Inpainting for Remotely Sensed Images With a Multichannel Nonlocal Total Variation Model

Qing Cheng, Huanfeng Shen, *Member, IEEE*, Liangpei Zhang, *Senior Member, IEEE*, and Pingxiang Li, *Member, IEEE*

***Abstract*—Filling dead pixels or removing uninteresting objects is often desired in the applications of remotely sensed images. In this paper, an effective image inpainting technology is presented to solve this task, based on multichannel nonlocal total variation. The proposed approach takes advantage of a nonlocal method, which has a superior performance in dealing with textured images and reconstructing large-scale areas. Furthermore, it makes use of the multichannel data of remotely sensed images to achieve spectral coherence for the reconstruction result. To optimize the proposed variation model, a Bregmanized-operator-splitting algorithm is employed. The proposed inpainting algorithm was tested on simulated and real images. The experimental results verify the efficacy of this algorithm.**

***Index Terms*—Inpainting, multichannel, nonlocal total variation (NLTV), remotely sensed image.**

I. Introduction

REMOTELY sensed images provide an unparalleled data source for land-surface mapping and monitoring. However, in some situations, such as detector failure or image damage, dead pixels will exist in the remotely sensed images. Dead pixels are those pixels whose measurement does not have any correlation with the true scene that is being measured [1]. The existence of dead pixels severely degrades the quality of the imagery. There are also some situations in which we need to remove or replace certain objects from the imagery for the sake of improving its application value. For example, we can remove pedestrians on a zebra crossing to reconstruct the zebra crossing on an aerial image or remove the map lettering and labels on a raster image map to obtain the original image data. The recovery of dead pixels and the removal of selected objects from remotely sensed images can be unified into one problem, i.e., image inpainting, which has been intensively studied in the field of digital image processing [2]–[4]. The purpose of image inpainting is to reconstitute the missing or damaged portions of the image, in order to make it more legible and to restore its unity.

To solve the inpainting problem of remotely sensed images, quite a few methods have been proposed. The approaches can be grouped into three categories. The first category comprises multitemporal-complementation-based approaches. These approaches consist of selecting the best measurement among a set of measurements acquired over a limited time period. Examples of these approaches are described in the works in [7]–[11]. In [7] and [8], in order to fill the Landsat-7 scan line corrector-off (SLC-off) gaps, the authors use the data from multiple Enhanced Thematic Mapper Plus (ETM+) scenes to provide complete ground coverage. The work in [9]–[11] studies the spectrotemporal relationships between a sequence of multitemporal images for the reconstruction of areas obscured by clouds or atmospheric disturbance, by statistics, training, or classification approaches.

The second category comprises the multispectral-complementation-based approaches. Most of these approaches make use of another clear and complete band of data to recover the contaminated band of data by modeling a relationship between the contaminated band and the reference band. Examples of this class of approaches can be found in [12]–[15]. The research in [12]–[14] describes how to restore the missing data of Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) band 6 by the use of other correlated bands, such as the commonly used Aqua MODIS band 7, or other image data. The work in [15] presents a haze-optimized-transformation method to detect the spatial distribution of haze and clouds in images and describes how to radiometrically adjust the visible-band imagery by the analysis of a visible-band space.

Both categories of approaches mentioned earlier need complementary information from other acquired images or spectral bands. However, in many cases, complementary images or bands cannot be acquired. Therefore, a third category of approaches is explored, which consists of filling in the missing data regions using the remaining parts in the image. The goal of the approaches in this category is to seamlessly synthesize a complete, visually plausible, and coherent image. Examples of these approaches are the recent works in [16]–[19]. The research in [16]–[18] describes how to synthesize the missing regions in remotely sensed images by propagating the geometrical structure from the remaining parts around the missing zone. In [19], the authors consider both destriping and inpainting as ill-posed inverse problems and develop a maximum *a posteriori* method based on a Huber–Markov model to solve these inverse

Manuscript received August 13, 2012; revised November 24, 2012; accepted December 23, 2012. This work was supported in part by the Major State Basic Research Development Program of China (973 Program) under Grant 2011CB707103 and in part by the National Natural Science Foundation of China under Grants 40930532, 41071269, and 41271376.

Q. Cheng, L. Zhang, and P. Li are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: chengqing@whu.edu.cn; zlp62@public.wh.hb.cn; pxli@lmars.whu.edu.cn).

H. Shen is with the School of Resource and Environmental Science, Wuhan University, Wuhan 460079, China (e-mail: shenhf@whu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2012.2237521

problems. Furthermore, some typical examples of digital image inpainting can be found in [2]–[6].

Since there is no need for auxiliary data, the third strategy of approaches is more attractive. Most prior researchers in this category have only made use of the local neighboring information to reconstruct the missing regions in remotely sensed images, which is far from sufficient. Moreover, few publications have considered the multiple bands of remotely sensed images as an ensemble to do the reconstruction. To remedy these weaknesses, this paper presents an efficient multichannel nonlocal inpainting approach for missing data synthesis. The proposed algorithm unites the advantages of nonlocal methods, which have a superior performance when dealing with textured images and large areas, and local methods, which are good at recovering geometric structures such as image edges. Furthermore, it takes advantage of the multichannel data of remotely sensed images to achieve spectral coherence for the reconstruction result. That is to say, our proposed method can achieve both the spatial and the spectral coherence. In order to optimize the proposed multichannel nonlocal total variation (NLTV) (MNLTV) inpainting model, a Bregmanized operator splitting (BOS) algorithm is employed.

The rest of this paper is organized as follows. In Section II, the basic image observation model and image inpainting model are described. In Section III, the proposed MNLTV inpainting model is formulated. The BOS optimization method is presented in Section IV. Section V contains the experimental results, and Section VI is the conclusion.

II. Basic Inpainting Model

A. Image Observation Model

Assuming that we have a multispectral image with some pixels missing, the degradation model can be written as

$$f = Au + \varepsilon \tag{1}$$

where $u = [u_1, u_2, \ldots, u_B]$ is the original true image, with the size of $M \times N \times B$, in which M represents the samples of the image, N stands for the lines of the image, and B is the number of bands. $f = [f_1, f_2, \ldots, f_B]$ is the observed degradation image, which is also of size $M \times N \times B$. A is a diagonal matrix with diagonal elements consisting of 0 and 1, with 0 representing the missing data. ε is additive noise with the same size as u and f. Our objective is to find the unknown target image u from the observed image f.

B. Image Inpainting Model

The multispectral image inpainting process is essentially an ill-posed inverse problem, which is similar to many other image processing problems, such as image denoising [20], destriping [19], superresolution reconstruction [21], [22], and others. The work in [23] and [24] provided approaches that achieve superresolution reconstruction and inpainting simultaneously. It is standard to use a regularization technique to make these inverse problems well posed. Regularization methods assume some prior information about the unknown image u, such as smoothness, sparsity [25], manifold [26], or small TV [31]. Based on a regularization technique, the inpainting problem for a multispectral image can be represented by the following model:

$$\hat{u} = \arg\min_u J(u) \quad \text{s.t. } Au = f \tag{2}$$

where $J(u)$ is the regularization item giving a prior model of the target image. The corresponding constrained problem for a noisy case is then written as

$$\hat{u} = \arg\min_u J(u) \quad \text{s.t. } \|Au - f\|^2 \le \sigma \tag{3}$$

where σ is the standard deviation of the noise ε.

III. MNLTV Method

Almost all the regularization methods mentioned earlier such as sparsity, manifold, and TV regularization belong to local methods which recover a pixel using only the local neighboring information; it is insufficient. In recent years, nonlocal methods for image denoising and inpainting have gained considerable attention. This is partly due to their superior performance in dealing with textured images. Local methods, on the other hand, have proved to be very effective for the recovery of geometric structure such as image edges. The synthesis of both types of methods is an important research area. Variation analysis is an appropriate tool for the unification of local and nonlocal methods. In recent research, single-channel nonlocal regularization has been developed for digital image processing [27], [28], and it has proved to be very effective. In order to take advantage of the multichannel data of remotely sensed images, an MNLTV inpainting model is presented in this paper.

