DS_i\rangle + \frac{\beta}{2}\|Q_i - DS_i\|_F^2), \end{aligned}$$

where $\langle :,:\rangle$ is the inner product of the corresponding matrix.

There are a large number of variables in the augmented Lagrangian function sorted out above, so it is difficult to solve the display solution of all variables at once. So we choose to optimize one variable at a time while keeping the others constant until the convergence condition is met. Below we show the steps to solve each variable.

(1) R sub-problem:

Fractions related to R are extracted from the augmented Lagrangian function and sorted as follows:

$$\begin{aligned} R &= \arg\min_R \ L(X,S,N,R,Q_i,W_j,\Lambda_i), \\ &= \arg\min_R \ \lambda_2\|R\|_1 + \frac{\beta}{2}\|R-S+\frac{W_2}{\beta}\|_F^2. \end{aligned}$$

The soft threshold contraction operator can be used to obtain the solution of R:

$$R = \text{Shrinkage}(S-\frac{W_2}{\beta}, \frac{\lambda_2}{\beta}), \tag{2.3}$$

where $\text{Shrinkage}(a,b) = \text{sign}(a)\max(|a| \ -b, 0)$.

(2) Q_i sub-problem:

Fractions related to Q_i are extracted from the augmented Lagrangian function and sorted as follows:

$$\begin{aligned} Q_i &= \arg\min_{Q_i} \ L(X,S,N,R,Q_i,W_j,\Lambda_i), \\ &= \arg\min_{Q_i} \ \lambda_3\|W\odot Q_i\|_{2,1} + \frac{\beta}{2}\|Q_i - DS_i + \frac{\Lambda_i}{\beta}\|_F^2. \end{aligned}$$

Let $DS_i - \dfrac{\Lambda_i}{\beta} = G$. The closed-form solution of Q_i is calculated by the following:

$$Q_i = \begin{cases} \dfrac{\|G\|_2 - \frac{W\lambda_3}{\beta}}{\|G\|_2} G & , \quad if \ \dfrac{W\lambda_3}{\beta} < \|G\|_2, \\ 0 & , \quad otherwise. \end{cases} \tag{2.4}$$

(3) X sub-problem:

We refer to the solution of [34].

Fractions related to X are extracted from the augmented Lagrangian function and sorted as follows:

$$\begin{aligned} X &= \arg\min_X \ L(X,S,N,R,Q_i,W_j,\Lambda_i), \\ &= \arg\min_X \ \lambda_1\|X\|_{w,*} + \frac{\beta}{2}\|Y-X-S-N+\frac{W_1}{\beta}\|_F^2, \end{aligned}$$

which is a matrix nuclear norm minimization problem, and it can obtain a closed-form solution by using the singular value thresholding operator. The solution of X is

$$\begin{aligned} X^* &= D_{\frac{\lambda_1 w}{\beta}}(Y-S-N+\frac{W_1}{\beta}), \\ &= U(\tilde{\Sigma}-\frac{\lambda_1 w}{\beta})_+ V^T. \end{aligned} \tag{2.5}$$

Among them, $U(\tilde{\Sigma})V^T$ is the singular value decomposition (SVD) of $Y-S-N+\frac{W_1}{\beta}$, and $(\tilde{\Sigma}-\frac{\lambda_1 w}{\beta})_+ = \max\{\tilde{\Sigma}-\frac{\lambda_1 w}{\beta},0\}$,

$$\tilde{\Sigma} = \begin{pmatrix} diag(\sigma_1(X^*),\sigma_2(X^*),\cdots,\sigma_n(X^*)) \\ 0 \end{pmatrix},$$

and

$$\sigma_i(x^*) = \begin{cases} 0 & , \quad if\ c_2<0, \\ \dfrac{c_1+\sqrt{c_2}}{2} & , \quad if\ c_2 \geq 0, \end{cases}$$

where

$$\begin{aligned} c_1 &= \sigma_i(Y-S-N+\frac{W_1}{\beta})-\varepsilon\ , \\ c_2 &= [\sigma_i(Y-S-N+\frac{W_1}{\beta})+\varepsilon]^2-4C, \end{aligned}$$

where ε is a very small positive number, and C is a compromise constant.

(4) S_i sub-problem:

Fractions related to S_i are extracted from the augmented Lagrangian function and sorted as follows:

$$\begin{aligned} S_i &= \arg\min_{S_i} \ L(X,S,N,R,Q_i,W_j,\Lambda_i), \\ &= \arg\min_{S_i} \ \frac{\beta}{2}\|Y_i-X_i-S_i-N_i+\frac{W_{1i}}{\beta}\|_F^2+\frac{\beta}{2}\|R_i-S_i+\frac{W_{2i}}{\beta}\|_F^2+\frac{\beta}{2}\|Q_i-DS_i+\frac{\Lambda_i}{\beta}\|_F^2, \end{aligned}$$

which is a least squared problem and equivalent to solve the following linear equations system. It is equivalent to

$$-\beta(Y_i-X_i-S_i-N_i+\frac{W_{1i}}{\beta})-\beta(R_i-S_i+\frac{W_{2i}}{\beta})-D^T\beta(Q_i-DS_i+\frac{\Lambda_i}{\beta})=0,$$

Then

$$(2I+D^TD)S_i = (Y_i-X_i-N_i+\frac{W_{1i}}{\beta})+(R_i+\frac{W_{2i}}{\beta})+D^T(Q_i+\frac{\Lambda_i}{\beta}),$$

Algorithm 1 WNGS with ADMM Optimization.

Input: Degraded image $Y \in R^{mn\times bt}$, the parameters $\lambda_1, \lambda_2, \lambda_3, \beta, \varepsilon, k_{max}$.
Initialization: Let $X = Y$, $S = N = R = W_1 = W_2 = 0$, $Q_i = \Lambda_i = 0$, and $k = 0$.
while stopping condition is not satisfied **do**
(a) Compute R via (2.3).
(b) Compute Q_i via (2.4).
(c) Compute X via (2.5).
(d) Compute S_i via (2.6).
(e) Compute N via (2.7).
(f) Update the three Lagrange multipliers W_1, W_2 and Λ_i via (2.8).
(g) Update W in Q_i via (2.9).
Check the stopping criterion $\frac{\|X^{k+1}-X^k\|_F}{\|X^k\|_F} \le \varepsilon$ and $k < k_{max}$.
end while

Using the fast Fourier transform, the solution can be obtained as follows:

$$S_i = \text{ifft}(\frac{A}{2+|\text{fft}(D)|^2}). \tag{2.6}$$

ifft is the inverse of the fast Fourier transform, and $A = \text{fft}((Y_i - X_i - N_i + \frac{W_{1i}}{\beta}) + (R_i + \frac{W_{2i}}{\beta}) + D^T(Q_i + \frac{\Lambda_i}{\beta}))$.

