This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2017.2703092, IEEE Signal Processing Letters

SPL-21724-2017

Missing Information Reconstruction for Single Remote Sensing Images Using Structure-Preserving Global Optimization

Qing Cheng, Huanfeng Shen, *Senior Member, IEEE*, Liangpei Zhang, *Senior Member, IEEE*, Zhenghong Peng

***Abstract*—Filling missing information or removing special objects is often required in the applications of high spatial resolution images. A novel single-image reconstruction method is presented in this letter to solve this task, without the use of any complementary data. Firstly, the spatial pattern of the image is obtained by the statistics of similar patch offsets in the known regions, which provide reliable information for reconstructing the image. The missing regions are then filled by combining a series of shifted pixels via global optimization. The proposed method was tested on a cloudy image for cloud removal and on a public image for military object concealment. The experimental results show that the proposed method can produce visually convincing and coherent reconstructed images, and the accuracy of the reconstruction is better than the existing non-complementation methods.**

***Index Terms*—Image reconstruction, missing information, single image, high spatial resolution, global optimization.**

I. Introduction

During remote sensing data acquisition, sensor failure often results in dead pixels [1] in images, and thick cloud cover is often present in passive data. Furthermore, we sometimes need to remove particular objects from remote sensing images, in applications such as military object concealment. In all the above cases, the remote sensing data suffer from the loss of information. In the last few years, a number of different approaches have been presented to reconstruct the missing information of remote sensing images. However, most of the existing approaches are based on the prerequisite that there are other complementary data at hand, such as multi-temporal images [2]–[4] or other spectral bands [5], which can provide auxiliary information for the reconstruction. However, the need for complementary data is sometimes a fatal limitation.

Another strategy is reconstructing the missing information using only the damaged image itself, which is also called the non-complementation approach [6]. The most common method in this category is interpolation, such as the interpolation methods [7-8] which are useful in remote sensing data processing. The second category of methods is the propagated diffusion methods, which propagate the local information from the exterior to the interior of the missing regions, using partial differential equations (PDEs). An example of such an approach is the recent work in [9]. The third classic approach is the variation-based methods, which take the information reconstruction as solving an ill-posed inverse problem, and use a regularization technique to make it well posed. An example of such an approach is the method proposed in [10]. Generally speaking, a common feature of the three approaches mentioned above is that they work well for recovering simple ground features or filling small regions, but they are not suitable for reconstructing large areas and textures, in which they lead to oversmoothing and blurring.

In the non-complementation category, another representative approach is the exemplar-based methods, which are aimed at recovering large missing regions and texture information [11]. However, these methods have been intensively studied in the field of digital image processing, but have rarely been applied to remote sensing images. More recently, He *et al.* [12] developed an advanced global optimization method for this category of digital image completion, which yields better results than the existing state-of-the-art methods. A similar idea could be applied to reconstruct large missing regions for remote sensing images with a high spatial resolution. However, as the structural patterns of remote sensing images are more complex than natural images, an improvement to the method of He *et al.* is proposed in this letter for the reconstruction of remote sensing images. The proposed method strengthens the maintenance of image structures and can achieve spatially coherent reconstructions for high spatial resolution images.

II. Method

A. Basic Idea

The basic idea of the proposed method is to fill each missing pixel by copying a known pixel. In order to select an appropriate known pixel and fill the missing area with reliable information, we should first analyze the spatial patterns of the structure and texture in the image. He *et al.* [12] found that when matching similar patches in an image, the statistics of the offsets (relative locations) of the similar patches are sparsely distributed, i.e., a majority of the patches have similar

This work was supported in part by the National Natural Science Foundation of China under Grants 41601357 and 41422108, Natural Science Foundation of Hubei Province under Grant 2016CFB333.