A. Nonlocal Filter

The nonlocal methods in image processing are generalized from the Yaroslavsky filter and patch-based methods. The idea is to restore an unknown pixel using other similar pixels. The resemblance is regarded in terms of a patch centered at each pixel, not just the intensity of the pixel itself. In order to restore a pixel, the nonlocal methods average the other pixels with structures (patches) similar to that of the current one. This idea was generalized to a famous neighborhood denoising filter, the nonlocal means (NL-means) by Buades *et al.* in [29]. More precisely, given a reference image f, Ω is its pixel domain. We define the NL-means solution NLM_u of the image u at point x as

$$\mathrm{NLM}_u(x) := \frac{1}{C(x)} \sum_{y \in \Omega} w_f(x,y) f(y) \tag{4}$$

where

$$w_f(x,y) = \exp\left\{ -\frac{\left(G_a^* \left| f(x+\cdot) - f(y+\cdot) \right|^2\right)(0)}{h^2} \right\} \tag{5}$$

$$C(x) = \sum_{y \in \Omega} w_f(x,y) \tag{6}$$

where G_a is the Gaussian kernel with standard deviation a, $C(x)$ is the normalizing factor, h is a filtering parameter, and $f(x+\cdot)$ can be known as a patch centered at a point x. The patch $f(x+\cdot)$ of size $m \times m$ (m is chosen as an odd number) is given as (7), and the size of the patch is according to the noise intensity

$$f(x+\cdot) = f(x+t), \qquad t = \left[-\frac{m-1}{2}, \ldots, \frac{m+1}{2}\right]^2. \quad (7)$$

We recall that

$$\left(G_a^* \left|f(x+\cdot) - f(y+\cdot)\right|^2\right)(0) = \sum G_a(t) \left|f(x+t) - f(y+t)\right|^2. \quad (8)$$

Following (5), we can compute a weight function $w_f(x,y)$ between two points x and y by using the difference of patches around each point. This choice of weight is very efficient in reducing noise while preserving the textures and contrast of natural images. It is to be noted that the missing points are not needed to be excluded from the convolution summation since this operation does not affect the final solution of the proposed variation model.

B. Nonlocal Operators

In order to formulate the NL-means filter in a variational framework, Gilboa and Osher [27] defined variational framework-based nonlocal operators. In the following, we give the definitions of the nonlocal functions introduced in [27]. Let $\Omega \subset R^2$, $x, y \in \Omega$, $u(x)$ be a real function $u : \Omega \to \mathrm{R}$, and $w(x,y)$ be a weight function. Furthermore, $w(x,y)$ is assumed to be nonnegative and symmetric.

Nonlocal gradient $\nabla_w u(x) : \Omega \to \Omega \times \Omega$ is defined as the vector of all partial derivatives $\nabla_w u(x,\cdot)$ at x such that

$$(\nabla_w u)(x,y) := (u(y) - u(x)) \sqrt{w(x,y)}, \qquad \forall y \in \Omega. \quad (9)$$

We denote vectors as $\vec{p} = p(x,y) \in \Omega \times \Omega$, and the nonlocal divergence $(div_w \vec{p})(x) : \Omega \times \Omega \to \Omega$ is defined as the adjoint of the nonlocal gradient

$$(div_w \vec{p})(x) := \sum_{y \in \Omega} (p(x,y) - p(y,x)) \sqrt{w(x,y)}. \quad (10)$$

The nonlocal H^1 norm and the NLTV are defined as follows:

$$J^w_{NLH^1}(u) := \sum_{x \in \Omega} |\nabla_w u(x)|^2 = \sum_{x \in \Omega} \sum_{y \in \Omega} \times (u(x) - u(y))^2 w(x,y) \quad (11)$$

$$J^w_{NLTV}(u) := \sum_{x \in \Omega} |\nabla_w u(x)| = \sum_{x \in \Omega} \times \sqrt{\sum_{y \in \Omega} (u(x) - u(y))^2 w(x,y)}. \quad (12)$$

In this paper, we are interested in NLTV because, analogous to classical TV, the L^1 norm is generally more efficient than the L^2 norm for sparse reconstruction.

C. MNLTV

For multichannel images, Blomgren and Chan [30] presented a multichannel TV (MTV) regularization by coupling the channels

$$J_{\mathrm{MTV}}(u) := \sum_{x \in M \times N} \sqrt{\sum_{j=1}^{B} |\nabla u_j(x)|^2} \quad (13)$$

$$\nabla u_j(x) = \sqrt{(u_j(x) - u_j(rx))^2 + (u_j(x) - u_j(bx))^2} \quad (14)$$

where rx and bx represent the nearest neighbor to the right and below the pixel, respectively. The work of Yuan *et al.* [31] demonstrated that this MTV model has a powerful spectrally adaptive ability in remotely sensed image processing.

Inspired by the previously mentioned work, we extend NLTV to MNLTV and propose an MNLTV regularization for multispectral images

$$J^w_{\mathrm{MNLTV}}(u) := \sum_{x \in M \times N} \sqrt{\sum_{j=1}^{B} |\nabla_w u_j(x)|^2} = \sum_{x \in M \times N} \sqrt{\sum_{j=1}^{B} \sum_{y \in M \times N} (u_j(x) - u_j(y))^2 w(x,y)}. \quad (15)$$

D. MNLTV Inpainting Models

According to the inpainting models mentioned in Section II, the corresponding MNLTV inpainting model is

$$\hat{u} = \arg\min_u J^w_{\mathrm{MNLTV}}(u), \qquad \text{s.t. } Au = f. \quad (16)$$

In a noisy case, the model is then written as

$$\hat{u} = \arg\min_u J^w_{\mathrm{MNLTV}}(u), \qquad \text{s.t. } \|Au - f\|^2 \le \sigma. \quad (17)$$

IV. Optimization

The Bregman methods for image processing introduced in [32] have been demonstrated to be efficient optimization methods for solving sparse reconstruction, such as the l^1 norm and TV. Recently, based on a Bregman method, a well-performing optimization algorithm called BOS [33] was developed to provide a general algorithm framework for equality-constrained convex optimization. In this paper, the BOS algorithm is extended and used to optimize the MNLTV inpainting model in (16) and (17). The basic idea of this optimization algorithm can be stated as follows.

First, the constraint problems in (16) and (17) are enforced with the Bregman iteration process

$$\begin{cases} u^{k+1} = \arg\min_u \left(\mu J^w_{\mathrm{MNLTV}}(u) + \frac{1}{2}\|Au - f^k\|^2\right) \\ f^{k+1} = f^k + f - Au^{k+1}. \end{cases} \quad (18)$$

where μ is a positive parameter $\mu > 0$; it is the regularization term scale. The first subproblem can sometimes be difficult and slow to solve directly since it involves the inverse of the operator A and the convex function J^w_{MNLTV} in (16). The forward–backward operator splitting technique is used to solve the unconstrained subproblem in (18) as follows: For $i \geq 0$, $u^{k+1,0} = u^k$

$$\begin{cases} v^{k+1,i+1} = u^{k,i} - \delta A(u^{k+1,i} - f^k) \\ u^{k+1,i+1} = \arg\min_u \\ \quad \times \left(\mu J^w_{\mathrm{MNLTV}}(u) + \frac{1}{2\delta}\|u - v^{k+1,i+1}\|^2\right) \end{cases} \tag{19}$$

where δ is a positive parameter such that $0 < \delta < (2/\|A\|)$. We can see that the efficiency of the BOS algorithm depends on solvers for the u subproblem in (19). Here, we extend the split Bregman method proposed by Goldstein and Osher in [34] to the multichannel nonlocal case. It has proved to be a fast and efficient algorithm to minimize the MNLTV function in the subproblem

$$\hat{u} = \arg\min_u \left(\mu\delta J^w_{\mathrm{MNLTV}}(u) + \frac{1}{2}\|u - v\|^2\right). \tag{20}$$

