A fast multi-temporal optical data cloud removal method based on isophote constraint

Xiaoyu Yu [1], Jun Pan [1], Jiangong Xu [1], Junli Li [2]

[1] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China – (yuxiaoyu72, panjun1215, dd_xjg)@whu.edu.cn
[2] National Key Laboratory of Ecological Security and Sustainable Development in Arid Region, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China – lijl@ms.xjb.ac.cn

Keywords: Optical satellite, Multi-temporal data reconstruction, Cloud removal, Isophote constraint.

Abstract:

The problem of cloud cover significantly affects the quality and utility of optical satellite imagery, posing challenges for image analysis and interpretation. In response, this research presents a fast and efficient image reconstruction method by exploiting the strength of isophote in linear structure propagation of ground features, providing users with a practical option for cloud removal task. In addition, this research solves the inapplicability problem when the cloud-contaminated regions overlap between different temporal images, and then extends the application to the field of multi-temporal data cloud removal. Through several experiments, the proposed method outperforms other comparative methods for both dual-temporal data reconstruction and multi-temporal data reconstruction task. Furthermore, the method proves to be efficient and robust even with limited auxiliary information. The proposed multi-temporal image reconstruction framework is available at: https://github.com/YuXiaoyu221/ICMIR.

1. Introduction

Research based on MODIS data indicates that clouds cover approximately 67% of the Earth's surface (King et al., 2013). As a result, missing pixels in optical remote sensing imagery due to cloud covering have become a significant issue, severely reducing data quality and availability (Yu et al., 2024a). To address this challenge, various approaches have been explored to mitigate the adverse effects and recover pixels contaminated by clouds in optical data. According to the sources of auxiliary information, these methods can be categorized into two groups: single image-based methods, and reference image-based methods (Shen et al., 2015; Tao et al., 2022).

Single image-based reconstruction methods utilize the remaining information from the unaffected areas (spatial-domain) or the unaffected bands (spectral-domain). It is assumed that the missing areas have similar statistical or geometric structures with the remaining areas within the same image (Guillemot and Meur, 2014). Hence, interpolation methods, such as nearest neighbor interpolation and bilinear interpolation (Atkinson et al., 1990), as well as geostatistical interpolation methods (Rossi et al., 1994; Yu et al., 2011), can be employed to predict the missing pixels using auxiliary information from the clean areas. The exemplar-based image inpainting method, which relies on the structure propagation (Criminisi et al., 2004), also demonstrates its effectiveness in filling missing pixels. In the case of multispectral and hyperspectral optical data, the abundant spectral information also aids in the recovery of thin clouds. This is because longer wavelength bands generally have a greater ability to penetrate clouds (Kulkarni and Rege, 2020). Spectral correlation forms the basis for filling pixels in affected bands using clean pixels from unimpacted bands (Kulkarni and Rege, 2020; Li et al., 2014). Accordingly, approaches such as the haze-optimized transformation model, homomorphic filtering algorithm, and principal components transform method have been proposed to improve data quality under various atmospheric conditions (Feng et al., 2004; Hong and Zhang, 2018; Xu et al., 2019; Zhang et al., 2002). Since the complementary information from the spatial and spectral domains within the same image is limited, these methods are not suitable for reconstructing complex scenes with extensive coverage or thick cloud layers (Guillemot et al., 2014; Xu et al., 2022b).

Reference image-based methods use additional information from a reference image that covers the same geographic region to assist in image reconstruction (Fang et al., 2021; Kang et al., 2016). Both SAR and optical images from other time periods can serve as references to provide compensating information (Duan et al., 2024; Ebel et al., 2022; Mao et al., 2023; Xu et al., 2022a). However, achieving high-fidelity, cloud-free optical imagery through SAR data, that is, mapping cross-domain images, remains a challenging due to the distinct imaging mechanisms of active and passive sensors (Denaro and Lin, 2020). Reference information from temporal optical images can be more direct and effective, and is extensively used in cloud removal task (Zhang et al., 2020; Zheng et al., 2023; Zou et al., 2023). For example, regression model is studied to estimate missing pixels based on clean pixels of reference image (Cao et al., 2020; Zeng et al., 2013). In addition to pixel-based recovery, global optimization framework is also employed in image reconstruction. On the basis of Poisson editing algorithm in image inpainting (Perez et al., 2003), some methods use the gradient information of reference image as guidance field to minimize the gradient difference between the target image and the reference image (Hu et al., 2019; Lin et al., 2013; Yang et al., 2009; Yu et al., 2024b).

The isophote-based information method has recently been presented, which has demonstrated the strength of isophote in the cloud removal task of optical satellite imagery (Yu et al., 2024c). However, the high time cost remains a challenge, which limits its application in many fields. Furthermore, this method is mainly introduced for single image reconstruction, and the cloud coverage areas between the target image and the reference image should not overlap, otherwise it will lead to unstable results with abnormal color transitions. To address these issues, this research introduces a new fast image reconstruction method with both high efficiency and good results, which is also applicable to multi-temporal images with overlapping cloud-contaminated areas. Moreover, several comparative experiments are conducted, and the experimental results further demonstrate the method effectiveness in multi-temporal data reconstruction.

2. Methodology

2.1 Background

Isophote constraint is first introduced in image reconstruction to avoided the unstable outcomes of the gradient-guided image inpainting framework. The Poisson editing algorithm, which uses gradient information as constraint, is widely recognized as a foundational framework in image composition, inspiring many image reconstruction methods (Hu et al., 2019; Lin et al., 2013; Perez et al., 2003; Yang et al., 2009). The method presented by Hu et al. (2019) is used as example for subsequent discussion.