Q. Cheng and Z. Peng are with the School of Urban Design, Wuhan University, Wuhan 430072, China (e-mail: qingcheng@whu.edu.cn; laopeng129@vip.sina.com)

H. Shen is with the School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China (e-mail: shenhf@whu.edu.cn).

L. Zhang is with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: zlp62@whu.edu.cn)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2017.2703092, IEEE Signal Processing Letters

offsets, forming several prominent peaks in the 2-D histogram statistics. Such dominant offsets indicate the essential features of the image, such as linear edges, textures, and repeated objects. Therefore, the dominant offsets can provide clues for reconstructing the missing regions. With a few dominant offsets, we can build a global optimization model to jointly assign each missing pixel a known pixel, from which it copies the content, thereby achieving the image reconstruction.

B. He et al.'s Algorithm

This letter proposes an algorithm taking its origins from the previous method developed by He *et al.* [12], which is considered one of the state-of-the-art image completion algorithms. Their original algorithm has three steps as follows.

1) Matching Similar Patches. We suppose that $P(\mathbf{x})$ is an image patch centered at pixel $\mathbf{x}=(x,y)$. For each patch P in a known region, we find another known patch which is the most similar to P, and calculate its relative position $\mathbf{s}$:

$$\mathbf{s}(\mathbf{x})=\arg\min_{\mathbf{s}}\left\|P(\mathbf{x}+\mathbf{s})-P(\mathbf{x})\right\|^2 \quad s.t.\ \ |\mathbf{s}|>\tau \tag{1}$$

where $\mathbf{s}=(u,v)$ is the 2-D coordinates of the offset, which represent the distance vector of a patch and its most similar patch. The similarity is measured by the sum of squared differences between two patches. Parameter τ is used to ignore the nearby patches which are likely to be similar but do not contribute to the repetitive patterns in the image.

2) Computing the Offset Statistics. After obtaining all the offsets $\mathbf{s}$ for all the known pixels, their 2-D histogram statistics $H(u,v)$ are computed:

$$H(u,v)=\sum_{\mathbf{x}}\delta(\mathbf{s}(\mathbf{x})=(u,v)) \tag{2}$$

where $\delta(\cdot)$ is equal to 1 if the argument is true, and 0 otherwise. $H(\mathbf{s})$ gives the amount of patches having the offset of $\mathbf{s}$. The K highest peaks of the histogram are then picked out. These peaks correspond to the K dominant offsets and describe how the similar features appear in the image.

3) Filling the Missing Regions by the Optimization Model. Within the specified K dominant offsets, each missing pixel is assigned an offset: the relative location (pixel) from where it copies the content. The missing region is then filled by combining a series of shifted pixels corresponding to these offsets. Therefore, the goal is to calculate a suitable offset map $L(\mathbf{x})=\mathbf{s}_i$ for all the pixels in the missing regions. The optimal offset map minimizes the following Markov random field (MRF) energy function:

$$E(L)=\sum_{\mathbf{x}\in\Omega}E_d(L(\mathbf{x}))+\sum_{(\mathbf{x},\mathbf{x}')|\mathbf{x}\in\Omega,\mathbf{x}'\in\Omega}E_s(L(\mathbf{x}),L(\mathbf{x}')) \tag{3}$$

where Ω is the missing region; $(\mathbf{x},\mathbf{x}')$ are the 4-connected neighbors; $L(\mathbf{x})$ means the label assigned to the missing pixel $\mathbf{x}$; and the label represents the pre-selected offsets $\{\mathbf{s}_i\}_{i=1}^{K}$. $L(\mathbf{x})=\mathbf{s}_i$ means that the pixel $\mathbf{x}$ is filled by copying the pixel at $\mathbf{x}+\mathbf{s}_i$.

The first term E_d in (3) is a data term. E_d is 0 if the label is valid (i.e., $(x+s_x, y+s_y)$ is a known pixel); otherwise, it is $+\infty$. The second term E_s is a smoothness term for penalizing the incoherent seams. He *et al.* [12] used image brightness continuity as the smoothness constraint, and presented a smoothness term as follows:

$$E_s(L(\mathbf{x}),L(\mathbf{x}'))=\left\|I(\mathbf{x}+L(\mathbf{x}))-I(\mathbf{x}+L(\mathbf{x}'))\right\|^2+\left\|I(\mathbf{x}'+L(\mathbf{x}))-I(\mathbf{x}'+L(\mathbf{x}'))\right\|^2 \tag{4}$$

where $I(\mathbf{x})$ is the RGB color of pixel $\mathbf{x}$. Note that a seam would appear between $\mathbf{x}$ and $\mathbf{x}'$ in the reconstructed image if their offsets are discontinuous. Thus, this smoothness term penalizes such a seam to ensure that the neighboring pixels in the reconstructed image are compatible.

