A Fast Method for Cloud Removal and Image Restoration on Time Series of Multispectral Images

Manuel Bertoluzza*, Claudia Paris†, Lorenzo Bruzzone‡
Dept. of Information Engineering and Computer Science, University of Trento, Trento, Italy
Email: *manuel.bertoluzza@unitn.it, †claudia.paris@unitn.it, ‡lorenzo.bruzzone@unitn.it

***Abstract*—Missing information in optical remote sensing data due to sensor malfunction or cloud cover reduces their usability. Although several methods in the literature allow an accurate image restoration, they are usually time-consuming and thus impractical for operational applications. In this paper, we propose a Fast Cloud Removal Approach (FCRA) that exploits the multitemporal information in an efficient way to sharply reduce the computational burden. In greater details, the proposed unsupervised method: (i) generates a short time series (TS) of cloud free images temporally close to the cloudy one (i.e., the target image), (ii) detects ground-clear pixels in the target image similar to the missing pixels using the multitemporal information (i.e., similar temporal patterns), and (iii) restores the occluded pixels using the most similar ground-clear spectral values from the same image. The method reduces the computational effort by focusing on short image TS and by detecting similar temporal patterns in a fast and effective way by using the KD-trees. Moreover, since the method restores the missing pixels with ground clear pixels belonging to the same image, no computational demanding approaches are needed to adjust the replacement information (i.e., regression). Experimental results obtained on Sentinel-2 images on simulated clouds acquired over two test areas located in Italy demonstrate the effectiveness of the proposed method.**

***Index Terms*—Image restoration, image time series, multitemporal analysis, satellite multispectral images, similarity search, KD-tree, Remote Sensing.**

I. Introduction

Nowadays sensors such as Sentinel-2 or Landsat 8 are generating unprecedented volumes of data at high spatial, spectral and temporal resolutions with a worldwide coverage. This leads to dense time series (TS) of optical images that increase the opportunity to monitor dynamic phenomena, thus enabling a wide range of new opportunities in the field of multitemporal applications [1]. To better exploit optical images, it is important to recover missing information in the data due to poor atmospheric conditions (e.g., thick clouds and related shadows) or defective sensors. In the literature, a large effort has been devoted to solve this problem. In particular, the gap filling methods can be grouped into four main categories: (i) spatial-based methods, (ii) spectral based methods, (iii) temporal-based methods, and (iv) hybrid methods [2]. Spatial-based and spectral-based methods exploit the self-similarity properties of the optical images to reconstruct missing information. Although these methods do not require auxiliary information sources, they are effective when restoring small portions of images (e.g., corrupted pixels) but may generate inaccurate results on large regions having complex textures. Hence, when addressing the presence of clouds, it is necessary to exploit the temporal information [3]. In particular, the most accurate results have been achieved by hybrid methods, which usually combine the temporal information with spatial/spectral features. In [4], the authors propose a cloud removal technique based on a modified neighborhood similar pixel interpolator (MNSPI) which predicts the value of a cloudy pixel from its neighboring similar pixels with the help of auxiliary cloud-free images. In [5], the authors propose a sophisticated cloud removal method which combines multitemporal and dictionary learning methods. Dictionaries of the cloudy target areas and the cloud-free reference areas are learned in the spectral domain, and then used to perform the restoration. Although these methods accurately solve the cloud restoration problem, they are time-consuming.

In this paper, we propose a fast and effective technique that can achieve a good trade-off between restoration accuracy and computational burden. In greater detail, the method: (i) automatically detects a short TS of cloud free images temporally close to the target one, (ii) for each non-valid pixel of the target image, it identifies the most similar pixels present in the scene by analyzing their temporal patterns, and (iii) restores invalid pixels using the most similar ground-clear one in the target image. To reduce the computational effort the method: (a) exploits a short TS of images, (b) performs the detection of similar temporal patterns using an efficient KD-tree search algorithm, and (c) restores the cloudy pixels without using time-consuming approaches to reconstruct the missing information (e.g., regression).

