Recovering Missing Pixels for Landsat-7 ETM + SLC-off Images with No Reference Images

Yanjun Chen[1,a], Xinke Li[2,b] and Yuhong Zhang[1,c,*]

[1]School of Mathematical Science Harbin Normal University Harbin, China
[2]College of Medical Informatics Chongqing Medical University
a. cheenyanjun@126.com, b. zmdcg@126.com, c. zhangyuhong77@163.com
**Yuhong Zhang*

Keywords: Interpolation, regression analysis, unsupervised classification, gap filling, no-reference-image recovering.

Abstract: A recovering method for the missing pixels without using any reference images is proposed to fill the gap filling for the Landsat-7 SLC-off images. The image is firstly preprocessed for the gap location, unsupervised classification and local average gray. Then, regression analysis is used to calculate the gray-level-change slope for the pixels on both sides of the stripe respectively in three directions of each missing pixel. Under the unsupervised classification criteria, we choose the filling direction of the current pixel of the gap by judging the classifications of the pixels on both sides of the gap in three directions. Next, different algorithms are used to fill the boundary pixels and the non-boundary pixels in the stripes. For the non-boundary pixels, cubic spline interpolation is used to calculate the gray value of the current pixel. Finally, all the stripes are filtered by an adaptive filter. The experimental results show that the recovery algorithm by the proposed method has better visual effect comparing with other three methods, and no stripe trace can be found in the unsupervised classification image, which lays a foundation for further research on geographic spatial analysis.

1. Introduction

Landsat-7 is the seventh in the US terrestrial satellite program and was launched on April 15, 1999. The images scanned after the scanning line corrector (SLC) of the Landsat-7 satellite ETM + sensor permanently failed on May 31, 2003, can only be seen in an effective continuous scan within a range of approximately 22 kilometers wide directly below the center of the image. From the center to both sides, the scanning patterns are no longer arranged in parallel, but overlap alternately [1]. The stripe width between overlapping scanning patterns gradually increased from 1 pixels to 14 pixels, forming a huge wedge-shaped gap from the middle of the image to both sides, resulting in about 25% data loss in the image [2]. The images acquired before and after SLC failure are designated as "SLC-on" and "SLC-off" image, respectively. Due to the huge research and practical value of the data, scholars have carried out various methods to fill the missing pixels in the SLC-off images of Landsat-7.

The filling methods for missing pixels in remote sensing images are mainly divided into two categories, one is the non-reference image method and the multi-reference-image method according to whether reference images are used.

Some non-reference-image methods use spatial or frequency domain filters to remove stripes [3~6]. Some of the no-reference-image methods adopt histogram matching algorithm, and most of them use direct interpolation algorithm [7~10]. Studies have shown that filtering methods are often inferior to direct interpolation methods [11].

The no-reference-image methods are mostly used for pixel filling of MODIS sensor images with stripes of 1 to 2 pixels wide. Since the gray value of the pixel in the stripes obtained from the pixels outside the stripes are not the true reflectance of the ground, a visible boundary exists between the filled stripes and the surrounding image [12].

Based on the earliest multi-reference-image method for stripes filling of Landsat-7 SLC-off images developed by the Landsat team of the joint United States Geological Survey / National Aeronautics and Space Administration (USGS / NASA), more methods were proposed by authors [13, 14]. Some multi-reference-image methods are based on Landsat-7 images, and others are based on the images from such satellite sensors as MODIS. Some methods involve multi- scale segmentation, filtering or image restoration.

Zhang et al. and Pringle et al. used Kriging or Co-Kriging technique from geostatistics to recover the stripes in 2007 and 2009 respectively [15~17]. Although the geostatistics methods can fill gaps with relatively satisfactory results for the homogeneous forest area, for heterogeneous landscapes with an area smaller than the size of the local moving window, the effect of histogram matching is not ideal.

In 2011, Chen proposed the Neighbor Similar Pixel Interpolation (NSPI) method to fill the stripe pixels, improving the filling effect of heterogeneous landscape areas [18]. In 2013, Zeng et al. developed a multi-reference-image recovery method based on weighted linear regression (WLR) and NSPI method. Zeng et al. Compared the Local Linear Histogram Matching (LLHM), NSPI and WLR methods. Experiments have shown that relatively satisfactory results can be obtained by their scheme [19].

In 2015, Sulong proposed a mean-filter method and an inverse distance weighting (IDW) interpolation method [20]. The clear border between the stripes recovered by this method and the background image appears. In 2017, comparing various filters such as morphological operations, mean filtering and median filtering, Jain et al. believed that the morphological opening operation filtering is superior, however the filling effect is not ideal [21].

The introduction above shows that gap filling methods at present in Landsat-7 SLC-off image are mainly based on multi-reference-image methods, which have the advantage that for most stripes the calculation values of pixels' gray scale are accurate and reliable. However, the reference images were obtained under different observation conditions, such as the sunlight, atmosphere, vegetation growth, land humidity, and the changing sensor status during shooting. Moreover, along with abrupt changes caused by artificial or earthquake, etc. will all cause a certain difference between the reference images and the target image. For some missing pixels without reference points in any reference images still need no-reference-image filling method. The methods using reference images captured by other sensors are limited by the different spectral bands and resolution of the sensors.

