Patch-based Information Reconstruction of Cloud-contaminated Multitemporal Images

Chao-Hung Lin, *Member*, *IEEE*, Kang-Hua Lai, Zhi-Bin Chen and Jyun-Yuan Chen

***Abstract*—Cloud covers which are generally present in optical remote-sensing images limit the usage of acquired images and increase the difficulty of data analysis. Thus, information reconstruction of cloud-contaminated images generally plays an important role in image analysis. This study proposes a novel method to reconstruct cloud-contaminated information in multitemporal remote-sensing images. Based on the concept of utilizing temporal correlation of multitemporal images, we propose a patch-based information reconstruction algorithm that spatiotemporally segments a sequence of images into clusters containing several spatially connected components, called patches, and then clones information from cloud-free and high-similarity patches to their corresponding cloud-contaminated patches. Besides, a seam that passes through homogenous regions is used in information reconstruction to reduce radiometric inconsistency, and the information cloning is solved using an optimization process with the determined seam. These processes enable the proposed method to well reconstruct missing information. Qualitative analyses of image sequences acquired by Landsat-7 Enhanced Thematic Mapper Plus (ETM+) sensor and a quantitative analysis of simulated data with various cloud contamination conditions are conducted to evaluate the proposed method. The experimental results demonstrate the superiority of the proposed method to related methods in terms of radiometric accuracy and consistency, particularly for large clouds in a heterogeneous landscape.**

***Index Terms*—cloud removal, image reconstruction, Landsat ETM+, Poisson equation**

I. Introduction

THE primary limitation of passive remote sensing sensors is their sensitivity to weather conditions during data acquisition. Land scenes are on average approximately 35% cloud-covered globally [1], significantly reducing the availability of cloud-free surface observations. Clouds in remote-sensing images can be regarded as information for measuring liquid water, or as contaminations that partially obstruct observation of landscapes. This study addresses the latter issue in which clouds obstruct land covers, thereby resulting in missing data for passive image sensors.

Data analysis, such as classification of land covers, generally requires a cloud-free image composed of several patches that are acquired at different times and at different conditions, such as atmospheric conditions, soil moisture, and vegetation phenology. These conditions cause the relations between land-cover classes and pixel intensities to vary over a data acquisition period. Thus, the approach of replacing the cloud-contaminated pixels with their corresponding cloud-free pixels and then linearly adjusting the intensity values of the replaced pixels has been proven inappropriate when the conditions of data acquisition change significantly [2]. Lin et al. [2] proposed a nonlinear scheme instead of linear one to mathematically formulate the reconstruction problem as a Poisson equation and then solve the equation using a global optimization process. In addition, instead of reconstructing information pixel-by-pixel [3], [4], which may contain radiometric inconsistency, Lin et al. proposed a patch-based scheme to ease this inconsistency problem. Although this method can yield good cloud-free results, it is sensitive to boundary conditions when solving the Poisson equation and also sensitive to the quality of selected patch.

To address the problems above, we propose a seam determination approach to select a seam passing through homogenous regions for providing good boundary conditions in reconstruction optimization. Besides, by utilizing temporal correlation, a clustering algorithm is proposed to segment a cloud-contaminated region into several clusters with similar temporal intensity variations. This segmentation enables the proposed method to handle clouds in a heterogeneous landscape and to select suitable cloud-free pixels. Compared with the cloud-removal methods by previous studies, the proposed method can yield better cloud-free images in terms of radiometric accuracy and consistency.

II. Related Work

Information reconstruction of images has become an active research topic in the fields of remote sensing, computer vision, and computer graphics because of its practical importance. Following the categorization suggested in [2], the reconstruction methods are classified into three categories: *inpainting*-based, *multispectral*-based, and *multitemporal*-based methods. In inpainting-based methods, the information is reconstructed using the techniques of image synthesis and inpainting [5]–[7], in which the information of cloud-contaminated regions is synthesized by propagating geometric flows inside such regions. The inpainting-based methods can yield a visually plausible result that is suitable for cloud-free visualization. However, the disregard for radiometric accuracy makes the results unsuitable for data analysis. Besides, the strategy of image synthesis generally makes approaches difficult in handling large clouds and clouds in a heterogeneous landscape.

Manuscript received xx xx, 2012; revised xx xx, 2012; accepted xx xx, 2012. This paper was supported in part by the National Science Council, Taiwan (contracts NSC 101-2221-E-006-257-MY2 and 101-2119-M-006-005), and was supported in part by the Central Geological Survey, Ministry of Economic Affairs, Taiwan.

The authors are with the Department of Geomatics, National Cheng-Kung University, Tainan 701, Taiwan (e-mail: linhung@mail.ncku.edu.tw; firoy7771@gmail.com; a77527@yahoo.com.tw; slanla@gmail.com).

In multispectral-based methods, multispectral data are utilized in cloud detection and information restoration [8]–[12]. Rakwatin et al. [8] proposed to reconstruct the missing data of Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) band 6 by using the techniques of histogram matching and least-squares fitting. Histogram matching corrects the detector-to-detector striping of functional detectors, and least-squares fitting reconstructs the missing parts based on a polynomial function derived from the relation between Aqua MODIS bands 6 and 7. Similarly, Roy et al. [9] introduced a method that utilizes the information observed by MODIS to predict Landsat ETM+ images. In general, the fusing of information from different sensors is constrained by spectral compatibility and spatial resolution. Although MODIS has comparable spectral bands with ETM+, it has a coarser spatial resolution. Feng et al. [10] regarded cloud removal as a denoising problem. Based on the statistical characteristics of images, an improved homomorphism filtering is applied to filter low-frequency components that potentially present clouds. Similarly, Wang et al. [11] filtered clouds in the infrared band using wavelet frequency analysis and then reconstructed the information of cloud-contaminated pixels in the other bands using a B-spline-based surface repair approach. Zhang et al. [12] and Pringle et al. [13] also proposed geostatistical methods, in which the missing information is reconstructed by kriging or cokriging interpolation techniques. Although the methods based on the ideas of denoising and intensity interpolation can effectively reconstruct information with good results, they tend to experience difficulties in large-region reconstruction.

Compared with the inpainting-based and multispectral-based methods, the multitemporal-based methods [3], [4], [14]–[20], which rely on both temporal and spatial coherences, cope better with large region reconstruction. In [14], to solve the scan-line corrector (SLC) failure in Landsat 7 ETM+ sensor, a joint United States Geological Survey Landsat team proposed a local and linear histogram-matching method to reconstruct information of data gaps in SLC-off images. This method performs histogram matching in a moving window of each missing pixel to derive a rescaling function. This function is then used to convert the radiometric values of an input scene into equivalent radiometric values of the gap-filled scene. This method is simple and can solve numerous missing-data problems when the input images are of high quality and have comparable seasonal conditions in data acquisition. Melgani and Benabdelkader [3], [4] proposed a contextual prediction approach to determine spectrotemporal relationships between the sequences of acquired images. The spectrotemporal relationships are inferred from cloud-free regions in the neighborhood of cloud-contaminated regions over the available temporal images. Liew et al. [15] adopted a threshold-based approach to identify the best cloud-free and non-shadow pixels in a given region. A cloud-free image is then generated by stitching or mosaicking the selected cloud-free pixels. Gabarda and Cristóbal [16] introduced a cloud-removal method based on image fusion that involves a 1D pseudo-Wigner distribution transformation and a pixel-wise cloud model. This method can be regarded as a denoising method and can be used to select noise-free pixels from images. Helmer and Ruefenacht [17] utilized the regression tree to detect and predict the intensity values of cloud-contaminated pixels from other image data that are acquired at different times. They then applied an improved histogram matching to match the temporally adjacent images. Similarly, Jiao et al. [18], Wang et al. [19], and Tseng et al. [20] adjusted the intensity values of cloud-contaminated images and their corresponding cloud-free images by first using the means and the standard deviations of pixel intensity values and then applying the wavelet-based fusion method to fuse the boundaries of cloud-contaminated regions. While the methods above [14]–[20] can yield good results for homogenous regions, it should be noted that these methods based on histogram matching and data fusion tend to experience difficulties with clouds in heterogeneous landscapes. Lin et al. [2] recently proposed an information-cloning algorithm that selects cloud-free patches using a quality assessment index and solves the multipatch reconstruction problem using a global optimization process. Thus, this method can potentially yield good cloud-free results. The proposed method is inspired by such work [2]; however, there are substantial differences between our method and Lin et al.'s method. First, a seam determination approach is proposed to provide a suitable boundary condition in reconstruction optimization. Second, a spatiotemporal segmentation approach is proposed for cloud-free and high-similarity patch selection. Third, the optimization is refined to avoid the propagation of error to the cloud-contaminated region. Because of these differences, the proposed method can yield better cloud-free results in terms of radiometric accuracy and consistency.