Using the gradient of reference data as constraint can well preserve image details. However, when the color differences between original image and reference image are non-linear, the result may be unstable. Figure 1 shows three groups of data to analyse the limitations of the gradient operator in image reconstruction. Figure 1(a) shows simulated non-natural images. Figure 1 (b) and (c) are cropped from satellite data. Five columns in each figure from left to right are respectively the original clean data, simulated data with missing area (denoted in white), reference data, the reconstruction result using gradient constraint (Hu et al., 2019), and the reconstruction result based on isophote constraint (Yu et al., 2024c).

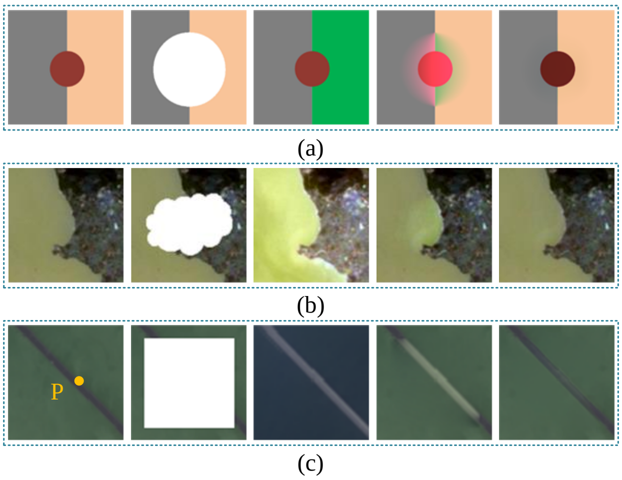

Figure 1. Limitation of gradient constraint. (a) simulated non-natural images. (b) and (c) are cropped from satellite data.

In the case of Figure 1 (a)-(b), the reconstruction results exhibit anomalous color transitions, where the color should be smooth and uniform. In Figure 1(c), the bridge in the gradient-based result is incorrectly filled with a mismatched color, inconsistent with the surrounding area. The isophote-based method achieves more accurate result and better preserves the original data characteristics. Taking Figure 1(c) as an example, the row containing point P is sampled to illustrate the profile more intuitively. Figure 2(a)-(e) are respectively the horizontal profile derived from the five images in Figure 1(c). The horizontal coordinate represents the row-coordinate, while the vertical coordinate corresponds to the grey value of each pixel. The red section between two dotted lines corresponds to the missing area, while the blue section represents the clean area. The arrow indicates the location of the bridge. In Figure 2(a), the bridge exhibits a lower grey value compared to the surrounding river, while in Figure 2(c), the bridge shows a higher grey value than the river. Figure 2(d) illustrates the result when the gradient from Figure 2(c) is used as a constraint to fill the missing area in Figure 2(b). Although the method preserves the details of Figure 2(c), it results in incorrect color variation. In contrast, Figure 2(e) exhibits a curve trend that more closely matches Figure 2(a) compared to Figure 2(d). Corresponding to the last column of Figure 1(c), the entire bridge maintains a uniform color, ensuring consistency across the image.

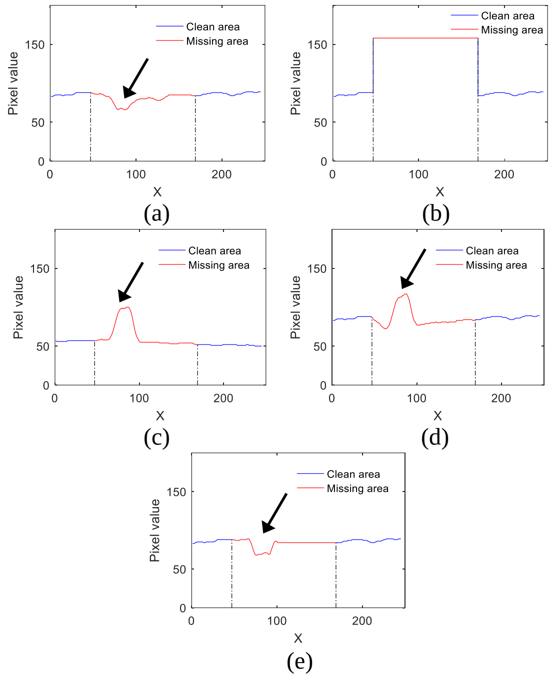

Figure 2. The horizontal profiles derived from Figure 1(c). (a) Original image. (b) Simulated image. (c) Reference image. (d) Result of gradient constraint. (e) Result of isophote constraint.

2.2 Motivation

It is shown that isophote can effectively capture the linear structure of ground features in an image. On this basis, the approach is verified in the cloud removal task of optical satellite imagery (Yu et al., 2024c). However, the unavoidable high time cost of the entire reconstruction process remains an obstacle, limiting its applicability in various real-world scenes. In addition, this approach is mainly used for single image reconstruction and does not take into account the overlap of cloud-contaminated areas between different images, which is very common especially in multi-temporal image reconstruction. Therefore, in view of these limitations, the main objectives of this research are as follows:

(1) Introduce a fast and efficient image reconstruction method by exploiting the strength of isophote in linear structure propagation of ground features, thus providing users with a more practical option for cloud removal task.

(2) Resolve the inapplicability problem when the cloud-contaminated regions overlap between different temporal images, and then extend the method application to the field of multi-temporal image reconstruction.

2.3 A fast multi-temporal data reconstruction method

2.3.1 Isophote calculation

The isophote represents the line connecting neighboring pixels with the same illumination intensity, i.e. similar grey values,

https://doi.org/10.5194/isprs-archives-XLVIII-3-2024-601-2024

and has been used in various image painting approaches to fill small holes (Ballester et al., 2001; Criminisi et al., 2004). In discrete images, the isophote of a pixel can be inferred from its local neighbors, which are aligned with the direction of the smallest gradient. To speed up the computation and make the equation system used to solve for missing pixels sparser, the isophote is determined by the local four-connected neighbors. As shown in Figure 3, the red arrow indicates the direction of the isophote at the central point *P*.