For the disadvantages of the above methods, this paper proposes a method for gap filling in Landsat-7 ETM + SLC- off images under the unsupervised classification criteria, in which no reference images are needed. This method makes full use of the spatial correlation between the current pixel to be filled and the pixels around it in the image.

2. Methods

The process of the proposed filling scheme can be divided into four parts as shown in Figure 1.

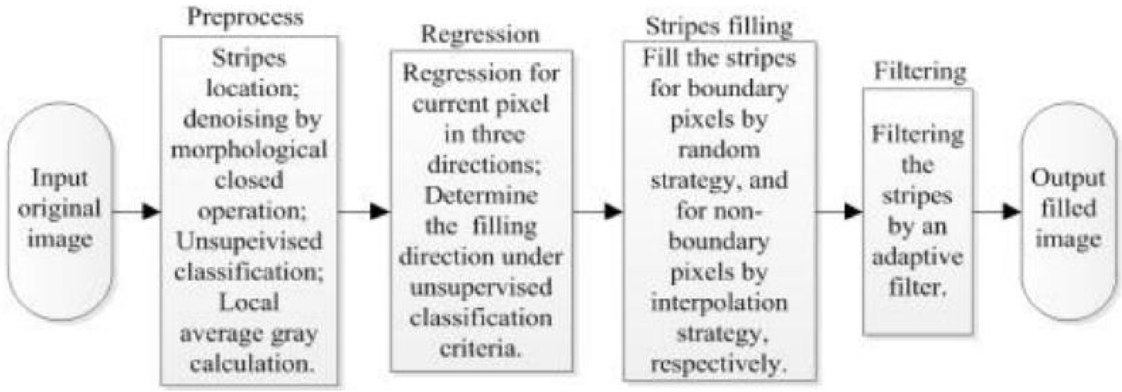

Figure 1: The recovery flow chart for missing pixels in Landsat-7 ETM+ SLC- off remote sensing image.

Firstly, the stripes must be located accurately after the original image is input. In the stripe graph, effective stripes region can be obtained using the morphological closed operation to remove the noise. And, five classes can be acquired using unsupervised classification criteria to determine the filling direction later. In the preprocessing stage, local average gray scale in 17×17 window area is also calculated for each stripe pixel. Next, we carry out regression analysis on the gray levels of six pixels (three pixels from each side of the stripe) in the vertical direction, 45º direction and 135º direction for the current pixel, and calculate the slope. Then, under the unsupervised classification criteria, the filling direction of the current pixel can be determined according to whether the pixels outside two sides of the stripe in the three directions belong to the same class. Thirdly, the stripe pixel are filled by different algorithms according to whether they are boundary points in the filling direction of the current pixel. For boundary points, they are filled by the pixel gray level randomly selected from nearest two pixels outside the stripe in the filling direction. For non-boundary points, the pixel gray level is obtained by the cubic spline interpolation of the six pixels' gray levels from both sides of the stripe in the filling direction. The closest gray level from the six pixels is selected to replace the interpolation to fill the current non-boundary pixels. Finally, we use a max-min filtering strategy to filter the stripe area, which can not only remove the noise pixels, but also avoid the overall smoothness for the target image, and the image texture information is preserved. The following details the key recovery steps for the missing pixels in the stripes.

2.1. Stripes Locating

To recover the missing pixels in the stripe area, the stripes must be located accurately. The stripes are wedge- shaped and gradually widened from the middle to both sides. Stripes consist of black, red and blue-green pixels can be found in the color image composed of 5, 4 and 3 bands captured by Landsat-7. The pixels in the upper side of the stripe in Figure 2(a) and the narrowest wedge in Figure 2(c) are red stripes. The 1-pixel-width area at the bottom edge of the stripe in Figure 2(a) and at the narrowest wedge area in Figure 2(b) are blue-green stripes. The gray value of these missing pixels in the stripes all needs to be recovered.