III. Overview of Proposed Scheme

Figure 1 shows the workflow of proposed information reconstruction scheme, which consists of six main steps including *cloud detection*, *image intensity normalization*, *multitemporal image segmentation*, *image quality assessment*, *seam determination*, and *information reconstruction*. In the proposed scheme, the method proposed by Hagolle et al. [21] is first adopted to detect clouds in the input images, and then a simple user interface with the functions of selection and erasion is provided to manually refine the detected results. In the subsequent step, an image intensity normalization process is performed to achieve consistency in the intensity range of input images. Thereafter, the pixels of input images are partitioned into several clusters with similar temporal intensity variations, and then sorted according to image similarity and amount of clouds. In the seam determination step, an optimization process based on dynamic programming is performed to search for an optimal seam for each cloud-contaminated region. The proposed information reconstruction algorithm is then performed to recover the missing information of a cloud-contaminated region by solving a Poisson equation with the obtained seam as boundary conditions. The spatiotemporal segmentation is described in Section IV-C, and the seam determination and information reconstruction are described in Sections IV-E and IV-F, respectively.

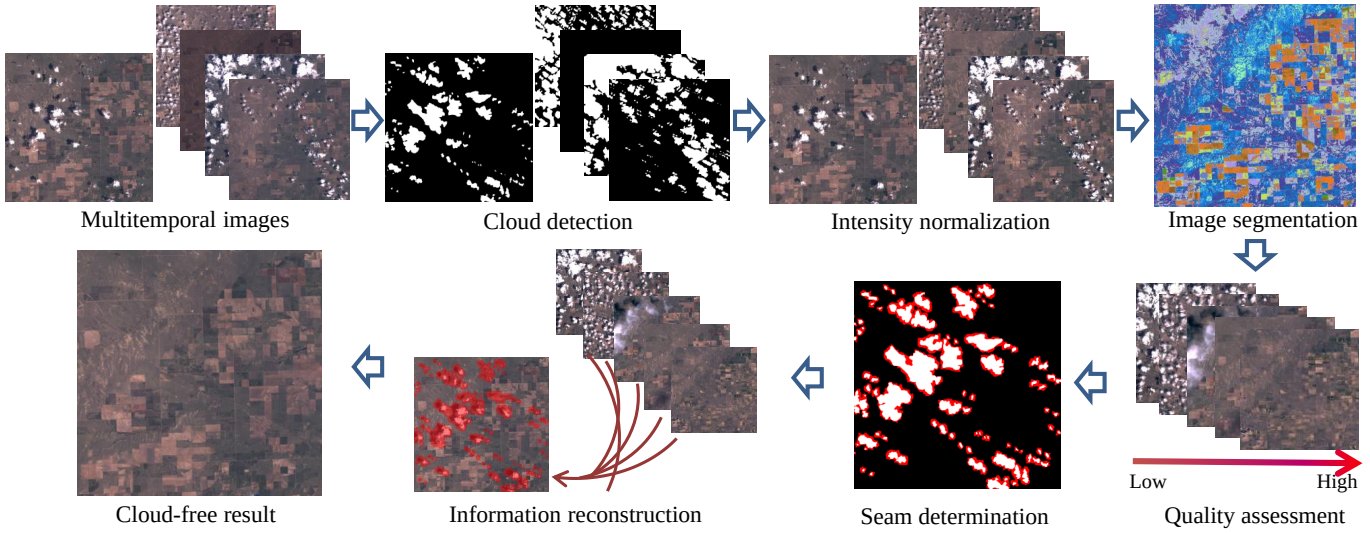


Fig. 1. Workflow of proposed information reconstruction method. The proposed method consists of six steps including cloud detection, image intensity normalization, multitemporal image segmentation, image quality assessment, seam determination, and information reconstruction. The inputs and outputs of the proposed method are cloud-contaminated and cloud-free images.

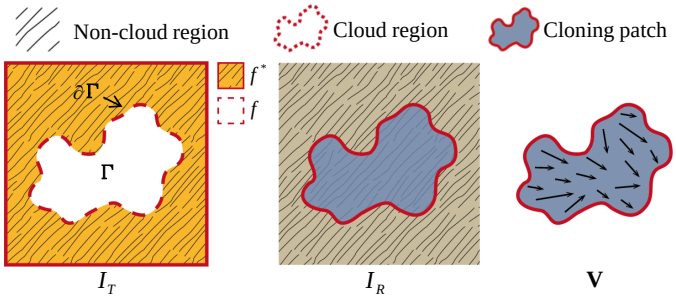


Fig. 2. Illustration of information reconstruction. The information of cloud-contaminated region $\mathbf{\Gamma}$ in target image I_T is reconstructed using a selected patch in reference image I_R with the aid of the gradient field $\mathbf{V}$ of the selected patch.

IV. Information Reconstruction of Cloud-Contaminated Regions

A. Review of Information Cloning Technique

The proposed method is inspired by the concept of image editing [22] and is an extension of the information cloning method [2]. Thus, this section begins with a brief introduction of the information cloning technique. In [2], the information reconstruction problem is mathematically formulated as a Poisson equation and then solved using global optimization. Given a cloud-contaminated image, which is called target image and denoted as I_T, and a set of corresponding images that are acquired in the same position but on different dates, which are called reference images and denoted as $\{I_{R_1}, \cdots, I_{R_n}\}$, the information of cloud-contaminated pixels in target image I_T is reconstructed by utilizing the information of its corresponding reference images $\{I_{R_1}, \cdots, I_{R_n}\}$.

As illustrated in Figure 2, the cloud-contaminated region in target image is denoted by $\mathbf{\Gamma}$, and the cloud boundary (or called seam) is denoted by $\partial\mathbf{\Gamma}$. Let f be an unknown image intensity function defined over $\mathbf{\Gamma}$, f^* be the function defined over the target image minus $\mathbf{\Gamma}$, and $\mathbf{V}$ be a vector field defined as the gradient of selected patches in reference images for guiding information reconstruction. To obtain the solution for unknown function f, the problem is formulated as an optimization equation with the boundary condition $f\,|_{\partial\mathbf{\Gamma}} = f^*\,|_{\partial\mathbf{\Gamma}}$ as follows:

$$\min_f \iint_{\mathbf{\Gamma}} \|\nabla f - \mathbf{V}\|^2 \quad \text{with} \quad f|_{\partial\mathbf{\Gamma}} = f^*|_{\partial\mathbf{\Gamma}}, \tag{1}$$

where $\nabla = \left(\frac{\partial}{\partial_x}, \frac{\partial}{\partial_y}\right)$ is gradient operator. The solution to (1) is the unique solution of the following Poisson equation with Dirichlet boundary condition:

$$\Delta f = div\,\mathbf{V} \text{over } \mathbf{\Gamma} \quad \text{with} \quad f|_{\partial\mathbf{\Gamma}} = f^*|_{\partial\mathbf{\Gamma}}, \tag{2}$$

where $\Delta = \frac{\partial^2}{\partial_{x^2}} + \frac{\partial^2}{\partial_{y^2}}$ is Laplacian operator and $div\mathbf{V} = \frac{\partial_{v_1}}{\partial_x} + \frac{\partial_{v_2}}{\partial_y}$ represents divergence of the vector field $\mathbf{V} = (v_1, v_2)$. Equations (1) and (2) are the fundamental formulas of the information reconstruction method in [2], which aim to derive the result f with a gradient that is as close to the guidance vector field $\mathbf{V}$ (i.e., the details of the selected patches in reference images) as possible under the boundary condition $f|_{\partial\mathbf{\Gamma}} = f^*|_{\partial\mathbf{\Gamma}}$. This boundary condition is used to enforce the seam $\partial\mathbf{\Gamma}$ between the scalar functions f and f^* to have radiometric consistency in reconstruction optimization. The minimization indicates that the scalar function f^* and the obtained function f have the same boundary intensity and the gradient of the obtained function f in the L_2-norm is close to the gradient field $\mathbf{V}$ of the selected patches. Therefore, this minimization has a good probability of consistently and accurately cloning the details of selected patches to the cloud-contaminated regions when a suitable boundary condition and an accurate gradient field are given. In this study, to improve the quality of information reconstruction, a spatiotemporal image segmentation approach is proposed to select cloud-free and high-quality patches and to provide an accurate vector field $\mathbf{V}$ (described in Section IV-C). Moreover, a seam

Fig. 3. Example of image intensity normalization. Left: target image; middle: reference image; right: result of image intensity normalization.

determination approach is proposed to select a seam that provide a good boundary condition in the optimization process (described in Section IV-E), and an intensity constraint is included in the optimization of information reconstruction to refine error propagations (described in Section IV-F).