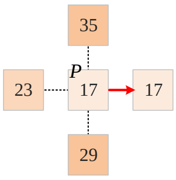

Figure 3. Isophote within the four-connected neighbors.

Assuming that q is a pixel in image I, N_q represents the four-connected neighboring pixels of q, and G_q is the gradient in each direction of the four neighbors, that is:

$$N_q = \begin{bmatrix} & q_{(-1,0)} & \\ q_{(0,-1)} & q & q_{(0,1)} \\ & q_{(1,0)} & \end{bmatrix}, \; G_q = \begin{bmatrix} & g_{(-1,0)} & \\ g_{(0,-1)} & 0 & g_{(0,1)} \\ & g_{(1,0)} & \end{bmatrix} \tag{1}$$

where $g_{(i,j)}$ is the variation of q and $q_{(i,j)}$: $g_{(i,j)} = q - q_{(i,j)}$. To highlight the direction of minimum gradient, the weight of the gradient in every direction is defined as:

$$w_{(i,j)} = \frac{1}{{g_{(i,j)}}^2 + \alpha} \tag{2}$$

Here, $\alpha = 0.01$, which used to avoid unstable result when $g_{(i,j)} = 0$. Then, isophote $\nabla I^{\perp}(q)$ can be calculated by adaptive gradient weighting:

$$\nabla I^{\perp}(q) = W_q \otimes G_q = \sum w_{(i,j)} \cdot \left(q_{(i,j)} - q\right) = W_q \otimes N_q \tag{3}$$

Consequently, greater weight is assigned to the direction of the smallest gradient, making the isophote heavily dependent on the minimum gradient.

2.3.2 Isophote-based missing region reconstruction

With the aim of minimizing the isophote differences between images during the reconstruction process, the isophote-constrained reconstruction algorithm can be expressed as:

$$\min_{I^*|_\Omega} \left(\sum_{q\in\Omega} \left(\nabla I^{*\perp}(q) - \nabla I_r^{\perp}(q) \right)^2 + \sum_{p\in\partial_\Omega} \left(I^*(p) - I_t(p) \right)^2 \right) \tag{4}$$

Here, $\nabla I^{\perp}$ is the isophote calculated by Equation (3). I_t, I_r, I^* are the target data, the reference data, and the recovered data, respectively. Ω is the missing area to be restored, and ∂_Ω is the outer boundary of Ω.

According to Equation (3), the isophote of Ω in I_r and I^* in Equation (4) can be represented as:

$$\begin{cases} \nabla I_r^{\perp}(\Omega) = W_\Omega * I_r(\Omega') \\ \nabla I^{*\perp}(\Omega) = W_{\Omega^*} * I^*(\Omega') \end{cases} \tag{5}$$

Since the isophote of a pixel involves its four-connected neighbors, Ω' also includes the outer neighboring area of Ω, that is: $\Omega' = \Omega \cup \partial_\Omega$. W_Ω denotes the isophote weight of the missing area Ω. Since the isophote differences between I_r and I^* are minimized, it can be approximated as:

$$\nabla I_r^{\perp}(\Omega) = W_\Omega * I^*(\Omega') \tag{6}$$

Assuming that there are n pixels in Ω, m pixels in the boundary area ∂_Ω, that is, $q_i\big|_{i=1}^{n} \in \Omega$, $q_i\big|_{i=n+1}^{n+m} \in \partial_\Omega$. Then, the terms in Equation (6) are represented as:

$$\begin{cases} \nabla I_r^{\perp}(\Omega) & = & [& \nabla I_r^{\perp}(q_1) & \nabla I_r^{\perp}(q_2) & \cdots & \nabla I_r^{\perp}(q_n) &]^{\mathrm{T}} \\ W_\Omega & = & [& W_{q_1} & W_{q_2} & \cdots & W_{q_n} &]^{\mathrm{T}} \\ I^*(\Omega') & = & [& I^*(q_1) & I^*(q_2) & \cdots & I^*(q_{n+m}) &]^{\mathrm{T}} \end{cases} \tag{7}$$

W_Ω is the weight matrix of size $n \times (n+m)$. W_{q_i} involves the weights related to q_i, and the size $1 \times (n+m)$:

$$W_{q_i} = \begin{bmatrix} w_{(q_i,q_1)} & w_{(q_i,q_2)} & \cdots & w_{(q_i,q_{n+m})} \end{bmatrix} \tag{8}$$

where $w_{(q_i,q_j)}$ represents the weight of pixel q_i and q_j. $w_{(q_i,q_j)}$ needs to be computed only when q_j is located within the four-connected neighbourhood of q_i; otherwise, $w_{(q_i,q_j)} = 0$.

The restored result should also maintain color consistency with the original data and achieve a seamless transition in the boundary. Therefore, incorporating the boundary condition $I^*(p) = I_t(p)\big|_{p\in\partial_\Omega}$, Equation (4) can be formulated as:

$$\begin{bmatrix} \nabla I_r^{\perp}(\Omega) \\ I_t(\partial_\Omega) \end{bmatrix} = \begin{bmatrix} W_\Omega \\ W_o \end{bmatrix} * I^*(\Omega') \tag{9}$$

$$I_t(\partial_\Omega) = \begin{bmatrix} I_t(q_{n+1}) & I_t(q_{n+2}) & \cdots & I_t(q_{n+m}) \end{bmatrix}^T \tag{10}$$

Here, W_o is a matrix of size $m \times (n+m)$. Only when $i = j$, $W_{o(q_i,q_j)} = 1$; otherwise, $W_{o(q_i,q_j)} = 0$. In this way, Equation (9) forms a system of linear equations with a sparse and symmetrical coefficient matrix, so that $I^*(\Omega')$ can be computed efficiently.

