

Blind cloud and cloud shadow removal of multitemporal images based on total variation regularized low-rank sparsity decomposition

Yong Chen^{a,b}, Wei He^{b,*}, Naoto Yokoya^{b,c,*}, Ting-Zhu Huang^a

^a School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, PR China

^b Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project (AIP), RIKEN, 103-0027 Tokyo, Japan

^c Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, 183-8538 Tokyo, Japan



ARTICLE INFO

Keywords:

Cloud/shadow removal
Multitemporal images
Low-rank and sparsity
Total variation
Alternating direction method of multipliers

ABSTRACT

Cloud and cloud shadow (cloud/shadow) removal from multitemporal satellite images is a challenging task and has elicited much attention for subsequent information extraction. Regarding cloud/shadow areas as missing information, low-rank matrix/tensor completion based methods are popular to recover information undergoing cloud/shadow degradation. However, existing methods required to determine the cloud/shadow locations in advance and failed to completely use the latent information in cloud/shadow areas. In this study, we propose a blind cloud/shadow removal method for time-series remote sensing images by unifying cloud/shadow detection and removal together. First, we decompose the degraded image into low-rank clean image (surface-reflected) component and sparse (cloud/shadow) component, which can simultaneously and completely use the underlying characteristics of these two components. Meanwhile, the spatial-spectral total variation regularization is introduced to promote the spatial-spectral continuity of the cloud/shadow component. Second, the cloud/shadow locations are detected from the sparse component using a threshold method. Finally, we adopt the cloud/shadow detection results to guide the information compensation from the original observed images to better preserve the information in cloud/shadow-free locations. The problem of the proposed model is efficiently addressed using the alternating direction method of multipliers. Both simulated and real datasets are performed to demonstrate the effectiveness of our method for cloud/shadow detection and removal when compared with other state-of-the-art methods.

1. Introduction

With the development of remote sensing technology, satellite images become effective resources for identifying the distributions of our ecological environment and have been extensively applied in many fields, such as unmixing (Yokoya et al., 2017; Zhao et al., 2013), classification (Xia et al., 2018), and object detection (Yokoya and Iwasaki, 2015). However, optical remote sensing images have been inevitably corrupted by cloud/shadow because optical imaging is severely affected by scattered cloud and illumination changes owing to cloud shadows. According to the research, cloud covers approximately 35% of the earth's surface in anytime (Ju and Roy, 2008), thereby extensively limiting the collection of cloud/shadow-free data, particularly for multitemporal image analysis. Cloud/shadow in remote sensing images is a negative problem, which not only reduces image quality but also obscures real information for subsequent applications. Therefore, removing the cloud/shadow in satellite images is critical in improving

image quality before their subsequent applications.

1.1. Related works

In the past decades, many cloud/shadow removal approaches have been proposed to reconstruct missing and degraded information. These approaches can be divided into three categories: self-complementation-based, multispectral-complementation-based, and multitemporal-complementation-based methods (Chen et al., 2017). Self-complementation-based methods take full advantage of the information from cloud/shadow-free regions in an image. The contaminated information of the cloud/shadow regions is supplemented by propagating the geometrical flow from the cloud/shadow-free regions. For instance, image inpainting techniques have been used extensively for information reconstruction (Maalouf et al., 2009; Lorenzi et al., 2011; Shen et al., 2014). However, self-complementation-based methods cannot reconstruct large-scale cloud/shadow regions, which are the actualities in

* Corresponding authors at: Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project (AIP), RIKEN, 103-0027 Tokyo, Japan.

E-mail addresses: chenyong1872008@163.com (Y. Chen), wei.he@riken.jp (W. He), naoto.yokoya@riken.jp (N. Yokoya), tingzhuhuang@126.com (T.-Z. Huang).

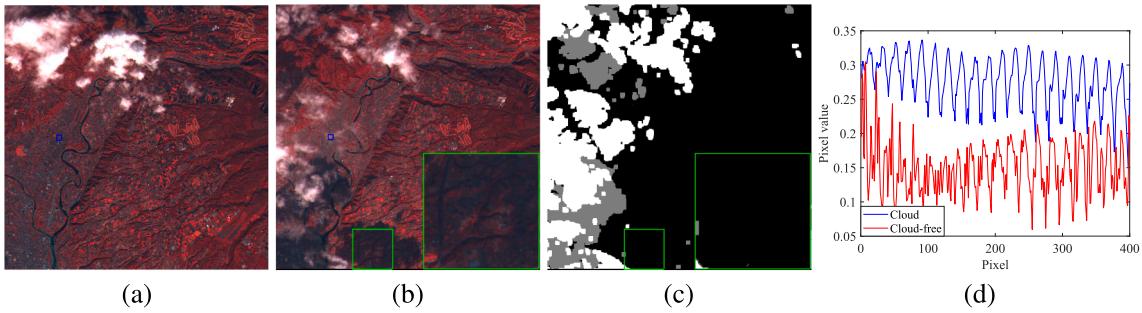


Fig. 1. (a) and (b) represent the degraded image of two adjacent time nodes, (c) the cloud/shadow mask of (b), (d) the pixel values distribution of cloud-free and thin cloud locations (blue box from (a) and (b)). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

optical remote sensing images.

To make full use of the auxiliary information, multispectral-complementation-based methods reconstruct the degraded bands by utilizing extra auxiliary clean bands (Zhang et al., 2002; Rakwatin et al., 2009; Shen et al., 2011; Wang et al., 2005; Zhang et al., 2009; Lorenzi et al., 2013; Li et al., 2019). For example, the missing data of moderate resolution imaging spectroradiometer (MODIS) band 6 is reconstructed on the basis of the strong correlation between MODIS bands 6 and 7 (Rakwatin et al., 2009; Shen et al., 2011). A haze-optimized transformation was proposed to rectify and compensate the thin cloud-contaminated visible bands of Landsat images (Zhang et al., 2002). Zhang et al. (2009) proposed a geostatistical method to interpolate the DN values of clouded pixels in remotely sensed multispectral images by employing the ordinary cokriging method. Malek et al. (2018) recently presented the strength of autoencoder networks (AE) to remove the cloud in a multispectral image. In summary, these methods can reconstruct large-scale regions of the cloud/shadow and obtain satisfactory results compared with self-complementation-based methods. However, when large-scale cloud/shadow regions appear in all bands, both self-complementation-based and multispectral-complementation-based methods have the difficulty in reconstructing the information because reference information is not sufficient (Chen et al., 2017).

Among the two categories, multitemporal-complementation-based methods are widely used to remove the cloud/shadow in remote sensing images because they can simultaneously capture the spatial, spectral and temporal coherence. Satellite sensors can collect multitemporal acquisitions of remote sensing images in the same region. Meanwhile, cloud/shadow varies with time node; thus, cloud/shadow-contaminated regions vary in different time nodes. Melgani (2006, 2008) proposed an unsupervised contextual prediction process, which efficiently used the spectral-temporal relationships to remove the cloud. Lin et al. (2013) proposed an information cloning method to eliminate clouds in multitemporal images, in which the changes in the land cover over a short time period are assumed to be minimal. Thereupon, Lin et al. (2014) again developed a patch-based information reconstruction algorithm for removing cloud in an image. Cheng et al. (2014) offered a method based on similar pixel replacement to remove clouds. The authors used a multitemporal image as a guide to locate similar pixels, and similar pixels were found by implementing a spatio-temporal Markov random field global function. To obtain a continuous cloud-free Landsat image, Chen et al. (2017) proposed a novel spatially and temporally weighted regression method to reconstruct cloudy areas. A multitemporal dictionary learning method was also proposed in Xu et al. (2016) for cloud removal. In addition, many other multitemporal-complementation-based methods are available to address the problem in cloud removal (Tseng et al., 2008; Zhang et al., 2010; Li et al., 2015; Li et al., 2016; Zhang et al., 2018). Majority of the existing multispectral-complementation-based methods can achieve satisfactory results. However, a cloud/shadow-free auxiliary image should be selected as a reference to guide cloud/shadow-free patches or to locate similar

pixels, which will remarkably depend on the availability of cloud/shadow-free images (Chen et al., 2017).

In recent, by considering the cloud/shadow regions as missing information in an image, low-rank matrix and tensor completion based methods are effectively used to reconstruct cloud/shadow-contaminated multitemporal images. Wang et al. (2016) developed a temporally contiguous robust matrix completion model to remove cloud and recover ground observation in satellite image sequences. This model used the temporal correlations of the image and solved by using augmented Lagrangian method with an inexact proximal gradient. An adaptive weighted tensor completion method was also proposed in Ng et al. (2017) to reconstruct remote sensing image with cloud coverage. Ji et al. (2017) suggested a nonlocal tensor completion method for the cloud removal of multitemporal remotely sensed images. Furthermore, the classic matrix completion (Candès and Tao, 2010) and tensor completion (Liu et al., 2013) also can be applied to cloud/shadow removal in multitemporal remotely sensed images. If the cloud/shadow locations are precisely provided, the low-rank matrix/tensor completion methods can efficiently reconstruct the cloud/shadow areas by simultaneously exploring the spatial, spectral, and temporal information.

1.2. Motivation

Although completion-based methods are popular and achieve satisfactory results in cloud/shadow removal, they also have several disadvantages. First, these methods belong to non-blind approaches because they require the cloud/shadow locations in the image. However, identifying an excellent method to accurately detect the cloud/shadow position is challenging. This phenomenon can be illustrated in Fig. 1, where Fig. 1(b) shows the degraded image and Fig. 1(c) presents the cloud/shadow mask detected using the state-of-the-art improved Fmask algorithm (Frantz et al., 2018). In the figure, black portion denotes the cloud/shadow-free locations, white portion is cloud locations, and gray portion indicates the cloud shadow locations. Based on the results, we can observe that the locations of the shadow in Fig. 1(b) (small green box) are not detected. In this case, completion-based methods always have the difficulty to eliminate the shadow in multitemporal images. Second, completion-based methods consider the pixel values of the cloud/shadow areas as missing information in the cloud/shadow areas. Generally, the areas that are covered by cloud/shadow also contain the available information. Fig. 1(b) shows the enlarged image of the shadow areas in the lower-right corner, and it also illustrates that the shadow areas contain useful image structure and texture. We cannot regard these areas as missing information and should make the use of these information. Finally, when the mask areas of thin cloud/shadow are the same in all temporal images, previous matrix/tensor completion methods cannot reconstruct the images by simply using the image low-rank prior. Therefore, the techniques in improving the practicability of the method and making use of cloud/

shadow information are the noteworthy problems.

In this paper, we propose a unified cloud/shadow detection and removal method for multitemporal remote sensing images using a low-rank sparsity decomposition framework to address the problems mentioned previously. In the proposed framework, the low-rank regularization is used to constrain the image component given that it has global correlations in spectral-temporal dimensions. The sparsity term is employed to capture the global sparse distributional property of the cloud/shadow component in the entire multitemporal images because the proportions of the cloud/shadow in many scenes are relatively low. The characteristics for low-rank image (surface-reflected) component and sparse cloud/shadow component can be captured simultaneously. Moreover, we adopt the spatial-spectral total-variation (TV) regularization to preserve the continuity of the sparse cloud/shadow component in each temporal image. An efficient alternating direction method of multipliers (ADMM) algorithm is introduced to optimize the proposed model. Finally, to better preserve the cloud/shadow-free information, a further step is to detect the sparse cloud/shadow locations using a threshold method in the extracted cloud/shadow component. Then, the information of the cloud/shadow-free locations is compensated to the low-rank clean image. Fig. 2 presents the framework of the proposed method for the blind cloud/shadow removal.

