A morphology-stitching method to improve Landsat SLC-off images with stripes

Aghamohamadnia M and Abedini A
Division of Remote Sensing, Department of Surveying and Geomatics, University of Tehran, Iran

Abstract: Stripes are artifacts in satellite images caused by various factors such as hardware defects. In some cases, these artifacts are introduced by some mitigating algorithms like Landsat SLC-off (Scan Line Corrector) gap-filling methods of LLHM (Local Linear Histogram Matching) and AWLHM (Adaptive Window Linear Histogram Matching), which leave stripes as a byproduct. To improve Landsat SLC-off images with stripes, we propose an algorithm involving some hypothetical stripe-crossing stitch lines using the mean pixel value of the stitch lines.

Key words: SLC-off; stripe; data analysis; morphology stitching; Landsat; LLHM; AWLHM

1 Introduction

Landsat Data Continuity Mission (LDCM) is one of the world's most enduring earth-observation missions; it holds one of the largest collections of images used in remote-sensing applications. In the Landsat 7 satellite, there is a very large amount of data loss caused by defective hardware in acquiring images with the onboard ETM+ sensor. This sensing instrument forms images with a small array of sensors that sweeps back and forth across its flight path, serving as a platform while flying over the earth's surface. This whisk-broom style sensor has a functioning mechanism known as Scan-Line Corrector (SLC), which keeps the whisk-broom module aligned with the cross-track motion of the satellite's ground track or path[1]. The Landsat 7 SLC failed on May 31, 2003 and caused about 22% pixels un-scanned in these images[2]. Additionally, the TM sensor onboard Landsat 5 has been suspended since November 2011. The next generation of LDCM, namely Landsat 8, was scheduled to be launched in 2013 (USGS). Since a very long period of this continuing program has gaps and some ETM+ data has been lost (since 2003), seeking improvement of the acquired images is necessary. Several approaches to estimate the missing values have been proposed in the past; they may be categorized according to the use of the techniques and the way they approach the phenomena: Basic and computationally-cheap image processing methods; outsource auxiliary data-set utilization; combination of various methods and stochastic-related image processing techniques; and selective, customized, and precise approaches.

Soon after Landsat 7 SLC failure, NASA and USGS published a report proposing some approaches to partially recover the un-scanned pixels. The proposed algorithms basically rely on filtering kernels, and thus the use of mean and variance of pixel values in the encompassing window as estimated pixel values in various SLC-off and SLC-on cases. Scramuzza et al[3] presented these methods, which are called global linear histogram matching (GLHM) and localized linear histogram matching (LLHM), respectively. In a similar effort, USGS[4] declared an adaptive window linear histogram match method (AWLHM), which is essentially the...
same as the method of Scramuzza et al\textsuperscript{3} in its processing algorithms, but different in window size.

Other researchers focused on the quality of the chosen auxiliary (fill) images, which may not be Landsat images, shortly before or after the master image of ETM+ was taken. For instance, Mobasheri and Sadeghi\textsuperscript{5} used data from the Indian remote sensing satellites (IRS) as filling image, whereas Boloorani et al\textsuperscript{6} and Chen et al\textsuperscript{7} incorporated EO-1/ALI and CBERS products. Roy et al\textsuperscript{2} used some much lower-resolution information observed by MODIS to calculate the reflectance of the missing pixels.

By combining various methods and stochastically related image-processing techniques, Maxwell et al\textsuperscript{8} developed a Multi-scale segmentation approach and Zhang et al\textsuperscript{9} and Pringle et al\textsuperscript{10} estimated gap area using geo-statistical techniques, of which kriging or co-kriging are well known. There have been efforts through single-source (such as Inpainting algorithm), multi-source, and hybrid methods. A single-source method uses same-image information to fill gaps; a multi-source method involves more than one image for reconstruction; and a hybrid method combines both of the above approaches. There is ongoing research in the field of Inpainting, also aimed at reconstruction of images in plausible ways. Inpainting methods may be classified according to whether they are based on PDE, exemplar inpainting, or blocks.

By using available data and a customized algorithm, Chen et al\textsuperscript{11} proposed a simple but effective method, in which information about the adjacent similar pixels are utilized for estimating the missing pixels over heterogeneous areas.

All of the above-mentioned algorithms have pros and cons in various applications. LLHM is good in uniform and semi-uniform areas, but lacks precision in disparate and complex regions. A simple algorithm is fast to process large volumes of images, but its global applications are compromised. By using a moving window, the estimation can be slightly improved\textsuperscript{4}. The multi-scale segmentation approach focuses on segments but not pixel-level details, thus it is poor in recovering small and narrow objects, such as streams and borders\textsuperscript{8}. The geo-statistical-interpolation methods estimate the values at the pixel-level poorly, and thus are not optimal for studying objects of small-size\textsuperscript{9-10}.