The idea is to reformulate the problem as

$$\hat{u} = \min_u \left(\mu\delta \sum_{x\in M\times N} \sqrt{\sum_{j=1}^{B} d_j(x)^2} + \frac{1}{2}\|u - v\|^2\right) \quad \text{s.t. } d_j = \nabla_w u_j. \tag{21}$$

By enforcing the constraint problem with a Bregman iteration process, the extended MNLTV split Bregman algorithm using the MNLTV norm is given by

$$\begin{cases} (u^{k+1}, d^{k+1}) = \min_u \left(\mu\delta \sum_{x\in M\times N} \sqrt{\sum_{j=1}^{B} d_j(x)^2} \right. \\ \qquad\qquad \left. + \frac{1}{2}\|u - v\|^2 + \frac{\lambda}{2}\|d - \nabla_w u - b^k\|^2\right) \\ b^{k+1} = b^k + \nabla_w u^{k+1} - d^{k+1} \end{cases} \tag{22}$$

where λ is the scale of penalty term $\|d - \nabla_w u - b^k\|^2$; it is usually inversely proportional to the value of $\mu \cdot \delta$. The solution of (22) is obtained by performing alternately with the following two subproblems:

$$\begin{cases} u^{k+1} = \min_u \left(\frac{1}{2}\|u - v\|^2 + \frac{\lambda}{2}\|d^k - \nabla_w u\| - b^k\right) \\ d^{k+1} = \min_d \left(\mu\delta \sum_{x\in M\times N} \sqrt{\sum_{j=1}^{B} d_j(x)^2} \right. \\ \qquad\qquad \left. + \frac{\lambda}{2}\|d - \nabla_w u^{k+1} - b^k\|^2\right). \end{cases} \tag{23}$$

The subproblem for u^{k+1} consists of solving the linear system

$$(u^{k+1} - v) - \lambda div_w(\nabla_w u^{k+1} + b^k - d^k) = 0. \tag{24}$$

As the linear function in (24) is strictly diagonal, we can solve u^{k+1} by a Gauss–Seidel algorithm.

The d^{k+1} subproblem equation in (23) can be solved using a shrinkage operator as follows:

$$d^{k+1} = shrink\left(\sqrt{\sum_{j=1}^{B}\left(\nabla_w u_j^{k+1} + b^k\right)^2}, \frac{\mu\delta}{\lambda}\right) \tag{25}$$

where

$$shrink(x, \tau) = \frac{x}{|x|}\max\left(|x| - \tau, 0\right). \tag{26}$$

The weights computed from the initial image are not generally sufficient to give a good estimation. Therefore, it is necessary to adapt the weight used in nonlocal regularization, according to (5), during the iteration.

The optimization procedure of the MNLTV inpainting model is described in Algorithm 1.

Algorithm 1. MNLTV Inpainting Algorithm.

Initialization: $u^0 = v^0 = 0$, $f^0 = f, nOuter, nInner$.
While $k < nOuter$ and $\|Au^k - f\| > tolerance$ do
  for $i = 0$ *to* $nInner$ do
    $u^{k+1,0} = u^k$;
    update $v^{k+1,i+1} = u^{k+1,i} - \delta A(u^{k+1,i} - f^k)$;
    solve $u^{k+1,i+1}$ by the split Bregman algorithm:

$$u^{k+1,i+1} = \arg\min_u \left(\mu\delta J^w_{\mathrm{MNLTV}}(u) + \frac{1}{2}\|u - v^{k+1,i+1}\|^2\right)$$

  end
  update the nonlocal weight according to u^{k+1} : $w^k = w(u^{k+1})$ by (5);
  update $f^{k+1} = f^k + f - Au^{k+1}$;
End

The convergence of the BOS optimization algorithm has been proved theoretically in [33]. Moreover, we experimentally verified that it has a very fast convergence rate. Then, the computational complexity of the proposed algorithm is provided. We assume that the number of the bands of the input image is L, and the total number of the pixels in each single band is N. Then, in Algorithm 1, the computational complexity for updating v is $O(NL)$. In the step of solving u by the split Bregman algorithm, it is resolved into two subproblems as in (23), and the u subproblem is just a fast Gauss–Seidel algorithm, with a linear computational complexity of $O(N^2L)$, while the d subproblem is an efficient soft threshold/shrinkage operator, with a linear computational complexity of $O(N^2L^2)$, and then, in the update of nonlocal weight w, the computational complexity is $O(N^2L)$. Finally, the computational complexity for updating f is $O(NL)$. Taking all the aforementioned parts into account, the total computational complexity for Algorithm 1 is $O(N^2L^2)$.

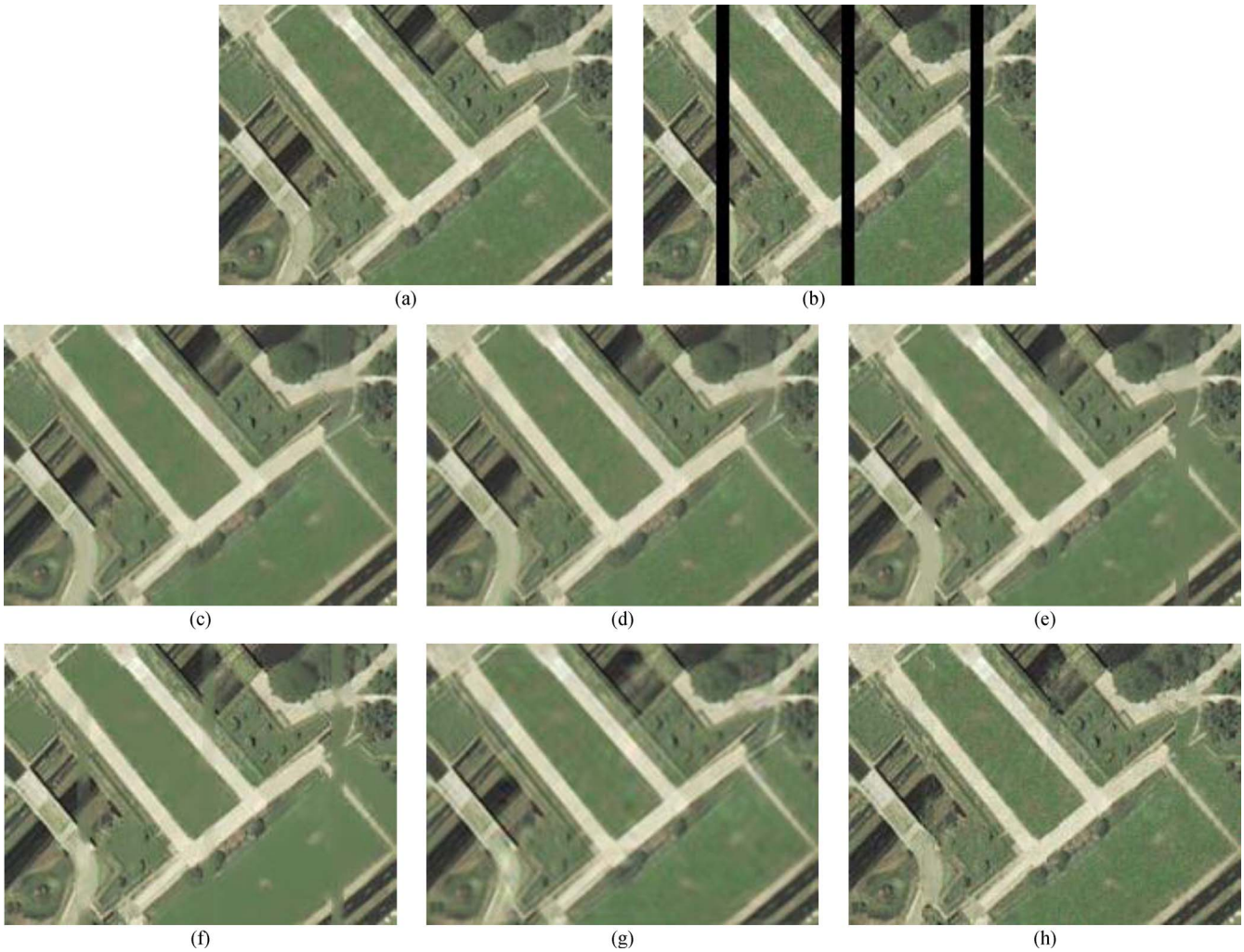


Fig. 1. Experimental results for the recovery of noise and vertical dead lines. (a) Original image. (b) Simulated image contaminated by noise and dead lines. Recovered images using the following: (c) Proposed algorithm, (d) NLTV inpainting algorithm, (e) MTV inpainting model, (f) NL-means inpainting method, (g) MCA algorithm, and (h) Criminisi method.