II. Proposed Image Restoration Method

Fig. 1 shows the workflow of the FCRA method. The method is based on three main steps: (i) pre-processing of multi-spectral images to generate a short TS of products used to detect similar temporal patterns, (ii) optimized detection of similar temporal pattern, and (iii) automatic restoration of cloudy pixels in the target image. Cloud and shadow masks are assumed to be available by one of the many methods presented in the literature [6].

A. Pre-Processing

The first step of the proposed method generates the short TS of cloud free images temporally close to the target one. By analyzing the dense TS of images, the method automatically retrieves cloud-free data by using the available cloud and related shadow masks. Starting from the target image, the

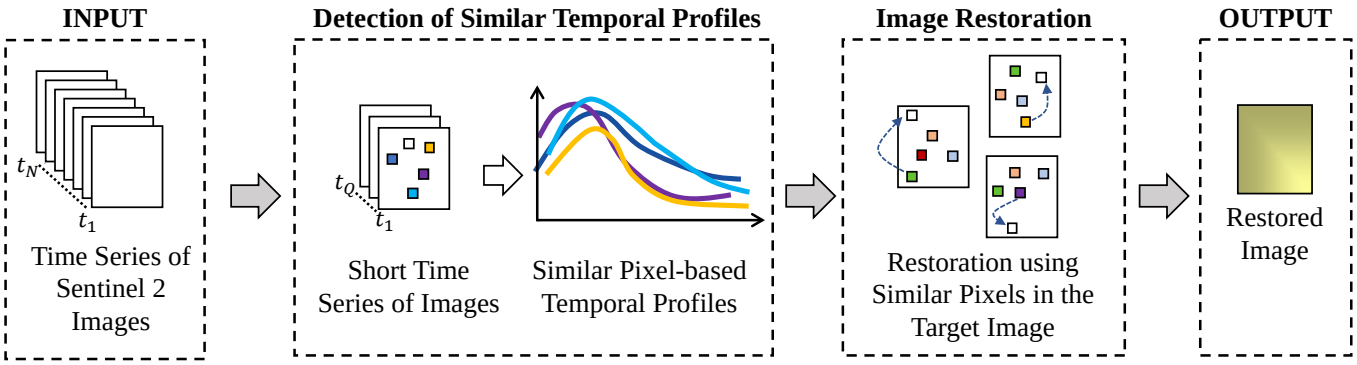

Fig. 1: Workflow of the proposed unsupervised multitemporal image restoration technique.

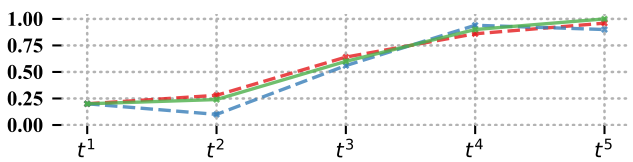

Fig. 2: Example of similar temporal patterns with $F = 5$: temporal profile of the jth pixel (green line) and two most similar profiles (red and blue dashed lines).

cloud masks are automatically evaluated to detect the images having clouds non overlapped to the clouds present in the target data. Temporally close images are preferred to increase the reliability of the similar patterns detection step.

Let $\boldsymbol{X}^q$ be a multispectral image of size $M \times G \times B$ acquired at time t^q, where $q \in [1, N]$ and N is the number of images in the dense initial TS. Let $\{\boldsymbol{X}^1, \boldsymbol{X}^2, \cdots, \boldsymbol{X}^Q\}$, with $(Q \ll N)$, be the set of selected co-registered images temporally close to the target image $\boldsymbol{X}^t$. Finally, let $C^q \in \mathbb{R}^{M \times G}$ be the binary mask indicating the cloud or cloud shadow pixels and valid pixels in the image $\boldsymbol{X}^q$.