2.3.3 Framework of multi-temporal data reconstruction

Based on the reconstruction procedure presented in the previous section, a framework for multi-temporal imagery reconstruction is built. The workflow is shown in Figure 4. The cloud areas in the multi-temporal data need to be pre-marked first using the cloud detection method in (Dong et al., 2019). The multi-temporal images are then processed temporal-by-temporal until the reconstructed images are generated.

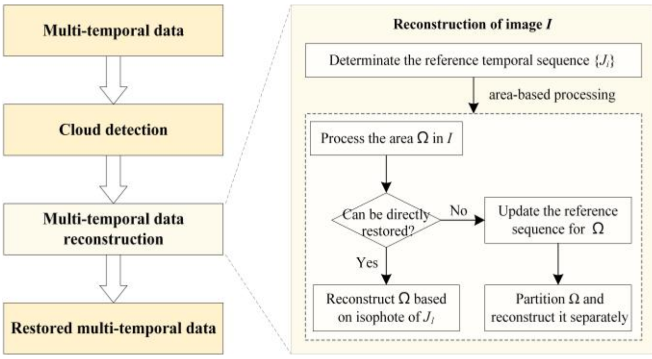

Figure 4. Workflow of the proposed method

Taking the image I in the multi-temporal data as an example, the reconstruction process can be concluded in two primary steps: the reference temporal sequence determination, and the area-based information reconstruction. The reference temporal sequence $\{J_i\}$ records which images are needed to minimize the cloud-contaminated areas in image I in a certain priority order. Subsequently, in the area-based reconstruction step, if the current region Ω can be fully restored using a single temporal in $\{J_i\}$, the isophote-based information reconstruction algorithm in Equation (4) is applied directly; otherwise, the region Ω is partitioned and restored separately. The two main steps are described in detail below.

(1) The reference temporal sequence determination

The reference temporals for the reconstruction of image I can be determined by considering the overlap of cloud cover between different images. If more than one temporal is available, that is, the cloud cover in image does not overlap with that of I, the temporal with the highest correlation is taken as the final reference temporal. In this case, image I can be restored directly according to Equation (4).

However, in some cases, image I cannot be restored using only one temporal. Then, the temporal that has the least cloud cover overlap with I is added to the reference sequence $\{J_i\}$ until the contaminated areas in I can be completely restored or all temporals are in the sequence.

(2) The area-based information reconstruction

For the current cloud region Ω, if it can be fully reconstructed using a single temporal in $\{J_i\}$, the isophote-based information reconstruction procedure is employed. Otherwise, the temporal sequence $\{J_i\}$ is reordered according to the proportion of the overlapping clouds with Ω. Therefore, the image with the least overlapping clouds is preferentially used to provide auxiliary information for partitional reconstruction, until the whole Ω is completed.

This process is illustrated in Figure 5. Figure 5(a) is a section of the target image I, and Figure 5(b) is the reference sequence for the region Ω consisting of $\{J_1, J_2\}$. J_1 is utilized first for reconstruction as it has the least overlapping clouds with Ω, and the partitional reconstructed result is shown in Figure 5(c), where the region Ω_2 remains cloud-contaminated. Figure 5(d) displays the final result by employing the reference information in J_2.

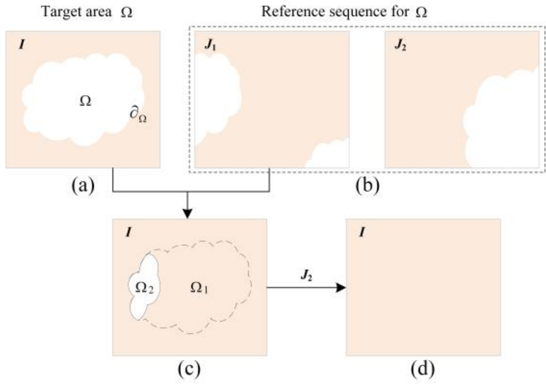

Figure 5. Area-based information reconstruction

Figure 6. The reconstruction of Ω_1.

A further explanation of the reconstruction process of Ω_1 is shown in Figure 6. In Figure 6(a), the point q lies on both the inner boundary of Ω_1 and the outer boundary of Ω_2. The local neighbourhood of q is illustrated in Figure 6(b). In this case, the isophote and the weight of q involved in Equation (7) are determined by the mean gradient, that is:

$$\nabla I^{\perp}(q) = W_q \otimes G_q = \sum_{i=1}^{5}\left(q_i - q\right) \tag{11}$$

3. Experiments and analysis

Three sets of experiments were performed. First, a dual-temporal data experiment was conducted to evaluate the performance of the presented method compared to the popular reference image-based reconstruction methods. Second, a multi-temporal experiment was conducted to verify the effectiveness of the method in the multi-temporal image reconstruction task. Finally, images with varying cloud cover were simulated to analyze the sensitivity to the size of the missing areas. All the following experiments were performed using Matlab R2016b with the Caffe framework on a Windows 11 system.

3.1 Dual-temporal data experiment

The dual-temporal experiment was conducted on Sentinel-2 data at 20 m resolution. The results of the presented method were compared with those of four baseline methods. Besides the gradient-based method (GBCR) (Hu et al., 2019), the comparison methods include the weighted linear regression method (WLR) (Zeng et al., 2013), the deep learning method (PSTCR) (Zhang et al., 2020), and the method integrating isophote and color-structure control (ICCSC) (Yu et al., 2024c).