C. The Improved Algorithm

It can be seen that the most significant step of He *et al.*'s algorithm is step 3, which determines how the missing region is filled and which known pixel is used to replace the missing pixel. As we all know, the structural patterns of remote sensing images are more complex than natural images, so the spatial and spectral continuity is a great challenge for the reconstruction of remote sensing images. Therefore, the smoothness term of the optimization model is crucial for remote sensing image reconstruction. However, in the optimization model of He *et al.*'s algorithm, the smoothness term (4) only stresses the continuity of the image brightness, but the continuity of the spatial structures is weak, which means that the edge structures may often not be well maintained, especially for curved edges in the images. In order to alleviate this limitation, we propose to combine the brightness continuity and the gradient continuity as the constraint to strengthen the maintenance of spatial structures for remote sensing images. We thus present the following smoothness term:

$$E_s(L(\mathbf{x}),L(\mathbf{x}'))= \tag{5}$$
$$\left(\left\|I(\mathbf{x}+L(\mathbf{x}))-I(\mathbf{x}+L(\mathbf{x}'))\right\|^2+\left\|I(\mathbf{x}'+L(\mathbf{x}))-I(\mathbf{x}'+L(\mathbf{x}'))\right\|^2\right)+$$
$$\alpha\left(\left\|\nabla I(\mathbf{x}+L(\mathbf{x}))-\nabla I(\mathbf{x}+L(\mathbf{x}'))\right\|^2+\left\|\nabla I(\mathbf{x}'+L(\mathbf{x}))-\nabla I(\mathbf{x}'+L(\mathbf{x}'))\right\|^2\right)$$

where $\nabla I(\mathbf{x})$ is the magnitude of the image gradient at location $\mathbf{x}$, and α is a weight parameter. Since the mean value of the image gradient is generally lower than the mean value of the image brightness, we used $\alpha=2$ to balance the brightness and gradient terms in most of the experiments. As both the brightness continuity and the gradient continuity are used for the smoothness, the structure is better preserved in the reconstructed image.

We optimize the energy function (3) by multi-label graph cuts [13]. The code of the multi-label graph cuts can be downloaded from http://vision.csd.uwo.ca/code/. A flowchart of the proposed method is presented in Fig. 1.

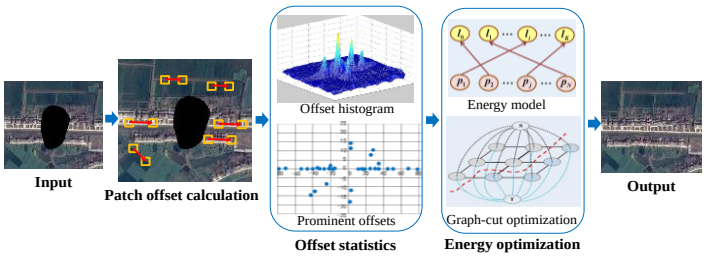


Fig. 1. The flowchart of the proposed reconstruction method.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2017.2703092, IEEE Signal Processing Letters

III. EXPERIMENTS

We undertook one simulated data experiment and two real data experiments to test and verify the efficacy of the proposed method. The reconstruction results were quantitatively evaluated by several statistical indices. The coherence index ($d_{\text{coherence}}$) [11] was used to measure the coherence between the filled region and the remaining region, which gives an assessment of the results from the visual quality aspect. The peak signal-to-noise ratio (*PSNR*) index was used to measure the consistency between the reconstructed image and the true intact image, which gives an assessment of the results from the pixel value fidelity aspect. The definitions of these evaluation indices are as follows:

$$d_{\text{coherence}} = \frac{1}{N}\sum_{P\in\Omega}\min_{Q\in\bar{\Omega}}\|P-Q\|^2 \tag{6}$$