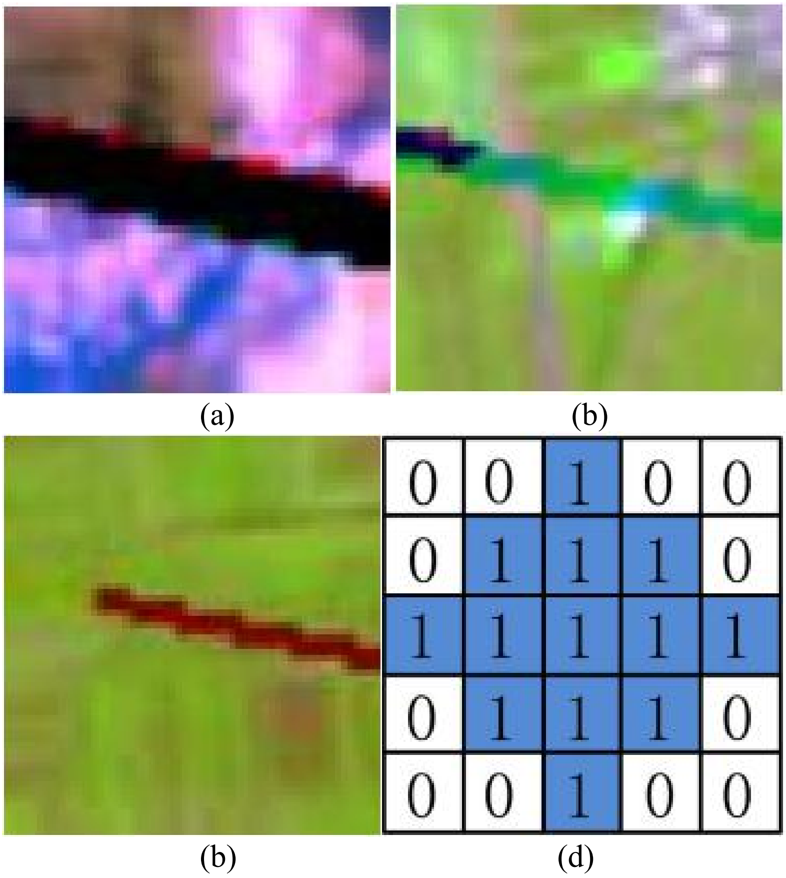

Figure 2: (a), (b), (c) Several typical kinds of stripe pixels; (d) Structural element of morphological closed operation.

We extracted the stripe image from the SLC-off image with the pixel value in the stripe area marked by 0 and other area marked by 255. The mistaken extraction for some small scattered non-striped regions needs to be removed through the morphological corrosion and dilation operations that can remove small areas and small bridges. We use a disc-shaped structural element shown in Figure 2(d) to process the stripe image by morphological closed operation. If A denotes a stripe image and B denotes the structural element, the morphological closed operation of B first expanding and then corroding A can be expressed as:

$$A \bullet B = (A \oplus B) \ominus B \tag{1}$$

For the area with gray value of 0, the closed operation can be used to remove small bridges and small points smaller than the structural element, and the 0-gray value area larger than the structural element do not change before and after the closed operation. Experiments show that the noise can be removed effectively by closed operations.

2.2. Unsupervised Classification

All the pixels in the original SLC-off remote sensing image including the ones in the stripes are divided into 5 classes. As shown in Figure 3.

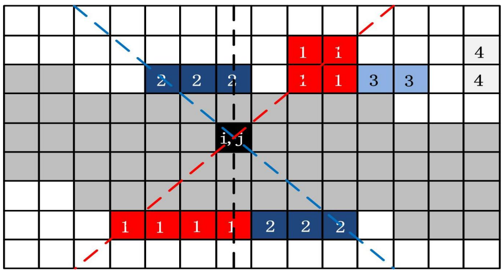

Figure 3: Determining of filling direction according to the classes of the pixels outside both sides of the stripe under the unsupervised.

The gray area in the middle shows a stripe composed of missing pixels. In the vertical direction, outside the stripe, the neighbor pixels up the stripe are blue class 2 and the neighbor pixels down the stripe are red class 1, which are not of the same class. It means that the pixels in vertical direction are not continuous, and the vertical direction can not be the filling direction. Along the 45º direction, both the upper right and lower left adjacent pixels outside the stripe belong to red class 1, so the 45º direction can be selected as a candidate for the filling direction. Similarly, the 135º direction can also be selected as a candidate for the filling direction.

2.3. Regression Analysis

From a typical SLC-off remote sensing image which consists of diverse land cover types, we cut an area with a 6- pixel-width stripe in it. In order to determine the filling direction, we calculate the slope of the pixels' gray value by a linear regression analysis for each candidate direction. For example, in the 45º direction, for the current pixel (i, j), 3 pixels each side (6 pixels in total) are searched out from the bottom left to the top right outside the stripe. The gray values are 85, 97, 109, 84, 109 and 121, respectively. Regression analysis is conducted for the 6 pixels' gray values. Similar calculations can be performed for other candidate directions.

We set 0.05 as the significance level of regression analysis. Figure 4 shows that the confidence interval of the pixel with grayscale of 84 does not contain 0, and the slope is 2.7429 after removing this pixel. We conduct regression analysis for all the stripe pixels, and remove the points whose residual confidence interval excludes 0, and recalculate the slope value. If the significance level is no more than 0.05, the slope is directly calculated by fitting.