V. Experiments

A. Experiments on Simulated Data

In the simulation experiments, we conduct three experiments to test and verify the efficacy of our proposed MNLTV inpainting algorithm. Results of the three experiments are shown in Figs. 1–4. The peak signal-to-noise ratio (PSNR) index is used to give a quantitative assessment of the results of the simulated experiments from the gray-level fidelity aspect. The structural similarity (SSIM) index [35] is used to give a quantitative assessment of the results from the structure-level fidelity aspect. Recently, some image quality assessment indices based on human vision system have been proposed, such as metric Q [36] and the multiscale geometric analysis-based indices [37], which provide a quantitative measure of image content (i.e., sharpness and contrast in visually salient geometric features such as edges). In this paper, we use the metric Q index to give a quantitative assessment of the inpainting result from a human vision aspect

$$\text{PSNR} = 10\log_{10}\left(\frac{255^2 * MN}{\|\hat{u}-u\|^2}\right) \tag{27}$$

$$\text{SSIM} = \frac{(2\mu_u\mu_{\hat{u}}+C_1)(2\sigma_{u\hat{u}}+C_2)}{(\mu_u^2+\mu_{\hat{u}}^2+C_1)(\sigma_u^2+\sigma_{\hat{u}}^2+C_2)} \tag{28}$$

$$Q = s_1\frac{s_1-s_2}{s_1+s_2} \tag{29}$$

where MN is the total number of pixels in the multichannel image, $\hat{u}$ and u represent the recovered image and the original clear image, and μ_u and $\mu_{\hat{u}}$ represent the average gray values of the original clear image and the recovered image, respectively. σ_u and $\sigma_{\hat{u}}$ represent the variances of the original clear image and the recovered image, respectively. $\sigma_{u\hat{u}}$ represents the covariance between the original clear image and the recovered image. C_1 and C_2 are two constants. s_1 and s_2 are two singular values of the gradient matrix of the recovered image.

The first simulated data test of the proposed algorithm is for the recovery of vertical dead lines in a QuickBird image with the resolution of 0.6 m. We use three bands of the red, green, and blue to undertake this test. Fig. 1(a) shows the original subimage. In this experiment, the simulated image is contaminated by dead lines of 7-pixel widths. Moreover, we simulate the additive noise by adding different variance

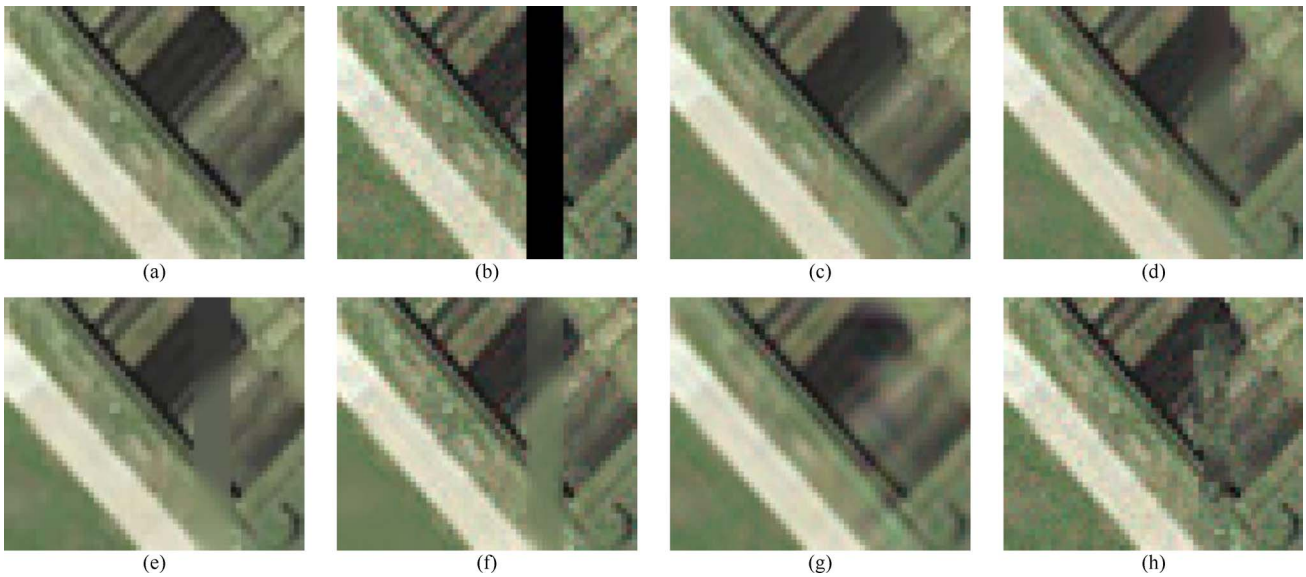


Fig. 2. (a)–(h) Detailed regions cropped from Fig. 1(a)–(h).

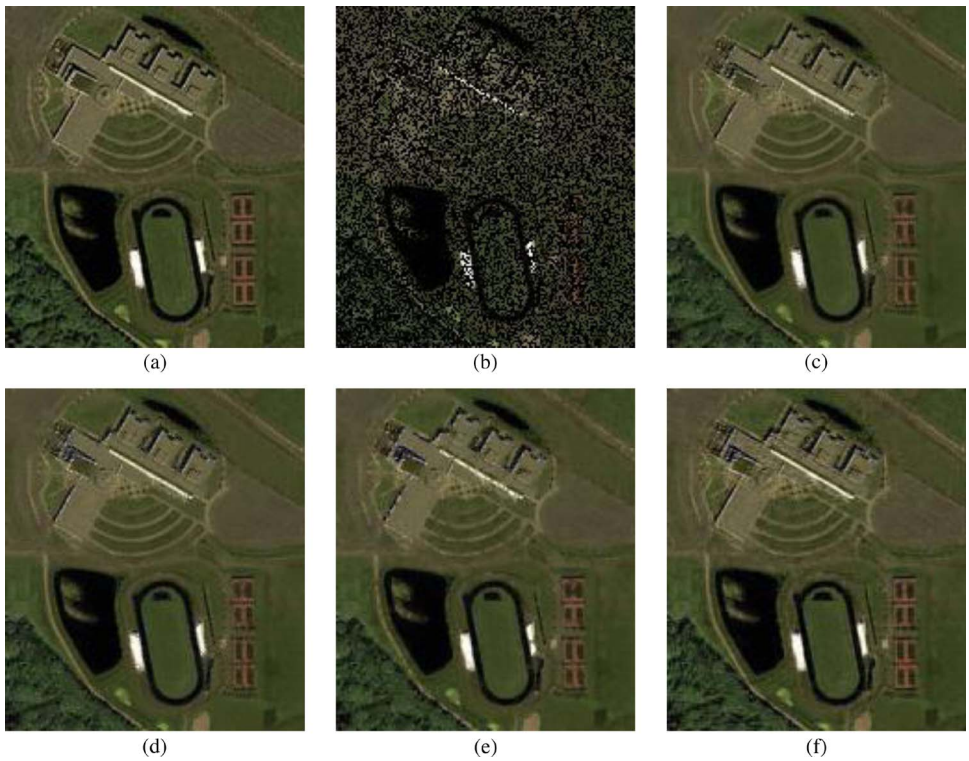


Fig. 3. Inpainting experiment for the recovery of random dead pixels. (a) Original image. (b) Simulated image contaminated by 50% dead pixels. Recovered images using the following: (c) Proposed algorithm, (d) NLTV inpainting algorithm, (e) MTV inpainting model, and (f) MCA algorithm.

zero-mean Gaussian noise in different bands. This is shown in Fig. 1(b). To make a comparative analysis, the proposed MNLTV inpainting algorithm is compared with the NLTV inpainting algorithm [27], the MTV inpainting algorithm [31], the NL-means inpainting method [29], the morphological component analysis (MCA) method [5], and the Criminisi method [6]. The recovery results of each method are shown in Fig. 1(c)–(h).