B. Image Intensity Normalization

As a preprocessing step, image intensity normalization is performed to achieve consistency in the mean and standard deviation of the intensity values of target and reference images. In this process, each reference image with the mean and standard deviation of intensity values $[\mu(I_R), \sigma(I_R)]$ is linearly transformed into the target image with those of intensity values $[\mu(I_T), \sigma(I_T)]$. The normalization is formulated as follows:

$$I_R(i,j) = \Big(I_R(i,j) - \mu(I_R)\Big) \times \frac{\sigma(I_T)}{\sigma(I_R)} + \mu(I_T) \qquad (3)$$

The cloud-contaminated pixels are excluded in the calculation of Equation (3), and such image normalization process can ease the difficulty caused by inconsistent intensity dynamic ranges of the input images. As shown in Figure 3, the dynamic range of the reference image (middle figure) is altered to make it closer to that of the target image (left figure), consequently reducing the difficulty in performing the subsequent steps of information reconstruction. Note that a biased correction may happen since the cloud-contaminated pixels are excluded in the calculation of image normalization. Fortunately, the bias is insignificant for most cases since the images are acquired by the same sensor and at the same position. In addition, the bias is eased by considering the cloud cover rate in the cloud-free patch selection (Section IV-D), and the bias is also eased by the Dirichlet boundary condition $f|_{\partial\mathbf{\Gamma}} = f^*|_{\partial\mathbf{\Gamma}}$ in Poisson equation.

C. Spatiotemporal Segmentation

To select a suitable guidance vector field (i.e., $\mathbf{V}$ in (1)) and to handle cloud covers in a heterogeneous landscape, a spatiotemporal segmentation is proposed to segment the input images into several patches with similar temporal intensity variations. As shown in Figure 4, a temporal variation map M_{var} is generated first according to pixel intensity variations during a defined period. The temporal variation of pixel (i,j) is defined as the average intensity variation in the sequence of target and reference images and is formulated as follows:

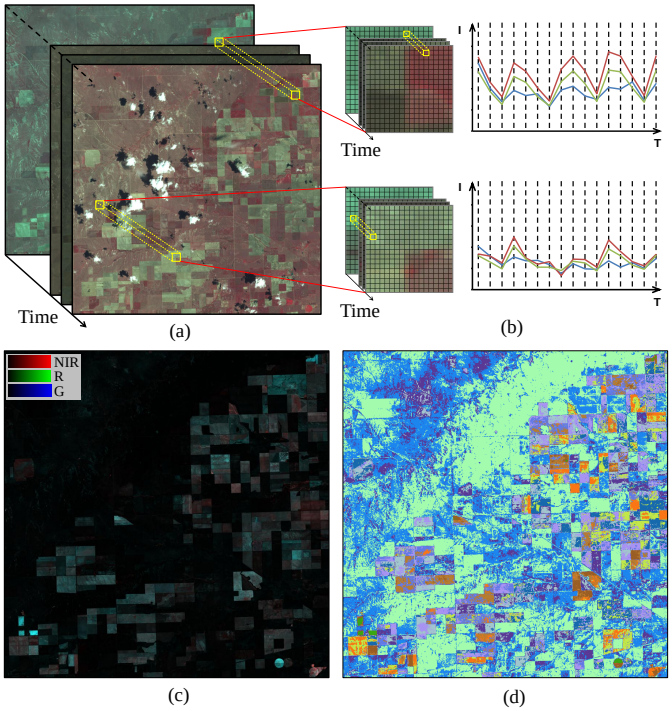


Fig. 4. Illustration of spatiotemporal segmentation. (a) Input multitemporal images displayed using false color (NIR-Red-Green). (b) Temporal variation of pixel intensity values. (c) Generated temporal variation map. The variations in the near infrared, red, and green spectral bans are visualized by RGB colors. (d) Segmentation result using $k = 12$. Each cluster is visualized by a color.

$$M_{var}(i,j) = \sum_{k=1}^{R_n} \left(\frac{\Big(I_{k+1}(i,j) - I_k(i,j)\Big)^2}{NumDay(I_{k+1}, I_k)} \right), \qquad (4)$$

where $NumDay(I_{k+1}, I_k)$ is a function that returns the number of days between the acquisition dates of two consecutive images I_{k+1} and I_k that are cloud-free at the position (i,j), and $R_n + 1$ images, including the reference and target images, are used in the calculation. Note that the cloud-contaminated pixels are excluded in the calculation of Equation (4) since their information is missing, and the intensity variation is normalized by the number of days between the acquisition dates instead of the number of available images because of the inconsistent positions of cloud-free pixels in temporal domain.

In the subsequent step, the temporal variation map is partitioned into several clusters with similar temporal variations, indicating that each cluster probably belongs to a landscape. Any clustering technique can be used here. For simplicity, the efficient and commonly used k-means clustering method [23] is adopted. With an initial set of k means, the clustering method adopts an iterative refinement strategy to group the data into k clusters $\{s_1, \cdots, s_k\}$ by minimizing the within-cluster sum of squares as follows:

$$\arg\min_{s} \sum_{i=1}^{k} \sum_{x_j \in s_i} \|x_j - u_i\|^2, \qquad (5)$$

where u_i is the mean of pixel intensity values in class s_i.

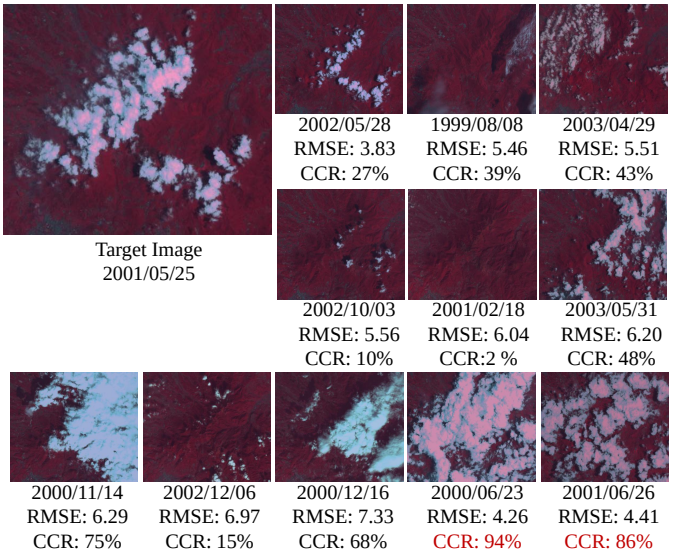


Fig. 5. Result of quality assessment. Left: reference images sorted according to RMSE and CCR. The images captured on June 23, 2000 and June 26, 2001 are filtered because of the CCR constraint.

Figure 4 shows a result of temporal variation map generation and clustering. Obviously, the croplands have a high temporal variation and the mountain areas have a low temporal variation, meeting the expectation. In addition, different landscapes are separated into different groups after applying the clustering algorithm to the variation map, enabling the proposed method to handle clouds in a heterogeneous landscape. Note that always obtaining accurate clustering is difficult and incorrect clustering may happen. Fortunately, the proposed method does not require a perfect clustering and the reconstruction quality is moderately sensitive to the clustering accuracy. It is because that the clustering result is used only in the selection of cloud-free patches and guidance vector field.

D. Image Quality Assessment

After the clustering, the pixels of target and reference images are grouped into several spatial fragmented clusters, and each cluster contains several connected components, called patches. For each patch within the cloud-contaminated region, a cloud-free high-similarity patch is selected from the reference image by utilizing quality assessment. In this study, the root mean square error (RMSE) is used to estimate patch quality and to select a cloning patch from the reference images. The RMSE is defined as $\sqrt{\sum_{i=1}^{m}\left(p_i^{I_T}-p_i^{I_R}\right)^2/m}$, where m is the total number of pixels, and $p_i^{I_T}$ and $p_i^{I_R}$ are the intensity values of i-th pixel in the target image I_T and the reference image I_R, respectively. To consider heterogeneous landscapes, the RMSE index between the target and reference images is calculated for only the pixels that belong to a cluster. In this manner, a suitable cloning patch can be selected by using only the pixels that belong to the cluster of this patch in quality estimation. In addition, the cloud cover rate (CCR) is considered in the patch selection. If the CCR is greater than a defined threshold (set to 80% for all experiments), the reference image will not be selected as a candidate in information reconstruction. As shown in Figure 5, the reference image acquired on May 28, 2002, which has the best quality, is selected as the information cloning patch. The CCRs of the reference images acquired on June 23, 2000 and June 26, 2001 are both greater than the defined threshold, and thus, these two images are not selected as candidates.