$$PSNR = 10\log_{10}\left(255^2 * N \Big/ \sum_{j=1}^{N}(I_{Oj}-I_{Rj})^2\right) \tag{7}$$

where Ω represents the missing region. $\bar{\Omega}$ represents the remaining region. P is the patch in the missing region. Q is the patch in the remaining region. N is the total number of pixels in the missing region. I_{Oj} and I_{Rj} are the original and reconstructed values of the *j*th missing pixel, respectively. The smaller the $d_{\text{coherence}}$ value is, the better the visual effect of the reconstruction. A higher *PSNR* value means that the reconstructed image is closer to the true intact image.

A. Simulated Data Experiment

This experiment was performed on an IKONOS image with a resolution of 0.82 m, as shown in Fig. 2(a). This color image is a data fusion product blending panchromatic and multispectral data, provided by the IKONOS satellite. In the original test image, we simulated a missing region with pixel values of zero. The simulated image is shown in Fig. 2(b). The proposed method was compared with He *et al.*'s method [11], the kriging interpolation method (ordinary kriging with exponential variogram model) [7], and the Content-Aware Fill method in Adobe Photoshop. The reconstruction results of each method are shown in Fig. 2(c)–(f).

From Fig. 2, it can be seen that the proposed method gives better reconstruction results than the other methods. In the result of the kriging interpolation method (Fig. 2(c)), the filled region suffers from serious oversmoothing and blurring. For the result of the Content-Aware Fill method (Fig. 2(d)), obvious artifacts are introduced in the reconstructed region. Furthermore, the road on the right is not well connected. For the result of the method of He *et al.* (Fig. 2(e)), the road on the left is not well recovered, and dislocation occurs in the filled region. This defect is mainly because the road edge is curved, and the method of He *et al.* cannot handle this curved structure very well. For the proposed method, the reconstruction result (Fig. 2(f)) is more satisfactory, with the roads well connected and the buildings reconstructed in an orderly manner. The filled region is spatially coherent with the surrounding features. The effectiveness of the proposed method was also evaluated by quantitative assessment. The $d_{\text{coherence}}$ and *PSNR* values of the simulated experiment results are listed in Table I. It can be seen that, for the result of the proposed method, the $d_{\text{coherence}}$ value is lower and the *PSNR* is higher than for the other methods. This indicates that the proposed method can provide a result with a better visual quality and higher fidelity.

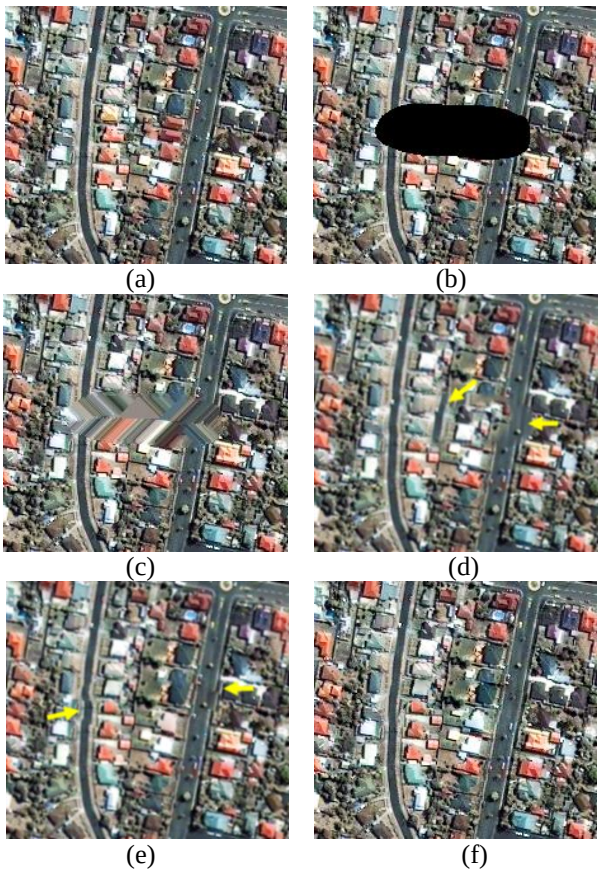


Fig. 2. Simulated data experiment. (a) Original IKONOS image. (b) Simulated image with missing region. Reconstructed images by the following methods: (c) kriging interpolation method; (d) Content-Aware Fill method; (e) He *et al.*'s method; and (f) the proposed method.