From Fig. 1 and its zoomed detailed regions in Fig. 2, it can be seen that the proposed MNLTV algorithm gives better denoising and inpainting results, compared to the results of the other methods. In the MNLTV result, on the one hand, the noise is suppressed more thoroughly, and on the other hand, the recovery of dead lines is spatially continuous, with more convincing visual quality. The result of the NLTV algorithm is not as sharp as the result of our proposed algorithm, and some noise still remains in the smooth regions because the noise intensity difference between different bands is not taken into consideration and an equal denoising strength is used in all bands. For the MTV algorithm, the result suffers from a "staircasing" effect; moreover, it is incapable of performing narrow and long

Fig. 4. Experimental results for the removal of map lettering. (a) Original image. (b) Simulated photographic map. Recovered images using the following: (c) Proposed algorithm, (d) NLTV inpainting algorithm, (e) MTV inpainting model, (f) NL-means inpainting method, (g) MCA algorithm, and (h) Criminisi method.

region connection. In the result of the NL-means inpainting method, the narrow and long regions also cannot be connected completely, and the denoising result is oversmoothed. The result using the MCA algorithm has a strong ripple effect and spectral distortion, and the recovery of the dead lines shows severe artifacts, particularly in the homogeneous region. From the result of the Criminisi method, we can see that it is incapable of denoising at all: The noise in the resulting image has the same intensity as the contaminated image, and a lot of spurious detailed information appears in the recovery result of the dead region due to its block filling process.

It is to be noted that all the test image data in our paper are red, green, and blue true color composite of the original acquired data.

Next, the proposed algorithm is tested with another type of inpainting problem, in which dead pixels are randomly distributed. Fig. 3(a) shows an original GeoEye-1 image with the resolution of 1.65 m. In this experiment, the simulated image is contaminated by 50% random dead pixels in all bands, as shown in Fig. 3(b). The proposed MNLTV algorithm is compared with the NLTV algorithm, MTV model, and MCA algorithm. The other two comparative methods used in the first experiment, the NL-means inpainting method and the Criminisi method, fail to work in this experiment because the missing data are distributed in almost every patch of the image and occupy a large-scale area. The ideas of patch matching and patch filling, which the other two methods are based on, cannot work in this situation. The recovery result of each method is shown in Fig. 3.

From Fig. 3, it can be seen that our proposed algorithm achieves a nice inpainting result. The image quality is very good, with most of the detailed information recovered well, and the ground objects can be clearly distinguished. Although the result of the NLTV algorithm is quite similar to that of the proposed MNLTV algorithm from a visual perspective, the quantitative measure of our proposed algorithm is improved, as shown in Table I. In the result of the MTV algorithm, we can see that it shows some blurring, and some edge information is not recovered completely. For the MCA algorithm, the ripple effect still exists.

The third simulation experiment involves the removal of unwanted map lettering on remotely sensed images. Remote sensing mapping, as a branch of cartography, provides a significant application for remotely sensed images and makes maps more intuitive and richer in content. This kind of map, which is known as a photographic map or image map, is made up of the appropriate remotely sensed image and several auxiliary feature notes on the image. Once the map is printed out or rasterized, the feature notes cannot be separated from the image, i.e., we cannot regain the original image. However, we sometimes need to get the original remotely sensed image from a map for other uses, particularly when the image data are rare or valuable. In such a case, the map lettering needs to be removed, and the original image needs to be reconstructed. The proposed MNLTV method can be used to efficiently undertake this task. In this experiment, figure notes were manually created on an aerial image in order to simulate a photographic map, as shown in Fig. 4(b). The original aerial image with a 0.2-m resolution is shown in Fig. 4(a). Fig. 4(c) shows the result of our proposed method. The results of the other comparative methods are shown in Fig. 4(d)–(h).

Fig. 4(c) shows that our proposed method is capable of removing the map lettering and recovering a clear image that is much closer to the actual image. Although it may not be exactly precise, the visual quality is convincing enough. In the results

TABLE I
PSNR, SSIM, AND METRIC Q VALUES OF THE THREE SETS OF SIMULATED EXPERIMENTAL RESULTS

		Contaminate image	MNLTV	NLTV	MTV	NL-means	MCA	Criminisi
Fig. 1	RSNR	15.68	33.04	31.84	30.61	29.21	31.15	27.88
	SSIM	0.7178	0.9414	0.9084	0.8944	0.8857	0.8715	0.8057
	Q	--	69.66	64.59	55.61	62.79	61.72	57.94
Fig. 3	RSNR	14.48	31.23	30.86	28.92	--	29.21	--
	SSIM	0.1912	0.8982	0.8853	0.8810	--	0.8651	--
	Q	--	51.83	48.53	36.09		45.32	
Fig. 4	RSNR	19.30	28.16	27.95	27.08	26.65	26.61	25.79
	SSIM	0.7580	0.9307	0.9152	0.8984	0.8961	0.8894	0.8865
	Q	--	52.58	50.95	49.24	49.89	51.21	50.48

Fig. 5. (a)–(h) Detailed regions cropped from Fig. 4(a)–(h).

of the NLTV inpainting algorithm and the NL-means inpainting algorithm, it appears that there is spectral distortion in some inpainted regions. The results of the other approaches all show the same flaws as the results of the foregoing experiments. For the convenience of visual judgment, a series of detailed regions cropped from Fig. 4(a)–(h) is shown in Fig. 5(a)–(h).

The effectiveness of the proposed MNLTV inpainting method can also be illustrated by the quantitative assessment. The PSNR, SSIM, and metric Q values of Figs. 1, 3, and 4 are shown in Table I to give an overall quantitative assessment of the results. It can be seen from Table I that, for the results in every simulation experiment, the PSNR values using the MNLTV model are higher than those of the other methods. All three groups of simulated data experimental results indicate that the proposed MNLTV method can provide a better and more robust inpainting result.

In nonlocal methods, the filtering parameter h in (5) can exert considerable influence on the inpainted image. Here, we test the sensitivity of this parameter. For each of the three sets of simulated experimental data, we select a series of different h's to implement the inpainting methods of NLTV and the proposed MNLTV. The relationship of the acquired PSNR value and h is shown in Fig. 6. From Fig. 6(a)–(c), it can be seen that the PSNR values using the MNLTV method fluctuate less than those using the NLTV method with the change in h. On two sides of the peak point, the PSNR values using the MNLTV method descend much more slowly, while the PSNR values using the NLTV method show a dramatic decline. This

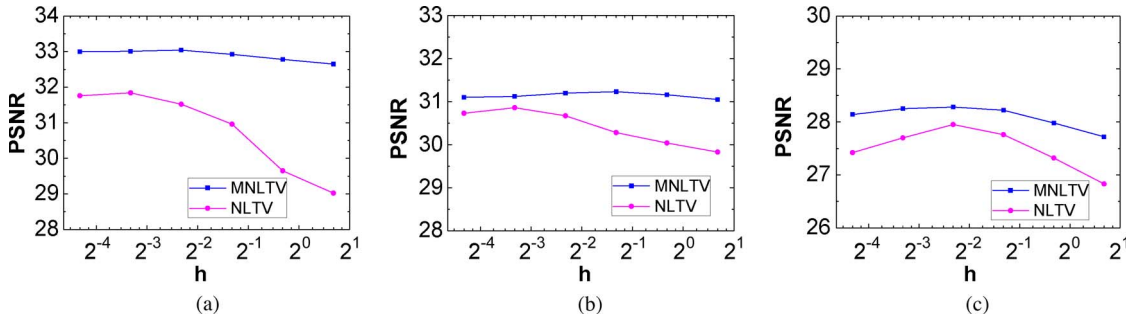


Fig. 6. PSNR values corresponding to different values of h for each of the three simulated experimental results: (a) First simulated experiment, (b) second simulated experiment, and (c) third simulated experiment. The x-axis represents the value of h, and the y-axis represents the PSNR values of the resulting images.

