A Multitemporal Approach for Cloud Removal from Satellite Images

JASNA AP\textsuperscript{1}, RAMZAN BASHEER\textsuperscript{2}

\textsuperscript{1,2} Department of ECE, MESCE, Kuttippuram

Abstract—Clouds are an obstruction for land surface observation, which results in regional information being blurred or lost. An efficient and effective cloud removal approach is proposed in this paper. The approach removes cloud contaminated portion of a satellite image and then reconstructs the information of missing data by taking advantage of the temporal correlation of multi-temporal images. Patch based information reconstruction is mathematically formulated as a Modified Poisson equation. The proposed approach was tested on images acquired by Landsat-7 Enhanced Thematic Plus sensor and the experiments verify that the proposed approach can achieve better results in terms of radiometric accuracy and consistency compared to other related approaches. In addition, quantitative and qualitative analysis on simulated data with single reference image and multiple reference images are conducted using quality index and visual inspection, respectively, to demonstrate the advantage of utilizing multi-temporal reference images. The significance of size of the patch used is also tested

Keywords—cloud removal; multitemporal images; patch based; Modified Poisson equation.

I. INTRODUCTION

The satellite images are used for different purposes like defense (change detection in regions), agriculture (for analysis of agriculture), environmental assessment etc. Quality of image is one of the most important factors in satellite images because every object in a satellite image is important for accurate processing [11]. Quality of satellite image is always questionable when there is higher influence of clouds. Land scenes are, on average, approximately 35% cloud covered globally, significantly reducing the availability of cloud-free surface observations [1]. Presences of these clouds in satellite images are unavoidable during image acquisition time. And it causes many problems in the study of satellite image based applications. Removing cloud from an image will be helpful for better analysis of satellite imaging applications. But removal of cloud covered region is a challenging task because each region in a satellite image is essential one.

Generally cloud cover does not always appear in the same location; therefore if multi-temporal images are acquired at different times over a specific region, the cloud free images may be generated by replacing the cloud area with cloud free areas from other multi-temporal images. The aim of this study is to remove clouds and reconstruct information of missing data by information cloning algorithm. Instead of reconstructing information pixel by pixel, we propose a patch-based approach that mathematically formulates the reconstruction problem as a Modified Poisson equation and then finds a closed form solution.

There are a number of cloud removal approaches that will ease the difficulties caused by cloud covers. In the past decade, a number of cloud removal approaches have been proposed. These approaches can be classified into three categories [1]: inpainting-based, multispectral-based, and multi-temporal-based. In the first category, the cloud-contaminated regions are synthesized using image synthesis and inpainting techniques. The information inside the cloud contaminated region is synthesized by propagating the spectrogeometrical information retrieved from the remaining parts of the image [2]. The synthesis approaches can yield a visually plausible result, which is suitable for cloud free visualization. But the lack of restoring information of cloud-contaminated pixels makes them unsuitable for future applications, and it is sensitive to the size of missing area. In multispectral
based approaches, multispectral data are utilized in cloud detection and removal. Chun et al. [3] regarded cloud removal as a denoising problem. Based on statistical characteristics of images, an improved homomorphism filtering is applied to filter out low-frequency components that potentially represent clouds. This method is not suitable for thick clouds and also such methods tend to have difficulty with large clouds.

Compared with the previous approaches, the multi-temporal based approaches [4]-[6] have a better ability to cope with large clouds. Here multi-temporal images are used. If multi-temporal images are acquired, the cloud-cover problem has a chance to be eased by reconstructing the information of cloud contaminated pixels under the assumption that the land covers change insignificantly over a short period of time. Li et al. [4], Wang et al. [5], and Tseng et al. [6] corrected the radiometric inconsistency of cloud-contaminated images and their corresponding temporal images using mean and standard deviation of pixel intensities first. Wavelet-based fusion method was then used to fuse boundaries of cloud contaminated regions. Based on the shortcomings in the related works, the aim and the main contribution of this study are to reconstruct information of cloud-contaminated regions using patch based information cloning, which can yield better results in terms of radiometric accuracy and consistency by using less computational time.

II. CLOUD REMOVAL ALGORITHM

2.1. Overview

The workflow of the proposed automatic cloud removal method is schematically shown in Figure. 1. Our method consists of three main processing steps: cloud detection, quality assessment, and reconstruction. In the first step, cloud detection approaches are adopted to detect clouds for the input images. A quality assessment based on structural similarity (SSIM) index [7] is then applied to sort the input reference images according to image similarity. In the third step, the proposed information cloning algorithm is performed to fill in the missing data after removing the cloud contaminated pixels from the target image.