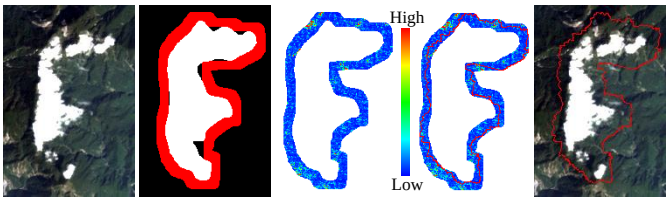


Fig. 6. Illustration of seam determination. From left to right: original image containing clouds, search space marked by red, cost map visualized by color ranging from blue (lowest cost) to red (highest cost), and obtained seam (red curve) displayed in the cost map and the original image.

E. Seam Determination

The boundary condition in Equation (1) significantly affects the quality of inward interpolation as well as information reconstruction. In this study, instead of using the detected cloud boundary [2], a seam that passes though homogenous regions is used. It is because non-homogeneous regions containing high-gradient content generally have low consistency in pixel intensity and gradient of multi-temporal images. Using a seam passing through nonhomogeneous regions as boundary conditions in reconstruction may result in inaccurate inward interpolation and information reconstruction.

The proposed seam determination approach consists of three steps: *search space determination*, *cost map generation*, and *optimal seam determination*, as shown in Figure 6. In addition to the enforcement of passing though homogenous regions, the seam must be close to the cloud boundary. Moreover, the extreme concaves are excluded from the search space to simplify seam determination. To meet the requirements above, the morphological operations, namely, dilation and erosion, are utilized to define a search space without extreme concaves. The search space is determined by first applying the dilation operator to the cloud-contaminated region, and then applying the erosion operator to the dilated region using a 3×3 structure element. The eroded region is defined as the search space. The number of dilations and erosions is set to 10, indicating that the search space has a width of approximately 10 pixels. Cloud-contaminated pixels are excluded from the search space, and the cloning patches for the search space are selected using the quality assessment mentioned in Section IV-D. The search space is then embedded into a regular, rectangular space to facilitate the calculation of seam optimization. Such embedding is achieved by iteratively stretching and elongating the eroded one-pixel-width circles to fit a rectangle, as shown in Figure 7. In the stretch step, the eroded one-pixel-width circles are stretched to lines with a cutting. In the elongation, the stretched lines are elongated to fit a rectangle, and thus, several empty pixels exist. These empty pixels are filled by

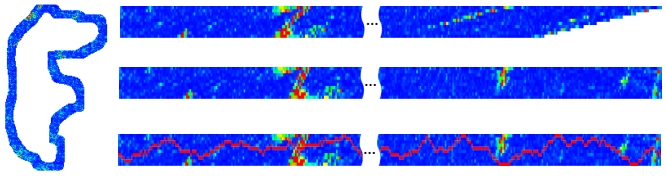


Fig. 7. Illustration of search space embedding. Left: defined search space in image space. Right: results of one-pixel-width circles stretching (top), elongating (middle), and seam determination (bottom).

linearly interpolating their neighboring pixels in the one-pixel-width circle. A result of search space embedding is shown in the middle figure of Figure 7. The regular search space facilitates the calculation of seam determination.

In the second step, a cost map defined over the search space is generated for seam optimization. The seam is proposed to pass through homogenous regions in the target image and to pass regions that have similar gradient magnitudes in the target and reference images. Thus, the cost map is defined as follows:

$$\text{Cost}M = \sum_{(i,j)\in\omega} \Big(\|\nabla I_T(i,j)\|^2 + \|\nabla I_T(i,j) - \nabla I_R(i,j)\|^2 \Big), \quad (6)$$

where $\nabla I_T(i,j)$ and $\nabla I_R(i,j)$ represent the gradients of the target and reference images, respectively, at position (i,j) and ω is the search space defined in the first step. In (6), the first part (i.e., $\|\nabla I_T(i,j)\|^2$) can force the seam to pass through homogenous regions, and the second part (i.e., $\|\nabla I_T(i,j) - \nabla I_R(i,j)\|^2$) can force the seam to pass regions that have similar gradient magnitude in the target and reference images.

In the last step, the optimal seam is determined using dynamic programming with the aid of cost map. By viewing pixel intensity values in the cost map as the costs for a seam to pass through, the minimum cost seam passing through the low-cost regions is obtained by implementing the following procedures. The cumulative minimum cost function $C_{min}(i,j)$ is defined as the cost to reach pixel (i,j) in the search space, and the cumulative minimum error C_{min} for all possible seams are calculated as follows:

$$C_{\text{min}}(i,j) = Cost(i,j) + Min\Big(C_{\text{min}}(i-1,j+1), C_{\text{min}}(i-1,j), C_{\text{min}}(i-1,j-1) \Big) \quad (7)$$

The minimum value of the last row in C_{min} indicates the end of the minimum cost seam. Thus, the optimal seam and the best boundary condition in information reconstruction can be obtained by tracing back to the first row. The obtained seam shown in Figure 7 is the global optimal seam that passes through the low-cost pixels.

F. Information Reconstruction

While the method proposed by Lin et al. [2] can yield good results for most cases, it should be noted that the process of interpolating inward prorogate errors from the boundaries to the cloud-contaminated region may lead to an unnatural result, especially for a reconstruction with a high-cost seam. To solve such problem, a pixel intensity constraint is included in the optimization to balance fitting the guidance vector field $\mathbf{V}$ and linearly replacing pixel intensity according to that of the selected cloning patch. Including the intensity constraint, the optimization in Equation (1) is reformulated as follows:

$$\min_f \iint_{\mathbf{\Gamma}} \Big(\|\nabla f - \mathbf{V}\|^2 + (f - f'_{cp})^2 \ \text{with} f|_{\partial\mathbf{\Gamma}} = f^*|_{\partial\mathbf{\Gamma}} \Big), \quad (8)$$

where f_{cp} is the intensity function defined over the selected cloning patch and f'_{cp} is the intensity adjustment of f_{cp} , which can be calculated using Equation (3). In the implementation, Equation (8) is discretized on a pixel grid. Let N_p be the pixel set of the 4-connected neighbors for a pixel p in target image I_T and $< p, q >$ be denoted as a pixel pair such that pixel q is one of the 4-connected neighbors of pixel p (i.e., $q \in N_p$). Let $f(p)$ be the value of the image function f at position/pixel p. The task is to compute the pixel intensity values in the cloud-contaminated region $\mathbf{\Gamma}$. The discretization of Equation (8) yields the following discrete optimization equation:

$$\min_f \sum_{<p,q>\cap\mathbf{\Gamma}\neq 0} \Big| \big(f(p) - f(q) - v_{pq}\big)^2 + \big(f(p) - f'(p)\big)^2 \Big| \quad \text{with} f(s) = f^*(s) \text{for all} s \in \partial\mathbf{\Gamma}, \quad (9)$$

where $v_{pq} = I_R(p) - I_R(q)$, which is the directional gradient of the reference image at position p. According to Equation (9), the following equation can be generalized:

$$|N_p| f(p) - \sum_{q\in N_p\cap\mathbf{\Gamma}} f(q) = \sum_{q\in N_p\cap\partial\mathbf{\Gamma}} f^*(q) + \sum_{q\in N_p} v_{pq} - |f(p) - f'(p)| \quad \text{for all} \quad p \in \mathbf{\Gamma}, \quad (10)$$

where $|N_p|$ is the number of neighbors q in N_p. Equation (10) is iteratively solved until the unknown function f is converged. As mentioned above, several patches in the reference images are used to reconstruct the information of a cloud-contaminated region. To generate a smooth guidance field, the guidance vectors on the patch boundary are calculated by averaging the gradients of neighboring patches as follows:

$$\mathbf{v}_{pq} = \Big| \big(P_{R_i}(p) - P_{R_i}(q)\big) + \big(P_{R_j}(p) - P_{R_j}(q)\big) \Big| / 2, \quad (11)$$

where P_{R_i} and P_{R_j} represent the neighboring patches in the reference images I_{R_i} and I_{R_j} , respectively.

V. Experimental Results and Discussion

The proposed algorithms were tested on a desktop PC with 3.0 GHz CPU and 4GB memory. For a dataset containing twenty reference images of 900x900 resolution and one target image with 100000 cloud-contaminated pixels, the average

TABLE I
LANDSAT ETM+ ACQUISITIONS FOR THE STUDY SITES.