B. Detection of Similar Temporal Patterns

Let $\mathbf{p}_j = [\mathbf{x}_j^1, \mathbf{x}_j^2, \cdots, \mathbf{x}_j^Q]$ bet the multitemporal pattern associated to the jth pixel of the short TS of images, where $\mathbf{x}_j^1 = [x_{j,1}^1, x_{j,2}^1, \cdots, x_{j,B}^1]$ represents the B spectral values of the j-th pixel in the first image of the TS. Let us define as $L = M \times G$ the number of pixels present in the TS and $F = Q \times B$ the number of points along the multitemporal patterns. To perform the restoration of the j-th cloudy pixel in the target image, i.e., $\mathbf{x}_j^t$, we aim to detect the most similar K temporal patterns present in the target image $\{\mathbf{x}_{j,1}^t, \mathbf{x}_{j,2}^t, \cdots, \mathbf{x}_{j,K}^t\}$. Such detection is performed by considering their temporal patterns in the short TS of images. In particular, the similarity between two temporal patterns $\mathbf{p}_j$ and $\mathbf{p}_i$ can be computed by using the Euclidean distance, i.e.,:

$$d(\mathbf{p}_j, \mathbf{p}_i) = \sqrt{\left(x_{j,1}^1 - x_{i,1}^1\right)^2 + \cdots + \left(x_{j,B}^Q - x_{i,B}^Q\right)^2} \quad (1)$$
$$\text{with } i, j = [1, L] \text{ and } i \neq j$$

where $\{\mathbf{x}_{j,1}^t, \mathbf{x}_{j,2}^t, \cdots, \mathbf{x}_{j,K}^t\}$ are obtained by selecting the patterns with the smallest distance value (see Fig. 2). It is worth noting that in the computation of the distance only the ground-clear samples present in the short TS of images are considered, leveraging on the information provided by the set of cloud masks $\{C^1, C^2, \cdots, C^Q\}$.

To make the search of similar temporal profiles scalable and avoid the comparison of $L \cdot (L-1)$ temporal profiles, the k-nearest neighbor search is performed by using the KD-trees search algorithm. KD-trees are well known space-partitioning data structures which allow the efficient organization of the multitemporal patterns $\mathbf{p}_j$ in a F-dimensional space. They have been widely employed in the literature for fast image patch matching [7] and clustering [8]. Note that since the j-th pixel is covered by a cloud in the target image, the detection of its similar pixels rely completely on the information provided by the cloud-free TS. Although the considered TS is short, its multitemporal information increases the probability of correctly detecting similar patterns with respect to bi-temporal images methods (which are typically used), where the result may be affected by land-cover changes (e.g., crop or vegetation phenology).

C. Image Restoration

The last step of the proposed method performs the restoration of invalid pixels in the target image $\boldsymbol{X}^t$. The restored target image $\hat{\boldsymbol{X}}^t$ can be computed pixel-wise by using the cloud mask C^t and the most similar temporal profiles found in the previous steps. Cloudy pixels in the target image $\boldsymbol{X}^t$ are restored by computing the band-wise median of the cloud-free multispectral pixels in the same image $\boldsymbol{X}^t$ corresponding to the most similar temporal profiles along the short TS, i.e.,:

$$\hat{\mathbf{x}_j}^t = \begin{cases} \text{Med}\{\mathbf{x}_{j,1}^t, \cdots, \mathbf{x}_{j,K}^t\} & \text{if cloud} \\ \mathbf{x}_j^t & \text{otherwise} \end{cases}.$$

A value of $K > 1$ is needed to be robust to possible errors in the cloud masks, i.e., to avoid to restore a cloudy pixel with another cloudy pixel at the cost of a lower restoration spectral accuracy due to the median applied to a high number of pixels. Please note that differently from the literature, we are not reconstructing the target image using pixels extracted

from cloud-free images acquired at a different date. In contrast, the radiometric values of the target data are used to reduce the probability of introducing bias in the statistics of the reconstructed images. It is worth noting that this step is fully automatic and can be easily parallelized by processing small tiles of the TS of images at the same time.