1.3. Contribution

The main contributions of this paper can be summarized as follows:

- (1) We propose a unified convex optimization model spatial-spectral TV regularized low-rank sparsity decomposition (TVLRSD) for the joint cloud/shadow detection and removal of multitemporal satellite images. Compared to existing matrix/tensor completion-based modeling methods, our method needn't cloud/shadow detection as preprocessing. Furthermore, our method is the first to exploit and model the underlying characteristics and priors hidden in the cloud/shadow component of time-series data.
- (2) We design a low-rank regularization to model the high correlation between the image component in spectral-temporal dimensions, and sparse regularization and spatial-spectral TV regularization to encode the spatial-spectral continuity of the cloud/shadow component. An efficient threshold method is proposed to detect the cloud/shadow from the sparse component, and the detected cloud/shadow is utilized to improve the reconstruction accuracy of TVLRSD.
- (3) We testify our method with different sensors, different numbers of spectral bands and temporal acquisitions, and different spatial

resolution of multitemporal satellite images. The whole cloud/shadow detection and removal results demonstrate the effectiveness and practicability of our approach.

The remainder of this paper is organized as follows. Section 2 describes the cloud/shadow degradation and reconstruction framework. The proposed method for multitemporal remote sensing image cloud/shadow detection and removal are discussed in Section 3. Section 4 presents the experimental results and discussions to demonstrate the effectiveness of the proposed method. Finally, we conclude this paper in Section 5.

2. Problem formulation

The observed cloud/shadow-contaminated multitemporal image datasets can be regarded as a fourth-order tensor $\mathcal{Y} \in R^{m \times n \times b \times t}$, where m and n are the spatial sizes of the image, b is the number of spectral bands in each time node, and t is the number of time nodes. To discuss the low-rank sparsity decomposition framework better, we reshape the fourth-order tensor as matrix $Y \in R^{mn \times bt}$, where each column represents an image and each row denotes a pixel along all bands and time nodes.

In our work, we assume that the cloud/shadow-contaminated multitemporal images Y consist of two components: clean image (surface-reflected) and cloud/shadow. The cloud/shadow effect is approximated as additive noise. Thus, we can formulate the cloud/shadow degradation model as

$$Y = X + S + N, \quad (1)$$

where $X \in R^{mn \times bt}$, $S \in R^{mn \times bt}$ and $N \in R^{mn \times bt}$ are the clean image component, cloud/shadow component, and residual, respectively.

We formulate the degradation process due to the thick and thin clouds and cloud shadows as the additive model. In this degradation modeling, we can simultaneously consider various underlying characteristics. It is common to approximate that thin cloud (or haze) is an additive component to the radiance signal at the sensor from the physical perspective (Richter and Schläpfer, 2005). To provide a visual example, Fig. 1(a) and (b) show two cloud/shadow-contaminated images that are collected by two adjacent time nodes. Fig. 1(d) displays the distribution pixel values of the cloud-free area in Fig. 1(a) and the thin cloud (or haze) area in Fig. 1(b) (blue box in Fig. 1(a) and (b)). As shown in Fig. 1(d), we can observe that the pixel values of the thin cloud area are not constant, and the profile shape is similar to that of the cloud-free version with large values. Thus the thin cloud areas contain useful surface-reflected information with the additive component of thin clouds. Meanwhile, the surface-reflected component in the

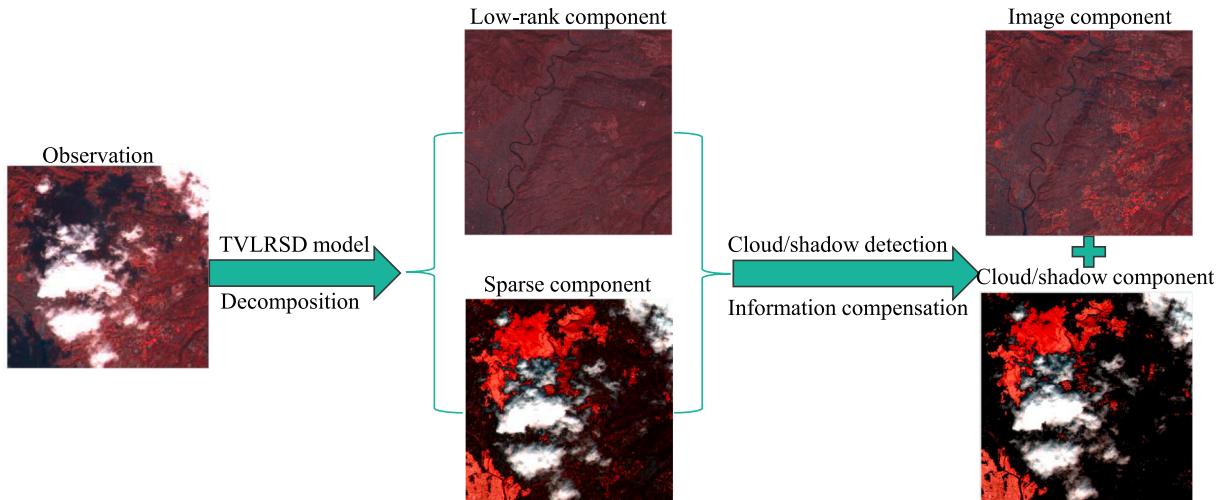


Fig. 2. The framework of the proposed method for blind cloud/shadow removal.

shadow areas is smaller than that of the shadow-free version due to the lack of a direct solar beam. We approximate the decrease as an additive component with being negative. As shown in Fig. 1(b), the shadow areas contain the image component owing to the diffuse solar flux. We use such information as a hint to image restoration. The surface-reflected information is not visible at the locations covered by thick clouds. We regard the difference between the thick-cloud-contaminated image and the cloud-free version as the sparse component in our model. This modeling is similar to sparse modeling based inpainting.

The first goal of our work is to simultaneously estimate the image component X and the cloud/shadow component S from the cloud/shadow-contaminated Y . However, this is a strong ill-posed inverse problem for simultaneously solving X and S from Y . To address this ill-posed inverse problem, we employ the regularization method which is an efficient strategy to solve it. Thus, the separation model can be generally formulated as

$$\min_{X,S,N} \lambda_1 J_1(X) + \lambda_2 J_2(S) + J_3(N), \quad s.t. \quad Y = X + S + N, \quad (2)$$

where λ_1 and λ_2 are two positive regularization parameters that are employed to balance the three regularization terms. $J_1(X)$, $J_2(S)$ and $J_3(N)$ are three regularization terms, which depict the underlying prior knowledge of the image component, cloud/shadow component, and residual, respectively. N represents the residual, which is assumed to obey the Gaussian distribution and specified as $\frac{1}{2} \|N\|_F^2$, where $\|\cdot\|_F$ is the Frobenius norm. Then, we focus on exploring the prior knowledge of the image and cloud/shadow components and design these priors with the approximated regularization terms.

3. Proposed methodology

By considering the cloud/shadow areas as the missing information, completion-based methods can be utilized to efficiently reconstruct the image component. Due to the reason that multitemporal images have strong correlation in spectral-temporal dimensions, the low-rank matrix/tensor completion can be employed to reconstruct the information in cloud/shadow areas. The low-rank completion methods by solving the following model

$$\min_X \|X\|_* \quad s.t. \quad \mathcal{P}_\Omega(X) = Y, \quad (3)$$

where $\mathcal{P}_\Omega(\cdot)$ is the projection operator, and Ω is the cloud/shadow-free areas. Model (3) only considers the low-rank property of the clean image and location information of the cloud/shadow. It fails to take full advantage of the information hidden in the thin clouds and shadow areas, and it needs to know the location in advance. To improve the practicability, we compensate the deficiency of the completion-based models in the low-rank sparsity decomposition framework, which can simultaneously consider the characteristics of the image and cloud/shadow components.

3.1. Preliminary regularization model of cloud/shadow removal

3.1.1. Low-rank sparse decomposition

(a) Low-rank prior of the image component X : As we know, multispectral images have a strong correlation in spectral dimension (Chen et al., 2018). In multitemporal images, the imaging target is the same scene. The only difference is the wavelength of light varies in each band, and the imaging time nodes are different. Thus, multitemporal images have a strong correlation not only in the spectral dimension but also in the temporal dimension. To illustrate the low-rank property of multitemporal images, we plot the distribution of the singular values of the image using a multitemporal Sentinel-2 dataset (see simulated dataset 1). Fig. 3(a) shows the obvious decaying trends of the curve, which indicate the low-rank property of the image. To explore this high spectral-temporal correlation of clean multitemporal images, we use low-rank prior to describe the reshaped matrix X . Thus the

regularization term $J_1(X)$ can be formulated as $J_1(X) = \|X\|_*$, where $\|X\|_* = \sum_i \sigma_i(X)$, and $\sigma_i(X)$ is the i -th singular value of X .

It is noteworthy that the cloud/shadow component has minimal correlation in the temporal dimension because the density and location of the cloud/shadow component are generally different in the temporal dimension. We can expect that the low-rank constraint of X would facilitate the separation of the image and the cloud/shadow components. Moreover, the low-rank prior is confirmed in Wang et al. (2016) for efficiently removing the cloud.

(b) Sparse prior of the cloud/shadow component S : When the proportions of cloud/shadow are relatively low in the entire multitemporal images, the sparse prior can be naturally introduced to describe the cloud/shadow component. To implement the sparse prior, ℓ_0 -norm, which is used to count the number of nonzero elements, is the best choice. However, due to the nonconvexity and NP-hard problem of ℓ_0 -norm, we use ℓ_1 -norm to replace the ℓ_0 -norm as a convex surrogate. Thus, the regularization term for the cloud/shadow component S can be designed as: $J_2(S) = \|S\|_1$, where $\|S\|_1 = \sum_{i=1}^m \sum_{j=1}^{b_i} |S_{i,j}|$.

When each temporal image is contaminated by heavy cloud/shadow, it seems that performing the sparse regularization is not proper. However, other regularization terms help to distinguish the cloud/shadow component, and we can adjust the parameter of the sparsity term to tackle heavy clouds/shadows. Moreover, the cloud/shadow component is intrinsically sparser than the image component in most cases. If the entire image is covered by considerably heavy clouds/shadows, accurately reconstructing the image component is remarkably difficult or even impossible.

In summary, we can summarize the discussion of the prior and regularization and obtain the low-rank sparse decomposition model as follows:

$$\min_{X,S,N} \frac{1}{2} \|N\|_F^2 + \lambda_1 \|X\|_* + \lambda_2 \|S\|_1, \quad s.t. \quad Y = X + S + N. \quad (4)$$

3.1.2. Spatial-spectral TV regularized low-rank sparsity decomposition model

In the above, we explore that the cloud/shadow component has sparse prior. However, the sparse prior is limited to explore the characteristics when the cloud/shadow component is comparative heavy. We need to explore another priors and design corresponding regularization term for the cloud/shadow component. As we known, the cloud component possesses a piece-wise smooth structure in the spatial domain (see Fig. 3(b)). Moreover, the cloud component is continuously presented in the spectral dimension since the locations are the same in the image. This phenomenon is similar to the foreground subtraction problem, in which the cloud component is regarded as the foreground in the image. The difference is that the cloud component is static, whereas the foreground subtraction also presents moving stage. In Cao et al. (2016), the authors explored the continuity characteristic of foreground component. The cloud component also has the continuity characteristic in the spectral dimension. It is noteworthy that the cloud shadow also possesses this characteristic. To preserve the spatial piece-wise smoothness and spectral continuity, we use a spatial-spectral TV regularization (Chen et al., 2017; He et al., 2016; He et al., 2018) to model the cloud/shadow component in each time node. By adding the spatial-spectral TV regularization to cloud/shadow component, we can formulate $J_2(S) = \lambda_2 \|S\|_1 + \lambda_3 \sum_{i=1}^t \|DS_i\|_1$, where S_i is the cloud/shadow component in the i -th time node, $\|DS_i\|_1 = |D_x S_i| + |D_y S_i| + |D_z S_i|$. D_x , D_y , and D_z represent the differential operators of spatial horizontal, spatial vertical, and spectral dimensions, respectively.