In Inpainting methods, the multi-source gap-filling process cannot cope well with sharp changes between two images, such as between conditions of sun glint, snow, and cloud. In a single-source Inpainting algorithm, it takes more computational time to reconstruct a large area. The Neighborhood Similar Pixel Interpolator (NSPI) approach does provide one or more reasonably clear auxiliary TM or ETM+ image(s), but has several potential limitations, such as cloud cover, when used in such areas as humid tropical forest ecosystems, which are covered by cloud about 90 percent of time\textsuperscript{11}.

In this paper, we present an algorithm which uses a gap-filling method with severe striping effects. It uses a moving neighbor-selection kernel associated with a designed morphological and geometrical selection scheme, and applies it to the first and second outer shields of every stripe. The proposed algorithm has a better linkage with the surrounding pixels of a striped pixel, incorporates geometrical and morphological characteristics of relatively small and narrow objects in an image, and has better precision and accuracy than NSPI. Also, it is simpler and more whole looking than the LLHM and AWLHM methods, while avoiding the complexity and computational expensiveness of Maxwell et al\textsuperscript{8}, Zhang et al\textsuperscript{9} and Pringle et al\textsuperscript{10}. It does not contradict with auxiliary data-using methods, and thus can be used as the core process for the above-mentioned algorithms of this category. Also, it is applicable to the output of other methods without a strict need of the original data set, and it may be used to "resurrect" the large amount of processed images in the ETM+ archive less expensively than some complex algorithms.

2 Algorithm

Although some methods have been proposed to cover the gaps in the ETM+ images, they do not cover all the gaps. This is because images are used in different applications by different users, who are focused only on certain aspects according to their needs. Also, there is a lack of interest in improving the weaknesses of the previous algorithms. Typically, a new method cannot make use of the many processed images already in ex-
istance.

In developing a new algorithm, a trade off between complexity and output quality should be considered. A simpler method would produce less homogeneous filling with lower level of adherence with the original image, and thus would introduce more obvious stripes in the output image. For instance, in the methods proposed by Scramuza et al and USGS, we may see improvements from LLHM to AWLHM. Thus, in this study, we tried to improve the striped images by adopting a customized approach, with steps described below.

2.1 LLHM and AWLHM algorithms

As a first step, gaps are located by the LLHM method. Then a linear transformation from one image to another is found. The pixel values of the SLC-Off image to be filled (the 'primary scene') can be generated by applying a corrective gain and bias to the pixel values of an SLC-On image (the 'fill scene'). Rather than performing a computationally expensive linear fit, the corrective gain and bias can be found by using the mean and standard deviations of the data. This transformation can be applied to the entire fill scene, giving a global linear histogram match. For greater precision and better-looking products, the corrective gains and biases may be calculated with a moving window around each pixel in the scene. This is the basis of the localized linear histogram match (LLHM), or Phase 1, method for gap filling in a SLC-off scene of a Landsat 7 image. The AWLHM, or Phase 2, method is an enhancement of the Phase 1 algorithm, which allows users to choose multiple scenes and to combine SLC-off scenes. It gives a more accurate result, but needs more computation time. If computational speed is more important than statistical error, then LLHM should be used.

2.2 Gap-filling methods

In the second step, LLHM and AWLHM are used to fill a gap scene. These methods use a moving window on the SLC-off and SLC-on images consecutively centering on every gap pixels, and implement linear histogram matching to find a linear transformation from one image to another. This method is effective in smooth areas, but lacks precision in areas with complex textures. Thus, our main purpose is to restore small features.

2.3 Stitching stripes to the background image

In SLC-off images, gaps are almost parallel. To estimate substitute values, it is reasonable to place the parallel values in their corresponding gap areas, because of the relative homogeneity and smoothness in the background and striped areas. However, the existence of the border lines between these areas in the image gives an unpleasant view for qualitative study and inaccuracy for quantitative analysis. In this paper, we propose the idea of stitch lines (with different degree of sensibility depending on the surrounding pixels) going through pixels of the outer shell of a stripe, from neighboring background candidate pixels to the inner neighboring pixels of the stripe. By applying these lines according to a statistical-similarity mechanism, it is possible to retrieve part of the texture on the border line along the stripes. By using the stitch lines, we may avoid the smoothing effect of some commonly used image-processing solutions, such as local-average filtering, and highlight the lost texture instead. Whether to apply a stitch line to a stripe's border pixel depends on the surrounding pixels of the target pixel in both the fill and the striped images. Also, by using this method, it is possible to make small and narrow objects bold and clear. The length of a stitch line may make some difference in the performance of this method. A longer line may incorporate more distant and irrelevant data and may cause greater smoothness. In this study, the line length is 5 pixels, or 2 on each side of the target pixel. The application is made firstly to the closest neighbors and then for the second-level expansion of the stitch line.