Site, country	Site location (latitude, longitude)	Path-row B coordinates	Acquisition date	CCR (%)
Colorado, USA	37.442N,102.807E	32/34	1999/07 - 2003/05	6%
Tel Aviv, Israel	31.952N, 35.892E	174/38	1999/11 - 2003/05	25%
Taichung, Taiwan	24.195N, 120.658E	118/43	2003/01- 2003/03	10%
Taipei, Taiwan	25.035N,121.494E	117/43	1999/08 - 2003/05	28%

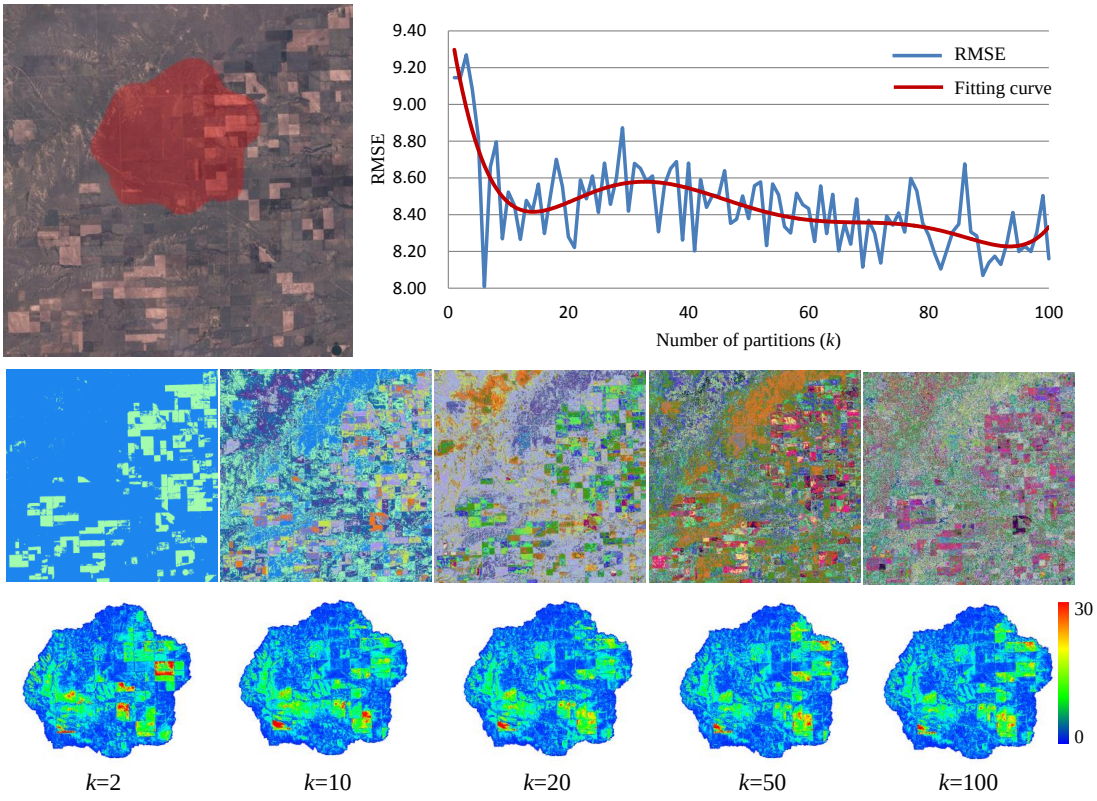


Fig. 8. Clustering result. Top: target image with a simulated cloud (left) and RMSE of the reconstruction results using various values of parameter k in clustering (right). The red curve is the polynomial fitting curve. Middle: clustering results using different values of k, including k = 2, 10, 20, 50, and 100 (from left to right). The partitioned clusters are visualized using different colors. Bottom: maps of reconstruction errors. The error is visualized using colors ranging from blue (lowest error; RMSE = 0) to red (highest error; RMSE = 30).

computation time for image intensity normalization is 1.2 seconds, for multitemporal image segmentation is 340.5 seconds, for seam determination is 1.9 seconds, and for reconstruction optimization is 256.2 seconds. In the experiments, Landsat-7 ETM+ images with various landscapes were used to test the feasibility and performance of the proposed method. For simplicity, only the 30-meter blue band (0.45-0.52 micron), green band (0.53-0.61 micron), red band (0.63-0.69 micron), and near-infrared band (0.78-0.90 micron) of ETM+ images were used only; the mid-infrared, thermal infrared, and panchromatic bands were not used. Four study sites, Colorado, Tel Aviv, Taichung, and Taipei, were selected, as shown in Table 1. The Colorado and Taichung Landsat acquisitions that have only a few cloud covers are suitable for analyzing parameter setting and conducting quantitative analyses. The Taipei and Tel Aviv Landsat acquisitions that contain approximately 28% cloud cover on average are suitable for testing the feasibility of the proposed method. In the Taipei and Tel Aviv Landsat acquisitions, the clouds and cloud shadows were identified using the unsupervised detection method proposed by Hagolle et al. [21]. Then, the detection was manually refined using a simple user interface with selection and erasion operations. The RMSE and average difference (AD) were used to estimate the quality of information reconstruction. The RMSE is defined as $\sqrt{\sum_{i=1}^{N}(f_i^{\text{Rec}} - f_i^{\text{Ori}})^2/N}$ where N is the total number of simulated cloud-contaminated pixels, and f_i^{Rec} and f_i^{Ori} are the reconstruction result and actual value (i.e., ground truth) of i-th pixel, respectively. A larger RMSE indicates a larger reconstruction error. The AD index is defined as $\sum_{i=1}^{N}(f_i^{\text{Rec}} - f_i^{\text{Ori}})\Big/N$ and is used to evaluate either the underestimation or overestimation of information reconstruction. A positive AD index indicates overestimation

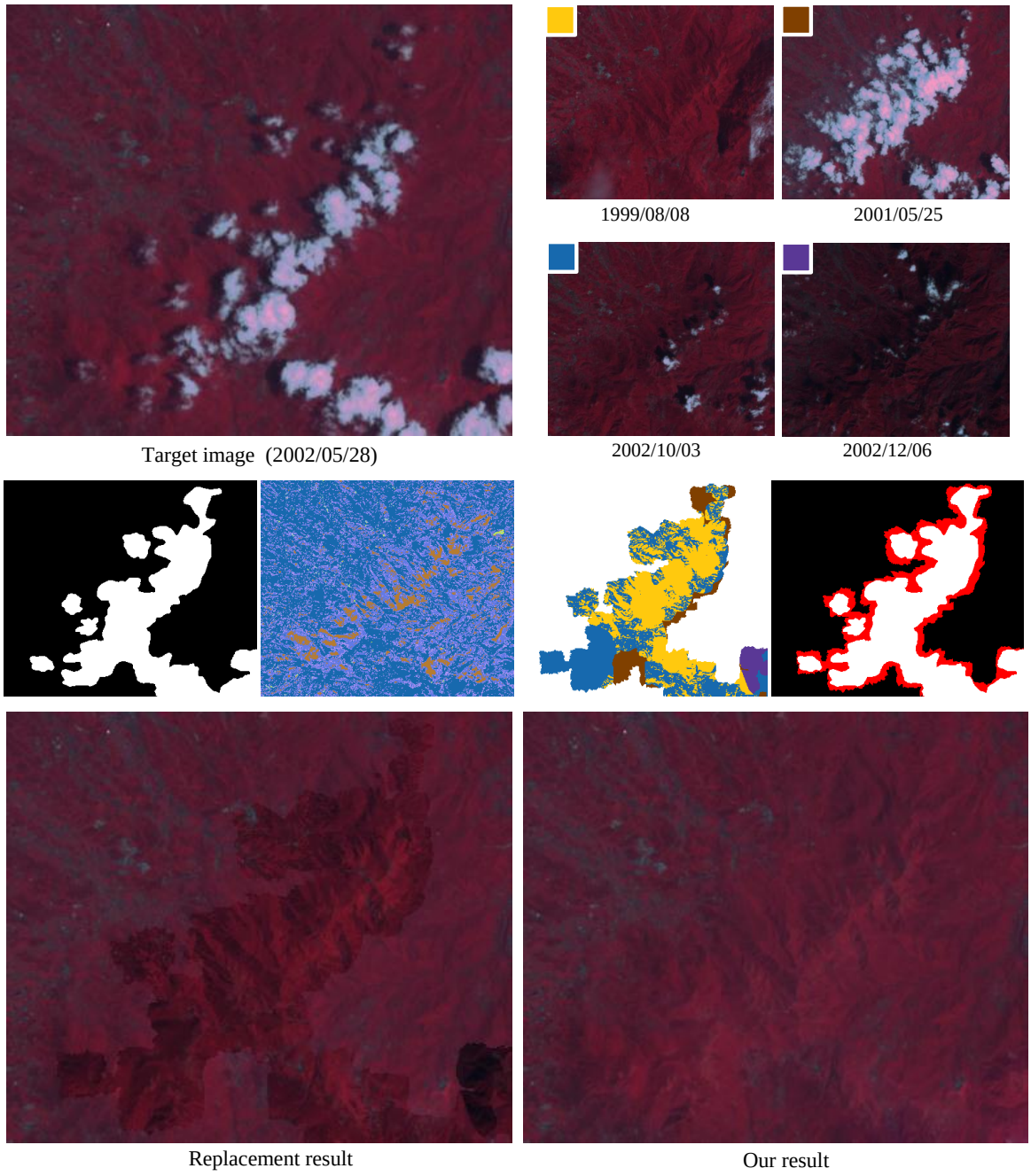


Fig. 9. Results of information reconstruction. Top: target image (left) acquired near Taipei City on May 28, 2002 and reference images (right). Middle: results of cloud detection, spatiotemporal segmentation, cloning patch selection (each reference image is represented by a color), and obtained seam. Bottom: patch replacement result and our result.