2.2. Cloud Detection

A satellite image in which there are a few clouds is selected as the target image and denoted as $I_T$. Then consider a set of its corresponding images captured at the same position but different times (multi-temporal images), called reference images and denoted as \{ $I_{R_1}, \ldots, I_{R_n}$ \}. The aim is to remove clouds and to reconstruct the information of missing data in the target image $I_T$ using the reference images \{ $I_{R_1}, \ldots, I_{R_n}$ \}. Two common methods for detecting clouded pixels and associated shadows are subtraction method and threshold method.

2.2.1. Subtraction method

In subtraction method, a cloud contaminated image and a cloud cleared image are subtracted to obtain the difference image. For similar areas the absolute values of differences will be zero or small, while in the areas of clouds the differences will be high and the cloud area will be detected. Consider a target image $I_T(i,j)$, that contains clouds, and a clear reference image $I_R(i,j)$, that is free from clouds. Then cloud in the target image can be detected by:

$$I_T(i, j) - I_R(i, j)$$ (1)

2.2.2. Threshold method

Here a threshold values is proposed for detecting clouds. First the reference image is coregistered with the main image. Image registration ensures the information from each image refers to the same physical structure in the environment. Here assume that images are already coregistered. Due to the
different solar irradiance and atmospheric effects, before the detection of clouds, the brightness of the reference image is corrected relative to that of the target image [5]. Correction of brightness can be performed using:

\[
I_T'(i, j) = \frac{\sigma_R}{\mu_R} \times [I_R(i, j) - \mu_R] + \mu_T
\]

where \(I_R(i,j)\) is the new brightness value of the reference image. \(\mu_R\) and \(\sigma_R\) are the mean and standard deviation of the reference image, and \(\mu_T\) and \(\sigma_T\) are those of target image respectively.

In general, clouds reflect the solar radiation in the visible and infrared spectra to a much higher degree than the ground. By setting a threshold \(C\), a cloud region can distinguish from the ground regions. If

\[
I_T(i, j) > C
\]

where \(I_T(i, j)\) is the brightness values of a pixel of the target image, it can be surmised that there is a cloud in the target image at the location \((i ,j)\). The threshold \(C\) can be easily determined by investigating the histogram of the image or by trial and error method.

**Figure 1.** Workflow of the proposed cloud removal method which consists of three main processing steps: cloud detection, image quality assessment, and information reconstruction.
2.3. Quality Assessment

Once the cloud-contaminated pixels in the target and reference images are identified, several cloud-free patches are selected from the reference images to reconstruct the information of cloud-contaminated regions in the target image. The cloud-free patches are selected based on image similarity. The simplest and most widely used similarity metric is the mean squared error (MSE), computed by averaging the squared intensity differences of target and reference image pixels, along with the related quantity of peak signal-to-noise ratio (PSNR). These conventional methods are simple and have clear physical meanings. But they are not very well matched to perceived visual quality, and also do not fully take structure similarity into account. To generate a satisfactory cloud-free image, the structural similarity (SSIM) index introduced by Wang et al. [8] is used to estimate the quality of working regions in the reference images. The SSIM index has three components, namely, illumination \( L(I_T, I_R) \), contrast \( C(I_T, I_R) \), and structure \( S(I_T, I_R) \),

\[
\begin{align*}
L(I_T, I_R) & = \frac{2\mu_I \mu_I + C_1}{\mu_I^2 + \mu_I^2 + C_1} \\
C(I_T, I_R) & = \frac{2\sigma_I \sigma_I + C_2}{\sigma_I^2 + \sigma_I^2 + C_2} \\
S(I_T, I_R) & = \frac{\sigma_{I, I} + C_3}{\sigma_I^2 + \sigma_I^2 + C_3}
\end{align*}
\]

where \( \mu_I \) and \( \sigma_I \) represent the mean intensity and the standard deviation of image \( I \), respectively, and \( \sigma_{I, I} \) represents the covariance coefficient between images \( I_T \) and \( I_R \). The constants \( C_1 \), \( C_2 \), and \( C_3 \) are used to avoid instability when the denominators are nearly zero. By combining these three similarity components, and following the parameter setting given in [8], following simplified form of SSIM index is obtained:

\[
SSIM(I_T, I_R) = \frac{(2\mu_I \mu_I + C_1)(2\sigma_I \sigma_I + C_2)}{(\mu_I^2 + \mu_I^2 + C_1)(\sigma_I^2 + \sigma_I^2 + C_3)}
\]