(5) N sub-problem:

Fractions related to N are extracted from the augmented Lagrangian function and sorted as follows:

$$\begin{aligned} N &= \arg\min_N \ L(X,S,N,R,Q_i,W_j,\Lambda_i), \\ &= \arg\min_N \ \frac{1}{2}\|N\|_F^2 + \frac{\beta}{2}\|Y - X - S - N + \frac{W_1}{\beta}\|_F^2, \end{aligned}$$

Consistent with the analysis of S_i in the above text, it is regarded as the least squares problem, and the solution is

$$N = \frac{\beta(Y - X - S + \frac{W_1}{\beta})}{1+\beta}. \tag{2.7}$$

(6) Update the Lagrange multipliers W_1, W_2, Λ_i:

$$\begin{aligned} W_1 &= W_1 + \beta(Y - X - S - N), \\ W_2 &= W_2 + \beta(R - S), \\ \Lambda_i &= \Lambda_i + \beta(Q_i - DS_i). \end{aligned} \tag{2.8}$$

(7) Update W in Q_i sub-problem:

$$W = \frac{1}{\|(DS + \frac{\Lambda}{\beta})\|_2 + \varepsilon}, \tag{2.9}$$

where ε is designed to avoid singularities.

See Algorithm 1 for the specific process of model solving. In our algorithm, we assume that we input a multi-time data set of size $m \times n \times b \times t$, and the complexity of our proposed

(a) 02/24/2020 (b) 08/03/2019 (c) 02/08/2020 (d) 09/17/2019

FIGURE 2. France images with bands 4, 3 and 2 taken on four different times by Landsat-8. (a) February 24, 2020. (b) August 3, 2019. (c) February 8, 2020. (d) September 17, 2019. ("MM/DD/YYYY" corresponds to the date the picture was taken)

(a) 03/29/2022 (b) 04/08/2022 (c) 06/07/2022 (d) 06/17/2022

FIGURE 3. Jiangsu images with bands 8, 4 and 3 taken on four different times by Sentinel-2. (a) March 29, 2022. (b) April 8, 2022. (c) June 7, 2022. (d) June 17, 2022. ("MM/DD/YYYY" corresponds to the date the picture was taken)

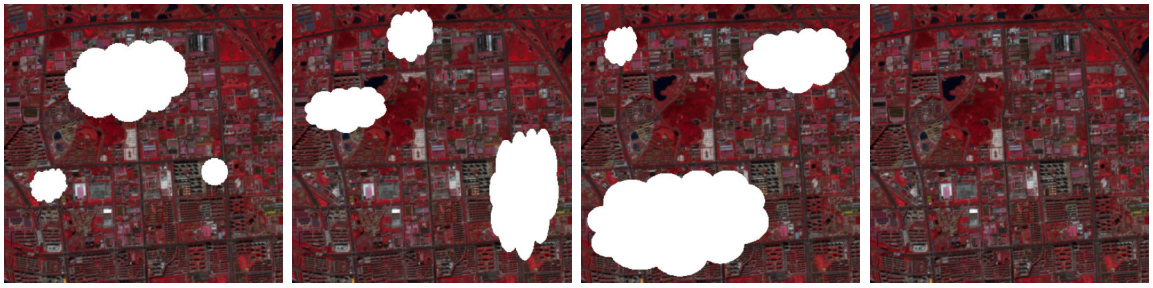
algorithm is analyzed as follows. In the subproblem of updating R, we use a fast soft threshold shrinkage operator to solve it, so that the computational complexity of updating R is $O(mnbt)$. In the subproblem of updating Q_i, we use the soft threshold shrinkage operator, so that the computational complexity of updating Q_i is $O(mnbt)$. In the subproblem of updating X, the singular value decomposition of the matrix of size is carried out, so that the computational complexity of updating X is $O(mn(bt)^2)$; in the subproblem of updating S_i, the operator is used to make the computational complexity of updating S_i be $O(mnbtlog(mnb))$. In the sub-problem of updating N, the computational complexity of updating N is $O(mnbt)$ because it is to calculate the matrix of size. In summary, the computational complexity of our proposed cloud removal algorithm is $O(3mnbt + mn(bt)^2 + mnbtlog(mnb))$ at each iteration.

3. EXPERIMENTS AND RESULTS

This part verifies the effectiveness of the proposed method through experiments. The proposed method is compared with HaLRTC, TNN, and TVLRSDC. When we run the code for the comparison method, we choose the parameters of the comparison method based on the papers code settings or the author's tendency. The experimental conditions in this chapter are as follows: The computer equipment is a desktop computer, the CPU is (Intel)Intel(R) Core(TM)

(a) Case1 (b) Case2 (c) Case3 (d) Original

FIGURE 4. Add different types of clouds to France dataset on February 24, 2020. (a) A small cloud. (b) A medium Cloud. (c) A big cloud. (d) Cloud free image on February 24, 2020.



(a) Case4 (b) Case5 (c) Case6 (d) Original

FIGURE 5. Add different types of clouds to Jiangsu dataset on March 29, 2022. (a) Two small clouds and one medium cloud. (b) Two small clouds and one medium cloud. (c) A large cloud, a medium cloud and a small cloud. (d) Cloud free image on March 29, 2022.

i7-9700 CPU @ 3.00GHz(3000 MHz), the memory is 8.00 GB(2666 MHz), the operating system is Windows10, and the test platform is MATLAB R2022a. In this part of the experiment, we respectively carried out the simulation experiment on the simulated data set and the real experiment on the real data set. In order to verify the effectiveness of the cloud removal method proposed by us in various situations, we selected the test data set and the real data set, especially the images collected from different satellites, images of different sizes and clouds of different sizes. In the real experiment, we selected the multitemporal images with clouds at one time node and the multitemporal images with clouds at four time nodes respectively.

3.1. **Simulated Experiment.** The experimental results of the simulation experiment demonstrate the visual effect of the image after the comparison method and our method to remove the cloud as well as PSNR, SSIM, CC and other indicators, while the experimental results of the real experiment mainly demonstrate the visual effect of the image after the comparison method and our method to remove the cloud/shadow. Both simulation and real experiment results show that our method is better than other advanced comparison methods in cloud removal.

(1) Datasets. In this part, we quantitatively test the validity of our proposed method with simulation data (see Fig. 2 and Fig. 3). Details of the test data set are given below.