III. Experimental Results

The experimental analysis was carried out over two study areas located in two different Sentinel-2 granules in Italy, i.e., 32TPQ and 32TPS. The images were atmospherically corrected and the cloud masks were generated by using Sen2Cor [9]. For the experiments, we considered the RGB and NIR spectral bands acquired by Sentinel-2 at 10m spatial resolution (i.e., $B = 4$). These bands were converted to surface reflectance and normalized in the range $[0, 1]$. The image size for both areas is 200×200px. Tab. I contains their acquisition dates. For each region of interest, a TS of $Q = 4$ cloud-free images was selected and for each study area synthetic gaps were simulated in a target image. For both of them, $\boldsymbol{X}^3$ has been selected as target image $\boldsymbol{X}^t$ where gap filling is performed. The number of pixels to restore are $4\,000$ and $6\,000$, respectively. This setup allows us to compare the original and reconstructed target images in a qualitative and quantitative way.

TABLE I: Acquisition dates of the five images acquired by Sentinel-2 over the two Regions of Interest (ROI).

	ROI 1 (32TPQ)	**ROI 2 (32TPS)**
$\boldsymbol{X}^1$	2017-06-03	2016-07-18
$\boldsymbol{X}^2$	2017-06-13	2017-07-16
$\boldsymbol{X}^3$	2017-06-23	2017-07-31
$\boldsymbol{X}^4$	2017-07-03	2017-08-02
$\boldsymbol{X}^5$	2017-07-23	2018-08-27

To draw comparisons, the basic patch replacement (PR) and the MNSPI methods [4] were applied to our dataset. The PR is fast and efficient, but achieves accurate restoration results if there is at least a Sentinel-2 image similar to the target one in the considered time series of images. Thus, the applicability of the method is strongly limited by the assumption that a cloud-free and temporally close acquisition is available for restoring cloudy target data. In contrast, MNSPI is effective in recovering the details of the missing data. However, due to the exhaustive search performed in the neighborhood of each cloudy pixels, it can be extremely time-consuming when recovering large cloudy areas.

Fig. 3 compares the image restoration results obtained by the proposed method using $Q = 150$ similar temporal patterns. The Root-Mean-Square Deviation (RMSD) computed per pixel (averaged on the 4 spectral bands) is provided. This qualitative comparison shows very good restoration performance. Both the spatial and spectral features are accurately recovered using the multitemporal information extracted from the TS of images. As expected, the reconstruction errors are mostly located in areas with a high frequency of changes.

TABLE II: Quantitative restoration results obtained using the MNSPI method [4], the simple Patch Replacement (PR) and the proposed method (FCRA) on ROI 1 (32TPQ). The Root Mean Square Error (RMSE), the Correlation Coefficient (CC) and the Peak Signal to Noise Ratio (PSNR) are reported per each of the four bands

Band	**Method**	**RMSE**	**CC**	**PSNR [dB]**
Blue (band 2)	MNSPI	0.05	0.89	26.5
	PR	0.09	0.51	20.9
	FCRA	0.04	0.88	26.2
Green (band 3)	MNSPI	0.05	0.85	25.3
	PR	0.10	0.47	19.9
	FCRA	0.05	0.87	25.5
Red (band 4)	MNSPI	0.07	0.83	22.7
	PR	0.13	0.48	17.6
	FCRA	0.06	0.88	24.0
NIR (band 8)	MNSPI	0.04	0.87	27.1
	PR	0.12	0.48	18.5
	FCRA	0.06	0.74	25.2

Note that this complex agricultural study area characterized by landscape heterogeneity was used as test case to assess the effectiveness of the proposed multitemporal approach. Due to the crop phenology, many land-cover changes occur in this study site among different acquisitions, thus increasing the level of challenge in cloud restoration.