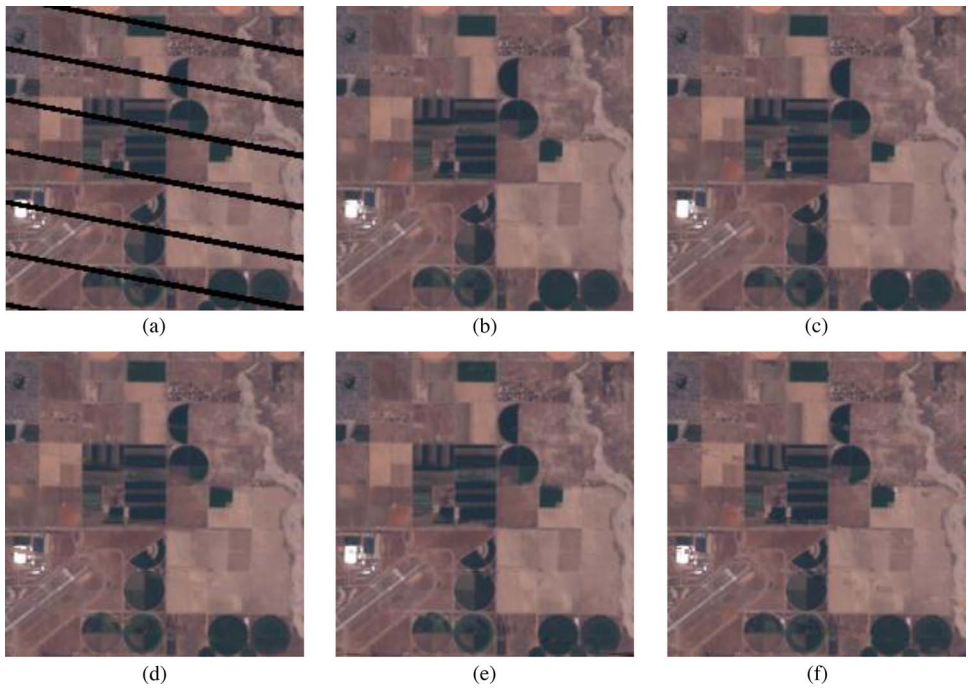


Fig. 7. Experimental results for filling the gaps in a Landsat-7 ETM+ SLC-off image. (a) Original ETM+ SLC-off image. Recovered images using the following: (b) Proposed algorithm, (c) MTV inpainting model, (d) NL-means inpainting method, (e) MCA algorithm, and (f) Criminisi method.

phenomenon can be explained as follows. From (5), we can see that h indicates the exponential decay rate that is also the decay rate of weight W. The decay of W has a significant effect on the nonlocal gradient $\nabla_w u_j(x)$ from (9) within each band. However, from (15), we can deduce the multichannel nonlocal gradient magnitude at point x

$$|\nabla_{mw} u(x)| = \sqrt{\sum_{j=1}^{B} |\nabla_w u_j(x)|^2}. \tag{30}$$

It can be seen that, for the MNLTV model, the nonlocal gradient magnitude of every point is determined by the combination of multichannel data, so it will be more steady and robust than just using a single band of data, as in the NLTV model.

B. Real Data Experiments

The SLC of the Landsat-7 ETM+ sensor failed in 2003, resulting in about 22% of the pixels per scene not being scanned. The SLC failure has seriously limited the scientific application of ETM+ data. Our first real data experiment involves filling the gaps in a Landsat-7 ETM+ SLC-off image. The test image is shown in Fig. 7(a), with a size of 200 $*$ 200 pixels and a 4–5-pixel width of dead lines, and the resolution is 30 m. The result of our proposed MNLTV method is shown in Fig. 7(b). The other comparative results are shown in Fig. 7(c)–(f).

From Fig. 7 we can see that the image using our proposed method has a better filling result than that of the other existing methods: It appears more spatially continuous, without stripes, and the edge information and the detailed information are well filled. However, in the results of the MTV and NL-means inpainting method, the narrow and long regions cannot be connected. The result using the MCA algorithm shows some artifacts in the homogeneous vegetation region. There are still visible stripes left in the result of the exemplar-based method. For the convenience of visual judgment, a series of detailed regions cropped from Fig. 7(a)–(h) are shown in Fig. 8(a)–(h).

In the second real data experiment, the proposed MNLTV algorithm is tested with the reconstruction of a zebra crossing

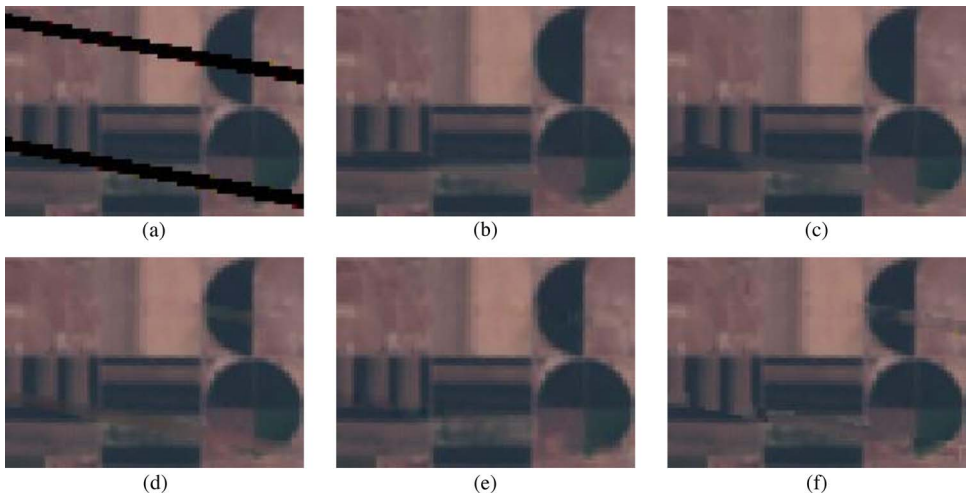


Fig. 8. (a)–(f) Detailed regions cropped from Fig. 7(a)–(f).