TABLE 1. Quantitative evaluation of cloud/shadow removal effects of various methods on France and Jiangsu datasets.

Case	Index	Target	HaLRTC	TNN	TVLRSDC	Ours
Case1	PSNR	12.178	34.330	32.810	48.927	**54.470**
	SSIM	0.9440	0.9692	0.9649	0.9991	**0.9995**
	CC	0.3835	0.9790	0.9658	0.9992	**0.9997**
	Time	-	**0.2798**	1.4937	0.8864	4.4561
Case2	PSNR	7.1246	28.169	27.342	40.324	**44.888**
	SSIM	0.8400	0.8970	0.8918	0.9961	**0.9978**
	CC	0.3188	0.9132	0.8821	0.9935	**0.9974**
	Time	-	**0.2929**	1.5360	1.5970	4.5996
Case3	PSNR	3.7068	24.485	23.930	35.129	**42.069**
	SSIM	0.6520	0.7636	0.7582	0.9888	**0.9958**
	CC	0.1575	0.7949	0.7625	0.9838	**0.9959**
	Time	-	**0.2938**	1.5026	2.9887	4.5800
Case4	PSNR	9.3305	33.330	36.641	43.680	**47.391**
	SSIM	0.8913	0.9583	0.9781	0.9965	**0.9977**
	CC	0.2811	0.9790	0.9899	0.9982	**0.9992**
	Time	-	**0.2044**	1.4950	1.2054	4.6575
Case5	PSNR	8.2935	32.262	35.661	42.903	**47.633**
	SSIM	0.8640	0.9448	0.9722	0.9974	**0.9985**
	CC	0.2895	0.9729	0.9876	0.9979	**0.9992**
	Time	-	**0.2106**	1.5381	1.3529	4.6959
Case6	PSNR	5.7701	27.164	31.693	41.733	**46.329**
	SSIM	0.7793	0.8864	0.9474	0.9970	**0.9978**
	CC	0.3133	0.9145	0.9706	0.9973	**0.9989**
	Time	-	**0.2177**	1.5076	2.0239	4.8831

France: Satellite RS image taken over France, by Landsat-8, and each time node contains four spectral bands (B1, B2, B3, and B4) with 30-m spatial resolution. The subimages of size $400 \times 400 \times 4$ of four time nodes are used in experiments.

Jiangsu: Satellite RS image taken over Jiangsu, China, by Sentinel-2, and each time node contains four spectral bands (B2, B3, B4, and B8) with 10-m spatial resolution. The subimages of size $400 \times 400 \times 4$ of four time nodes are used in experiments.

In the above two data sets, we selected one of the time node images of each data set to add simulated cloud/shadow as cloud pollution images, and three of them were used as reference images (see Fig. 4 and Fig. 5). On the France dataset, a small cloud is added to the cloudless image with the time of February 24, 2020 as the Case 1 of the France dataset; Add a medium cloud to the cloudless image with the time of August 3, 2019 as Case 2 of the France dataset; Add a large cloud to the cloudless image dated February 8, 2020 as Case 3 for the France dataset. Thus, we realized the addition of clouds of different sizes to the France dataset, which

https://theia.cnes.fr/atdistrib/rocket/#search?collection=LANDSAT
https://scihub.copernicus.eu/dhus/#/home

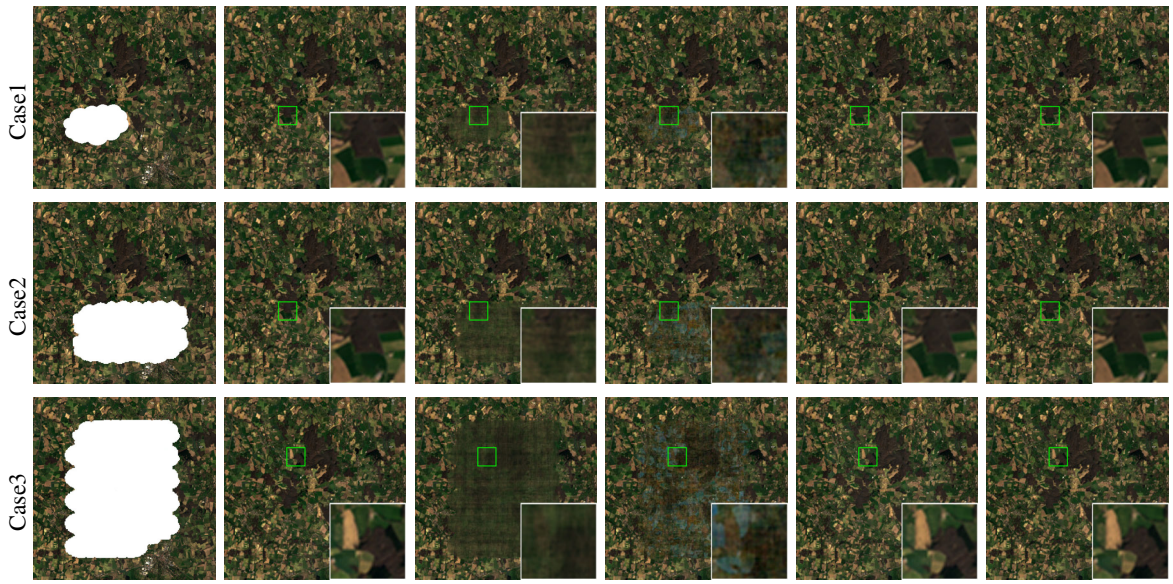


FIGURE 6. (a) Pseudocolor images of simulated observed images in Cases 1-3 of the France dataset. (b) Clean images. (c)-(f) Cloud removal results by all methods.

can be used to detect the de-cloud effect of our proposed method under the image processing cloud pollution of different sizes.

In Jiangsu data set, the combination of two small clouds and one medium cloud is added to the cloudless image on March 29, 2022, which is regarded as Case 4 in Jiangsu data set. The combination of two small clouds and one medium cloud is added to the cloudless image on April 8, 2022, which is regarded as Case 5 of Jiangsu data set. Add a combination of a large cloud, a medium cloud and a small cloud to the cloudless image on June 7, 2022 as Case 6 of Jiangsu data set. Thus, we realize adding different combinations of clouds to the Jinagsu dataset, which can be used to test the cloud removal effect of our proposed method in processing images polluted by different combinations of clouds.