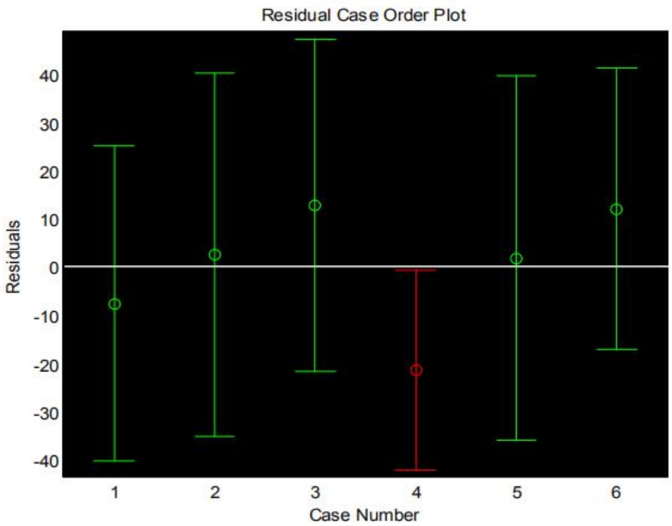

Figure 4: Residual diagram that the noisy confidence interval excluding 0 exist.

2.4. Determining the Filling Direction

Obtaining the slope of each direction by linear regression analysis, we use unsupervised classification result as criteria to determine the filling direction.

(1) If in only one direction the adjacent pixels to the both sides of the stripe belong to the same class, the direction can be selected as the filling direction.

(2) If in two or more directions the adjacent pixels to both sides of the stripe belong to the same class, the slopes in these directions are sorted from small to large.

(3) If in all of the three directions the adjacent pixels to both sides of the stripe do not belong to the same class, step(2) will be conducted.

(4) The directions of the minimum and the sub minimum slopes in the results of (2) are selected as candidate filling directions. Equation (2) can be used to determine the filling direction finally,

$$|G_1 - G_0| \geq |G_2 - G_0| \tag{2}$$

Here, G1 and G2 are the average gray value of 6 pixels in the direction of slope minimum and sub minimum respectively, and G0 is the average gray value of local 17×17 window. If (2) is true, the first direction, i.e. we select the direction with the minimum slope as the filling direction. Otherwise, the sub minimum slope direction is selected as the filling direction. It is worth noting that using (2) can help to highlight the bright or dark texture.

2.5. Stripe Pixels' Gray Value Calculating

When the width of the stripe is no more than 4 pixels, all the pixels are defined as boundary pixels, and for the stripe with more than four pixels width, the upper two pixels and the lower two pixels are defined as boundary pixels. The boundary pixels are filled by randomly selected pixels from adjacent two points above or below the stripe in the filling direction.

For the non-boundary pixels of the stripe, we use the cubic spline interpolation strategy to fill. The 6 adjacent pixels outside both sides of the stripe in the filling direction are involved as the nodes for the interpolation calculation. The current pixel coordinates are substituted into the curve equation, and the corresponding gray-value interpolations are obtained as shown in Figure 5. For all the non-boundary stripe pixels, the cubic spline interpolation algorithm can be performed in their respective filling directions to obtain the interpolation.

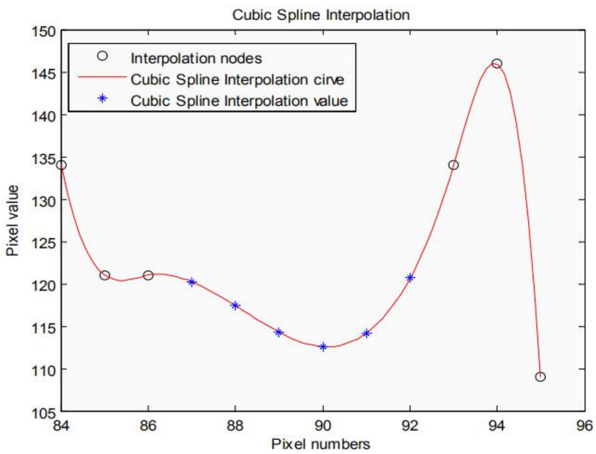

Figure 5: Cubic spline interpolation of the current pixel to be filled.

Since the interpolation may be a gray value not exist in the remote sensing image, we find the gray level closest to the interpolation from the six pixels to fill the current pixel, as shown in (3),

$$N(i,j,m_0)=\min\left\{\left|P(i,j)-N(i,j,m)\right| \mid m=1\cdots 6\right\} \quad (3)$$

Here $P(i, j)$ is the interpolation gray value of the current pixel in the stripe, and $N(i, j, m)$ are the gray level of the 6 nodes participating in interpolation calculation. $N(i, j, m0)$ is the gray level of the node closest to the interpolation of the current pixel. The gray value of the current pixel $P(i, j)$ is replaced by $N(i, j, m0)$.

2.6. Filtering the Stripes

A 3×3 window is used to filter the stripes. We use double thresholds to determine whether the pixel at the center of the window is a noise point. If the gray level of central pixel is higher than 3×3 local average gray level Pl(1+K) or lower than the 3×3 local average gray level Pl(1- K), where K is the empirical value, this point needs to be removed as noise. Sorting the gray levels of the surrounding 8 neighborhood pixels from small to large, we take the fourth or fifth value to replace the current bright pixel's gray value. Since we only filter the noise points, our filter is different from the median filter, which filters all pixels. Our filter avoid the image smoothing effectively.