In the experiments, the constants \( C_1 \) and \( C_2 \) are set to \((K_1 L)^2\) and \((K_2 L)^2\) respectively, where \( L \) is the dynamic range of pixels (i.e., 255 for an 8-bit channel) and \( K_1 = K_2 = 0.01 \). The SSIM index ranges from 1.0 (the most similar patches) to -1.0 (the most dissimilar patches). To accurately estimate image similarity and to select suitable cloning patches, the SSIM index between the target and reference images is calculate for the cloud free regions only. And for large clouds in heterogeneous
landscape, select several patches from the reference images instead of using only one patch, for reconstructing the cloud contaminated region. In this manner, many reference images may be selected, and many patches may be embedded into the cloud contaminated regions. For a reference image, if the patch contains cloud then the patches will not be selected as candidate. Figure 4 shows the reference images sorted by SSIM index for the marked region in the target image.

2.4. Reconstruction

The information from the selected patches are used to reconstruct the corresponding cloud contaminated regions. Instead of reconstructing the information pixel by pixel, the problem is mathematically formulated as a Poisson equation. If \( \Gamma \) indicates the cloud contaminated region in the target image \( I_T \), and \( \partial \Gamma \) is its boundary. Let the unknown image intensity function defined over the cloud contaminated region \( \Gamma \) (i.e., the unknown that is to be calculated) is denoted as \( f \). Let \( f^* \) be the image intensity function defined over the target image \( I_T \) minus the cloud-contaminated region \( \Gamma \), and let \( V \) be a guidance vector field defined over the cloud-contaminated region \( \Gamma \) and it is defined as the gradient of the selected patches. \( V \) is used to guide the reconstruction process to optimize the pixel intensities in the cloud contaminated regions. To find an accurate and optimized reconstruction result (i.e., the solution of the unknown function \( f \)), the problem is formulated as an optimization equation [8] with the boundary condition

\[
\min_{\Gamma} \iint_{\Gamma} |\nabla f - v|^2
\]

where \( \nabla \) is the gradient operator. The solution to (5) is the unique solution of the following Poisson equation with Dirichlet boundary conditions:

\[
\Delta f = \text{div}V \text{over} \Gamma
\]

where \( \Delta \) is the Laplacian operator and \( \text{div}V \) is the divergence of the vector field \( V = (v_1, v_2) \). Equations (5) and (6) are the fundamental formulations of information reconstruction. The equation minimization indicates that the gradient of the unknown function \( f \) is close to the gradient field \( V \) of the selected patches.
III. FAST POISSON RECONSTRUCTION

Although the Poisson image reconstruction is very effective for various applications, it requires large computational cost because it is an iterative technique. Furthermore, the Poisson image reconstruction approach has a limitation which is the color of the source region will be totally adapted to the target region. A Fast Poisson reconstruction [9] algorithm is proposed which optimizes whole image region in contrast to the Poisson image editing which only optimizes the source image region. The proposed fast Poisson reconstruction has a closed form solution and a color preserving parameter [10]. This color preserving parameter can control the color adaptation level. If the color adaptation parameter is large, the color of the source and destination is perfectly preserved in the reconstructed result. In the proposed fast Poisson reconstruction, we minimize the following energy functional:

$$\mathcal{E}[f] = \int_T \| \nabla f - v \|^2 dt + \epsilon \int_T \| f - \tilde{f} \|^2 dt$$  \hspace{1cm} (7)

where $\tilde{f}$ is the naive composed image, $T$ is the whole image region, and $\epsilon$ is the color preserving parameter. The closed form solution of (7) is derived from the functional derivative $dg[f]/df$ as:

$$(\text{div}V - \Delta f) + \epsilon(\tilde{f} - f) = 0$$  \hspace{1cm} (8)

The discretized version of (7) and (8) are given by:

$$G(F_i) = \sum_i \| \nabla F_i - V_i \|^2_G + \sum_i \| F_i - \tilde{F}_i \|^2$$  \hspace{1cm} (9)

$$(U_i - \Delta F_i) + \epsilon(F_i - \tilde{F}_i) = 0$$  \hspace{1cm} (10)

where $F_i$ and $\tilde{F}_i$ are discretized value of $f$ and $\tilde{f}$ at $i$-th pixel, $V_i$ is discretized value of $v$, and $U_i$ is the discretized version of $\text{div}V$ at $i$-th pixel, respectively. The solution of (10) is effectively obtained using discrete cosine transform which is given below as

$$F_w = \frac{U_w + \tilde{F}_w}{\epsilon + L_w}$$  \hspace{1cm} (11)

where $F_w$ is the discrete-cosine-transformed reconstructed image of at the frequency $w$, $U_w$ is the discrete cosine transformed component of $U_i$, $\tilde{F}_w$ is the discrete cosine transformed naive composed image, $L_w$ is the discrete cosine-transformed Laplacian operator $\Delta$.