TABLE II
ACCURACY OF RECONSTRUCTION RESULTS GENERATED USING THE PROPOSED SCHEME WITH THE CLOUD BOUNDARY (METHOD A) AND THE OPTIMAL SEAM (METHOD A') AS BOUNDARY CONDITIONS. THE TEST DATA IS SHOWN IN FIGURE 12.

Quality Index		Red band		Green band		Blue band		NIR band	
		RMSE	AD	RMSE	AD	RMSE	AD	RMSE	AD
R_A	MethodA	4.22	-1.62	2.69	-1.14	2.44	-0.90	2.03	-0.24
	MethodA'	4.04	-0.87	2.58	-0.83	2.40	-0.79	1.94	0.01
R_B	MethodA	13.75	1.81	8.49	0.60	7.37	0.65	7.86	-2.57
	MethodA'	13.03	1.62	8.03	0.31	6.97	0.47	7.59	-2.47

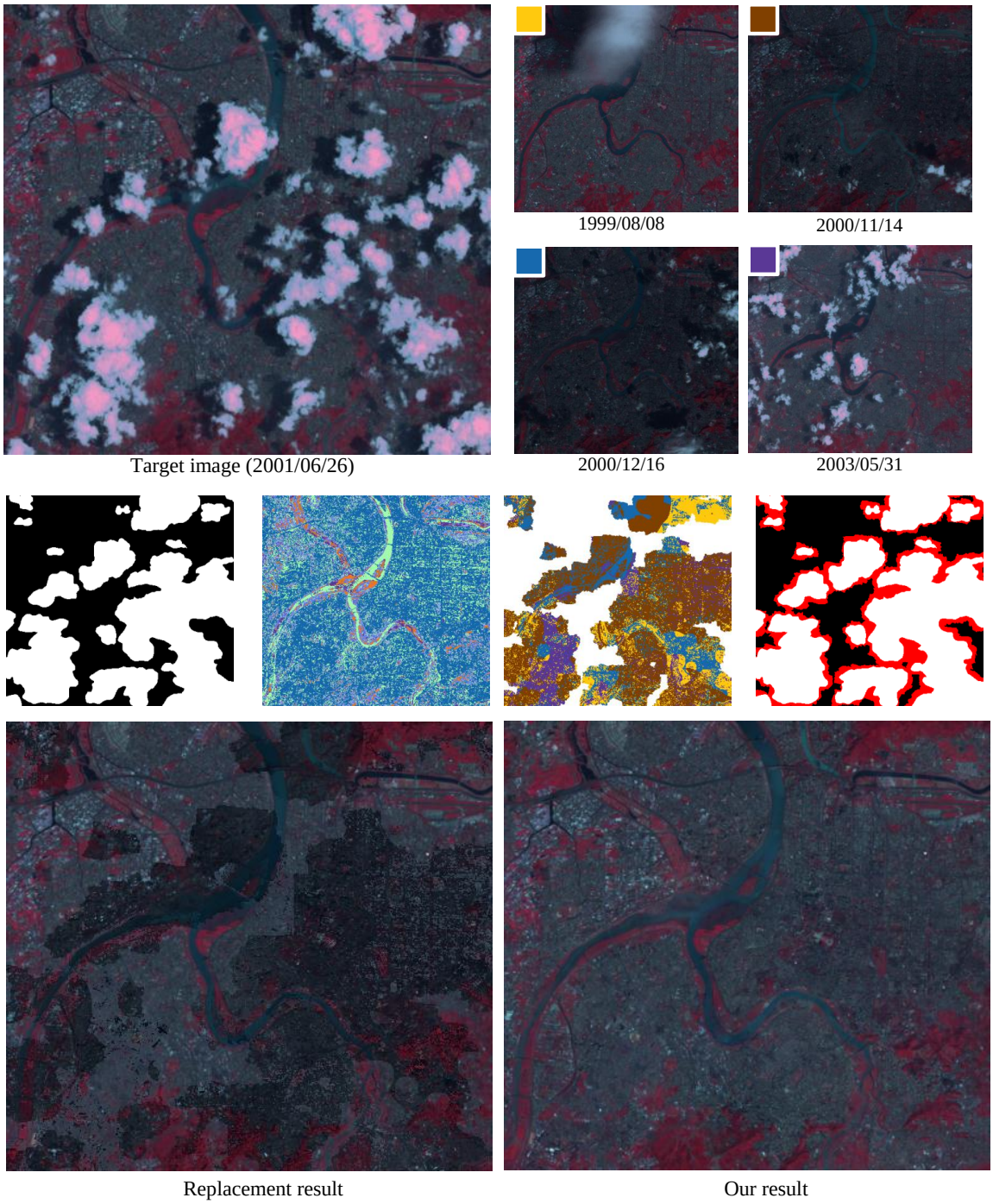


Fig. 10. Results of information reconstruction. Top: target image (left) acquired near Taipei City on June 26, 2001 and reference images (right). Middle: results of cloud detection, spatiotemporal segmentation, cloning patch selection, and obtained seam. Bottom: patch replacement result and our result.

TABLE III
ACCURACY OF RECONSTRUCTION RESULTS GENERATED USING THE PROPOSED SCHEME WITHOUT (METHOD B) AND WITH (METHOD B') SPATIOTEMPORAL SEGMENTATION (k = 20). THE TEST DATA IS SHOWN IN FIGURE 8.

Quality Index	Red band		Green band		Bule band		NIR band	
	RMSE	AD	RMSE	AD	RMSE	AD	RMSE	AD
MethodB	21.01	-0.90	13.34	-0.39	10.58	-0.35	9.22	2.33
MethodB'	13.62	0.58	9.02	0.57	6.91	-0.04	8.97	1.71