Figure 7 shows the experimental data in true color composition. Figure 7 (a) and (b) are the simulated data with clouds and the original clean data, respectively. Since the cloud areas in two images do not overlap, the images can be reconstructed using temporal complementary information. Figure 7 (c)-(g) display the outcomes of WLR, PBCR, PSTCR, ICCSC, and the proposed method, respectively. Apparently that all methods fully restore the missing areas. However, in the magnified areas of the orange rectangles, Figure 7 (c) exhibits inconsistent color and unclear ground features. Some details in Figure 7 (e) also appear blurred. The results in Figure 7 (d) exhibit darker colors compared to Figure 7 (b). Overall, Figure 7(f)-(g) shows both uniform color and clear details. In the area marked by the red circle, Figure 7(f) is more consistent with the corresponding area in Figure 7(b), while the other areas of Figure 7(f)-(g) remain very similar.

		SSIM	PSNR	RMSE	Time/m
WLR	I_1	0.9783	32.10	0.0237	19.25
	I_2	0.9641	30.53	0.0298	
PBCR	I_1	0.8738	26.13	0.0642	4.49
	I_2	0.9660	29.34	0.0342	
PSTCR	I_1	0.9797	32.16	0.0246	21.23
	I_2	0.9654	30.67	0.0303	
ICCSC	I_1	**0.9857**	**33.85**	**0.0209**	38.66
	I_2	**0.9720**	**31.72**	**0.0261**	
Ours	I_1	0.9825	33.09	0.0223	**1.16**
	I_2	0.9691	31.39	0.0267	

Table 1. Quantitative analysis. 'Time/m' denotes that the time unit is minutes, and bold indicates the best values.

Quantitative evaluation was also performed, with the results presented in Table 1. Structural Similarity Index Metric (SSIM), Peak Signal-to-Noise Ratio (PSNR), and Root Mean Square Error (RMSE) were used to assess the accuracy of the reconstruction. In general, higher SSIM and PSNR values and lower RMSE indicate a better result, i.e. the reconstructed image is more similar to the original cloud-free image with less distortion. The time cost reflects the efficiency of the algorithm. The quantitative statistics in Table 1 further demonstrate the strength of the proposed method over WLR, PBCR and PSTCR. Compared to ICCSC, the method can achieve similar results in much less time.

3.2 Multi-temporal data experiment

Four temporal Sentinel-2 data were used in the multi-temporal data experiment. In addition to PSTCR, a tensor-completion-based multi-temporal data reconstruction algorithm (RCTVCR) is used for comparison (Xu et al., 2024). Figure 8 shows the experimental data with a size of 1801×1801. Figure 8(a) is the simulated cloud-contaminated multi-temporal images, and Figure 8(b) is the original clean data. The four rows in each figure correspond to the four temporals. The multi-temporal data reconstruction results of PSTCR, RCTVCR and the proposed method are shown in Figure 8(c)-(e), respectively.

From a global perspective, the proposed method performs effectively for each temporal of the multi-temporal data. In contrast, PSTCR works well for temporal-1 and temporal-4, while RCTVCR performs better for temporal-1, temporal-3 and temporal-4. In the temporal-2 result of PSTCR, i.e. the second row in Figure 8(c), the area marked by the red circle was incorrectly reconstructed with an abnormal color. And in the temporal-3 result of PSTCR, i.e. the third row in Figure 8(c), the area marked with red circle was filled with a noticeable black color. In the temporal-2 result of RCTVCR, i.e. the second row in Figure 8(d), there are obvious color inconsistencies and discontinuities, mainly due to radiometric variations between different temporal images. This problem is resolved in the results of the proposed method. In addition, the quantitative evaluation results are shown in Table 2. The best values for each temporal image are indicated in bold. In Table 2, the proposed method achieves good reconstruction results for the multi-temporal data with a low time cost.

		SSIM	PSNR	RMSE	Time/m
PSTCR	I_1	0.9811	32.88	0.0228	36.23
	I_2	0.9608	31.04	0.0358	
	I_3	0.9837	33.81	0.0205	
	I_4	0.9856	32.23	0.0245	
RCTVCR	I_1	0.9801	33.56	0.0210	0.58
	I_2	0.9270	23.27	0.0695	
	I_3	0.9875	34.76	0.0190	
	I_4	**0.9898**	33.69	**0.0208**	
Ours	I_1	**0.9864**	**36.07**	**0.0157**	0.64
	I_2	**0.9626**	**31.66**	**0.0262**	
	I_3	**0.9887**	**35.50**	**0.0139**	
	I_4	0.9882	**33.71**	0.0225	

Table 2. Quantitative analysis. 'Time/m' denotes that the time unit is minutes, and bold indicates the best values.

3.3 Sensitivity to the size of missing areas

Data with different cloud cover percentages were simulated to assess the sensitivity of the image reconstruction method to the size of missing regions. The original data, as shown in Figure 9(a), were cropped from Landsat 8 imagery of 30 m resolution. The simulated cloud cover percentages ranged from 10% to 90%. Figure 9 display the simulated images and the corresponding restored results.

The top of Figure 9(a) shows the original target image without simulated clouds, while the bottom displays the reference image that provides complementary information. Despite the obvious radiometric differences between the target image and the reference image in Figure 9(a), the restored results of different sizes of missing regions always show a high consistency with the original clean image, even when the size of missing regions reaches 90%. According to the zoomed details of the marked orange rectangles, the restoration results did not appear degradation in quality as the size of the cloud area increased. Furthermore, Figure 9(b)-(e) maintain a smooth transition at the boundary of the missing region, and Figure 9(f)-(g) keep a consistent color with the original data in Figure 9(a). The values of the evaluation metrics are also calculated for quantitative analysis, and the results are shown in Table 3. SSIM, PSNR, and RMSE values exhibit a minimal fluctuation as the proportion of missing regions varies from 10% to 90%, but remain relatively stable and satisfactory. This indicates that the isophote-based method is robust to changes in the size of missing areas. In terms of time cost, the method demonstrates generally good efficiency. However, as the proportion of missing pixels increases, the computational cost rises accordingly, leading to longer running time.