(2) Evaluation Indexes. We select peak signal-to-noise ratio (PSNR), the structural similarity (SSIM), and the correlation coefficient (CC) as evaluation index of picture recovery quality, which are defined as follows:

$$\mathrm{PSNR}_i = 10 \times \log \frac{pq}{\left\| X_i - \hat{X}_i \right\|_F^2},$$

$$\mathrm{SSIM}_i = \frac{\left(2\mu_{X_i}\mu_{\hat{X}_i} + c_1\right)\left(2\sigma_{X_i\hat{X}_i}\right)}{\left(\mu_{X_i}^2 + \mu_{\hat{X}_i}^2 + c_1\right)\left(\sigma_{X_i}^2 + \sigma_{\hat{X}_i}^2 + c_2\right)},$$

$$\mathrm{CC} = \frac{\sum_{j=1}^{k}\left(X_j - \mu_X\right)\left(\hat{X}_j - \mu_{\hat{X}}\right)}{\left(\sum_{j=1}^{k}\left(X_j - \mu_X\right)^2 \sum_{j=1}^{k}\left(\hat{X}_j - \mu_{\hat{X}}\right)^2\right)^{\frac{1}{2}}}.$$

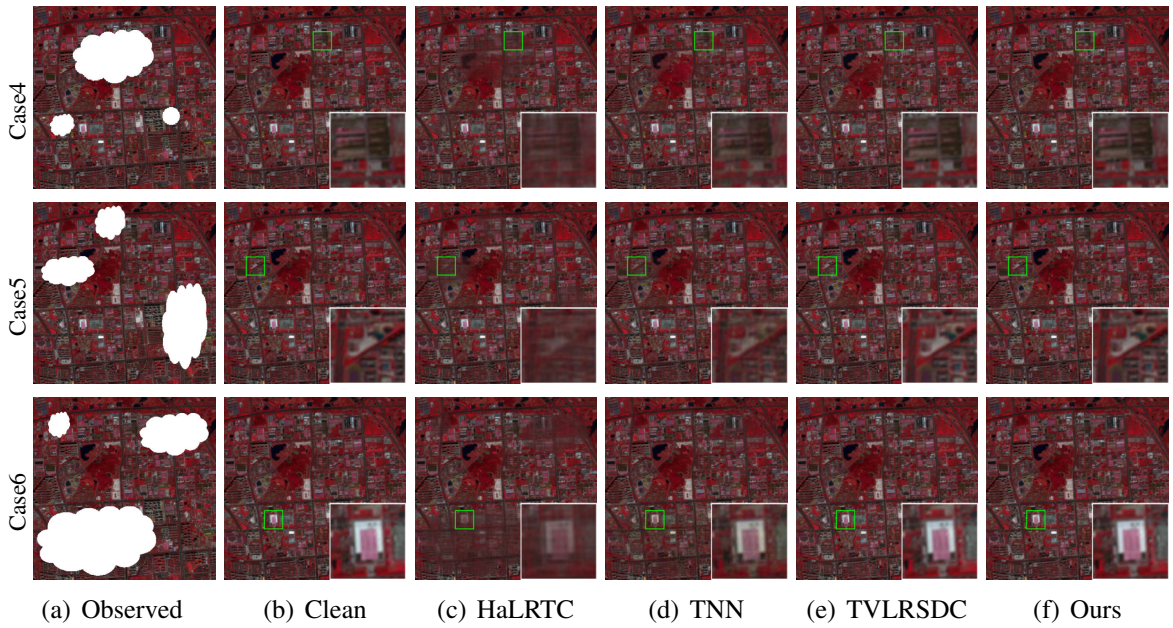


FIGURE 7. (a) Pseudocolor images of simulated observed images in Cases 4-6 of the Jiangsu dataset. (b) Clean images. (c)-(f) Cloud removal results by all methods.

where p and q denote the sizes of each band; X_i and $\hat{X}_i$ denote the i-th band of the original and reconstructed images; μ_{X_i} and μ_{X_i} denote the average values of X_i and $\hat{X}_i$; $\sigma^2_{X_i}$ and $\sigma^2_{\hat{X}_i}$ stand for the variances; $\sigma_{X_i\hat{X}_i}$ is the covariance between X_i and $\hat{X}_i$; c_1 and c_2 are default constants; X_j and $\hat{X}_j$ are the original and the reconstructed values of the j-th contaminated pixels, respectively; k denotes the number of contaminated pixels; and μ_X and $\mu_{\hat{X}}$ are the average values.

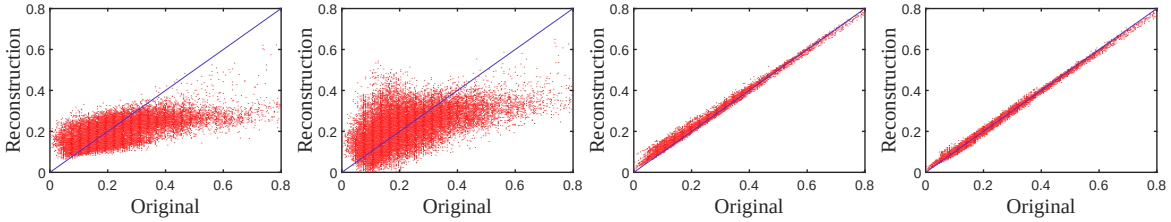
(3) Results on Simulated Datasets. We made a quantitative comparison between our proposed method and the comparison method, and the results obtained are shown in the Table 3.1. The higher indicators are shown in bold.

It can be seen from the data in the table that under different conditions of the two simulated data sets, our method has obtained satisfactory results under the three indexes of PSNR, SSIM and CC. No matter in the case of covering small clouds, medium clouds, large clouds and composite clouds, our method can well remove the clouds on the France and Jiangsu data sets and restore the original image well.

At the same time, it should be pointed out that our method is inferior to other comparison methods in terms of usage time, but the effect of cloud removal is superior to other comparison methods.

In terms of vision, the cloud removal results of all methods in different situations of France and Jiangsu data sets are shown in Fig. 6 and Fig. 7. In order to observe the effect of different methods of cloud removal more clearly, part of the image without cloud and the image with cloud removal were locally enlarged.