After fully analyzing the priors and regularization terms of X and S , we can obtain the low-rank sparsity decomposition with the spatial-spectral TV regularization model. By combining the regularization terms of the image low-rank prior and cloud/shadow prior into the regularization model (2), our final TVLRSR model can be formulated as

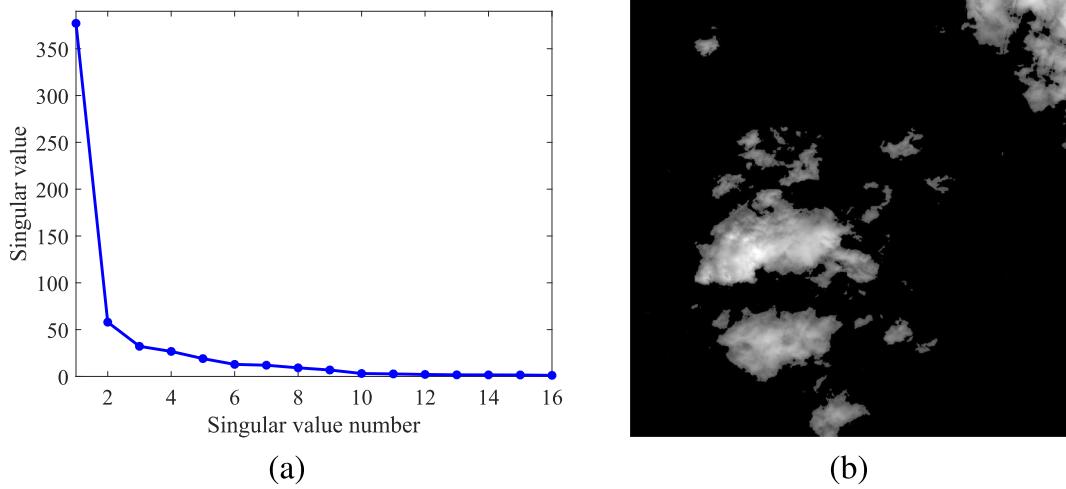


Fig. 3. (a) The distribution of the singular values of the Sentinel-2 dataset, (b) real cloud component.

$$\begin{aligned} \min_{X,S,N} & \frac{1}{2} \|N\|_F^2 + \lambda_1 \|X\|_* + \lambda_2 \|S\|_1 + \lambda_3 \sum_{i=1}^t \|DS_i\|_1, \\ \text{s. t. } & Y = X + S + N, \end{aligned} \quad (4)$$

where λ_1 , λ_2 , and λ_3 are three positive regularization parameters.

It is worth noting that the proposed model can fully capture the spectral-temporal correlation of the image component, and the smoothness and continuity of the cloud/shadow component. Specifically, this model does not require the location information of the cloud/shadow. By optimizing the image component, cloud/shadow component and residual components alternatively, we can separate them progressively to remove the cloud/shadow.

3.2. Image and cloud/shadow components separation using ADMM algorithm

In model (4), we cannot directly separate the image component X and the cloud/shadow component S from Y since there are three unknown variables. To effectively separate these two components, we employ the ADMM framework (Boyd et al., 2011) to optimize the proposed TVLRSD model. The ADMM algorithm has also been widely used for image processing tasks, such as image destriping (Chen et al., 2017) and image deraining (Jiang et al., 2019). It can transform the complex multi-variable optimization problem into several easily solved subproblems. In the ADMM framework, we alternately optimize one variable while fixing the other variables in an iterative manner.

The image and cloud/shadow components separation framework of ADMM iterations is shown in Fig. 4. In the framework, the degraded image is decomposed into low-rank image component and the sparse cloud/shadow component. By using the ADMM algorithm, we fix the cloud/shadow component and employ the low-rank prior (image subproblem in Fig. 4) to extract the image component from the observation. When the image component is separated, the sparse and spatial-spectral TV regularized priors are employed (cloud/shadow subproblem in Fig. 4) to extract the cloud/shadow component from the difference between the observation and the image component. By solving the subproblem via ADMM iterations, the image component and the cloud/shadow component can be estimated step-by-step from the observation, so as to achieve the detection and removal of cloud/shadow simultaneously. The details of the optimized implementation are presented in the supplementary material.

3.3. Cloud/shadow detection and information compensation

Although model TVLRSD can efficiently separate the image

component and the cloud/shadow component, it is inevitable that the cloud/shadow-free information will be degraded by subtracting a nonzero false cloud/shadow component from the surface-reflected component in cloud/shadow-free areas. To improve the results of the cloud/shadow removal, we propose a simple cloud/shadow detection method and compensate the information of cloud/shadow-free areas as much as possible.

In the first step, we have extracted the cloud/shadow component (i.e., the sparse component S by ADMM algorithm). In the sparse component S , the pixel values of cloud locations are positive and larger than other cloud/shadow-free locations because of the high reflective nature of clouds. The pixel values of cloud shadow tend to be negative due to the decrease of global flux on the ground. Suppose the estimated sparse component $S_i \in R^{m \times n \times p}$ from a certain time node, where m and n are the spatial sizes of the image, and p is the number of spectral bands in this time node, it can be separated into the addition of a positive component and a negative component. The threshold procedures will be executed to segment shadow and cloud locations, respectively. If the positive component is larger than a threshold value, then we regard this pixel as a cloud element. On the contrary, it is a shadow pixel when the negative component is less than a threshold value. The cloud/shadow detection step is summarized in detail in the supplementary material, and a more detailed example will be presented in the experimental part (see Fig. 14). After detecting the cloud/shadow locations in each time node, the information of cloud/shadow-free locations is compensated to the final result by replacing the low-rank image component X with the original image Y in the cloud/shadow-free areas. We summarize the two steps of blind cloud/shadow removal and information compensation in Algorithm 1. The framework of the proposed method is presented in Fig. 2.

4. Experiments and results

In this section, to demonstrate the performance of our proposed method, we tested it in both simulated and real experiments over different multitemporal remote sensing images. We compared the results with popular completion-based methods: the matrix completion (MC) (Candès and Tao, 2010), the tensor completion method (HaLRTC) (Liu et al., 2013), the temporally contiguous robust matrix completion method (ALM-IPG) (Wang et al., 2016), and the adaptive weighted tensor completion method (AWTC) (Ng et al., 2017). We reshaped the multitemporal dataset $\mathcal{Y} \in R^{m \times n \times b \times t}$ into $\mathcal{Y} \in R^{mn \times bt}$ to perform the MC and ALM-IPG methods. The four-order multitemporal dataset $\mathcal{Y} \in R^{m \times n \times b \times t}$ was reshaped to three-order $\hat{\mathcal{Y}} \in R^{m \times n \times bt}$ dataset to perform the tensor-based completion methods HaLRTC and AWTC. For the

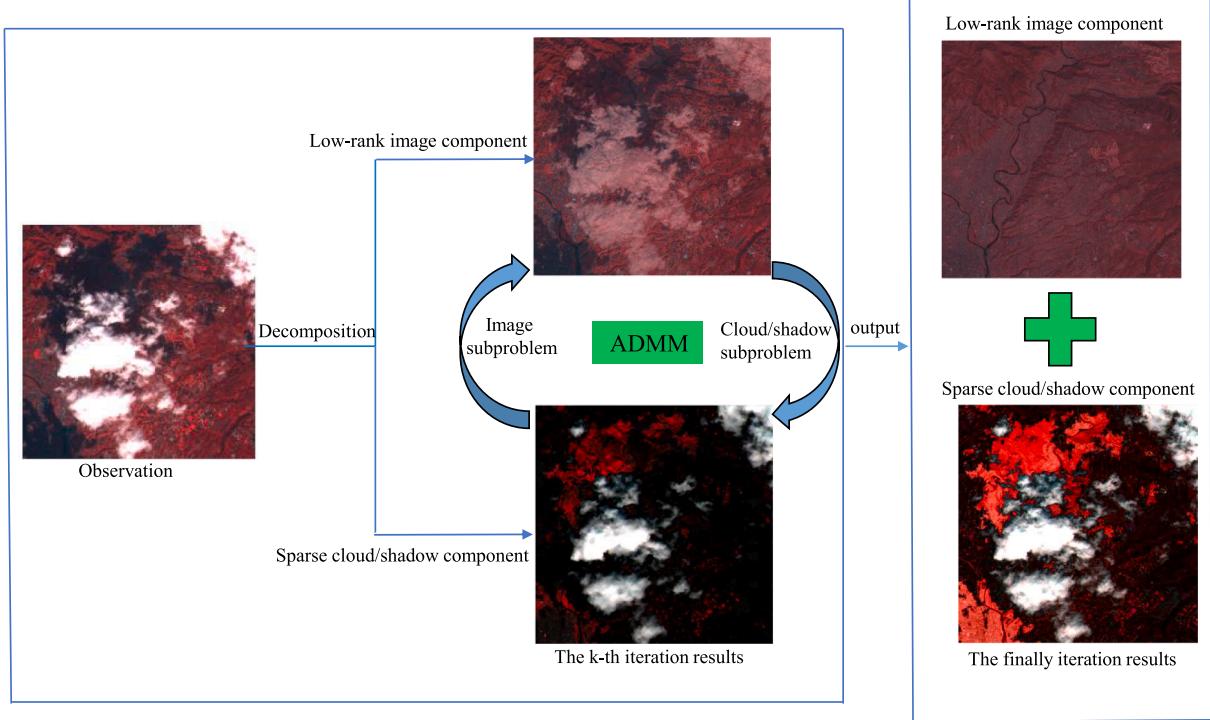


Fig. 4. The framework of ADMM iteration for separating the image component and cloud/shadow component.

weights of three different unfoldings on HaLRTC method, we set them to be 1/3 according to the authors' rules in their paper.

Algorithm 1. TVLRSD with information compensation

```

1: Step 1: Perform ADMM algorithm of TVLRSD model to separate the three components.
2: Output: Image component X, cloud/shadow component S, and residual N
3: Step 2: Perform cloud/shadow detection for each time node  $S_t$ .
4: for each  $k = 1: t$  do
5:   Perform cloud/shadow detection.
6: endfor
7: Output:  $\Omega$ : cloud/shadow locations
Output: Recovered  $X(\hat{\Omega}) = Y(\hat{\Omega})$ , where  $\hat{\Omega}$  is a complement to  $\Omega$ 
```

4.1. Simulated experiments

In the real experiments, it is hard to use quantitative analysis to evaluate the performance of cloud/shadow removal without ground-truth data and reference of cloud/shadow masks. We perform simulated experiments to quantitatively evaluate the effectiveness of our method with matrix/tensor completion methods for cloud/shadow removal and information reconstruction.

4.1.1. Simulated datasets

In the simulated experiments, we have ground-truth clean remote sensing images, and add different kinds of cloud/shadow to simulate the degraded multitemporal images. The advantage of simulated experiments is that we have ground-truth to evaluate the reconstruction results of different methods. To test different types of optical satellite sensors with different spatial resolutions, two multitemporal remote sensing images were selected as the ground-truth datasets, i.e., Sentinel-2¹ with 4 bands (NIR, red, green, and blue) at a ground sampling

distance (GSD) of 10 m, and Landsat-8² with 3 bands (red, green, and blue) at a GSD of 30 m. The details of the two datasets are given in Table 1.

4.1.2. Experimental setting

In this part, we mainly show the superiority of our method than existed completion-based methods for reconstructing the information in the cloud/shadow locations. It is worth noting that precise cloud/shadow masks are given for completion-based methods. We simulated different kinds of cloud and cloud shadow, and added them to the clean multitemporal remote sensing images. Two factors, *i.e.*, the feature value and distribution of the cloud/shadow are taken into consideration in the simulation. For the first Sentinel-2 dataset, we generated the distributions of the cloud and cloud/shadow randomly for each temporal image. We adopted different methods to generate the values of the cloud and cloud/shadow. In case 1, the cloud component was simulated by adding a constant value to the cloud areas as a toy example. In case 2, we utilized a mask to extract the cloud component from the real cloud data and then added the extracted cloud component to the image. To demonstrate that our method can also efficiently remove cloud shadow, in case 3, we also simulated the cloud shadow on the basis of case 1. Meanwhile, in case 4, the cloud shadow was also added on the basis of case 2.