Table I $d_{\text{coherence}}$ and *PSNR* values of the simulated experiment results

		Kriging interpolation	Content-Aware Fill	He *et al.*	Proposed
Fig. 2	$d_{\text{coherence}}$	102.13	49.24	26.74	22.53
	PSNR	12.43	13.02	13.81	14.22

B. Real Data Experiment

Removing clouds and recovering the ground information for cloud-contaminated images is often necessary in remote sensing applications. The first real data experiment involved cloud removal for a GeoEye-1 image with a resolution of 1.65 m. The test image is shown in Fig. 3(a), and the reconstruction results are shown in Fig. 3(b)–(e).

From Fig. 3, it can be seen that the result of the kriging interpolation method (Fig. 3(b)) is oversmoothed. For the results of the Content-Aware Fill method and the method of He *et al.* (Fig. 3(c) and 3(d)), the river is not connected and breaks in several places, which clearly makes no sense. For the result of the proposed method, we can see in Fig. 3(e) that the river is well connected, the residential district and other

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2017.2703092, IEEE Signal Processing Letters

ground features are restored in a rational manner. Overall, the reconstructed image has a convincing visual quality.

In the second real data experiment, the proposed method was tested for the concealment of military objects in a remote sensing image. With high-resolution images providing an abundant data source for detailed terrain observation, lots of important military objects can be clearly exposed in the public images, which can pose a severe threat to national security. This security threat to the military can be avoided if we are able to remove the military objects and reconstruct a visually convincing image with unnoticed modifications before the remote sensing images are made public. Therefore, military object removal (concealment) is often required for public images. The proposed method can be used to efficiently undertake this task. This experiment was performed on a QuickBird image with a resolution of 0.61 m, as shown in Fig. 4(a). This color image is a data fusion product blending panchromatic and multispectral data, provided by the QuickBird satellite. We can clearly see that there are important military targets in the region labeled with yellow lines. This test involved removing the region labeled with yellow lines and reconstructing an image without any military targets. Fig. 4(b)–(e) shows the reconstruction results of the four methods.

From Fig. 4(b), it can be seen that obvious oversmoothing and blurring is again introduced in the result of the kriging interpolation method. For the Content-Aware Fill method, lots of chaotic artifacts occur in the filled region, as shown in Fig. 4(c). In the result of He *et al.*'s method (Fig. 4(d)), a seam line arises between the building and the forested area, which is unreasonable. In the result of the proposed method (Fig. 4(e)), the different ground features are arranged reasonably and link naturally, and the filled region has a convincing visual quality. That is to say, the military objects are effectively removed, and this modification is difficult to detect.

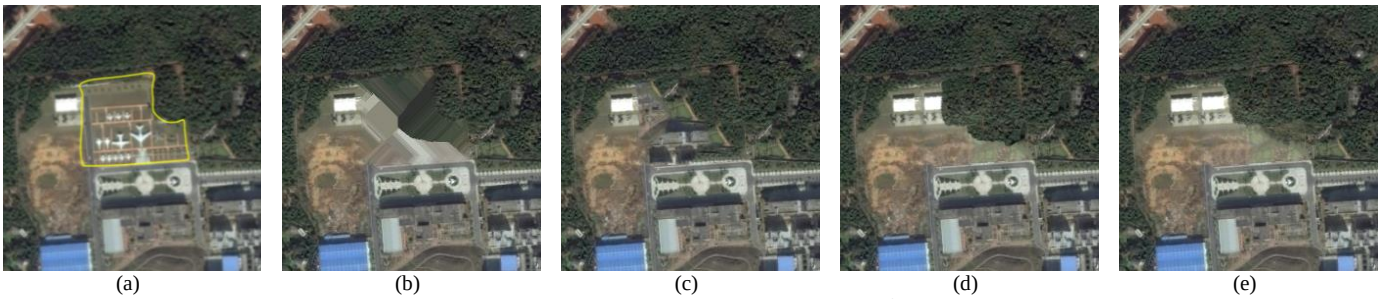
Finally, the reconstruction results in the two real data experiments were quantitatively evaluated by the $d_{\text{coherence}}$ index, as listed in Table II. It can be seen that the $d_{\text{coherence}}$ values obtained using the proposed method are the lowest. It is therefore suggested that the proposed method can provide more visually convincing results than the other methods.