3. Experimental Results

In this section, a simulation image was filled by three typical methods and the proposed method in this paper, respectively. Then, we recovered the typical Landsate-7 SLC-off remote sensing image which contains band 5, 4 and 3 of Zhalong wetland in China with the proposed method, and showed the recovering results.

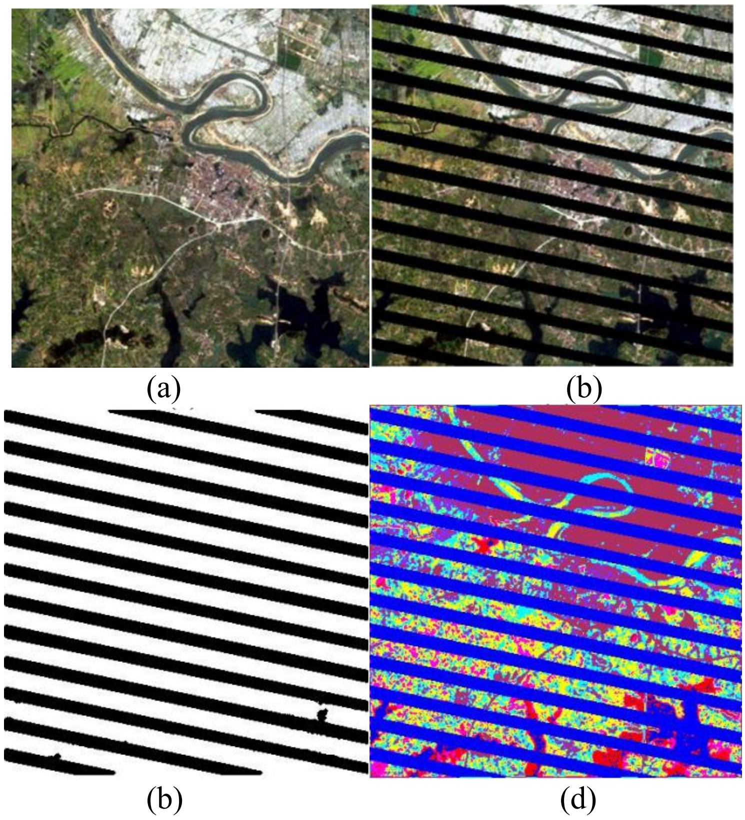

Figure 6: Simulation image: (a) Original image with no stripe; (b) Original image added stripes; (c) Stripes image; (d) Unsupervised classification.

3.1. Simulated SLC-off Image Filling

In the simulation experiment, stripes with 14-pixel width were added to the original image of Figure 6(a), and the simulation image of Figure 6(b) was obtained. In the simulation image, the stripes area accounts for about 42.05%, which is much more than 25% in SLC-off image.

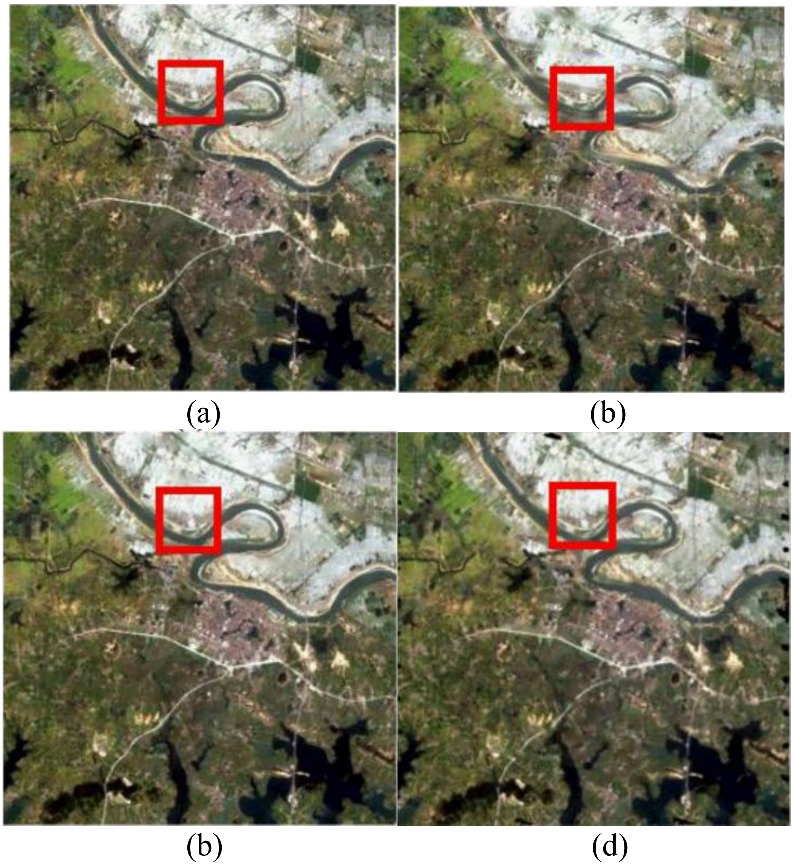

Figure 7: Filling results: (a) LLHM; (b) NSPI; (c) WLR; (d) Proposed method.