The qualitative evaluation is confirmed by the quantitative results reported in Tab. II, which shows the image restoration performance obtained by the proposed techniques in terms of Root Mean Square Error, Correlation Coefficient and Peak Signal to Noise Ratio. These indexes show a low reconstruction error and demonstrate the validity of the proposed multitemporal image restoration approach. Good results are achieved by the proposed FCRA, which is able to accurately restore the images with a computational efficiency similar to PR method. In such a complex scenario, the PR method leads to poor restoration results. Thus, this fast and efficient method achieves a reliable result if there is at least one similar Sentinel-2 image in the considered time series of images. However, it is not reasonable to assume a free-cloud temporal close acquisition for each cloudy target data. At the contrary, the proposed approach is able to properly recover the missing information due to clouds present along the TS of images. The best overall results are obtained by MNSPI which performs an exhaustive search to detect the most similar pixels at the cost of a significantly higher computational burned. This is confirmed by the quantitative evaluation presented in Tab. III, which reports the computation times adjusted for thread number obtained to restore four simulated clouds presented in the scene in tile 32TPQ ($4\,000$ cloudy pixels, representing an area of almost 0.4 Km^2).

IV. Conclusion

In this paper, a fast and efficient cloud removal method for multispectral images has been presented. The method assumes a short TS of images is available to accurately detected similar