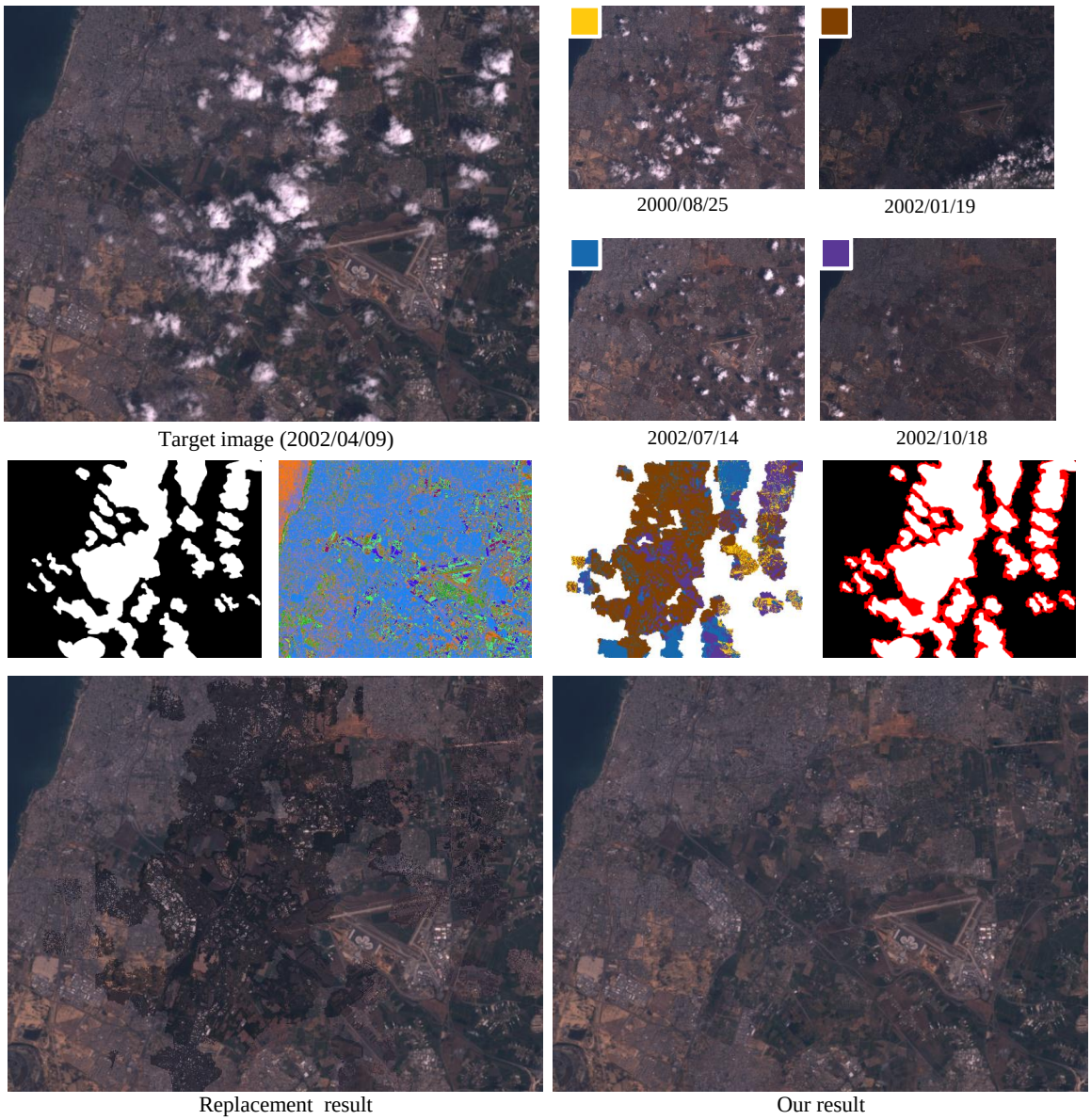


Fig. 11. Results of information reconstruction. Top: target image (left) acquired near Tel Aviv on April 09, 2002 and reference images (right). Middle: results of cloud detection, spatiotemporal segmentation, cloning patch selection, and seam determination. Bottom: patch replacement result and our result.

and a negative value indicates underestimation.

A. Parameter setting

The number of groups used in the k-means clustering is the only parameter that needs to be set in the proposed method. To test the sensitivity of the reconstruction results to such parameter and to find a suitable value, our method was tested using various parameter values on a dataset that contains a simulated cloud, as shown in Figure 8. Parameter k is determined by comparing the reconstruction results that use variousis determined by comparing the reconstruction results that use various k values. In the experiment, RMSE was used to estimate the reconstruction quality. From the RMSE statistics and the polynomial fitting curve (red curve) shown in Figure 8, a small parameter k (i.e., undersegmentation) decreases reconstruction accuracy; however, a larger parameter k (i.e., oversegmentation) does not necessarily yield better results. The RMSE curve is generally convergent when k is greater than 20, indicating that indicating that $k = 20$ is a suitable setting for most cases.

B. Information reconstruction results

Figures 9, 10, and 11 illustrate how the proposed method works for reconstruction of cloud-contaminated images. In Figure 9, the target image containing mountains was acquired near Taipei City on May 28, 2002, and four reference images whose acquisition dates are close to that of the target image are selected to reconstruct information. The results of cloud detection, spatiotemporal segmentation, cloning patch selection, and seam determination are shown in the middle section of Figure 9. Compared with the patch replacement, our method can seamlessly clone the details of multiple patches to

the corresponding cloud-contaminated regions, indicating that the radiometric inconsistency is greatly eased. In Figures 10 and 11, the target images that contain urban areas and rivers were acquired near Taipei City on June 26, 2001 and near Tel Aviv on April 09, 2002, respectively. The ability of radiometric consistent reconstruction was also demonstrated in these two examples. Note that to select thin clouds which sometimes occur on the cloud boundaries, a dilation operation with 5 structure element is applied to the detected clouds. Therefore, the selected cloud-contaminated regions in Figures 9, 10, and 11 are partially larger than the real cloud-contaminated regions. This problem might be solved using the state-of-the-art cloud detection method.

C. Performance of the proposed approaches

In this study, a seam optimization approach is proposed to select boundary conditions for information reconstruction and a spatiotemporal segmentation algorithm is proposed to select cloud-free cloning patches. To demonstrate the performance of these two approaches, the proposed scheme without (denoted as Method A) and with (Method A') the optimal seam, and the proposed scheme without (Method B) and with (Method B') the spatiotemporal segmentation were compared. Colorado Landsat acquisitions that contain a simulated cloud were tested for the comparisons. In this experiment, a cloud is simulated by partly obscuring a cloud-free image of the acquisitions, as shown in Figures 8 and 12, and the reconstructed images are compared with the original cloud-free image by using RMSE and AD measurements. The comparison results are shown in Tables II and III. From the RMSE and AD indices, these two approaches improve reconstruction quality because of the use of better cloning seam and cloning patches. Based on the visual comparisons, these two approaches also improve radiometric consistency. In Table II, as expected, the quality of homogeneous region (denoted by R_A) reconstruction was better than that of the non-homogenous region (denoted by R_B) because the non-homogenous region contains larger temporal variations.

D. Information reconstruction of simulated cloud-contaminated images

Reconstruction of simulated cloud-contaminated pixels was conducted to quantify reconstruction accuracy and to compare our method with the related methods. In this experiment, an image that contains three simulated regions with different cloud contaminations was tested, as shown in Figure 13. Several related methods, including patch replacement, intensity adjustment suggested in [15]–[20], histogram matching [14], information cloning method proposed by Lin et al. [2], and the proposed method were evaluated. To ensure fair comparison, only one reference image was used, and thus, the process of spatiotemporal segmentation is omitted from our method. In the method of intensity adjustment and histogram matching, information reconstruction is achieved by adjusting pixel intensity values according to a transformation function that is determined using mean and standard deviation of intensity values or histogram matching. However, discontinuity may occur

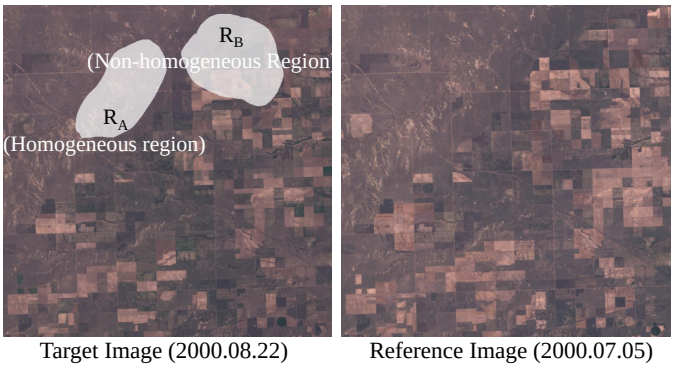


Fig. 12. Left: target image that contains two simulated cloud-contaminated regions, namely, homogeneous region R_A and non-homogeneous region R_B. The acquisition date is August 22, 2000. Right: reference image. The acquisition date is July 5, 2000.

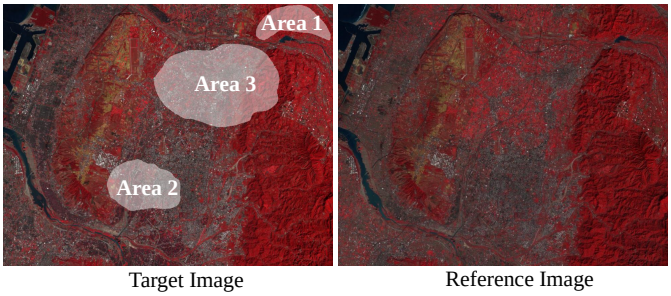


Fig. 13. Left: target image that contains three simulated regions with different cloud contaminations. Area 1 contains mountains; Area 2 contains mountains and urban areas; Area 3 contains mountains, urban areas, and croplands. The acquisition date is March 3, 2003. Right: reference image. The acquisition date is January 30, 2003.