To analysis the influence of the cloud/shadow distribution for the final results, we added thin clouds in which the locations and shapes are the same in all temporal images to the Landsat-8 dataset. The cloud/shadow removal problem becomes challenging, since no complementary (or cloud/shadow-free) information is available from other temporal images when we only utilized the prior for the image component. We investigate two cases (*i.e.*, cases 5 and 6) by adding a constant value and realistic clouds. Table 2 summarizes different settings. To evaluate the performance of cloud/shadow removal results, we used two quantitative indices, namely peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) (Wang et al., 2004), in the

¹ Sentinel-2 imagery is available at scihub.copernicus.eu/dhus/.

² Landsat-8 imagery is available at <https://espa.cr.usgs.gov/>.

Table 1
Multitemporal image for simulated experiments.

| Dataset | Source | Location | GSD | Image size | Spectral | Time nodes |
|-----------|------------|----------|------|------------|----------|------------|
| Dataset 1 | Sentinel-2 | Tokyo | 10 m | 400 × 400 | 4 | 4 |
| Dataset 2 | Landsat-8 | Munich | 30 m | 512 × 512 | 3 | 4 |

simulated experiments. The larger PSNR and SSIM values mean better cloud/shadow removal results. TVLRSR represents that we only employ the ADMM algorithm to obtain the result. TVLRSDC represents the proposed TVLRSR with information compensation guided by cloud/shadow detection results in [Algorithm 1](#).

4.1.3. Results on Sentinel-2 Dataset

[Figs. 5 and 6](#) show the cloud removal results of time node 1 and time node 2 under different comparison methods in cases 1 and 2, respectively. From the results, HaLRTC failed to reconstruct the cloud regions in these two cases as shown in [Figs. 5\(d\) and 6\(d\)](#). The reason is that the correlation of the dataset is not effective in both two spatial dimensions, thus the low-rank constraints of the two spatial dimensions degrade the final results. MC, ALM-IPG, and AWTC can reconstruct the most of the missing information covered by clouds, but the image details were smoothed and distorted, which can be observed in the enlarge box of [Figs. 5 and 6](#). Moreover, the results of MC and ALM-IPG presented a black block shown in [Figs. 5\(c\) and \(e\)](#). The main reason is that the

cloud is overlapped in all temporal images in this area. According to [Fig. 5\(g\)](#), the proposed TVLRSDC method completely reconstructed the cloud areas and preserved the original details in the image since the proposed model can efficiently capture the prior information for both the image and cloud components. Although the proposed method failed to completely recover the information in [Fig. 6\(g\)](#), our result was also superior to the completion-based methods because the comparison methods obtained the blurring artifacts.

[Figs. 7 and 8](#) show cloud/shadow removal results of case 3 and case 4, respectively. The results indicated that the completion-based methods obtained the same observation results compared with cloud removal cases. Moreover, these methods also cannot completely preserve the original information in cloud/shadow areas, but they can recover the texture information in these areas. The cloud shadow existed in the image only influenced the percentage rate of missing information, thus the results were similar to the cloud removal. Based on the results shown in [Figs. 7\(g\) and 8\(g\)](#), the proposed TVLRSDC method can efficiently reconstruct the information in cloud areas. Furthermore, the proposed method eliminated the cloud shadow and preserved the most of information even the locations of cloud shadow areas were not provided.

The above visual comparisons have demonstrated the superiority of the proposed TVLRSDC method for cloud/shadow removal. In the next, we will show the quantitative evaluation of PSNR and SSIM values to again demonstrate the efficiency of the proposed method. [Table 3](#) presents the results of the PSNR and SSIM values in simulated

Table 2
The characteristic of the simulated cloud/shadow component in different settings.

| Dataset | Case | Type | Distribution | Cloud | Cloud shadow |
|------------|--------|-------------------|-----------------------------------|-------|--------------|
| Sentinel-2 | Case 1 | Constant value | Different in different time nodes | ✓ | ✗ |
| | Case 2 | Real cloud | Different in different time nodes | ✓ | ✗ |
| | Case 3 | Constant value | Different in different time nodes | ✓ | ✓ |
| | Case 4 | Real cloud/shadow | Different in different time nodes | ✓ | ✓ |
| Landsat-8 | Case 5 | Constant value | The same in different time nodes | ✓ | ✗ |
| | Case 6 | Real cloud | The same in different time nodes | ✓ | ✗ |

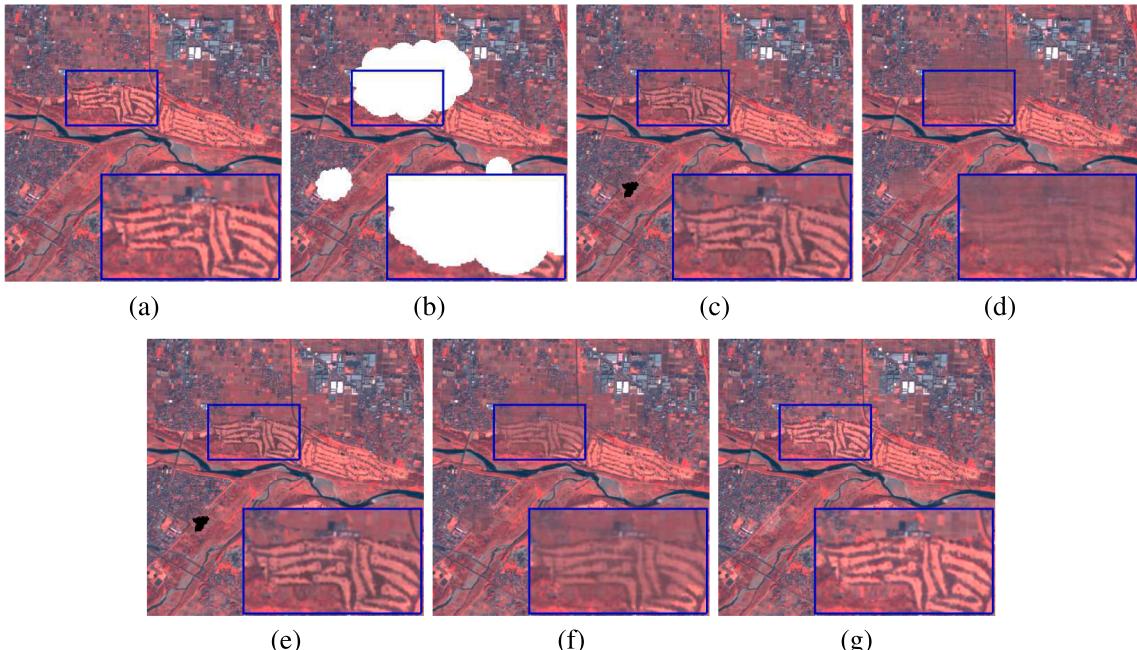


Fig. 5. The time node 1 cloud removal results of Sentinel-2 data under case 1. (a) Clean, (b) simulated cloud image, (c) MC, (d) HaLRTC, (e) ALM-IPG, (f) AWTC, and (g) TVLRSDC.

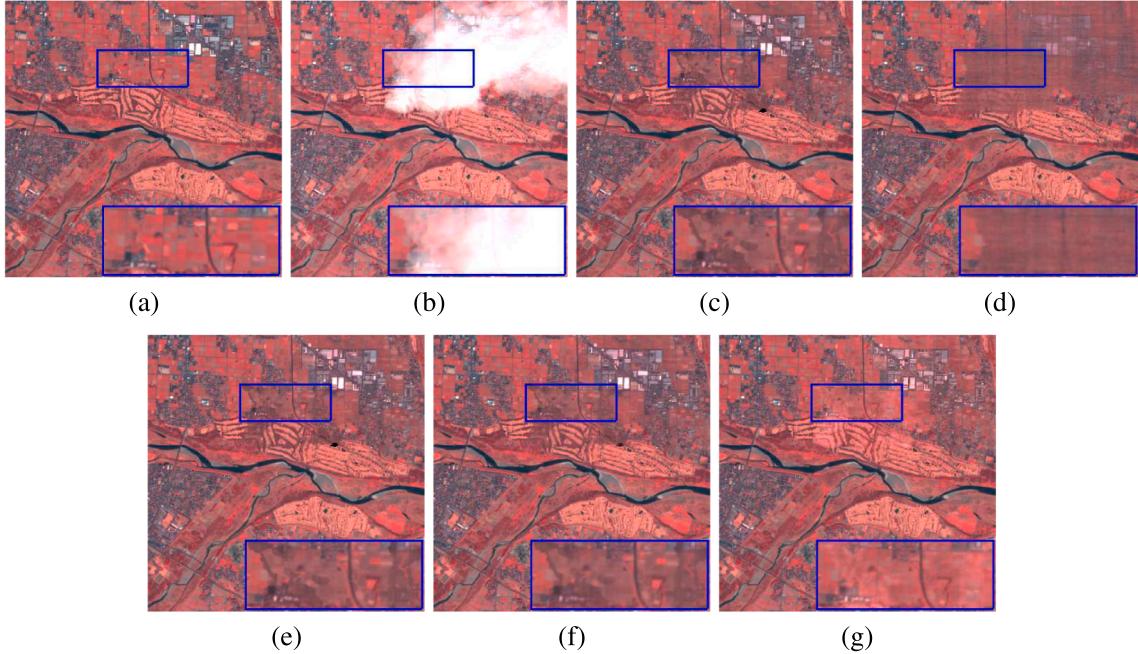


Fig. 6. The time node 2 cloud removal results of Sentinel-2 data under case 2. (a) Clean, (b) simulated cloud image, (c) MC, (d) HaLRTC, (e) ALM-IPG, (f) AWTC, and (g) TVLRSDC.

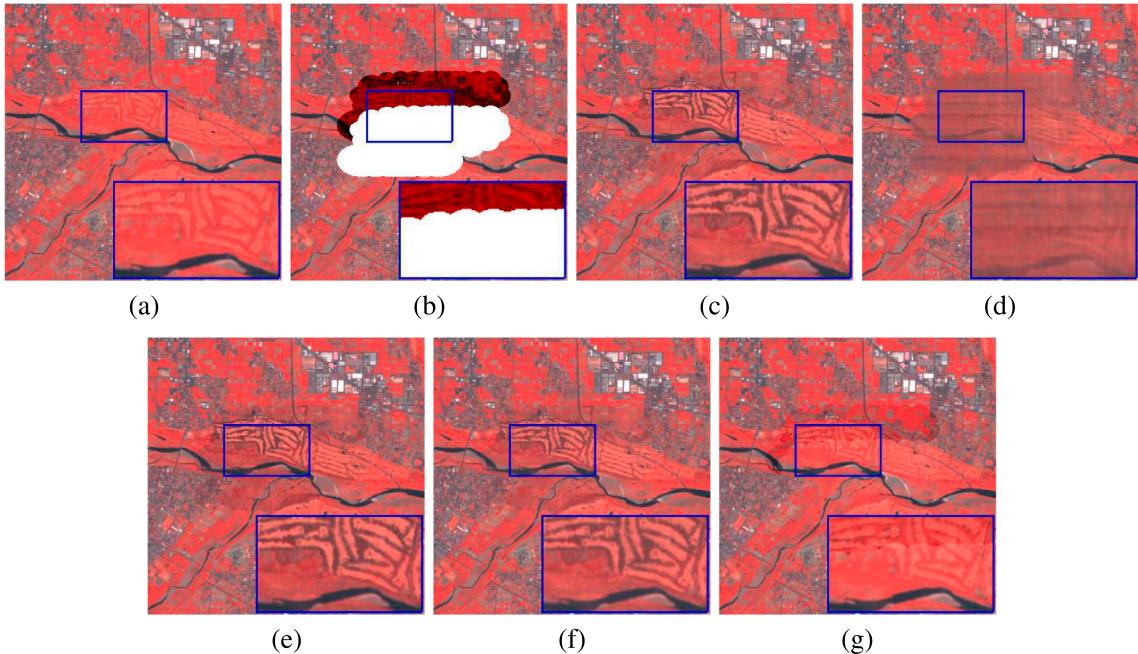


Fig. 7. The time node 3 cloud/shadow removal results of Sentinel-2 data under case 3. (a) Clean, (b) simulated cloud/shadow image, (c) MC, (d) HaLRTC, (e) ALM-IPG, (f) AWTC, and (g) TVLRSDC.

experiments of Sentinel-2 data under four different cases, where the bold values indicate the best results. Since there was a reference mask in simulated experiments, we used the results of the combination of TVLRSR with reference mask (TVLRSRDR) as the upper bound of our compensation method. In the table, we can observe that the evaluation indices achieved by the proposed TVLRSRDR were much higher than the other completion-based methods in case 1 and case 3. In case 2, the PSNR of our result was lower than those of completion-based methods since the information compensation was not completely precise. However, the proposed method obtained better visual results as shown in Fig. 6(g). It's worth noting that the evaluation results of TVLRSRDR were

much higher than that of TVLRSR, thereby demonstrating the effectiveness of information compensation. Although TVLRSRDR obtained lower PSNR value than AWTC in case 4, the SSIM value of our method was the highest among all the comparison methods. These results illustrated that the proposed TVLRSRDR with information compensation can preserve most of the details in the cloud/shadow areas. The TVLRSRDR results also illustrated that the proposed TVLRSRDR can reconstruct better information in the cloud/shadow areas.