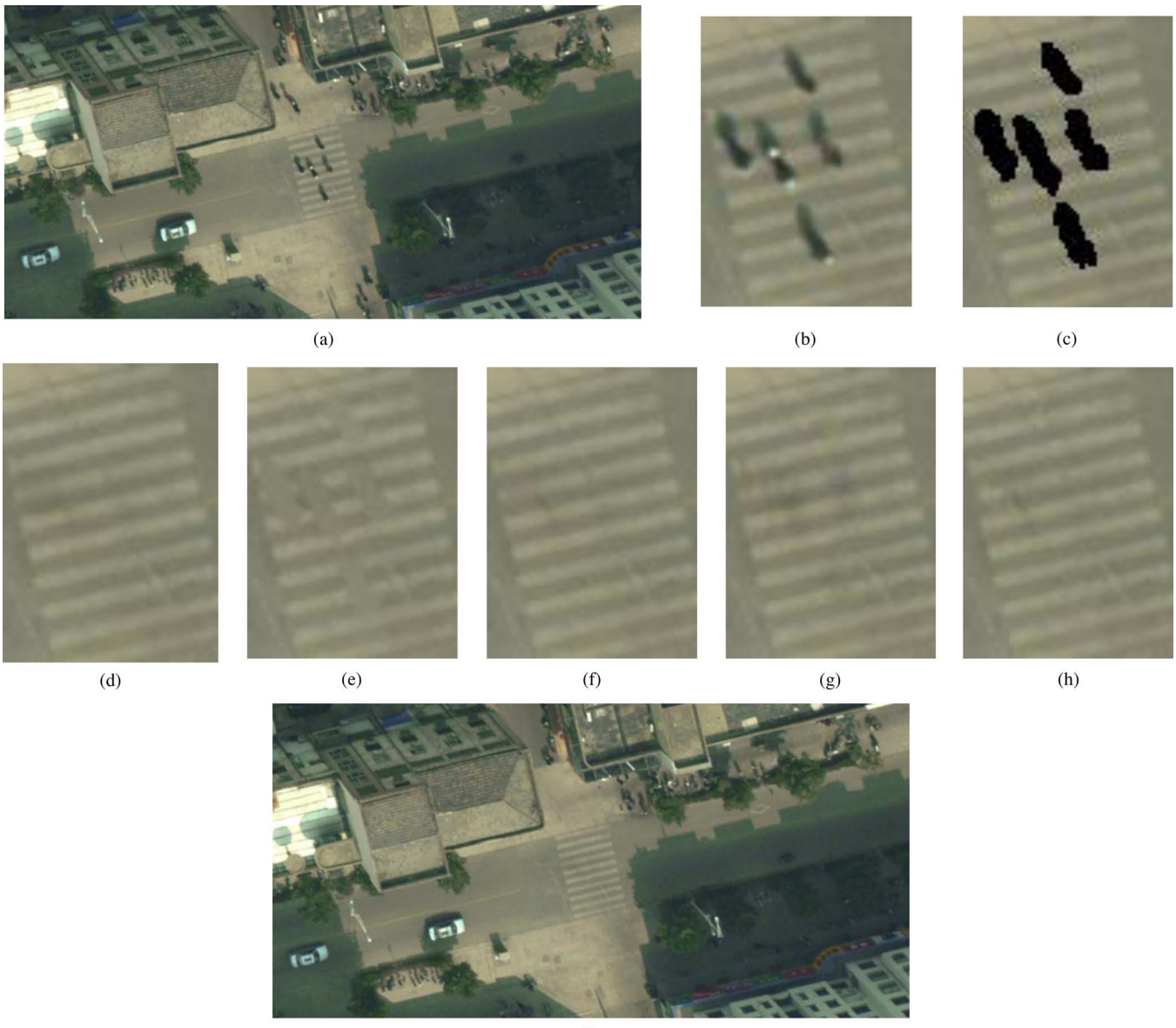


Fig. 9. Experimental results for the reconstruction of a zebra crossing. (a) Original image. (b) Zebra crossing contaminated with pedestrians. (c) Mask of the area to be reconstructed. Reconstructed zebra crossings using the following: (d) Proposed algorithm, (e) MTV inpainting model, (f) NL-means inpainting method, (g) MCA algorithm, (h) Criminisi method, and (i) Reconstructed image.

Fig. 10. Experimental results for the removal of map lettering. (a) Original photographic map. Recovered images using the following: (b) Proposed algorithm, (c) MTV inpainting model, (d) NL-means inpainting method, (e) MCA algorithm, and (f) Criminisi method.

on an aerial image. With high-resolution images providing an important data source for terrain observation, how to extract the required geoinformation rapidly and accurately has become an important research field. Traffic signs are important land objects in high-resolution images, such as the zebra crossing, which marks a safe area for people to pass across roads. However, in many cases, zebra crossings in high-resolution images are contaminated with other features, such as pedestrians, as shown in Fig. 9(a), which seriously interfere with the recognition and extraction of the zebra crossing. Therefore, it is valuable and necessary to reconstruct zebra crossings in images, not only for enriching the information of a map but also for prompting warnings in vehicle navigation, which helps to reduce the occurrence of traffic accidents.

This experiment is performed on an aerial image with a resolution of 0.1 m, as shown in Fig. 9(a). Our goal is to remove the pedestrians on the zebra crossing and reconstruct a complete zebra crossing. Fig. 9(c) shows the portion of the pedestrian-contaminated region that should be reconstructed (the black region). The region to be reconstructed is manually segmented by the user, for the moment. A further investigation of the pedestrian-contaminated area segmentation could help to automate the entire reconstruction process. The reconstructed results are shown in Fig. 9(d)–(h). It can be seen that we finally get an image with a complete zebra crossing in it, as shown in Fig. 9(i). The zebra crossing recovery results shown in Fig. 9 indicate that the proposed MNLTV algorithm is able to perform repetitive texture reconstruction well. The missing regions reconstructed using our proposed method achieve both the spatial and the spectral consistency with the surrounding texture.

In the third real data experiment, the proposed inpainting method is tested on an image map. The aim is to reconstruct image data without map lettering. The map that we used as the test data is an area of Washington, DC. Fig. 10(a) shows a part of this map, with a size of $350 * 300$ pixels. The reconstructed result of our proposed MNLTV method is shown in Fig. 10(b). The other four comparative results are shown in Fig. 10(c)–(f). A series of detailed regions cropped from Fig. 10(a)–(f) is shown in Fig. 11(a)–(f). From the results, it can be seen that our proposed MNLTV method is able to remove the map lettering and reconstruct different ground objects well; however, it still produces blurring to a certain extent in the regions where the ground features are complex. Another competitive result for this image is that of the Criminisi method.

The metric Q values of Figs. 7, 9, and 10 are shown in Table II to give an overall quantitative assessment of the three sets of real data experimental results.

VI. Conclusion

In this paper, we present an MNLTV inpainting model to deal with the remotely sensed image reconstruction problem. The proposed algorithm was applied to Landsat-7 ETM+ SLC-off image gap filling, zebra crossing reconstruction from an aerial image, and map lettering removal from a photographic map. All the simulated data experiments and real data experiments indicate that the reconstruction results using our proposed

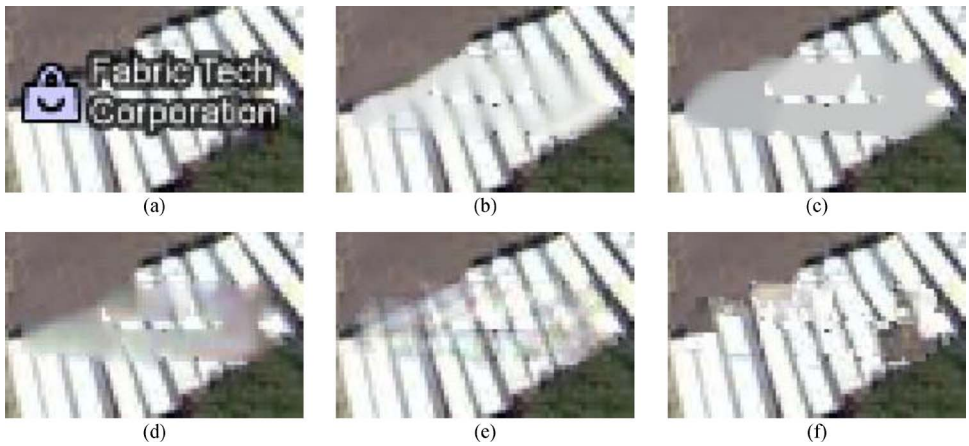


Fig. 11. (a)–(f) Detailed regions cropped from Fig. 10(a)–(f).