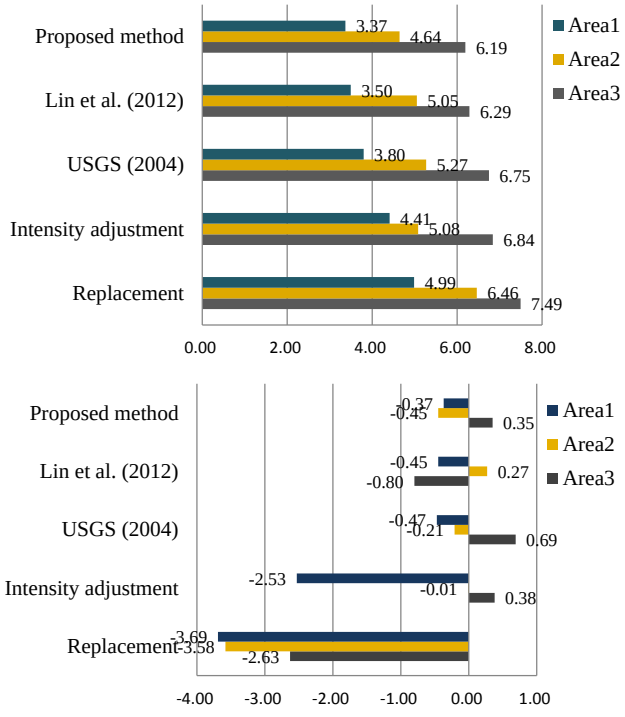


Fig. 14. Accuracy of reconstruction results generated using the methods of patch replacement, intensity adjustment, method proposed by USGS [14], Lin et al.'s method [2], and our method without spatiotemporal segmentation. The test data is shown in Figure 13.

on the cloud boundaries even though a smoothing process is applied. In [2], the radiometric inconsistency is eased by error propagation. However, the reconstruction quality is sensitive to boundary conditions and selected cloning patches. In this study, seam optimization and spatiotemporal segmentation are adopted to optimize information reconstruction. Therefore, as can be seen in Figures 14, our result is closer to the actual image, i.e., the ground truth, compared with the results generated by related methods, indicating that better reconstructions are obtained.

VI. Conclusions and Future Work

In this study, a novel information reconstruction method for cloud-contaminated images was introduced. Using temporal correlation, multi-temporal images were segmented into several patches that have similar temporal variations. Patches in the reference images were then sorted using the RMSE index to select cloning patches, and information of selected patches was seamlessly cloned to corresponding cloud-contaminated patches. The multi-patch information reconstruction was solved using an optimization process with the optimal seam. These processes enable the proposed method to well reconstruct information of cloud-contaminated regions. The major improvement is that our method makes better use of appropriate spatiotemporal information to reconstruct information, and thus, our method can potentially yield better results in terms of radiometric accuracy and consistency compared with related methods. Experiments on the sequences of multi-temporal images acquired using a Landsat-7 ETM+ sensor demonstrate that the proposed method can process clouds in various landscapes. In addition, quantitative and qualitative analyses on a simulated data with different cloud contamination conditions show that the method is superior to related cloud-removal methods. However, the proposed method cannot accurately reconstruct information when land covers change significantly over a short period, which is a limitation of multitemporal-based methods. In the future, a more comprehensive approach that fills the gaps in Landsat-7 ETM+ SLC-off and cloud-contaminated images, and automatic cloud detection and reconstruction approaches for thin clouds are being planned for development.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions. The authors would also like to thank the U.S. Geological Survey (USGS) Center for providing images.

References

[1] J. Ju and D. P. Roy, "The availability of cloud-free landsat etm+ data over the conterminous united states and globally," *Remote Sensing of Environment*, vol. 112, no. 3, pp. 1196–1211, 2008.

[2] C. Lin, P. Tsai, K. Lai, and J. Chen, "Cloud removal from multitemporal satellite images using information cloning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 1, pp. 232–241, Jan. 2013.

[3] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 2, pp. 442–455, Feb. 2006.

[4] S. Benabdelkader and F. Melgani, "Contextual spatiospectral postreconstruction of cloud-contaminated images," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 204–208, Apr. 2008.

[5] F. Chen, Z. Zhao, L. Peng, and D. Yan, "Clouds and cloud shadows removal from high-resolution remote sensing images," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 6, pp. 4256–4259, 2005.

[6] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 2, pp. 442–455, Feb. 2006.

[7] L. Lorenzi, F. Melgani, and G. Mercier, "Inpainting strategies for reconstruction of missing data in vhr images," *Geoscience and Remote Sensing Letters, IEEE*, vol. 8, no. 99, pp. 914–918, 2011.

[8] P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Restoration of aqua modis band 6 using histogram matching and local least squares fitting," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 2, pp. 613–627, 2009.

[9] D. Roy, J. Ju, P. Lewis, C. Schaaf, F. Gao, M. Hansen, and E. Lindquist, "Multi-temporal modis–landsat data fusion for relative radiometric normalization, gap filling, and prediction of landsat data," *Remote Sensing of Environment*, vol. 112, no. 6, pp. 3112–3130, 2008.

[10] F. Chun, M. Jian-wen, D. Qin, and C. Xue, "An improved method for cloud removal in aster data change detection," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 5, 2004, pp. 3387–3389.

[11] Z. Wang, J. Jin, J. Liang, K. Yan, and Q. Peng, "A new cloud removal algorithm for multi-spectral images," in *Proceeding of SAR and Multispectral Image Processing (MIPPR)*, vol. 6043, 2005, pp. 230–240.

[12] C. Zhang, W. Li, and D. Travis, "Restoration of clouded pixels in multispectral remotely sensed imagery with cokriging," *International Journal of Remote Sensing*, vol. 30, no. 9, pp. 2173–2195, 2009.

[13] M. Pringle, M. Schmidt, and J. Muir, "Geostatistical interpolation of slc-off landsat etm+ images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, no. 6, pp. 654–664, 2009.

[14] (2004) Phase 2 gap-fill algorithm: Slc-off gap-filled products gap-fill algorithm methodology. USGS. [Online]. Available: http://landsat.usgs.gov/documents/L7SLCGapFilledMethod.pdf

[15] M. Li, S. Liew, and L. Kwoh, "Producing cloud free and cloud-shadow free mosaic from cloudy ikonos images," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 6, 2003, pp. 3946–3948.

[16] S. Gabarda and G. Cristóbal, "Cloud covering denoising through image fusion," *Image and Vision Computing*, vol. 25, no. 5, pp. 523–530, 2007.

[17] E. Helmer and B. Ruefenacht, "Cloud-free satellite image mosaics with regression trees and histogram matching," *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 9, pp. 1079–1089, 2005.

[18] Q. Jiao, W. Luo, X. Liu, and B. Zhang, "Information reconstruction in the cloud removing area based on multi-temporal chris images," in *Proceeding of Remote Sensing and GIS Data Processing and Applications*, vol. 6790, 2007, p. 679029.

[19] B. Wang, A. Ono, K. Muramatsu, and N. Fujiwara, "Automated detection and removal of clouds and their shadows from landsat tm images," *IEICE transactions on information and systems*, vol. 82, no. 2, pp. 453–460, 1999.

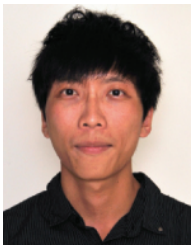
[20] D.-C. Tseng, H.-T. Tseng, and C.-L. Chien, "Automatic cloud removal from multi-temporal spot images," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 584–600, 2008.

[21] O. Hagolle, M. Huc, D. Pascual, and G. Dedieu, "A multi-temporal method for cloud detection, applied to formosat-2, venμs, landsat and sentinel-2 images," *Remote Sensing of Environment*, vol. 114, no. 8, pp. 1747–1755, 2010.

[22] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Transactions on Graphics*, vol. 22, no. 3, pp. 313–318, 2003.

[23] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1967, pp. 281–297.

Chao-Hung Lin (M'12) received his MS and PhD degree in computer science and information engineering from National Cheng-Kung University, Taiwan in 1998 and 2004, respectively. Since 2006, he has been a member of the faculty of the Department of Geomatics at National Cheng-Kung University. He currently is an associate professor. He leads the Digital Geometry Processing Laboratory (http://dgl.geomatics.ncku.edu.tw) and co-leads the Computer Graphics Laboratory, National Cheng-Kung University (http://graphics.csie.ncku.edu.tw). His current research interests include remote sensing, point cloud processing, digital map generation, information visualization, and computer graphics. He served as an editorial board member of the International Journal of Computer Science and Artificial Intelligence.



Kang-Hua Lai received his BS degree in civil engineering from National Chung Hsing University and received his MS degree in geomatics from National Cheng-Kung University, Taiwan in 2010 and 2012, respectively.

He is currently a member of the Digital Geometry Processing Laboratory (http://dgl.geomatics.ncku.edu.tw), and his research interests include remote sensing and image processing.

Zhi-Bin Chen received his BS and MS degree in geomatics from National Cheng-Kung University, Taiwan in 2010 and 2012, respectively.

He is currently a member of the Digital Geometry Processing Laboratory (http://dgl.geomatics.ncku.edu.tw), and his research interests include remote sensing and image processing.

Jyun-Yuan Chen received his BS and MS degree in geomatics from National Cheng-Kung Universiry, Taiwan in 2005 and 2007 respectively.

He is currently a Ph.D. student and a member of the Digital Geometry Processing Laboratory. His research interests include point cloud processing, information visualization, computer graphics, remote sensing, and image processing.