4.1.4. Results on Landsat-8 dataset

Landsat-8 dataset mainly contains mountains, hills, and rivers.

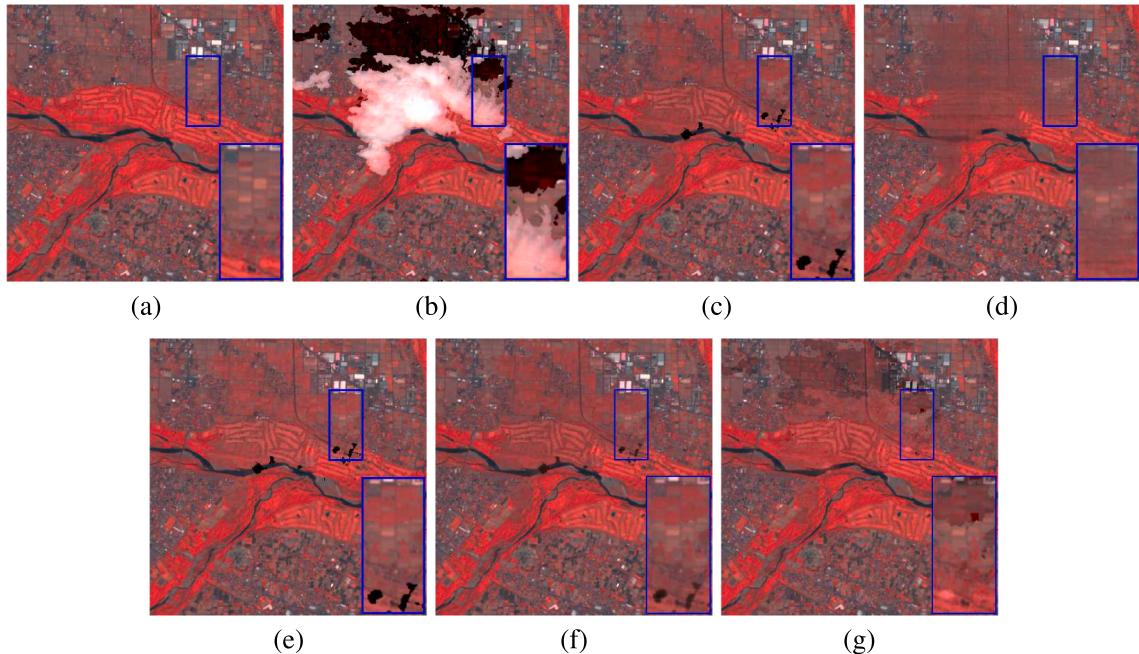


Fig. 8. The time node 4 cloud/shadow removal results of Sentinel-2 under case 4. (a) Clean, (b) simulated cloud/shadow image, (c) MC, (d) HaLRTC, (e) ALM-IPG, (f) AWTC, and (g) TVLRSDC.

Figs. 9(a) and 10(a) show the simulated cloud images, where we can observe that the image was seriously degraded by the simulated cloud component. Figs. 9,10(b) and (d) show the results of the matrix-based completion methods, and we can observe that the cloud regions cannot be reconstructed completely. The results of the tensor-based completion methods are shown in Figs. 9,10(c) and (e). With the help of another two unfoldings, HaLRTC and AWTC reconstructed some artifact information. The proposed method, in contrast, can effectively reconstruct a lot of information in the cloud regions and preserve some information in cloud-free regions as shown in Figs. 9,10(f). From these visual results, we obtained the following observations. First, completion-based methods fail to reconstruct the information in the cloud regions when the cloud lies in the same locations in all temporal images. The main reason is that there are not auxiliary information in the cloud locations, and completion-based methods only consider the image low-rank prior in three dimensions. Second, our method can efficiently tackle the above cloud problems since we simultaneously consider the image component prior and cloud component prior. Thus the iterative estimation manner of image and cloud components would benefit from each other.

Table 4 presents the quantitative evaluation results for simulated Landsat-8. In case 5, the proposed TVLRSDC method obtained the highest values among the comparison methods. TVLRSDC achieved lower PSNR values than the tensor completion based methods in case 6,

but we achieved the best SSIM value than the comparison methods. These results indicated that the tensor completion methods cannot preserve the structure in the cloud areas in these cases. The proposed method better reconstructed the information in cloud regions, which can be demonstrated by the upper limit result of our method TVLRSDR. In case 5, the information compensation step greatly improved the PSNR value. ALM-IPG obtains similar results as MC. The reason is that ALM-IPG reduces to MC when the temporal continuity is not evident.

From the simulated experiments, we can draw the following conclusions. First, our proposed method demonstrated the best cloud/shadow removal performance in most of the cases with different settings. Second, the proposed TVLRSDC performed much better than the other comparison methods particularly when the locations of the cloud and cloud shadow are overlapped. This advantage is achieved by utilizing the surface information underlying in the thin cloud and cloud shadow of multitemporal images for the information reconstruction.

4.2. Real experiments

To demonstrate the performance and show the superiority that our method needn't cloud-shadow detection as preprocessing, we used three multitemporal remote sensing datasets, which are severely contaminated by real cloud/shadow. Moreover, two additional real experiments were reported in the supplementary material. These datasets

Table 3
Quantitative evaluation PSNR (dB) and SSIM results of Sentinel-2.

| Case | Index | MC | HaLRTC | ALM-IPG | AWTC | TVLRSD | TVLRSDC | TVLRSDR |
|--------|-------|--------|--------|---------|---------------|--------|---------------|---------|
| Case 1 | PSNR | 36.405 | 38.739 | 36.488 | 39.890 | 43.821 | 46.197 | 46.197 |
| | SSIM | 0.9789 | 0.9688 | 0.9789 | 0.9777 | 0.9956 | 0.9980 | 0.9980 |
| Case 2 | PSNR | 34.960 | 33.036 | 34.932 | 35.224 | 32.409 | 34.803 | 35.793 |
| | SSIM | 0.9469 | 0.9114 | 0.9466 | 0.9468 | 0.9425 | 0.9659 | 0.9672 |
| Case 3 | PSNR | 38.297 | 35.742 | 38.545 | 38.351 | 40.308 | 42.752 | 42.758 |
| | SSIM | 0.9723 | 0.9456 | 0.9721 | 0.9721 | 0.9934 | 0.9960 | 0.9961 |
| Case 4 | PSNR | 32.588 | 31.332 | 32.731 | 33.538 | 30.096 | 32.282 | 32.814 |
| | SSIM | 0.9359 | 0.8898 | 0.9363 | 0.9380 | 0.9333 | 0.9583 | 0.9604 |

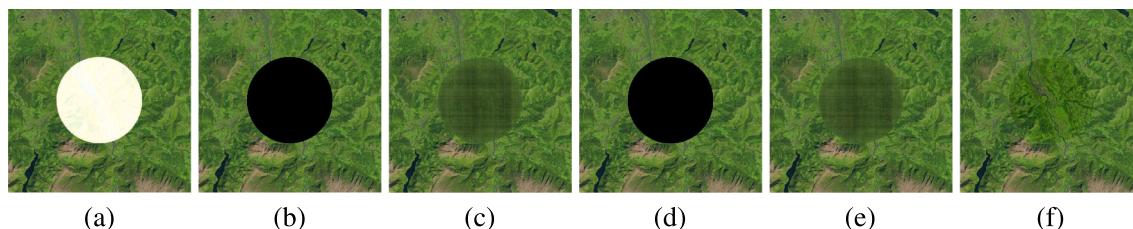


Fig. 9. The time node 4 cloud removal results of Landsat-8 under case 5. (a) Simulated cloud image, (b) MC, (c) HaLRTC, (d) ALM-IPG, (e) AWTC, (f) TVLRSDC.

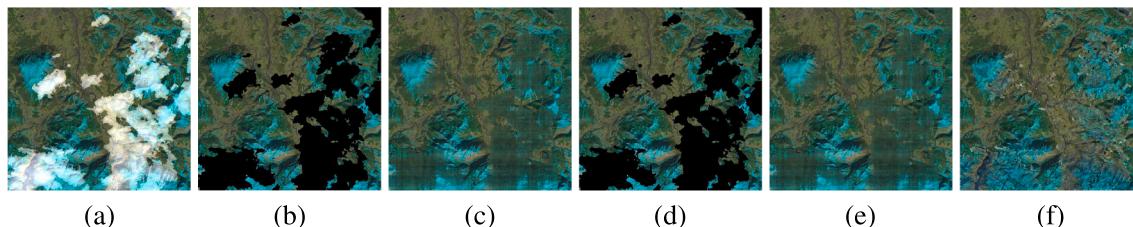


Fig. 10. The time node 2 cloud removal results of Landsat-8 under case 6. (a) Simulated cloud image, (b) MC, (c) HaLRTC, (d) ALM-IPG, (e) AWTC, and (f) TVLRSDC.

Table 4
Quantitative evaluation PSNR (dB) and SSIM results of simulated Landsat-8.

| Case | Index | MC | HaLRTC | ALM-IPG | AWTC | TVLRSD | TVLRSDC | TVLRSDR |
|--------|-------|--------|--------|---------|---------------|--------|---------------|---------|
| Case 5 | PSNR | 18.142 | 26.319 | 18.142 | 26.387 | 24.917 | 34.859 | 34.895 |
| | SSIM | 0.8145 | 0.8767 | 0.8144 | 0.8768 | 0.9452 | 0.9903 | 0.9904 |
| Case 6 | PSNR | 14.501 | 23.021 | 14.501 | 23.037 | 20.953 | 22.440 | 25.158 |
| | SSIM | 0.6049 | 0.7573 | 0.6049 | 0.7573 | 0.8055 | 0.8395 | 0.8915 |

were selected from different sensors, with different numbers of spectral bands and temporal acquisitions, and different spatial resolution of multitemporal satellite images.

4.2.1. Real datasets

Three real datasets were chosen to testify the performance of the proposed method on the multitemporal images with different resolutions and sensors. Two datasets were selected from Sentinel-2 over Tokyo. One of the dataset was acquired in 2017 with a GSD of 10 m ([Fig. 11](#)) for bands 2, 3, 4, and 8, mainly containing the mountains, roads, and village. The other Sentinel-2 dataset was also collected in 2017 with heterogeneous urban areas, and we use bands 5, 6, 7, 8A, 11, and 12 with a GSD of 20 m ([Fig. 12](#)). The third dataset was collected by the SPOT-5 satellite sensor³ over the urban city of Beijing on 2015 with GSD of 10 m ([Fig. 13](#)), in which the band of 20 m is upsampled into the same spatial size of 10 m. The details of the three datasets are listed in [Table 5](#).

4.2.2. Experimental setting

Our blind cloud/shadow removal method can be directly utilized to process the three real datasets. However, the compared completion methods need the locations of the cloud and cloud shadow in advance. Therefore, we adopted the state-of-the-art improved Fmask algorithm ([Frantz et al., 2018](#)) and the Sen2Cor tool ([Louis et al., 2016](#)) to detect the cloud/shadow of the two Sentinel-2 images, respectively. For the SPOT-5 dataset, the mask can be directly downloaded from the data website, and it is detected by MAJA processor.⁴ These cloud/shadow masks generated by different methods provided the prior information for the compared completion methods. We compare these masks with

the ones detected by the proposed TVLRSDC to demonstrate the advantage of the proposed method. The HaLRTC method was not chosen for comparison because the equal weights of three dimensions are not suitable for cloud removal shown in simulated experiments.