TABLE II
METRIC Q VALUES OF THE THREE SETS OF REAL DATA EXPERIMENTAL RESULTS

	MNLTV	MTV	NLM	MCA	Criminisi
Fig. 7	35.51	34.12	33.86	34.74	33.70
Fig. 9	41.77	34.43	40.11	39.68	39.41
Fig. 10	96.71	93.98	93.40	94.35	95.18

algorithm are very effective, no matter whether the image is noisy or not, how the missing data are distributed, or whether the reconstructed region is homogeneous, edge or texture. Furthermore, through an analysis of the filtering parameter, it was demonstrated that the proposed MNLTV model is more robust than the NLTV model. Nevertheless, there may still be room for improvement of our proposed method. In the proposed MNLTV model, the nonlocal weighting function calculation [see (5)] is restricted to within a single band, and our future work will focus on extending the weight calculation from a single band to multiple bands. Moreover, some regularization parameters such as μ in (18) and λ in (22) are set fixedly by manual selection, and we will try to make them to be automatic and self-adapting in the future work.

REFERENCES

[1] B. M. Ratliff, J. S. Tyo, J. K. Boger, W. T. Black, D. L. Bowers, and M. P. Fetrow, "Dead pixel replacement in LWIR microgrid polarimeters," *Opt. Exp.*, vol. 15, no. 12, pp. 7596–7609, Jun. 2007.

[2] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. ACM SIGGRAPH Conf. Comput. Graph.*, New Orleans, LA, 2000, pp. 417–424.

[3] T. F. Chan and J. Shen, "Mathematical models for local nontexture inpaintings," *SIAM J. Appl. Math.*, vol. 62, no. 3, pp. 1019–1043, 2002.

[4] A. Bugeau, M. Bertalmío, V. Caselles, and G. Sapiro, "A comprehensive framework for image inpainting," *IEEE Trans. Image Process.*, vol. 19, no. 10, pp. 2634–2645, Oct. 2010.

[5] M. Elad, J. L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Appl. Comput. Harmon. Anal.*, vol. 19, no. 3, pp. 340–358, Nov. 2005.

[6] 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, Sep. 2004.

[7] J. Storey, P. Scaramuzza, G. Schmidt, and J. Barsi, "Landsat 7 scan line corrector-off gap-filled product development," in *Proc. Pecora 16 Conf.*, 2005, pp. 23–27.

[8] M. Pringle, M. Schmidt, and J. Muir, "Geostatistical interpolation of SLC-off Landsat ETM+ images," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 6, pp. 654–664, Nov. 2009.

[9] S. Lee and M. M. Crawford, "An adaptive reconstruction system for spatially correlated multispectral multitemporal images," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 4, pp. 494–508, Jul. 1991.

[10] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 442–455, Feb. 2006.

[11] A. B. Salberg, "Land cover classification of cloud-contaminated multitemporal high-resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 377–387, Jan. 2011.

[12] P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Restoration of Aqua MODIS band 6 using histogram matching and local least squares fitting," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 613–627, Feb. 2009.

[13] H. Shen, C. Zeng, and L. Zhang, "Recovering reflectance of AQUA MODIS band 6 based on within-class local fitting," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 1, pp. 185–192, Mar. 2011.

[14] I. Gladkova, M. D. Grossberg, F. Shahriar, G. Bonev, and P. Romanov, "Quantitative restoration for MODIS band 6 on Aqua," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2409–2416, Jun. 2012.

[15] 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. 2/3, pp. 173–187, Oct. 2002.

[16] 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.

[17] F. Zhou, Z. Wang, and F. Qi, "Inpainting thick image regions using isophote propagation," in *Proc. IEEE ICIP*, 2006, pp. 689–692.

[18] L. Lorenzi, F. Melgani, and G. Mercier, "Inpainting strategies for reconstruction of missing data in VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 914–918, Sep. 2011.

[19] 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.

[20] A. B. Hamza and H. Krim, "A variational approach to maximum a posteriori estimation for image denoising," in *Proc. EMMCVPR*, Sep. 2001, vol. 2134, pp. 19–33.

[21] K. Zhang, X. Gao, D. Tao, and X. Li, "Multi-scale dictionary for single image super-resolution," in *Proc. IEEE CVPR*, 2012, pp. 1114–1121.

[22] Q. Yuan, L. Zhang, H. Shen, and P. Li, "Adaptive multiple-frame image super-resolution based on U-curve," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3157–3170, Dec. 2010.

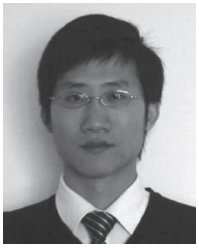
[23] M. K. Ng, H. Shen, S. Chaudhuri, and A. C. Yau, "Zoom-based super-resolution reconstruction approach using prior total variation," *Opt. Eng.*, vol. 46, no. 12, pp. 127 003-1–127 003-11, 2007.

[24] M. K. Ng, H. Shen, L. Zhang, and E. Lam, "A total variation based super-resolution reconstruction algorithm for digital video," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 74585, pp. 1–16, 2007.

[25] N. Guan, D. Tao, Z. Luo, and J. Shawe-Taylor, "MahNMF: Manhattan Non-Negative Matrix Factorization," *arXiv:1207.3438*, 2012.
[26] B. Geng, D. Tao, C. Xu, L. Yang, and X. S. Hua, "Ensemble manifold regularization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 6, pp. 1227–1233, Jun. 2012.
[27] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *Multiscale Model. Simul.*, vol. 7, no. 3, pp. 1005–1028, 2008.
[28] G. Peyré, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," *Proc. Eur. Conf. Comput. Vis. (ECCV)*, pp. 57–68, 2008.
[29] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.
[30] P. Blomgren and T. F. Chan, "Color TV: Total variation methods for restoration of vector-valued images," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 304–309, Mar. 1998.
[31] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral-spatial adaptive total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3660–3677, Oct. 2012.
[32] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489, 2005.
[33] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 253–276, Jul. 2009.
[34] L. Bregman, "The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex optimization," *USSR Comp. Math. Math.*, vol. 7, no. 3, pp. 200–217, 1967.
[35] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[36] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3116–3132, Dec. 2010.
[37] X. Gao, W. Lu, D. Tao, and X. Li, "Image quality assessment based on multiscale geometric analysis," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1409–1423, Jul. 2009.

Qing Cheng received the B.S. degree in the geographic information system from Wuhan University, Wuhan, China, in 2010, where she is currently working toward the Ph.D. degree in the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing.

Her research interests include image restoration, inpainting, and remote sensing image enhancement.



Huanfeng Shen (M'11) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

In July 2007, he joined the School of Resources and Environmental Science, Wuhan University, where he is currently a Full Professor. His research interests include image processing (for quality improvement), remote sensing application, data fusion, and assimilation. He has published more than 60 research papers. He has been supported by several talent programs, including the New Century Excellent Talents by the Ministry of Education of China (2011) and the Hubei Science Fund for Distinguished Young Scholars (2011).

Liangpei Zhang (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is currently with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, as the Head of the Remote Sensing Division. He is also a "Chang-Jiang Scholar" Chair Professor appointed by the Ministry of Education, China. He has more than 260 research papers and is the holder of five patents. He is currently the Principal Scientist for the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang regularly serves as a Cochair of the series The International Society for Optical Engineering Conferences on Multispectral Image Processing and Pattern Recognition, Conference on Asia Remote Sensing, and many other conferences. He edits several conference proceedings, issues, and the Geoinformatics Symposiums. He also serves as an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, International Journal of Ambient Computing and Intelligence, International Journal of Image and Graphics, International Journal of Digital Multimedia Broadcasting, Journal of Geo-spatial Information Science, and Journal of Remote Sensing. He is a fellow of The Institution of Electrical Engineers, Executive Member (Board of Governor) of the China National Committee of International Geosphere-Biosphere Programme, Executive Member for the China Society of Image and Graphics, and others.

Pingxiang Li (M'06) received the B.S., M.S., and Ph.D. degrees in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1986, 1994, and 2003, respectively.

Since 2002, he has been a Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. His research interests include photogrammetry and synthetic aperture radar image processing.