It can be seen from the France data set Cases 1-3 that all the methods can remove the white cloud, but only the TVLRSDC method and our method can restore the original image effect more completely after removing the cloud. Especially in Case 2 and Case 3, although the



(a) HaLRTC (b) TNN (c) TVLRSDC (d) Ours

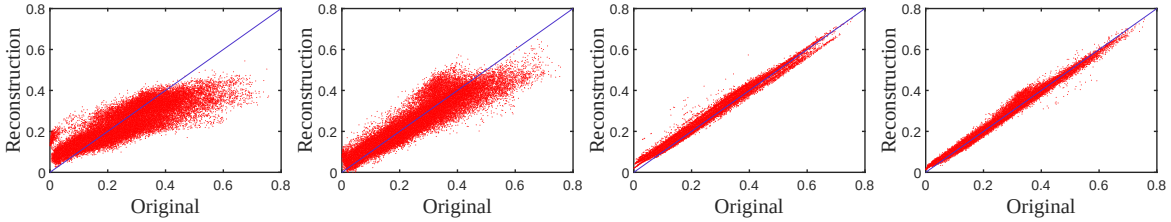
FIGURE 8. Scatter diagrams between the original and reconstructed pixels of the cloud-contaminated regions in Case 1 of the France dataset.

HaLRTC and TNN methods remove clouds, the image recovery effect is poor, and the TNN method even appears serious chromatic aberration, which indicates that HaLRTC and TNN have limitations when removing a large area of clouds, while TVLRSDC and our method can cope with it better. Meanwhile, compared with TVLRSDC, our method can recover more details of the original image information.

Similarly, in Cases 4-6 of Jinagsu dataset, it can be seen that all comparison methods, including the method proposed by us, can completely remove clouds, but HaLRTC and TNN cannot restore the original information of the image after removing clouds, and the part with clouds will become fuzzy after removing clouds. TVLRSDC and our method can recover the original image information better. Compared with TVLRSDC's cloud removal and restoration, our method is more similar to the original clean image. For example, in Case 6, the image obtained by our cloud removal method is brighter and more similar to the original clean image.

In order to further measure the effect of our method and other comparison methods in simulating cloud removal on cloud data sets, we made scatter plots between the original pixels and the reconstructed pixels after cloud removal on the two data sets, as shown in Fig. 8 and Fig. 9. The figure shows the scatter plot of France data set Case 1. It can be seen that the scatter plot of subgraph (a) HaLRTC and subgraph (b) TNN method deviates from the diagonal, and the scatter plot of subgraph (c) TVLRSDC and subgraph (d) under our method almost overlaps with the diagonal. Compare the scatter plot of subgraph (c) and subgraph (d). The scatter graph corresponding to TVLRSDC still has some deviation from the diagonal at the beginning, which indicates that the proposed method has obvious advantages in the reconstruction of multitemporal image information. The picture shows the scatter plot of Jiangsu data set Case 4. Similarly, similar to the French data set Case 1, the scatter plot of subgraph (a) HaLRTC and subgraph (b) TNN method deviates from the diagonal, and the scatter plot corresponding to subgraph (c) TVLRSDC is distributed around the diagonal, but some of them deviate from the diagonal in the late period. The corresponding methods in subgraph (d) are almost distributed around the diagonal line, again showing the advantages of our proposed method in information reconstruction.

3.2. **Real Experiment.** In this section, we test the validity of our proposed approach with real data sets. The details of the real data set are given below.



(a) HaLRTC (b) TNN (c) TVLRSDC (d) Ours

FIGURE 9. Scatter diagrams between the original and reconstructed pixels of the cloud-contaminated regions in Case 4 of the Jiangsu dataset.

(1) Datasets. Tokyo: Satellite RS image taken over Tokyo, Japan, by Sentinel-2, and each time node contains six spectral bands (B5, B6, B7, B8A, B11 and B12) with 20-m spatial resolution. The subimages of size $1500\times1500\times6$ of four time nodes are used in experiments.

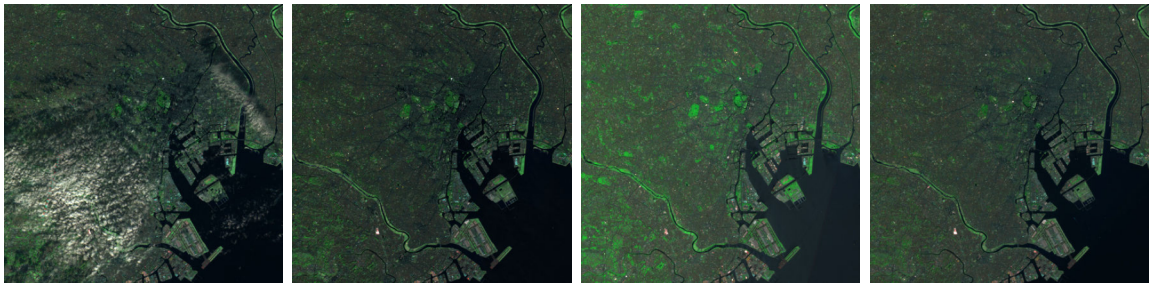
Jinhua[4]: Satellite RS image taken over Jinhua, China, by Sentinel-2, and each time node contains six spectral bands (B5, B6, B7, B8A, B11 and B12) with 20-m spatial resolution. The subimages of size $1500\times1500\times6$ of four time nodes are used in experiments.

The Tokyo data set we have selected is one with clouds on the first Jinhua data set and one with no clouds on the other three Jinhua data sets. On the basis of Jinhua data sets, we have selected data sets with clouds on four time nodes. On the basis of Jinhua data sets, we have Jinhua data sets with different distributions on images. The details are in Fig. 10 and Fig. 11.

(2) Results on Real Datasets. With all kinds of Jinhua data sets, we have carried out cloud stripping on both Tokyo and Jinhua data sets, and the results obtained are shown in Fig. 12 and Fig. 13. First, the graph shows the cloud removal effect of all methods on the Tokyo data set. As in the simulation experiment, we partially enlarged the details of the de-cloud image. From subgraph (b) to (e), it can be seen that HaLRTC and TNN methods did not completely remove the cloud/shadow, and cloud/shadow can still be seen in the partially amplified part. In contrast, the method proposed by us in TVLRSDC and subgraph (d) removed the cloud and better restored the information of the original image.