4.2.3. Real experimental results

We selected one time node reconstruction result for comparison in each dataset (more results are provided in the [supplementary material](#)). [Figs. 11–13](#) show the cloud/shadow removal results on time node 10, time node 1, and time node 6 for the first Sentinel-2 dataset, the second Sentinel-2 dataset, and the SPOT-5 dataset, respectively. As clearly shown in (a) of each figure, the images were simultaneously degraded by clouds and cloud shadows. The reconstructed results of different methods are presented in [Figs. 11–13\(b\)–\(e\)](#). From the results, the following results can be observed. First, as displayed in [Fig. 11](#), the completion-based methods failed to recover several shadow areas, and our method can simultaneously remove the cloud/shadow and reconstruct the information in the image. The reason is that the improved Fmask fails to detect these shadow areas, and our method is blind and can automatically reconstruct the information by separating the clean image and cloud/shadow components alternatively. Second, in [Fig. 12](#), the completion-based methods can efficiently eliminate the shadow in the image, but some discrete clouds left in the image. In contrast, the proposed TVLRSDC better removes the cloud/shadow and preserves the cloud/shadow-free information. This phenomenon illustrated that Sen2Cor cannot detect the cloud locations completely. Finally, all methods precisely reconstructed the information in cloud/shadow areas in [Fig. 13](#). This is because the mask for the SPOT-5 images is of high quality. From these observations, we can conclude that the results of completion-based methods heavily depend on the accuracy of the cloud/shadow detection results. The proposed method achieved the satisfactory cloud/shadow removal results in a stable manner for different sensors, resolutions, and cloud conditions, due to the fact that the proposed method not only consider the image prior but also the cloud/

³ SPOT-5 imagery is available at <https://take5.theia.cnes.fr/atdistrib/take5/client/#/home>.

⁴ <http://www.cesbio.ups-tlse.fr/multitemp/?p=6203>.

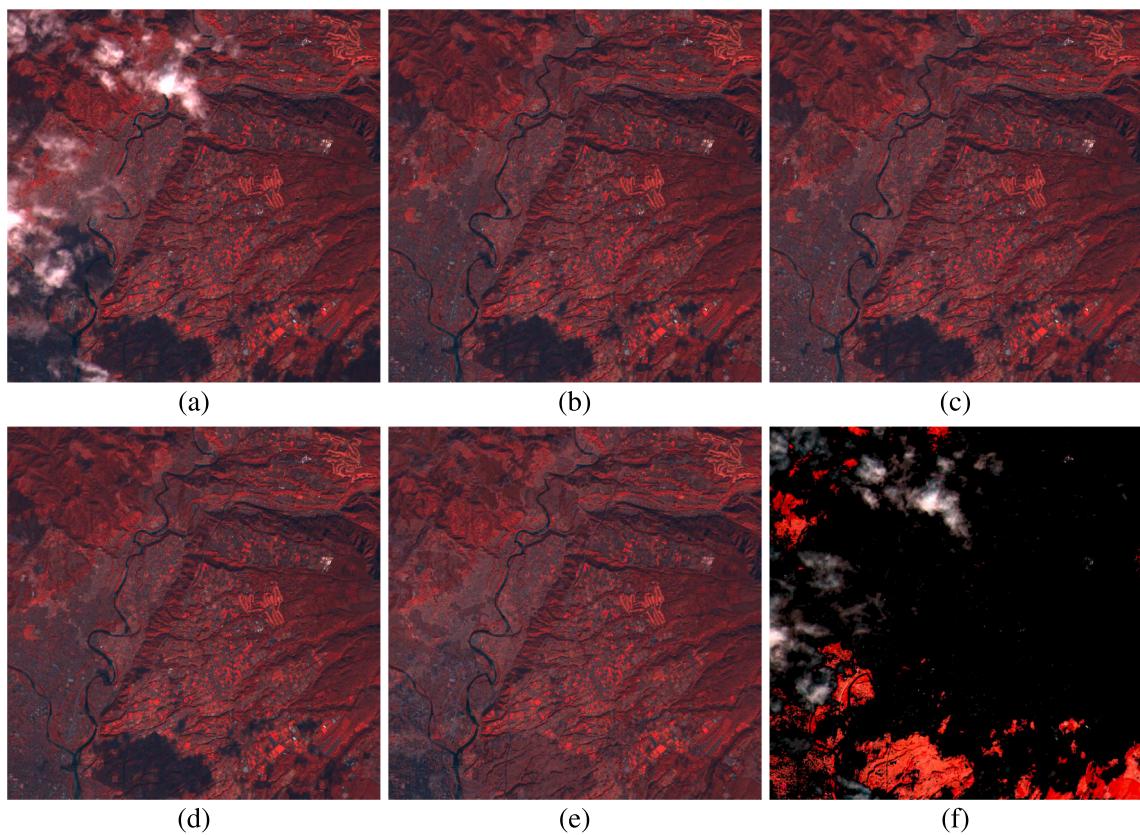


Fig. 11. Cloud/shadow removal results for the first real Sentinel-2 dataset on time node 10. (a) Cloud/shadow image, (b) MC, (c) ALM-IPG, (d) AWTC, (e) TVLRSDC, and (f) estimated cloud/shadow component by TVLRSDC.

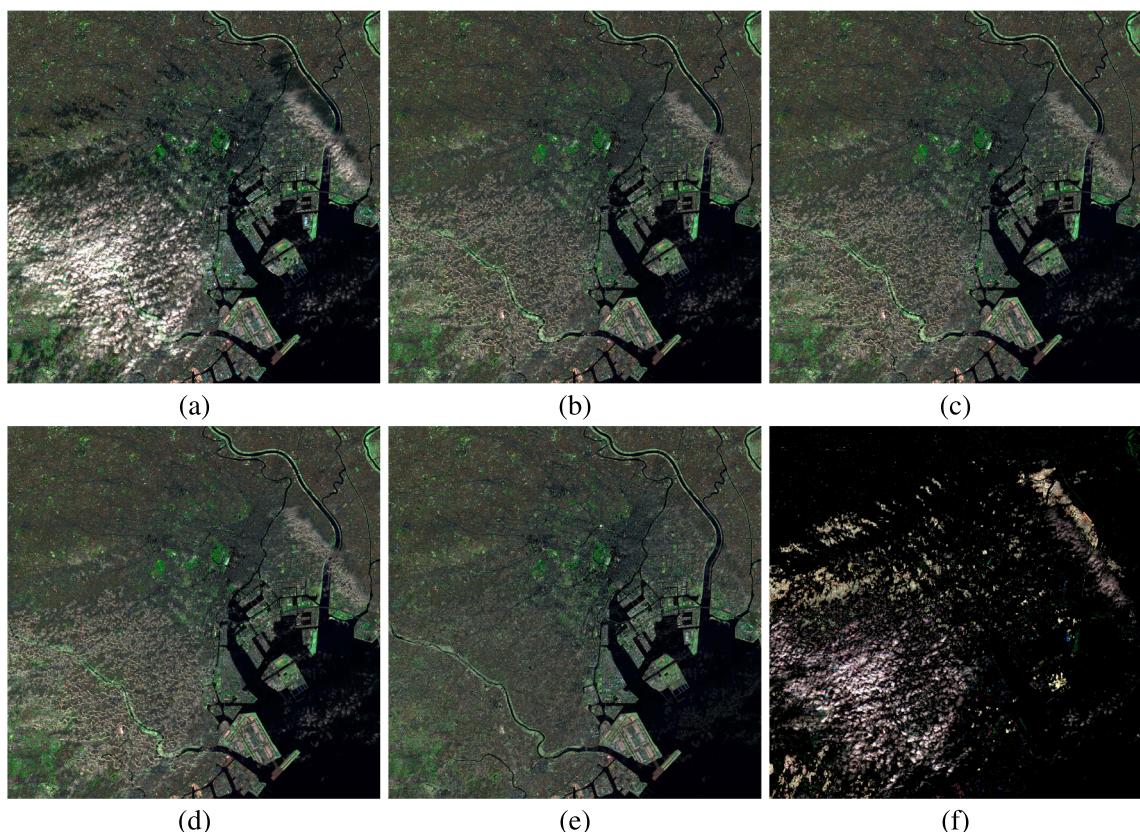


Fig. 12. Cloud/shadow removal results for the second real Sentinel-2 dataset on time node 1. (a) Cloud/shadow image, (b) MC, (c) ALM-IPG, (d) AWTC, (e) TVLRSDC, and (f) estimated cloud/shadow component by TVLRSDC.

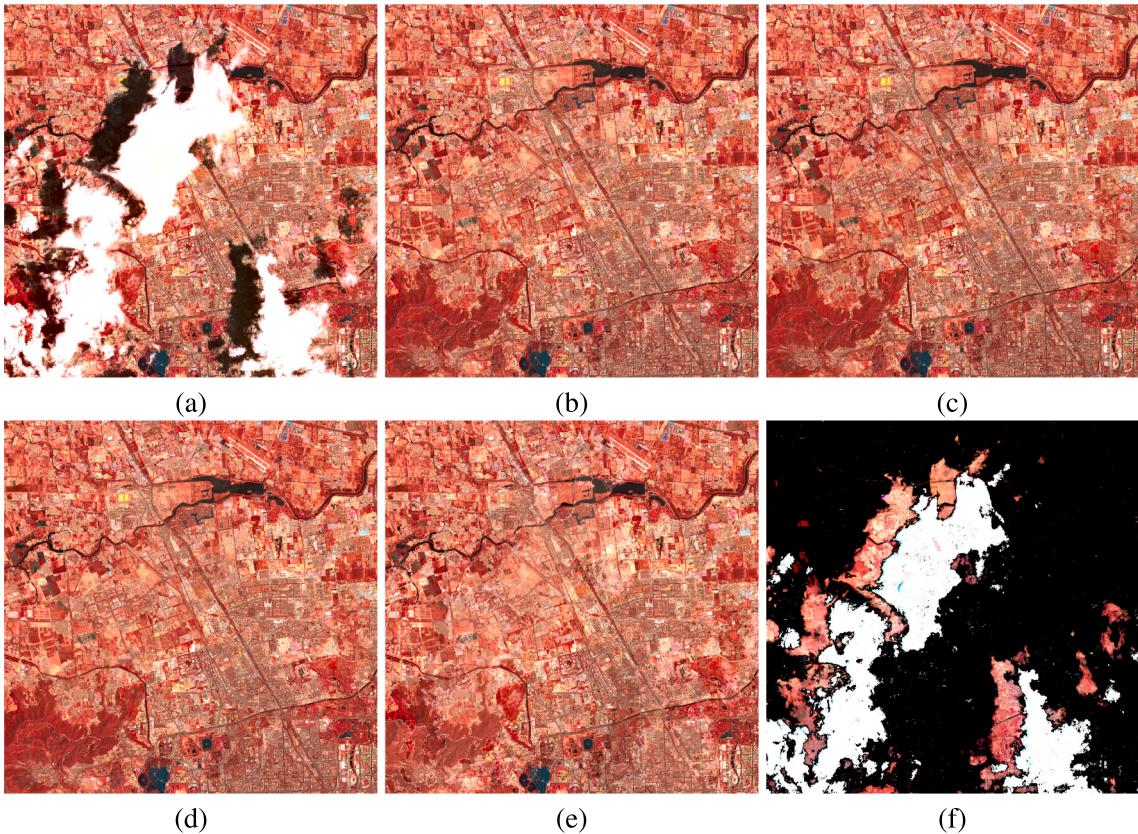


Fig. 13. Cloud/shadow removal results for the real SPOT-5 dataset on time node 6. (a) Cloud/shadow image, (b) MC, (c) ALM-IPG, (d) AWTC, (e) TVLRSDC, and (f) estimated cloud/shadow component by TVLRSDC.