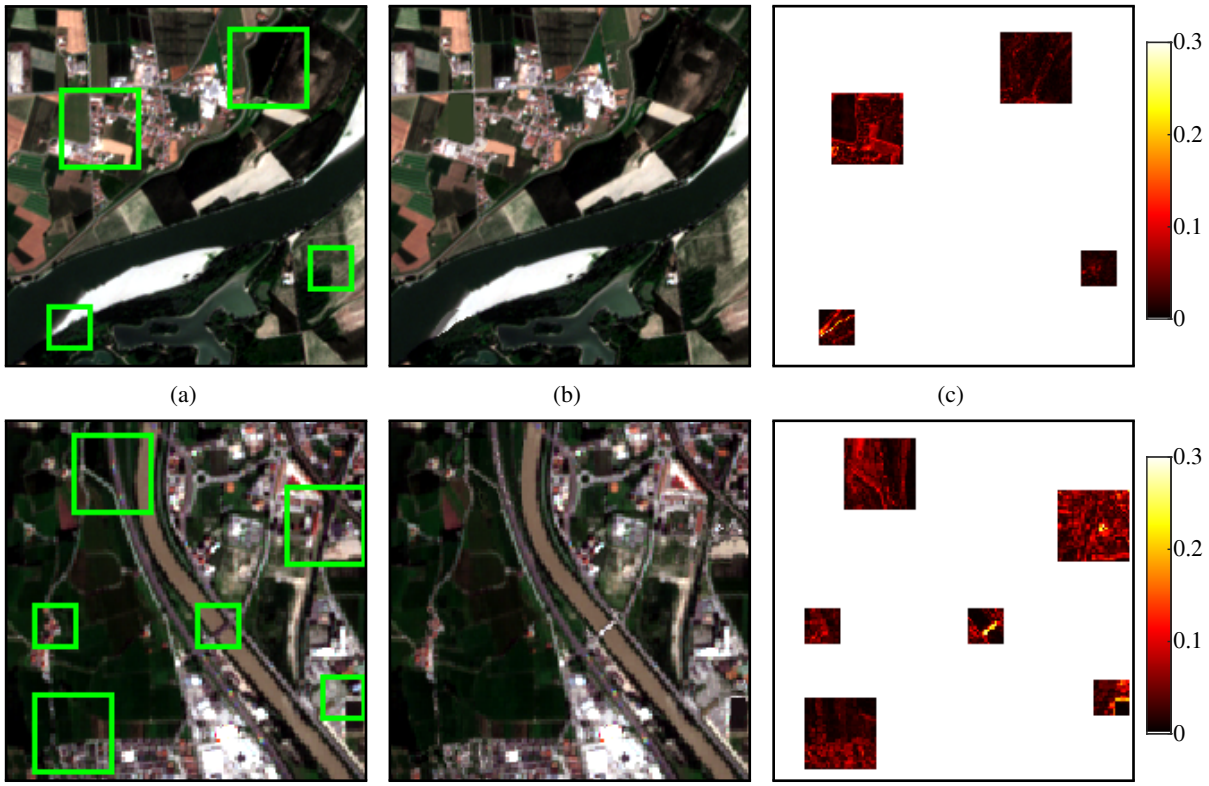

Fig. 3: True-color composites of the restoration results obtained on the two test areas, (a-c) 32TPQ and (d-f) 32TPS. (a,d) original image with synthetic gaps highlighted in green. (b,e) restored. (c,f) Restoration error, 4-band spectral Root-Mean-Square Deviation (RMSD).

TABLE III: Runtimes to restore all the four simulated gaps of size 4 000 pixels ($\approx$ 0.4 Km^2) over four spectral bands in ROI 1 (32TPQ) compared with 36 000 profiles

Method	Time (seconds)
MNSPI	16.8
PR	1.0
FCRA (proposed)	2.9

temporal patterns present in the scene that can be used to recover any missing information in the target cloudy image. The method was tested on a simulated Sentinel-2 cloudy image and compared with the PR and MNSPI methods. The results obtained demonstrate that the FCRA achieves accurate restoration results, which are very similar to the ones obtained at pixel level by the MNSPI method, while strongly reduces the computational effort. As future work, we plan to test the proposed approach on other multispectral sensor characterized by different spatial and spectral resolution (e.g., Landsat 8, Sentinel-3) and to extend the experimental analysis on more simulated clouds. Moreover, we plan to improve the scalability of the proposed approach by using search trees implementing fast approximate nearest neighbor in high dimensional spaces.

REFERENCES

[1] P. Jönsson, Z. Cai, E. Melaas, M. A. Friedl, and L. Eklundh, "A Method for Robust Estimation of Vegetation Seasonality from Landsat and Sentinel-2 Time Series Data," *Remote Sens.*, vol. 10, no. 4, p. 635, 2018.

[2] H. Shen, X. Li, Q. Cheng, C. Zeng, G. Yang, H. Li, and L. Zhang, "Missing information reconstruction of remote sensing data: A technical review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 3, no. 3, pp. 61–85, 2015.

[3] D.-C. Tseng, H.-T. Tseng, and C.-L. Chien, "Automatic cloud removal from multi-temporal SPOT images," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 584–600, nov 2008. [Online]. Available: https://doi.org/10.1016/j.amc.2008.05.050

[4] X. Zhu, F. Gao, D. Liu, and J. Chen, "A modified neighborhood similar pixel interpolator approach for removing thick clouds in landsat images," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 521–525, 2012.

[5] M. Xu, X. Jia, M. Pickering, and A. J. Plaza, "Cloud removal based on sparse representation via multitemporal dictionary learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2998–3006, 2016.

[6] V. Lonjou, C. Desjardins, O. Hagolle, B. Petrucci, T. Tremas, M. Dejus, A. Makarau, and S. Auer, "Maccs-atcor joint algorithm (maja)," in *Remote Sensing of Clouds and the Atmosphere XXI*, vol. 10001. International Society for Optics and Photonics, 2016, p. 1000107.

[7] C. Xiao, M. Liu, N. Yongwei, and Z. Dong, "Fast exact nearest patch matching for patch-based image editing and processing," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 8, pp. 1122–1134, aug 2011. [Online]. Available: https://doi.org/10.1109/tvcg.2010.226

[8] C. Xiao and M. Liu, "Efficient mean-shift clustering using gaussian kd-tree," in *Computer Graphics Forum*, vol. 29, no. 7. Wiley Online Library, 2010, pp. 2065–2073.

[9] D. Frantz, E. Haß, A. Uhl, J. Stoffels, and J. Hill, "Improvement of the fmask algorithm for sentinel-2 images: Separating clouds from bright surfaces based on parallax effects," *Remote Sens Environ*, 2018.