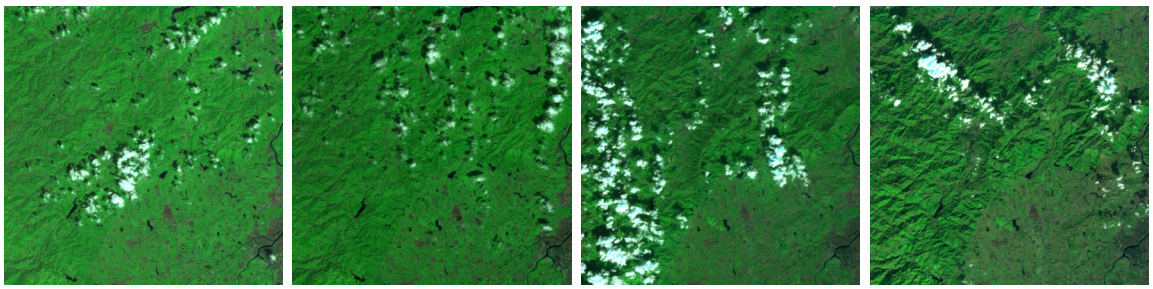
Figure shows discloud effects of all methods on Jinhua data set. In the first, second and fourth time node images on Jinhua data set, the partial magnifies are still very fuzzy after removing clouds by HaLRTC and TNN, indicating that original information on images has not been recovered. Although the partial features on TVLRSDC are not fuzzy after removing clouds, there are still traces of clouds where there were clouds. The method we propose can remove the cloud, but also ensure that the original cloud areas are less visible. In terms of de-cloud effect on images on 3rd time node under Jinhua data set, the method we have proposed can minimize the influence of cloud shadows on images compared with other comparison methods. In the image of the third time node, although TVLRSDC method can remove the cloud, it can not completely remove the cloud shadow. Our proposed method has advantages in removing or diluting the cloud shadow.

[4] https://scihub.copernicus.eu/dhus/#/home



(a) 01/18/2017 (b) 05/08/2017 (c) 12/14/2017 (d) 03/14/2018

FIGURE 10. Tokyo images with bands 12, 8A and 5 taken on four different times by Sentinel-2. (a) January 18, 2017. (b) May 8, 2017. (c) December 14, 2017. (d) March 14, 2018. ("MM/DD/YYYY" corresponds to the date the picture was taken)



(a) 09/10/2022 (b) 09/20/2022 (c) 10/30/2021 (d) 12/24/2021

FIGURE 11. Jinhua images with bands 12, 8A and 5 taken on four different times by Sentinel-2. (a) September 10, 2022. (b) September 20, 2022. (c) October 30, 2021. (d) December 24, 2021. ("MM/DD/YYYY" corresponds to the date the picture was taken)

(a) Observed (b) HaLRTC (c) TNN (d) TVLRSDC (e) Ours

FIGURE 12. Pseudocolor images with bands 12, 8A and 5 of the real observed image on the Tokyo dataset. (a) Observed. (b)-(e) Cloud removal results by all methods.

4. DISCUSSION

4.1. **Parameter Analysis.** The cloud removal model proposed by us mainly includes four parameters, which are regularization parameters λ_1, λ_2, λ_3 and Lagrangian multiplier β. We used PSNR as an evaluation index to analyze every parameter in the France and Jiangsu data sets we

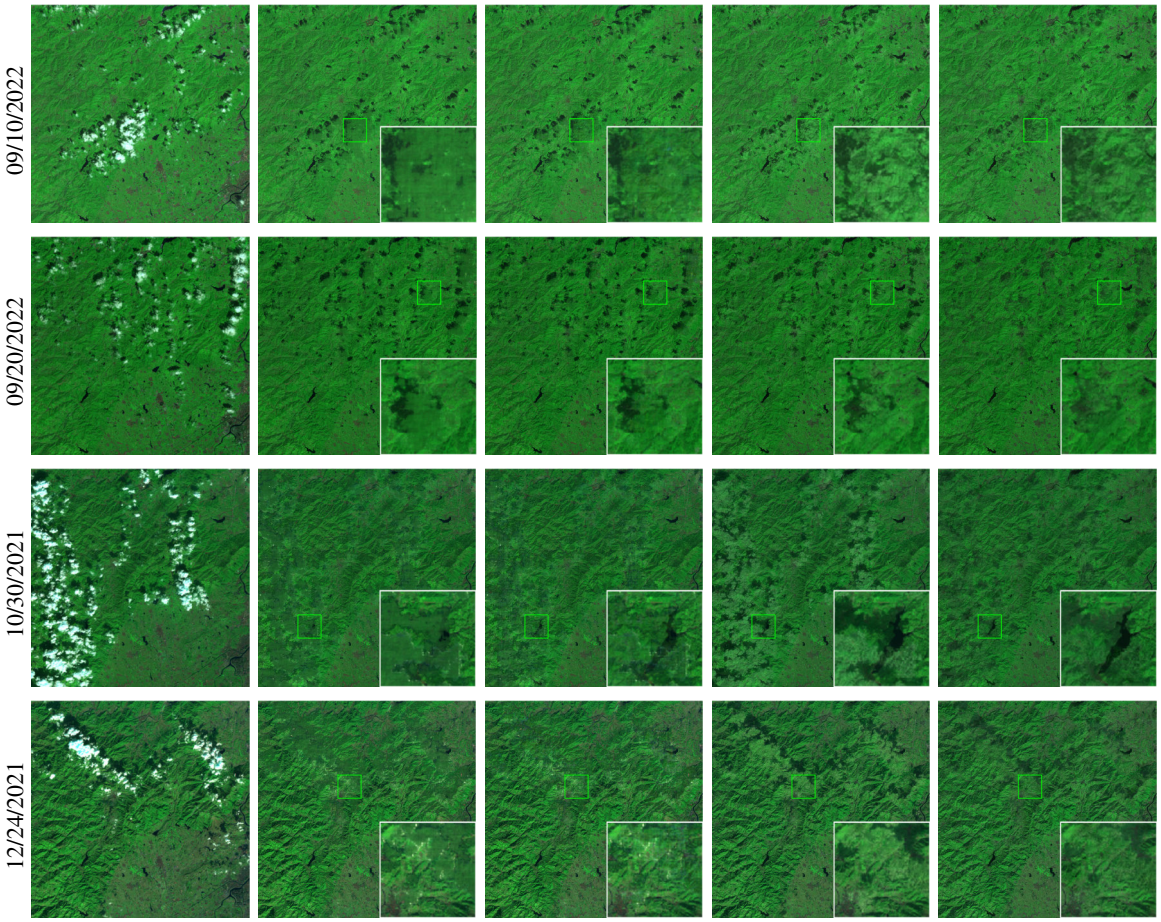


FIGURE 13. Pseudocolor images with bands 12, 8A and 5 of the real observed image on the Jinhua dataset. (a) Observed. (b)-(e) Cloud removal results by all methods.

selected. In order to more clearly show the influence of different parameters on the PSNR value, we drew a line graph of the changes of PSNR in Fig. 14 when different parameters changed.

We set the change range of λ_1 to be 1 to 10, the change range of λ_2 and λ_3 to be 0.001, and the change range of β to be 0.1 to 1. Finally, in the simulation experiment of this paper, we set λ_1 to be 1 and λ_2 to be 0.001. Set λ_3 to 0.005 and β to 0.1 to get the highest PSNR value.