We utilized morphological closing operation on the stripe area, most of the noise were removed as shown in Figure 6(c). Figure 6(d) shows the result of unsupervised classification, in which all pixels were divided into five classes.

LLHM, NSPI, WLR methods and the proposed method were used to fill the stripes in Figure 6(b). Figure 7(a), (b), (c), (d) show the results, respectively. Comparing the red rectangular area, we can see that the proposed method has fewest filling errors. The simulation results prove that the proposed method has relatively satisfied performance.

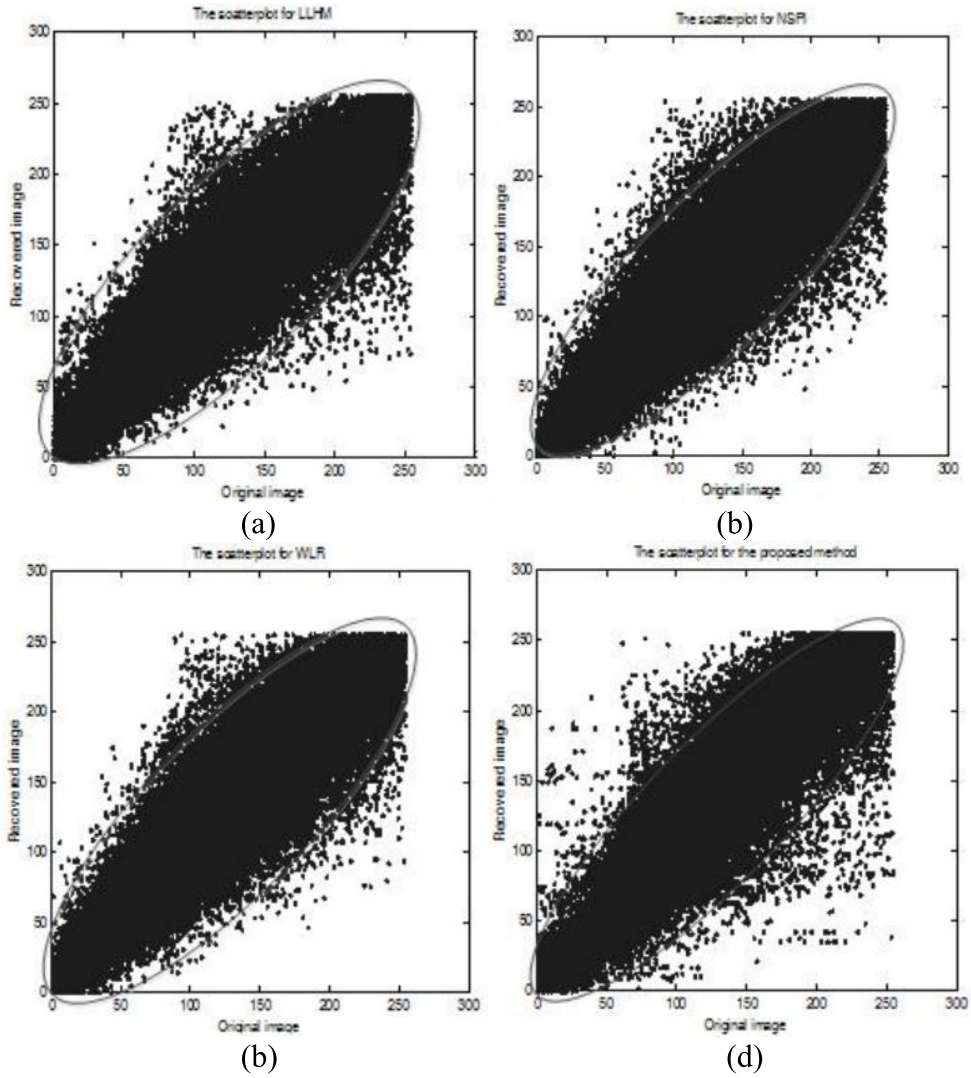

Figure 8: Scatter plots: (a) LLHM; (b) NSPI; (c) WLR; (d) Proposed method.

The gray values scatter plots of the recovered images versus the original image for the four methods are shown in Figure 8(a), (b), (c) and (d), respectively. Little difference can be found between them. Most of the scatter coordinates fall in the elliptical region with 45° line as the symmetry axis, which implicates linear relation for the pixels' gray levels before and after filling.

3.2. SLC-off Image Filling

Figure 9(a) shows a part of an original SLC-off image, which is a 400×400 pixels size partial color image of Zhalong Wetland (46º52′~47º32′N, 123º47′~124º37′E) in China, and Figure 10(b) shows the filling result by proposed method. Figure 9(c) and (d) enlarged the filtering results of the rectangular area in Figure 9(b). We can see that the image texture is preserved while most of the noise is removed in our filtering strategy.