Table II $d_{\text{coherence}}$ values of the real data experiment results

		Kriging interpolation	Content-Aware Fill	He *et al.*	Proposed
Fig. 3	$d_{\text{coherence}}$	69.36	28.94	15.74	12.61
Fig. 4	$d_{\text{coherence}}$	81.25	37.67	19.85	16.72

(a) (b) (c) (d) (e)

Fig. 3. The first real data experiment. (a) Original cloud-contaminated GeoEye-1 image. Cloud removal results by the following methods: (b) kriging interpolation method; (c) Content-Aware Fill method; (d) He *et al.*'s method; and (e) the proposed method.



(a) (b) (c) (d) (e)

Fig. 4. The second real data experiment. (a) Original QuickBird image. Object removal results by the following methods: (b) kriging interpolation method; (c) Content-Aware Fill method; (d) He *et al.*'s method; and (e) the proposed method

IV. Conclusion

This letter has presented a missing information reconstruction method for single remote sensing images. Based on He *et al.'s* method, which can predict a reliable spatial pattern for the missing regions and fill the missing pixels via global optimization, the proposed method uses an edge-preserving reconstruction model to strengthen the maintenance of image structures. The proposed method was applied to a cloudy image for cloud removal and to a public image for military object concealment, obtaining satisfactory results in both cases. There are, however, some limitations to the proposed method. For those remote sensing images with multiple ground features and complex spatial patterns, the statistics of similar patch offsets may not be able to reliably predict the spatial pattern of the missing regions. In such a case, the reconstruction accuracy may be low. This is also a common defect of all the non-complementation methods.

REFERENCES

[1] H. Shen and L. Zhang, "A MAP-based algorithm for destriping and inpainting of remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, May 2009.

[2] Q. Cheng, H. Shen, L. Zhang, Q. Yuan, and C. Zeng, "Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model," *ISPRS J. Photogramm.*, vol. 92, pp. 54–68, 2014.

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

[4] M. Xu, X. Jia, M. Pickering, and A. J. Plaza, "Cloud removal based on sparse representation via multitemporal dictionary learning," *IEEE Trans. Geosci. Remote Sens.*, DOI.10.1109/TGRS.2015.2509860, 2016.

[5] Y. Zhang, B. Guindon, and J. Cihlar, "An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images," *Remote Sens. Environ.*, vol. 82, no. 2, pp. 173–187, 2002.

[6] 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 Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 61–85, 2015.

[7] C. Zhang, W. Li, and D. Travis, "Gaps - fill of SLC - off Landsat ETM+ satellite image using a geostatistical approach," *Int. J. Remote Sens.*, vol. 28, no. 22, pp. 5103–5122, 2007.

[8] M. Žukovic and D. T. Hristopulos, "A Directional gradient-curvature method for gap filling of gridded environmental spatial data with potentially anisotropic correlations," *Atmos. Environ.*, vol. 77, pp. 901–909, Oct. 2013.

[9] A. Maalouf, P. Carré, B. Augereau, and C. Fernandez-Maloigne, "A bandelet-based inpainting technique for clouds removal from remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2363–2371, Jul. 2009.

[10] Q. Cheng, H. Shen, L. Zhang, and P. Li, "Inpainting for remotely sensed images with a multichannel nonlocal total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 175–187, 2014.

[11] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1200–1212, Sept. 2004.

[12] K. He and J. Sun, "Image completion approaches using the statistics of similar patches," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 12, pp. 2423–2435, 2014.

[13] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, 2001.