4.2. **Convergence Behavior.** Meanwhile, we drew a line graph of the number of iterations and relative change (RelCha) through simulation and real experiments to analyze the convergence of our proposed model algorithm, and the resulting graph was shown in the Fig. 15.

We define RelCha in the k th iteration as :

$$\text{RelCha} = \frac{\left\|X^k - X^{k-1}\right\|_F}{\left\|X^{k-1}\right\|_F}.$$

As can be seen from the figure, for different data sets, whether simulated or real, RelCha values gradually tend to zero, which indicates the convergence of the developed algorithm.

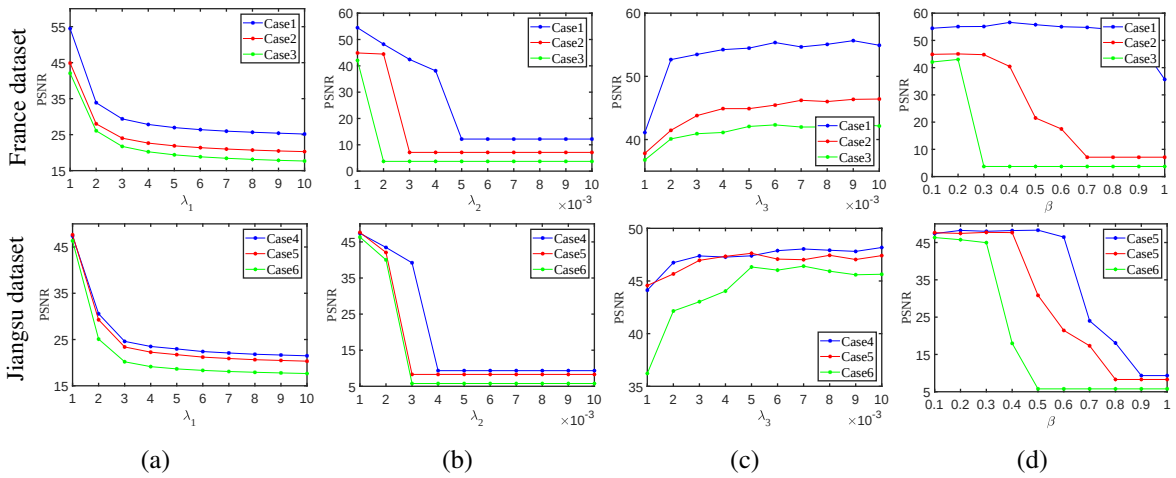


FIGURE 14. Sensitivity analysis of the parameters on France dataset and Jiangsu dataset. (a) PSNR versus regularization parameter λ_1. (b) PSNR versus regularization parameter λ_2. (c) PSNR versus penalty parameter λ_3 . (d) PSNR versus penalty parameter β.

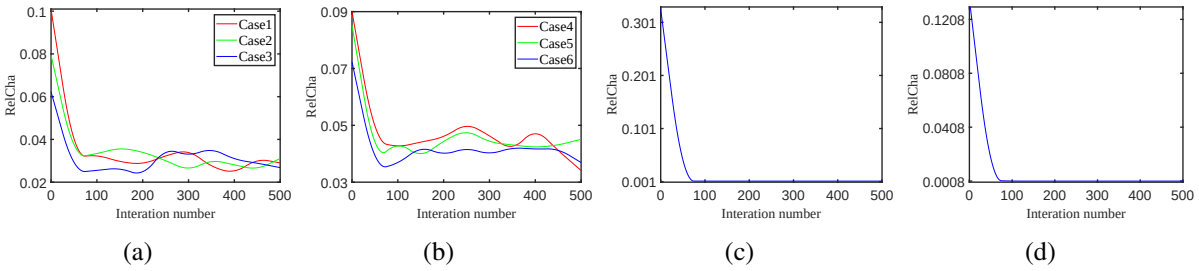


FIGURE 15. RelCha versus iteration number in different cases of two datasets. (a) Cases 1-3 of the France dataset. (b) Cases 4-6 of the Jiangsu dataset. (c) Real experiments of the Tokyo dataset. (d) Real experiments of the Jinhua dataset.

5. CONCLUSION

In this paper, we propose a WNGS model to reconstruct multi-temporal RS images destroyed by clouds and shadows. In this model, weighted nuclear norm is used to describe the global correlation of data, and group sparse regularization describes the smoothing information and time series information of data. Different from the completion-based method, this model does not depend on the accuracy of mask, and can effectively reconstruct multi-temporal remote sensing images without providing mask. This paper presents an ADMM algorithm for WNGS model and discusses its parameter setting. Experimental results show that compared with the existing methods, the proposed method significantly improves the recovery factor in terms of visual effect and quantitative evaluation.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grants U21A20455, 61972265, 11871348, by the Natural Science Foundation of Guangdong Province

of China under grant 2020B1515310008, 2023A1515011691, by the Educational Commission of Guangdong Province of China under grant 2019KZDZX1007, and by the Shenzhen Basis Research Project under grant JCYJ20210324094006017.

REFERENCES

[1] N. Yokoya, X.X. Zhu, A. Plaza, Multisensor coupled spectral unmixing for time-series analysis, IEEE Transactions on Geoscience and Remote Sensing, 55 (2017), 2842-2857.

[2] X.L. Zhao, F. Wang, T.Z. Huang, M.K. Ng, R.J. Plemmons, Deblurring and sparse unmixing for hyperspectral images, IEEE Transactions on Geoscience and Remote Sensing, 51 (2013), 4045-4058.

[3] D. Hong, N. Yokoya, J. Chanussot, X.X. Zhu, An augmented linear mixing model to address spectral variability for hyperspectral unmixing, IEEE Transactions on Image Processing, 28 (2018), 1923-1938.

[4] L. Gao, Z. Wang, L. Zhuang, H. Yu, B. Zhang, J. Chanussot, Using low-rank representation of abundance maps and nonnegative tensor factorization for hyperspectral nonlinear unmixing, IEEE Transactions on Geoscience and Remote Sensing 60 (2021), 1-17.

[5] J. Yao, D. Meng, Q. Zhao, W. Cao, Z. Xu, Nonconvex-sparsity and nonlocal-smoothness-based blind hyperspectral unmixing, IEEE Transactions on Image Processing, 28 (2019), 2991–3006.