Table 5
Multitemporal image for real experiments.

| Dataset | Source | Location | GSD | Image size | Spectral | Time nodes |
|-----------|------------|----------|------|-------------|----------|------------|
| Dataset 1 | Sentinel-2 | Tokyo | 10 m | 1000 × 1000 | 4 | 13 |
| Dataset 2 | Sentinel-2 | Tokyo | 20 m | 1500 × 1500 | 6 | 10 |
| Dataset 3 | SPOT-5 | Beijing | 10 m | 2000 × 2000 | 4 | 10 |

shadow prior in multitemporal images.

To demonstrate the efficiency of our cloud/shadow extraction, we present the cloud/shadow components estimated by the proposed TVLRSDC method in Figs. 11–13(f). From Figs. 11–13(a), we can observe that the images were degraded by not only cloud but also large areas of cloud shadows. By visually comparing the cloud/shadow component and distributed locations of the cloud/shadow in the observed image, we can observe that the clouds and cloud shadows were successfully extracted. The results showed that the completion-based methods cannot remove the clouds or shadows due to the fact that the detection methods failed to precisely detect these areas. In comparison, our method recovered the surface information in cloud/shadow areas in a blind manner since the proposed TVLRSDC method can estimate the cloud/shadow component accurately. Based on this good results, our method can be widely applied to practical cloud/shadow removal problems.

4.2.4. Cloud/shadow mask detection

To further analysis why our proposed method can achieve better cloud/shadow removal results, we present the cloud/shadow detection results via our method and other famous cloud/shadow detection tools. Recently, many cloud/shadow detection methods have been proposed (Frantz et al., 2018; Zhai et al., 2018; Li et al., 2017). In the real experimental results, we adopted improved Fmask (Frantz et al., 2018) to

detect the cloud/shadow in the first Sentinel-2 dataset, and utilized Sen2Cor for the second Sentinel-2 dataset. Due to the fact that the improved Fmask and Sen2Cor tool failed to detect some locations of the clouds and cloud shadows, the completion-based methods failed in reconstructing the missing information in these areas. However, the high-precision cloud detection results in the SPOT-5 dataset helped these completion-based methods successfully recover the cloud/shadow area. Since our model unified the cloud/shadow detection and removal together, all the cloud/shadow areas were efficiently reconstructed. Fig. 14 presents one illustration of the cloud/shadow location detection from the sparse component S . Firstly, the spare component was divided into two components, i.e., the positive component standing for the clouds and the negative component for the cloud shadows. Secondly, two thresholdings were utilized to segment the final cloud and shadow components, respectively.

Fig. 15 shows the cloud/shadow mask detection results of three different tools and our method for the three real datasets. From the visual observation, the image of first Sentinel-2 dataset was severely degraded by cloud/shadow. It can be observed that some cloud shadow areas were not detected by the improved Fmask algorithm. We outlined some examples in the green box of Fig. 15. In the second row, Sen2Cor successfully detected the cloud shadow, but failed to detect some discrete clouds. For the SPOT-5 dataset, all cloud/shadow locations are covered in the reference mask, however, some cloud/shadow-free areas were regarded as cloud/shadow. The last column of Fig. 15 presents the cloud/shadow mask results of our method. Compared to the improved Fmask method, the proposed method can detect clouds with higher accuracy. For instance, the cloud shadow areas, where the improved Fmask method failed to detect in the green box, were successfully detected by the proposed method. Our method also demonstrated better cloud/shadow detection results compared to Sen2Cor and the reference provided for the SPOT-5 dataset. This advantage is achieved by

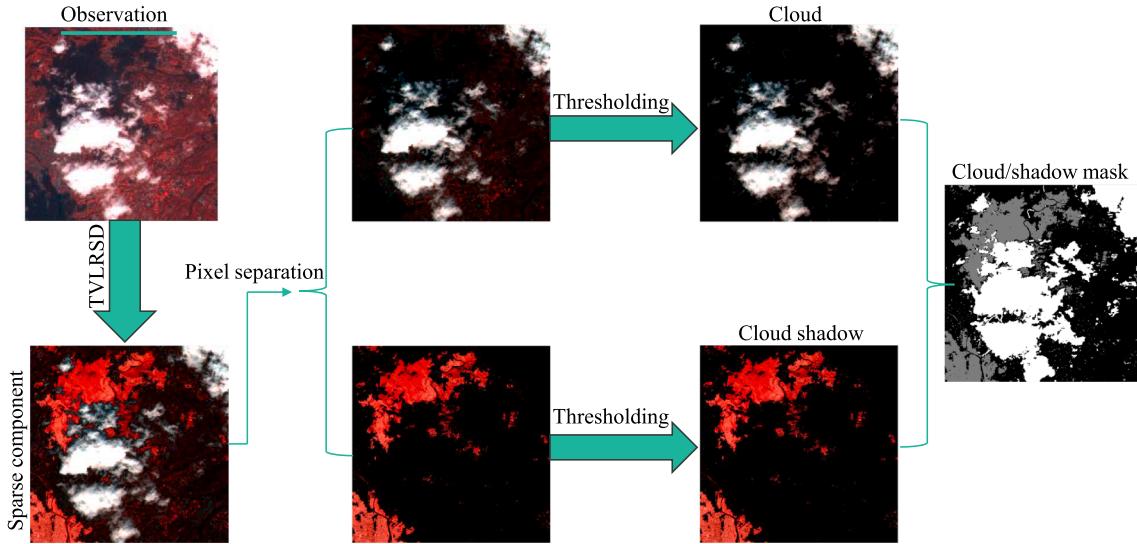


Fig. 14. The framework of the proposed method for cloud/shadow mask detection.

exploiting the surface information under thin clouds and cloud shadows in multitemporal images, which can help our model to iteratively separate the surface information and cloud/shadow.

In summary, the proposed method can separate the low-rank clean image and sparse cloud/shadow components. In addition, the detected cloud/shadow locations with higher accuracy can further improve the image quality obtained in the first step. From the above results, the proposed blind cloud/shadow removal method achieves the best results compared with the other state-of-the-art methods under investigation.

4.3. Discussion

We have proposed a unified convex optimization model for blind cloud and cloud shadow removal and detection of multitemporal images, which is different from previous completion related methods (Candès and Tao, 2010; Liu et al., 2013; Wang et al., 2016; Ng et al., 2017). For these completion related methods, they only considered the image low-rank prior and regarded the cloud/shadow areas as missing information. A number of simulated experiments were conducted to demonstrate the superiority of our TVLRSDC method compared to the state-of-the-art completion-based methods for information reconstruction and cloud/shadow detection. These observations can be found in Figs. 5–10. The completion-based methods destroyed and blurred the image details. In contrast, the proposed blind method successfully reconstructed the missing information in these areas. This is due to the fact that we consider the cloud/shadow prior and make full use of degraded surface information in thin cloud and cloud shadow areas of multitemporal images in our model. The detection of cloud/shadow can further help the information compensation of cloud/shadow-free areas.

Our method is also different from the completion-based methods, in which the cloud/shadow should be given in advance. That is to say, the results of the completion-based methods heavily depended on the given cloud/shadow mask. This phenomenon can be observed in real data results shown in Figs. 11–13. In Figs. 11 and 12, the locations of cloud and cloud shadow cannot be successfully detected by either improved Fmask or Sen2Cor. Then, the completion-based methods failed to reconstruct these areas in the final results. From the results of our method for real data, the cloud/shadow are successfully removed, and the information is reconstructed. In our framework, the low-rank regularization was used to reconstruct the information in cloud/shadow areas, and the sparse and TV regularizations were employed to extract the cloud/shadow component. The image and cloud/shadow components can be separated iteratively. Moreover, the cloud/shadow detection and

information compensation helped to preserve the cloud/shadow-free information. Different from the previous mask detection works such as (Frantz et al., 2018; Li et al., 2017), our method used a simple thresholding strategy to detect the cloud/shadow mask shown in Fig. 14. Compared to the masks extracted by improved Fmask and Sen2Cor, our method achieved better cloud and cloud shadow detection results as shown in Fig. 15. The simulated and real experimental results demonstrated that our proposed method achieves better results stably both in the information recovery and the cloud and cloud shadow detection for different kinds of datasets.

The proposed method can perform satisfactory results in most of the datasets, but there are some limitations. First, with the increment of spatial size, spectral bands, and temporal dimension, the computation cost and storage are too much. Second, if some specific areas are corrupted by the thick cloud/shadow all the time, these kinds of thick cloud/shadow are hard to be removed by our method, since in this case, the thick cloud component will be regarded as the low-rank image component. Third, it is very hard to find the perfect parameters for our proposed method. Due to the suboptimal selection of the parameters, there are still a few cloud/shadow-free areas that are destroyed and regarded as cloud/shadow masks.

5. Conclusion

In this paper, we proposed a method to remove and detect the cloud/shadow in multitemporal remote sensing images via the low-rank sparsity decomposition framework. Different from previous cloud/shadow removal methods, we simultaneously consider the characteristics of the surface-reflected image component and the cloud/shadow component. In the low-rank sparsity decomposition framework, the image component and the cloud/shadow component can be estimated iteratively. This strategy can benefit and complement each other. The low-rank prior that models the spectral-temporal correlations of multitemporal images can discriminatively separate the clean image and the cloud/shadow component, whereas the sparse and spatial-spectral TV regularization on the cloud/shadow component can help to isolate the cloud/shadow from the image component. We designed an ADMM-based algorithm to solve the proposed model. To prevent the degradation in the cloud/shadow-free areas, we further developed a thresholding method to detect the cloud/shadow areas to perform image restoration only in the cloud/shadow areas. Compared with completion-based methods in simulated and real experiments, the proposed method achieved better cloud/shadow removal results.