Table I

Reconstruction accuracy of cloud removal results generated using a single reference image and multiple reference images

<table>
<thead>
<tr>
<th>Quality index</th>
<th>Single reference image (R_C)</th>
<th>Multiple reference images (R_A, R_B, R_C, R_D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>4.71</td>
<td>2.9</td>
</tr>
<tr>
<td>PSNR</td>
<td>26.52</td>
<td>27.57</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.94</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Figure 5. Cloud removal result. (Top) Target and reference images. The target image is captured near Colombo on September 29, 2002. (Bottom) (a) Result of patch replacement, (b) information reconstruction using poisson method, (c) our result.

Figure 6. Cloud removal result. (Top) Target and reference images. The target image is captured near Mexico on September 29, 2002. (Bottom) (a) Result of patch replacement, (b) information reconstruction using poisson method, (c) our result.
Figure. 7. Results of cloud removal using a single reference image and multiple reference images. (a) Result of using a single reference image (RC). (b) Result of using multiple reference images (RA;RB;RC;RD).

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The images captured using Landsat-7 ETM+ sensor near the Colombo in September 2002, and images captured using Landsat-4 near Mexico in September 2014 were used to test the feasibility of the proposed approach. The cloud removal results are shown in Figures. 6 and 7, where 4 images are selected as reference images. For reconstruction of cloud contaminated region we use cloud free patches from reference images with highest rank. And if a reference patch contains cloud, then that patch is not considered for reconstruction. The Figures 6(a) and 7(a) shows the result of patch replacement. The reconstruction using Poisson method is shown in Figures 6(b) and 7(b), and results of our proposed method are shown in Figures. 6(c) and 7(c). The computational time for our method is less compared to the Poisson method. The time required for image reconstruction of Figures. 6 using Poisson method was 52.42 sec (for 100 iteration), but our method only required 4.12sec, and for reconstruction Figure. 7 time for reconstruction was 260.83(for 100 iteration) sec and 16.08sec respectively.

We conducted an experiment of information reconstruction using single and multiple reference images to demonstrate the advantages of utilizing multi-temporal images. The results are shown in Figure. 8. To quantitatively compare the results, the standard and commonly used measurements, root mean square error (RMSE), peak signal to noise ratio (PSNR), and SSIM index are used. The results are shown in Table 1. In this experiment for single reference reconstruction we use R_C as reference image. And for multiple reference reconstruction we take all the four reference images R_A to R_D. From Table. 1 its clear that by using multiple reference images we get better reconstruction accuracy than using single reference image. Also an experiment of removing cloud contaminated data and reconstructing the missing data was conducted to quantify the reconstruction accuracy with different patch size. The results are shown in Figure. 9. and Table 2. In Figure. 9(a), the size of selected patch is 25 X 25 pixels, and for Figure. 9(b) the size is 15 X 15. It is apparent that by reducing the size of the patch, the reconstruction accuracy will be higher and has better visual quality.

V. CONCLUSION

In this paper a cloud removal algorithm has been introduced. The cloud contaminated portion in a satellite image is detected, removed and then information of missing data is reconstructed using multi-temporal reference images. Our approach uses the correlation of multi-temporal images for reconstruction and is patch based, in contrast to conventional pixel based reconstruction methods. Experiments are conducted on images acquired by Landsat-7 ETM+ and Landsat- 4 sensors. The advantages of using multi-temporal images are experimentally analyzed and also the significance of
size of the patch used is also tested. Experimental results show that our method has better results in terms of radiometric accuracy and consistency with less computational time, compared to related approaches.

Table II

<table>
<thead>
<tr>
<th>Quality index</th>
<th>Patch size 25 X 25</th>
<th>Patch size 15 X 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>9.66</td>
<td>8.73</td>
</tr>
<tr>
<td>PSNR</td>
<td>19.03</td>
<td>19.60</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.882</td>
<td>0.91</td>
</tr>
<tr>
<td>COMPUTATIONAL TIME (Sec)</td>
<td>4.12</td>
<td>6.648</td>
</tr>
</tbody>
</table>

REFERENCES