	SSIM	PSNR	RMSE	Time/m
10%	0.9607	30.15	0.0497	0.06
20%	0.9554	30.06	0.0502	0.11
40%	0.9482	29.99	0.0506	0.37
60%	0.9581	30.16	0.0496	1.40
80%	0.9467	29.97	0.0507	2.52
90%	0.9436	29.60	0.0531	2.95

Table 3. Quantitative statistics of Figure 9. 'Time/m' denotes that the time unit is minutes.

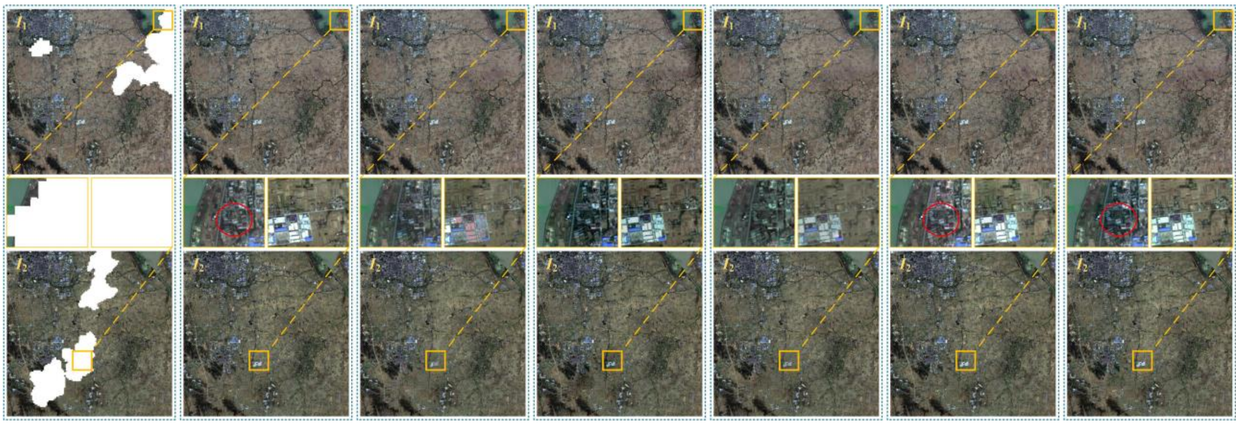

Figure 7. Dual-temporal experiment. (a) Simulated data. (b) Original data. (c) WLR. (d) PBCR. (e) PSTCR. (f) ICCSC. (g) Ours.

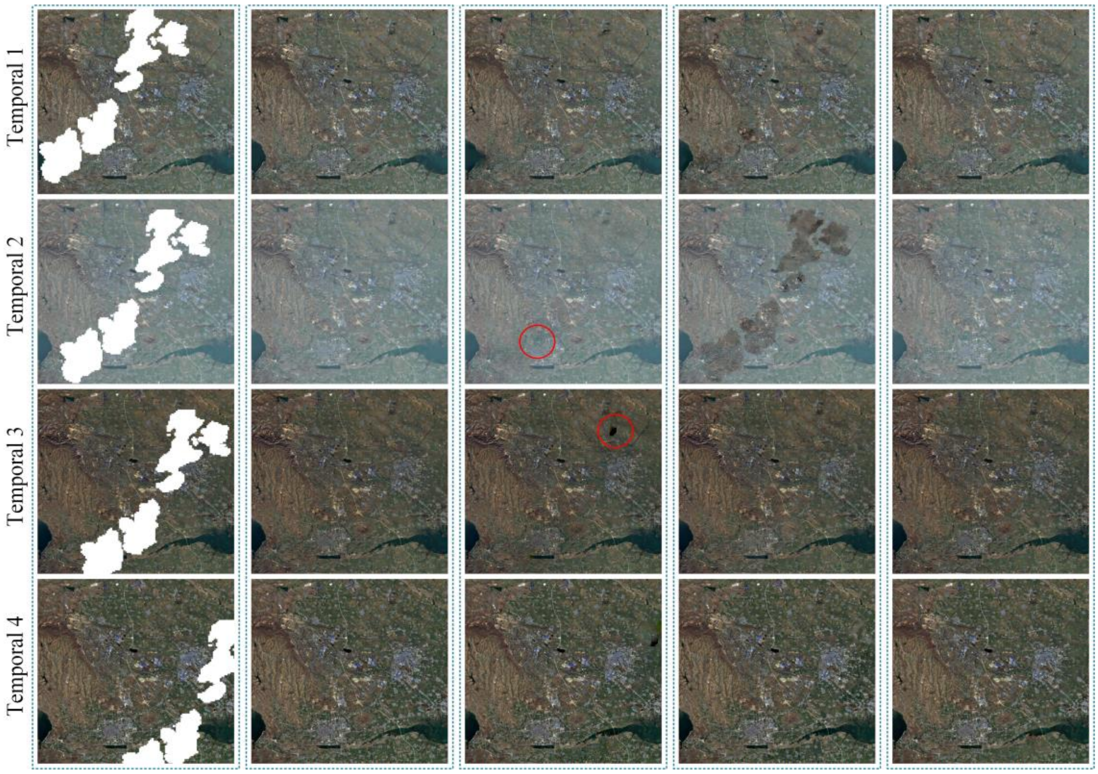

Figure 8. Multi-temporal experiment. (a) Simulated data. (b) Original data. (c) PSTCR. (d) RCTVCR. (e) Ours.