[6] B. Zhang, L. Zhuang, L. Gao, W. Luo, Q. Ran, Q. Du, Pso-em: A hyperspectral unmixing algorithm based on normal compositional model, IEEE Transactions on Geoscience and Remote Sensing, 52 (2014), 7782-7792.

[7] D. Hong, L. Gao, N. Yokoya, J. Yao, J. Chanussot, Q. Du, B. Zhang, Morediverse means better: Multimodal deep learning meets remote-sensing imagery classification, IEEE Transactions on Geoscience and Remote Sensing, 59 (2020), 4340-4354.

[8] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, J. Chanussot, Graph convolutional networks for hyperspectral image classification, IEEE Transactions on Geoscience and Remote Sensing, 59 (2020), 5966-5978.

[9] D. Hong, Z. Han, J. Yao, L. Gao, B. Zhang, A. Plaza, J. Chanussot, Spectralformer: Rethinking hyperspectral image classification with transformers, IEEE Transactions on Geoscience and Remote Sensing, 60 (2021), 1-15.

[10] N. Yokoya, A. Iwasaki, Object detection based on sparse representation and hough voting for optical remote sensing imagery, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8 (2015), 2053-2062.

[11] K. Wang, Y. Wang, X.L. Zhao, J.C.W. Chan, Z. Xu, D. Meng, Hyperspectral and multispectral image fusion via nonlocal low-rank tensor decomposition and spectral unmixing, IEEE Transactions on Geoscience and Remote Sensing, 58 (2020), 7654-7671.

[12] Y. Chen, J. Zeng, W. He, X.L. Zhao, T.Z. Huang, Hyperspectral and multispectral image fusion using factor smoothed tensor ring decomposition, IEEE Transactions on Geoscience and Remote Sensing, 60 (2022), 1-17.

[13] W. He, N. Yokoya, X. Yuan, Fast hyperspectral image recovery of dual-camera compressive hyperspectral imaging via non-iterative subspace-based fusion, IEEE Transactions on Image Processing, 30 (2021), 7170-7183.

[14] S. Li, R. Dian, L. Fang, J.M. Bioucas-Dias, Fusing hyperspectral and multispectral images via coupled sparse tensor factorization, IEEE Transactions on Image Processing, 27 (2018), 4118-4130.

[15] R. Dian, S. Li, B. Sun, A. Guo, Recent advances and new guidelines on hyperspectral and multispectral image fusion, Information Fusion, 69 (2021), 40-51.

[16] X.X. Liu, J. Lu, L.X Shen, C. Xu, Y.S. Xu, Multiplicative noise removal: Nonlocal low-rank model and its proximal alternating reweighted minimization algorithm, SIAM Journal on Imaging Sciences, 13 (2020),1595-1629.

[17] J. Lu, K. Qiao, X. Li, Z. Lu, Y. Zou, L(0)-minimization methods for image restoration problems based on wavelet frames, Inverse Problems, 35 (2019), 64001.

[18] C. Xu, X. Liu, J. Zheng, L. Shen, Q. Jiang, J. Lu, Nonlocal low-rank regularized two-phase approach for mixed noise removal, Inverse Problems, 37 (2021), 085001.

[19] Y.R. Zou, H.X. Hu, J. Lu, X.X. Liu, Q. T. Jiang, G, H. Song, A nonlocal low rank regularization method for fractal image coding, Fractals, 29 (2021), 2150125.

[20] X.Y Zeng, L.X. Shen, Y.S. Xu, Lu, J, Matrix completion via minimizing an approximate rank, Analysis and Applications, 17 (2019),689-713.
[21] J. Lu, J. Tian, Q. Jiang, X. Liu, Z. Hu, Y. Zou, Rician noise removal via weighted nuclear norm penalization, Applied and Computational Harmonic Analysis, 53 (2021), 180-198.
[22] H. Zhang, W. He, L. Zhang, H. Shen, Q. Yuan, Hyperspectral image restoration using low-rank matrix recovery, IEEE Transactions on Geoscience and Remote Sensing, 52 (2013), 4729-4743.
[23] W. He, H. Zhang, L. Zhang, H. Shen, Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration, IEEE Transactions on Geoscience and Remote Sensing, 54 (2015), 178-188.
[24] W. He, H. Zhang, H. Shen, L. Zhang, Hyperspectral image denoising using local low-rank matrix recovery and global spatial-spectral total variation, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11 (2018), 713-729.
[25] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, L. Zhang, Weighted nuclear norm minimization and its applications to low level vision, International Journal of Computer Vision, 121 (2017), 183-208.
[26] T.Y. Ji, D. Chu, X.L. Zhao, D. Hong, A unified framework of cloud detection and removal based on low-rank and group sparse regularizations for multitemporal multispectral images, IEEE Transactions on Geoscience and Remote Sensing, 60 (2022), 1-15.
[27] J. Lin, T.Z. Huang, X.L. Zhao, Y. Chen, Q. Zhang, Q. Yuan, Robust thick cloud removal for multitemporal remote sensing images using coupled tensor factorization, IEEE Transactions on Geoscience and Remote Sensing, 60 (2022), 1-16.
[28] Q. Zhang, Q. Yuan, C. Zeng, X. Li, Y.Wei, Missing data reconstruction in remote sensing image with a unified spatial-temporal-spectral deep convolutional neural network, IEEE Transactions on Geoscience and Remote Sensing, 56 (2018), 4274-4288.
[29] Q. Zhang, Q. Yuan, J. Li, Z. Li, H. Shen, L. Zhang, Thick cloud and cloud shadow removal in multitemporal imagery using progressively spatio-temporal patch group deep learning, ISPRS Journal of Photogrammetry and Remote Sensing, 162 (2020), 148-160.
[30] A. Meraner, P. Ebel, X.X. Zhu, M. Schmitt, Cloud removal in sentinel-2 imagery using a deep residual neural network and sar-optical data fusion, ISPRS Journal of Photogrammetry and Remote Sensing, 166 (2020), 333-346.
[31] O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better, in Proceedings of the IEEE/CVF International Conference on Computer Vision, (2019), 8878-8887.
[32] R. Li, J. Pan, Z. Li, J. Tang, Single image dehazing via conditional generative adversarial network, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (2018), 8202–8211.
[33] Y. Chen, W. He, N. Yokoya, T.Z. Huang, Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition, IEEE Transactions on Cybernetics, 50 (2019), 3556-3570.
[34] Y. Chen, W. He, N. Yokoya, T.Z. Huang, Blind cloud and cloud shadow removal of multitemporal images based on total variation regularized low-rank sparsity decomposition, ISPRS Journal of Photogrammetry and Remote Sensing, 157 (2019), 93-107.