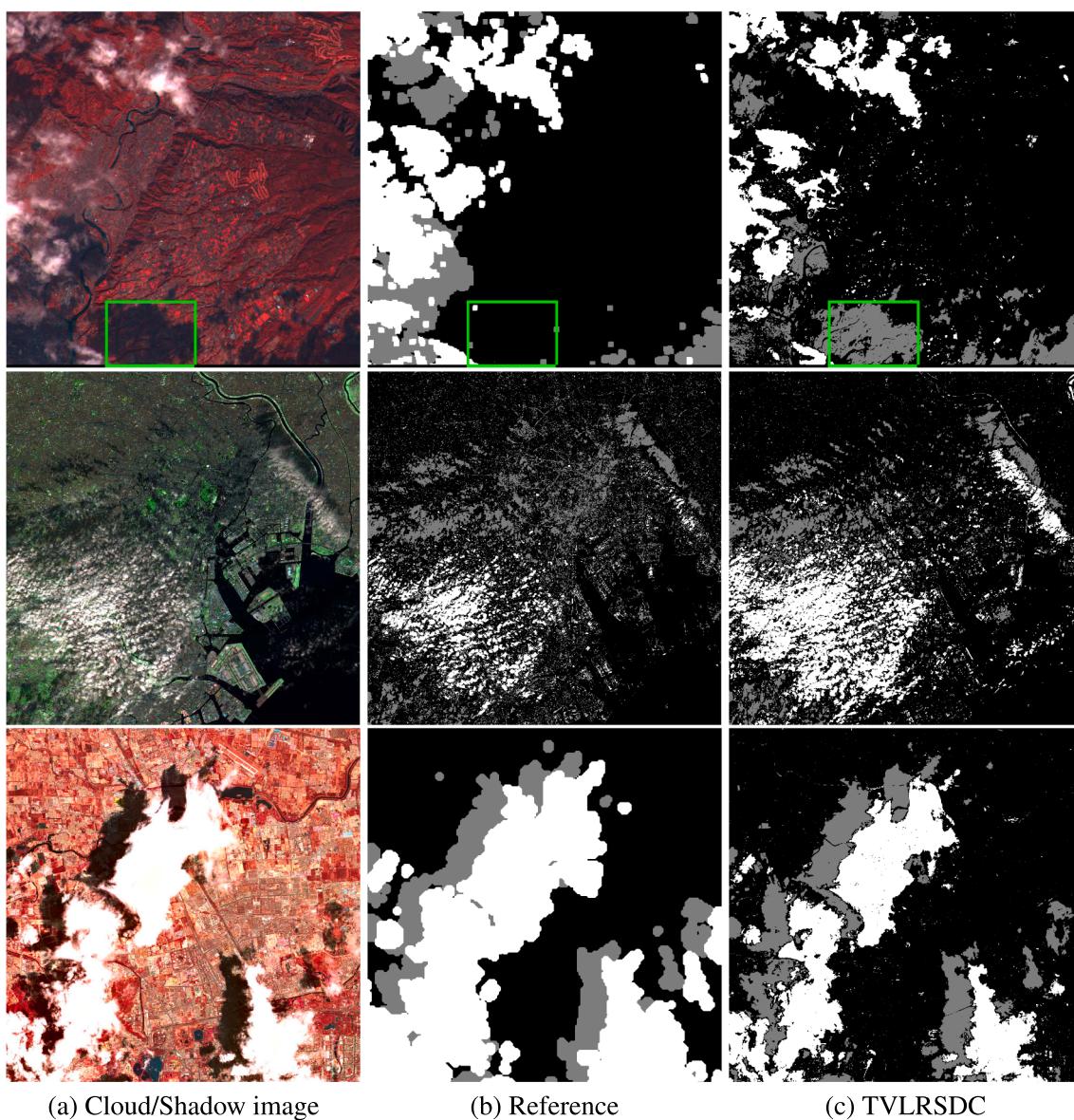


Fig. 15. Comparison of cloud/shadow mask detection results for three real datasets. (a) Cloud/shadow images, (b) masks detected by improved Fmask (first Sentinel-2 dataset), Sen2Cor (second Sentinel-2 dataset), and MAJA processor (SPOT-5 dataset), and (c) masks obtained by TVLRSDC.

Meanwhile, the proposed method can also obtain a competitive cloud/shadow detection result compared with the state-of-the-art improved Fmask algorithm and Sen2Cor tool.

In the future, we will try to incorporate physics-based priors into the image decomposition framework to further enhance its capability for cloud/shadow removal in multitemporal images. Moreover, by regarding the multitemporal images as four-order tensor, the tensor-based optimization model will be considered in future work.

Acknowledgments

The authors would like to thanks the editors and the anonymous reviewers for providing the valuable comments, which helped to greatly improve the manuscript. This work was supported by the NSFC (61772003) and the Japan Society for the Promotion of Science (KAKENHI 18K18067 and 19K20308).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the

online version, at <https://doi.org/10.1016/j.isprsjprs.2019.09.003>.

References

- Benabdelkader, S., Melgani, F., 2008. Contextual spatiotemporal postreconstruction of cloud-contaminated images. *IEEE Geosci. Remote Sens. Lett.* 5 (2), 204–208.
- Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* 3 (1), 1–122.
- Candès, E.J., Tao, T., 2010. The power of convex relaxation: near-optimal matrix completion. *IEEE Trans. Inf. Theory* 56 (5), 2053–2080.
- Cao, W., Yao, W., Jian, S., Meng, D., Yang, C., Cichocki, A., Xu, Z., 2016. Total variation regularized tensor RPCA for background subtraction from compressive measurements. *IEEE Trans. Image Process.* 25 (9), 4075–4090.
- Chen, Y., Huang, T.Z., Zhao, X.L., Deng, L.J., Huang, J., 2017. Stripe noise removal of remote sensing images by total variation regularization and group sparsity constraint. *Remote Sens.* 9 (6), 559.
- Chen, Y., Huang, T.Z., Deng, L.J., Zhao, X.L., Wang, M., 2017. Group sparsity based regularization model for remote sensing image stripe noise removal. *Neurocomputing* 267, 95–106.
- Chen, B., Huang, B., Chen, L., Xu, B., 2017. Spatially and temporally weighted regression: a novel method to produce continuous cloud-free landsat imagery. *IEEE Trans. Geosci. Remote Sens.* 55 (1), 27–37.
- Chen, Y., Huang, T.Z., Zhao, X.L., 2018. Destripping of multispectral remote sensing image using low-rank tensor decomposition. *IEEE J. Sel. Top. Appl. Earth Observ. Remote*

- Sens. 11 (12), 4950–4967.
- Cheng, Q., Shen, H., Zhang, L., Yuan, Q., Zeng, C., 2014. Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal mrf model. *ISPRS J. Photogramm. Remote Sens.* 92, 54–68.
- Frantz, D., Haß, E., Uhl, A., Stoffels, J., Hill, J., 2018. Improvement of the fmask algorithm for sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote Sens. Environ.* 215 (15), 471–481.
- He, W., Zhang, H., Zhang, L., Shen, H., 2016. Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration. *IEEE Trans. Geosci. Remote Sens.* 54 (1), 178–188.
- He, W., Zhang, H., Shen, H., Zhang, L., 2018. Hyperspectral image denoising using local low-rank matrix recovery and global spatial-spectral total variation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 11 (3), 713–729.
- Jiang, T.X., Huang, T.Z., Zhao, X.L., Deng, L.J., Wang, Y., 2019. Fastderain: a novel video rain streak removal method using directional gradient priors. *IEEE Trans. Image Process.* 28 (4), 2089–2102.
- Ji, T.Y., Yokoya, N., Zhu, X.X., Huang, T.Z., 2017. Nonlocal tensor completion for multitemporal remotely sensed images' inpainting. *IEEE Trans. Geosci. Remote Sens.* 56 (6), 3047–3061.
- Ju, J., Roy, D.P., 2008. The availability of cloud-free landsat ETM+ data over the conterminous united states and globally. *Remote Sens. Environ.* 112 (3), 1196–1211.
- Li, X., Shen, H., Zhang, L., Li, H., 2015. Sparse-based reconstruction of missing information in remote sensing images from spectral/temporal complementary information. *ISPRS J. Photogramm. Remote Sens.* 106, 1–15.
- Li, X., Shen, H., Li, H., Zhang, L., 2016. Patch matching-based multitemporal group sparse representation for the missing information reconstruction of remote-sensing images. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 9 (8), 3629–3641.
- Li, Z., Shen, H., Li, H., Xia, G., Gamba, P., Zhang, L., 2017. Multi-feature combined cloud and cloud shadow detection in GaoFen-1 wide field of view imagery. *Remote Sens. Environ.* 191, 342–358.
- Li, X., Wang, L., Cheng, Q., Wu, P., Gan, W., Fang, L., 2019. Cloud removal in remote sensing images using nonnegative matrix factorization and error correction. *ISPRS J. Photogramm. Remote Sens.* 148, 103–113.
- Lin, C.H., Tsai, P.H., Lai, K.H., Chen, J.Y., 2013. Cloud removal from multitemporal satellite images using information cloning. *IEEE Trans. Geosci. Remote Sens.* 51 (1), 232–241.
- Lin, C.H., Lai, K.H., Chen, Z.B., Chen, J.Y., 2014. Patch-based information reconstruction of cloud-contaminated multitemporal images. *IEEE Trans. Geosci. Remote Sens.* 52 (1), 163–174.
- Liu, J., Musialski, P., Wonka, P., Ye, J., 2013. Tensor completion for estimating missing values in visual data. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (1), 208–220.
- Lorenzi, L., Melgani, F., Mercier, G., 2011. Inpainting strategies for reconstruction of missing data in VHR images. *IEEE Geosci. remote sens. lett.* 8 (5), 914–918.
- Lorenzi, L., Melgani, F., Mercier, G., 2013. Missing-area reconstruction in multispectral images under a compressive sensing perspective. *IEEE Trans. Geosci. Remote Sens.* 51 (7), 3998–4008.
- Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., Gascon, F., 2016. Sentinel-2 sen2cor: L2a processor for users. *Living Planet Symp.* 2016, 1–8.
- Maalouf, A., Carré, P., Augereau, B., Fernandez Maloine, C., 2009. A bandelet-based inpainting technique for clouds removal from remotely sensed images. *IEEE Trans. Geosci. Remote Sens.* 47 (7), 2363–2371.
- Malek, S., Melgani, F., Bazi, Y., Alajlan, N., 2018. Reconstructing cloud-contaminated multispectral images with contextualized autoencoder neural networks. *IEEE Trans. Geosci. Remote Sens.* 56 (4), 2270–2282.
- Melgani, F., 2006. Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *IEEE Trans. Geosci. Remote Sens.* 44 (2), 442–455.
- Ng, M.K., Yuan, Q., Yan, L., Sun, J., 2017. An adaptive weighted tensor completion method for the recovery of remote sensing images with missing data. *IEEE Trans. Geosci. Remote Sens.* 55 (6), 3367–3381.
- Rakwatin, P., Takeuchi, W., Yasuoka, Y., 2009. Restoration of aqua MODIS band 6 using histogram matching and local least squares fitting. *IEEE Trans. Geosci. Remote Sens.* 47 (2), 613–627.
- Richter, R., Schläpfer, D., 2005. Atmospheric/topographic correction for satellite imagery, DLR report DLR-IB 565.
- Shen, H., Zeng, C., Zhang, L., 2011. Recovering reflectance of AQUA MODIS band 6 based on within-class local fitting. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 4 (1), 185–192.
- Shen, H., Li, H., Qian, Y., Zhang, L., Yuan, Q., 2014. An effective thin cloud removal procedure for visible remote sensing images. *ISPRS J. Photogramm. Remote Sens.* 96, 224–235.
- Tseng, D.C., Tseng, H.T., Chien, C.L., 2008. Automatic cloud removal from multi-temporal SPOT images. *Appl. Math. Comput.* 205 (2), 584–600.
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612.
- Wang, Z., Jin, J., Liang, J., Yan, K., Peng, Q., 2005. A new cloud removal algorithm for multi-spectral images, MIPPR 2005: SAR and Multispectral Image Processing, vol. 6043, 60430W, pp. 1–11.
- Wang, J., Olsen, P.A., Conn, A.R., Lozano, A.C., 2016. Removing clouds and recovering ground observations in satellite image sequences via temporally contiguous robust matrix completion. In: CVPR, pp. 2754–2763.
- Xia, J., Ghamisi, P., Yokoya, N., Iwasaki, A., 2018. Random forest ensembles and extended multientinction profiles for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* 56 (1), 202–216.
- Xu, M., Jia, X., Pickering, M., Plaza, A.J., 2016. Cloud removal based on sparse representation via multitemporal dictionary learning. *IEEE Trans. Geosci. Remote Sens.* 54 (5), 2998–3006.
- Yokoya, N., Iwasaki, A., 2015. Object detection based on sparse representation and hough voting for optical remote sensing imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8 (5), 2053–2062.
- Yokoya, N., Zhu, X.X., Plaza, A., 2017. Multisensor coupled spectral unmixing for time-series analysis. *IEEE Trans. Geosci. Remote Sens.* 55 (5), 2842–2857.
- Zhai, H., Zhang, H., Zhang, L., Li, P., 2018. Cloud/shadow detection based on spectral indices for multi/hyperspectral optical remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.* 144, 235–253.
- Zhang, Y., Guindon, B., Cihlar, J., 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of landsat images. *Remote Sens. Environ.* 82 (2–3), 173–187.
- Zhang, C., Li, W., Travis, D.J., 2009. Restoration of clouded pixels in multispectral remotely sensed imagery with cokriging. *Int. J. Remote Sens.* 30 (9), 2173–2195.
- Zhang, X., Qin, F., Qin, Y., 2010. Study on the thick cloud removal method based on multi-temporal remote sensing images. *Int. Conf. Multimedia Technol.* 1–3.
- Zhang, Q., Yuan, Q., Zeng, C., Li, X., Wei, Y., 2018. Missing data reconstruction in remote sensing image with a unified spatial-temporal-spectral deep convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* 56 (8), 4274–4288.
- Zhao, X.L., Wang, F., Huang, T.Z., Ng, M.K., Plemmons, R.J., 2013. Deblurring and sparse unmixing for hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* 51 (7), 4045–4058.