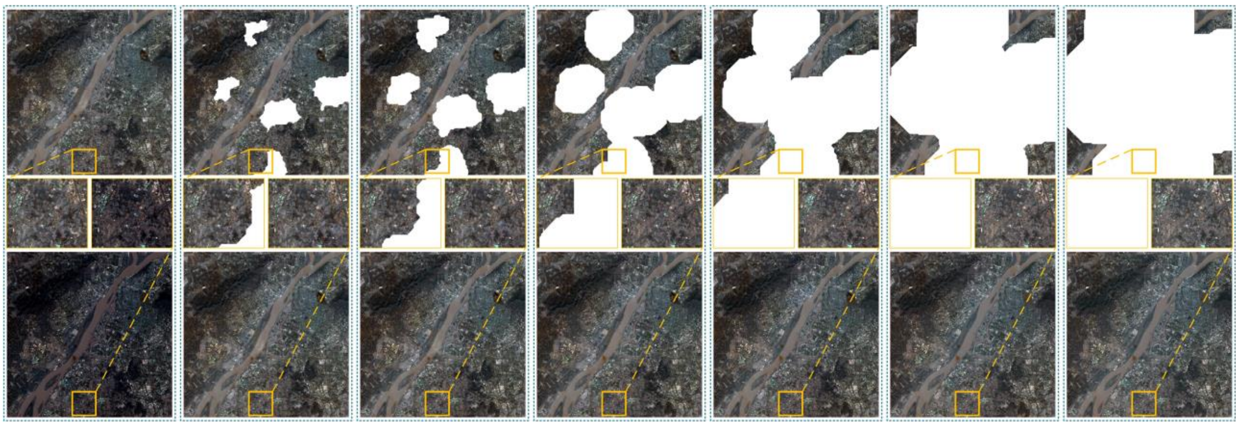

Figure 9. Experiment on different sizes of missing regions. (a) Original images. The percentages of cloud areas in (b)-(g) are approximately 10%, 20%, 40%, 60%, 80%, 90%, respectively.

3.4 Discussion

According to Figure 8 and Table 2, the proposed method demonstrates its effectiveness in cloud removal for multi-temporal data. However, when comparing the four temporal results of the method, Temporal-2 shows both lower SSIM and PSNR values compared to other temporals. Although the Temporal-2 result of the proposed method in Figure 8 exhibits visually satisfactory effects and achieves higher reconstruction accuracy than the other methods, the radiometric variations between the temporal-2 image and the other temporal images still negatively impact the method's overall performance. Therefore, the method is more suitable when radiometric differences and temporal changes between multi-temporal data are minimal not significant.

4. Conclusions

To address the challenges posed by unavoidable cloud cover in multi-temporal optical satellite data and to improve data usability, this research explores an efficient multi-temporal data reconstruction method that introduces isophote information as a constraint during the process. Through several experiments, the proposed method outperforms other comparative methods for both dual-temporal and multi-temporal data reconstruction. In the experiment with varying cloud cover, the proposed method produces high quality images with superior reconstruction accuracy even with limited auxiliary information. In conclusion, the isophote constraint shows significant potential in the multi-temporal image reconstruction task.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant 41971422 and Grant 42090011, and the Tianshan Talent-Science and Technology Innovation Team (2022TSYCTD0006).

References

Atkinson, P.M., Curran, P.J., Webster, R., 1990. Sampling Remotely Sensed Imagery for Storage, Retrieval, and Reconstruction. *The Professional Geographer*, 42(3), 345-353.

Ballester, C., Caselles, V., Verdera, J., Bertalmio, M., Sapiro, G., 2001. A variational model for filling-in gray level and color images, Proceedings Eighth *IEEE International Conference on Computer Vision*. ICCV 2001, pp. 10-16 vol.11.

Cao, R., Chen, Y., Chen, J., Zhu, X., Shen, M., 2020. Thick cloud removal in Landsat images based on autoregression of Landsat time-series data. *Remote Sensing of Environment*, 249, 112001.

Criminisi, A., Perez, P., Toyama, K., 2004. Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, 13(9), 1200-1212.

Denaro, L.G., Lin, C.H., 2020. Hybrid Canonical Correlation Analysis and Regression for Radiometric Normalization of Cross-Sensor Satellite Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 976-986.

Dong, Z., Wang, M., Li, D., Wang, Y., Zhang, Z., 2019. Cloud Detection Method for High Resolution Remote Sensing Imagery Based on the Spectrum and Texture of Superpixels. *Photogrammetric Engineering and Remote Sensing*, 85(4), 257-268.

Duan, C., Belgiu, M., Stein, A., 2024. Efficient Cloud Removal Network for Satellite Images Using SAR-Optical Image Fusion. *IEEE Geoscience and Remote Sensing Letters*, 21, 1-5.

Ebel, P., Xu, Y., Schmitt, M., Zhu, X.X., 2022. SEN12MS-CR-TS: A Remote-Sensing Data Set for Multimodal Multitemporal Cloud Removal. *IEEE Trans. Geosci. Remote Sensing*, 60, 1-14.

Fang, Z., Yu, X., Pan, J., Fan, N., Wang, H., Qi, J., 2021. A Fast Image Mosaicking Method Based on Iteratively Minimizing Cloud Coverage Areas. *IEEE Geoscience and Remote Sensing Letters*, 18(8), 1371-1375.

Feng, C., Ma, J.W., Dai, Q., Chen, X., 2004. An improved method for cloud removal in ASTER data change detection, *IEEE International Geoscience & Remote Sensing Symposium*, pp. 3387-3389.

Guillemot, C., Le, Meur, O., 2014. Image Inpainting : Overview and Recent Advances. *Signal Processing Magazine*, 31(1), 127-144.