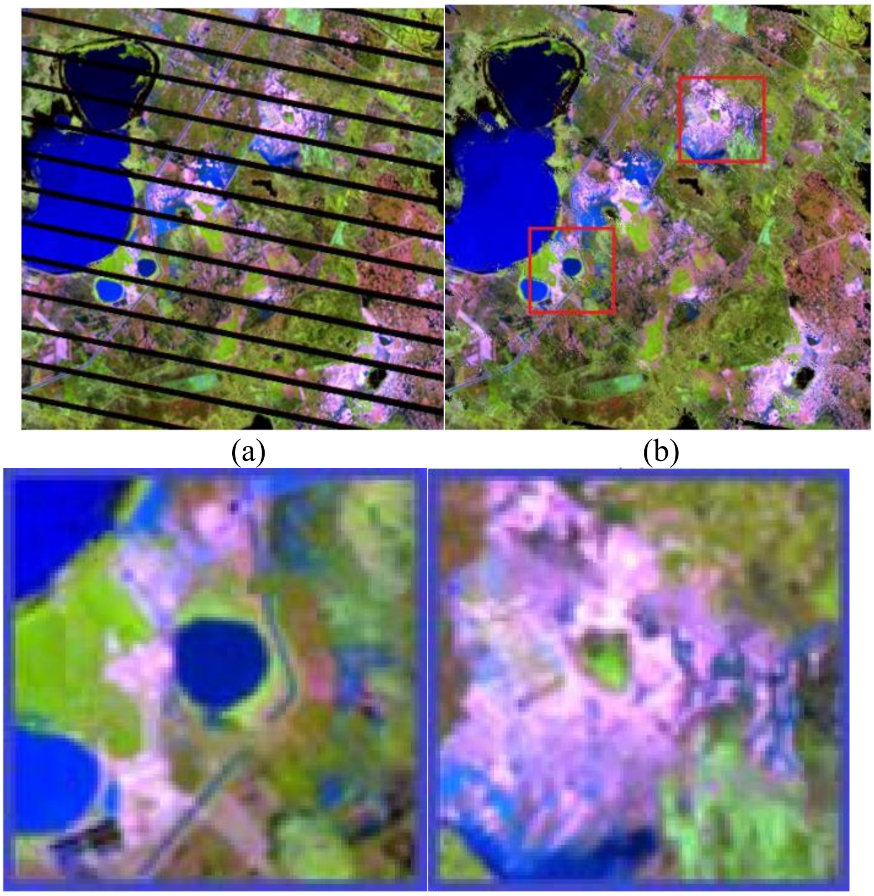

Figure 9: SLC-off image filling: (a) Original image; (b) The filling result; (c) and (d) Filtering the red rectangular area.

3.3. Results Evaluating

The simulation experiments can directly reflect the accuracy of the calculation results, and the scatter plots can evaluate the overall effect from the statistical aspect as shown in Figure 9. In order to quantitatively evaluate the filling effect, we compared the root mean square error (RMSE) in (4) for the four methods.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(R_i - O_i)^2}{n}} \tag{4}$$

Here, R_i is the gray level of the current pixel in recovering image, O_i is the gray level of the current pixel in the original image and n is the number of the recovered pixels.

We compared the root mean square error (RMSE) of the recovered pixels in the simulation image. 120931 stripe pixels are involved to calculate RMSE. Table I shows that the RMSE of the proposed method is better than LLHM and NSPI method, only slightly worse than WLR method.

Table 1: Comparison for several methods.

Method	LLHM	NSPI	WLR	The proposed method
RMSE	21.6273	18.0075	17.4006	17.5904

4. Conclusions

In this paper, we have proposed an interpolation based filling algorithm to recover missing pixels in Landsat-7 SLC- off image without reference images. The algorithm first locates the stripes and removes the non stripe pixels noises by morphological closing operation. Next, regression analysis is performed to calculate the slope of pixel gray change in three directions. The filling direction is determined by comparing the slope in three direction under unsupervised classification criteria. In the filling direction, cubic spline interpolation is used to calculate the gray value of the current pixel by the gray of the pixels on both sides neighbor to the stripe. And then a max-min adaptive filter removes most of the bright and the dark pixels in the recovered stripes. In order to avoid introducing new gray values not existing in the original image, the closest gray value is selected from the original image to replace the calculated gray value to fill the current pixel. The advantages of this method are no reference images needed, accurate recovery and no introduction of new gray values. Experiments show that the proposed method can effectively fill the stripe area.The future research should focus on the algorithm of highlighting such texture as roads and rivers, and improving recovering speed.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (NSFC) under Grant [No. 41771195].

References

[1] Maganioti, A.E., Chrissanthi, H.D., Charalabos, P.C., Andreas, R.D., George, P.N. and Christos, C.N. (2010) Cointegration of Event-Related Potential (ERP) Signals in Experiments with Different Electromagnetic Field (EMF) Conditions. Health, 2, 400-406.