Guillemot, C., Meur, O.L., 2014. Image Inpainting : Overview and Recent Advances. *IEEE Signal Processing Magazine*, 31(1), 127-144.

Hong, G., Zhang, Y., 2018. Haze removal for new generation optical sensors. *International Journal of Remote Sensing*, 39(5), 1491-1509.

Hu, C., Huo, L., Zhang, Z., Tang, P., 2019. Automatic Cloud Removal from Multi-Temporal Landsat Collection 1 Data Using Poisson Blending, IGARSS 2019 - 2019 *IEEE International Geoscience and Remote Sensing Symposium*, pp. 1661-1664.

Kang, Y., Pan, L., Chen, Q., Zhang, T., Zhang, S., Liu, Z., 2016. Automatic mosaicking of satellite imagery considering the clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-3, 415-421.

King, M., Platnick, S., Menzel, W.P., Ackerman, S.A., Hubanks, P.A., 2013. Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites. *IEEE Transactions on Geoscience & Remote Sensing*, 51(7), 3826-3852.

Kulkarni, S.C., Rege, P.P., 2020. Pixel level fusion techniques for SAR and optical images: A review - ScienceDirect. *Information Fusion*, 59, 13-29.

Li, X., Shen, H., Zhang, L., Zhang, H., Yuan, Q., 2014. Dead Pixel Completion of Aqua MODIS Band 6 Using a Robust M-Estimator Multiregression. *IEEE Geoscience and Remote Sensing Letters*, 11(4), 768-772.

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 Sensing*, 51(1), 232-241.

Mao, Y., Van Niel, T.G., McVicar, T.R., 2023. Reconstructing cloud-contaminated NDVI images with SAR-Optical fusion using spatio-temporal partitioning and multiple linear regression.

ISPRS Journal of Photogrammetry and Remote Sensing, 198, 115-139.

Perez, P., Gangnet, M., Blake, A., 2003. Poisson image editing. *Acm Transactions on Graphics*, 22(3), 313-318.

Rossi, R.E., Dungan, J.L., Beck, L.R., 1994. Kriging in the shadows: Geostatistical interpolation for remote sensing. *Remote Sensing of Environment*, 49(1), 32-40.

Shen, H., Li, X., Cheng, Q., Zeng, C., Yang, G., Li, H., Zhang, L., 2015. Missing Information Reconstruction of Remote Sensing Data: A Technical Review. *IEEE Geoscience and Remote Sensing Magazine*, 3(3), 61-85.

Tao, C., Fu, S., Qi, J., Li, H., 2022. Thick Cloud Removal in Optical Remote Sensing Images Using a Texture Complexity Guided Self-Paced Learning Method. *IEEE Trans. Geosci. Remote Sensing*, 60, 1-12.

Xu, F., Shi, Y., Ebel, P., Yu, L., Xia, G.-S., Yang, W., Zhu, X.X., 2022a. GLF-CR: SAR-enhanced cloud removal with global–local fusion. *ISPRS Journal of Photogrammetry and Remote Sensing*, 192, 268-278.

Xu, M., Jia, X., Pickering, M., Jia, S., 2019. Thin cloud removal from optical remote sensing images using the noise-adjusted principal components transform. *ISPRS Journal of Photogrammetry and Remote Sensing*, 149, 215-225.

Xu, S., Wang, J., Wang, J., 2024. Fast Thick Cloud Removal for Multi-Temporal Remote Sensing Imagery via Representation Coefficient Total Variation. *Remote Sensing*, 16(1), 152.

Xu, Z., Wu, K., Wang, W., Lyu, X., Ren, P., 2022b. Semi-supervised thin cloud removal with mutually beneficial guides. *ISPRS Journal of Photogrammetry and Remote Sensing*, 192, 327-343.

Yang, W., Zheng, J., Cai, J., Rahardja, S., Chen, C.W., 2009. Natural and Seamless Image Composition With Color Control. *IEEE Transactions on Image Processing*, 18(11), 2584-2592.

Yu, C., Chen, L., Su, L., Fan, M., Li, S., 2011. Kriging interpolation method and its application in retrieval of MODIS aerosol optical depth, 2011 19*th International Conference on Geoinformatics*, pp. 1-6.

Yu, X., Pan, J., Chen, S., Wang, M., 2024a. A flexible multi-temporal orthoimage mosaicking method based on dynamic variable patches. *Information Fusion*, 108, 102350.

Yu, X., Pan, J., Wang, M., Xu, J., 2024b. A curvature-driven cloud removal method for remote sensing images. *Geo-spatial Information Science*, 27(4), 1326-1347.

Yu, X., Pan, J., Xu, J., Wang, M., 2024c. Missing information reconstruction integrating isophote constraint and color-structure control for remote sensing data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 208, 261-278.

Zeng, C., Shen, H., Zhang, L., 2013. Recovering missing pixels for Landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method. *Remote Sensing of Environment*, 131, 182-194.

Zhang, Q., Yuan, Q., Li, J., Li, Z., Shen, H., Zhang, L., 2020. 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, 148-160.

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 Sensing of Environment*, 82(2), 173-187.

Zheng, W.J., Zhao, X.L., Zheng, Y.B., Lin, J., Zhuang, L., Huang, T.Z., 2023. Spatial-spectral-temporal connective tensor network decomposition for thick cloud removal. *ISPRS Journal of Photogrammetry and Remote Sensing*, 199, 182-194.

Zou, X., Li, K., Xing, J., Tao, P., Cui, Y., 2023. PMAA: A Progressive Multi-Scale Attention Autoencoder Model for High-Performance Cloud Removal from Multi-Temporal Satellite Imagery, ECAI 2023: 26th *European Conference On Artificial Intelligence*, pp. 3165-3172.