[2] Bootorabi, F., Haapasalo, J., Smith, E., Haapasalo, H. and Parkkila, S. (2011) Carbonic Anhydrase VII—A Potential Prognostic Marker in Gliomas. Health, 3, 6-12.

[1] USGS (2003). Preliminary assessment of the value of Landsat-7 ETM+ data following scan line corrector malfunction. Available online at http://landsat.usgs.gov/documents/ SLC_off_Scientific_Usability.pdf.

[2] Arvidson, Terry , et al, "Landsat-7 Long-Term Acquisition Plan," Photogrammetric Engineering & Remote Sensing vol. 72, pp. 1137– 1146, October 2006.

[3] J. J. Simpson, J. R. Stitt, and D. M. Leath, "Improved finite impulse response filters for enhanced destriping of geostationary satellite data,"Remote Sens. Environ., vol. 66, no. 3, pp. 235–249, December 1998.

[4] J. S. Chen, Y. Shao, H. D. Guo, W. Wang, and B. Zhu, "Destriping CMODIS data by power filtering," IEEE Trans. Geosci. Remote Sens.,vol. 41, pp. 2119–2124, September 2003.

[5] Z. D. Yang, J. Li, W. P. Menzel, and R. A. Frey, "Destriping for MODIS data via wavelet shrinkage," in Proc. SPIE-Applications with Weather Satellites, vol. 4895, pp. 187–199, 2003.

[6] J. S. Chen, H. Lin, Y. Shao, and L. M. Yang, "Oblique striping removal in remote sensing imagery based on wavelet transform," Int. J. Remote Sens., vol. 27, no. 8, pp. 1717–1723, April 2006.

[7] B. K. P. Horn and R. J. Woodham, "Destriping Landsat MSS imagesby histogram modification," Comput. Graph. Image Process., vol. 10, pp. 69–83, January 1979.

[8] M. Wegener, "Destriping multiple detector imagery by improved histogram matching," Int. J. Remote Sens., vol. 11, pp. 859–875, may 1990.

[9] Rakwatin P, Takeuchi W, Yasuoka Y, "Stripe noise reduction in MODIS data by combining histogram matching with facet filter," IEEE Transactions on Geoscience and Remote Sensing, vol. 45, pp. 1844–1856, June 2007.

[10] HUANG X, ZHOU R, LUO H, "Method for Removing the Stripe Noises with Neighboring-region Interpolation Algorithm in Modis Images," Geospatial Information, 2008, 1.

[11] Tsai F, Chen W W, "Striping noise detection and correction of remote sensing images," IEEE Transactions on Geoscience and remote sensing, vol. 46, pp. 4122–4131, December 2008,

[12] Desai M, Ganatra A, "Survey on gap filling in satellite images and inpainting algorithm," International Journal of Computer Theory and Engineering, vol. 4, pp. 341, March 2012,

[13] USGS (2004). Phase 2 gap-fill algorithm: SLC-off gap-filled products gap-fill algorithm methodology. landsat.usgs.gov/documents/L7SLCGapFilledMethod.pdf.
[14] Liu, Xingpo, and Yongsheng Ding, "Auxiliary pixel data selection for recovering Landsat ETM+ SLC-off images," The Egyptian Journal of Remote Sensing and Space Science, vol. 21, pp. 273 – 284, March 2018.
[15] Zhang, C., Li, W., & Travis, D., "Gaps-fill of SLC-off Landsat ETM plus satellite image using a geostatistical approach," International Journal of Remote Sensing, vol. 28, pp. 5103−5122, 2007.
[16] Pringle, M. J., Schmidt, M., & Muir, J. S., "Geostatistical interpolation of SLC-off Landsat ETM plus images," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 64, pp. 654−664, 2009.
[17] Yin, Gaohong, Gregoire Mariethoz, and Matthew F. McCabe, "Gap- filling of Landsat-7 imagery using the direct sampling method," Remote Sensing, vol 9, pp. 12, January 2017.
[18] Chen, Jin, et al., "A simple and effective method for filling gaps in Landsat ETM+ SLC-off images," Remote sensing of environment, vol.115, vol.4, pp. 1053–1064, Apr. 2011.
[19] Chao Zeng, Huanfeng Shen, and Liangpei Zhang, "Recovering missing pixels for Landsat ETM+ SLC-off imagery using multi- temporal regression analysis and a regularization method," Remote Sensing of Environment, vol. 131, pp. 182–194, April 2013.
[20] Sulong, Ghazali, and Asmaa Sadiq, "Single and multi-source methods for reconstruction the gaps in Landsat-7 ETM+ SLC-off images," Research Journal of Applied Sciences, Engineering and Technology, vol.11, no. 4, pp. 423–428, October 2015.
[21] Jain, Madhavi, and A. P. Dimri, "Efficacy of Filtering Techniques in Improving Landsat SLC-off Thermal Infrared Data," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, pp. 271–284, January 2017.