2.4 Extraction of stripe-border pixel values by using morphology operators

In this step, we use reliable and original adjacent values in the background image to estimate the unreliable inner pixels along the stripes. The border is between these two groups of pixels and we want to pass a stitch line through the border pixels. So, we need to extract the values of these border pixels first with a simple and
fast algorithm. By using the mask layer of the ETM+ image and the input image, the morphology operators is implemented according to:

$$E = I \times (M - (\Theta M))$$  \hspace{1cm} (1)

where $E$ is extracted border image of the stripes, $M$ is mask image, $I$ is the image itself, $\times$ is dot product operator, and $\Theta$ is morphological erosion. The Mask image is a binary image, having values of 1 in gap areas and zero in no-gap areas. The morphological kernel is a 3x3 matrix as shown in figure 1(c).

$$\begin{array}{c}
(a) \quad \text{Gap filled (stripe) pixel} \\
(b) \quad \text{Center pixel} \\
(c) \quad \text{Non-gapped neighbor pixel}
\end{array}$$

Figure 1 Schematic view of consecutive steps in this algorithm. Stripe induced by (a) gap filling image, (b) Stripe morphological border extraction (stripe’s outer shell pixels in blue squares), (c) Neighbor pixels identification, (d) First stage similar pixel, (e) First stage stitch line, (f) Second stage similar pixel, (g) Second and First stages stitch line.

### 2.5 Stitch-line length

In this step, we select pixels adjacent to each border pixel with a process based on radiometric similarity. There are $(n+1)$ pixels forming the stitch line with a border pixel at the center. Since this center value is fixed and the adjacent values are selected to be proportional to this value, it is reasonable to expand this line starting from this value outwardly. So our first values to use are the adjacent values. Selection is made among two groups of pixels as drawn in figure 1(c) (in orange and purple colors) without involving any borderline pixel. After detecting candidate pixels on each side of the border line, the next selection is based on the nearest value to the center value. By this selection we get two new pixels shown in figure 1(d) (in green color) on both sides of center pixel and a stitch-line length of 3 pixels as shown in figure 1(e). We then reiterate the previous process with center at each of the two adjacent pixels. In detecting candidate values of neighbors, the previously selected pixel values should not be included. The result is the addition of two more pixels (in yellow, Fig. 1(f)) to the stitch line, which has now 5 pixels and passes through the background adjacent pixels (Fig. 1(g)).

### 2.6 Estimation of borderline pixel values

In this last step, we form a line with a length of $(2n+1)$ pixels, where $n$ is the longitude level of the stitch line. We can estimate the best values for the borderline pixels by using the following equation with incorporated weights for the pixels in the stitch line.

$$X = \frac{1}{2g} \left( \frac{\sum_{i=1}^{2n} \frac{1}{D_i} P_i}{\sum_{i=1}^{2n} \frac{1}{D_i}} \right)$$  \hspace{1cm} (2)

where $X$ is the estimated value of borderline pixel, $P_i$ is the value of the $i^{th}$ adjacent pixel, $g$ is the old borderline pixel value, $n$ is the total number of adjacent pixels on each side of the borderline pixel, $D_i$ is the distance of a pixel from the center pixel, and more weight is granted to a closer pixel (Fig. 2).

### 3 Test results

#### 3.1 Algorithm implementation

The test data used in this study was from a single-band (Band 4) of a Landsat ETM+ SLC-Off image of north Tehran in Iran acquired in 2012 (Fig. 3(a)) and a SLC-on image of exactly the same area acquired by Landsat 5 satellite’s sensor MSS in 2002 (Fig. 3(b)). These images
were spatially co-registered and then imported into LLHM and AWLHM algorithms for gap filling.

The calculated results are shown in figures 3(c) and 3(d). As mentioned before, the stripes are more visible in the result of LLHM than AWLHM. Also, the complexity of texture may influence the clarity of stripes and make them bolder.

Although the resultant AWLHM image (Fig. 3(d)) is quite different from the imported data (Fig. 3(a)), a significant quantitative improvement is observable through statistical analysis. After this modification, the border pixels with new values are counted as pixels in a new non-striped region. It is possible to repeat this algorithm on the new borderline between new striped and non-striped regions to seek additional but smaller improvement.

3.1 Validation and evaluation

To evaluate the results qualitatively and quantitatively to see if they have made any improvement, we used an intact (no gap) part of the SLC-Off scene, which has both urban and homogenous rural areas, with its compatible fill scenes as our validation data. We also used an artificially masked image (Fig. 4) to introduce gaps arbitrarily. Following the above-mentioned steps, we first filled the gaps and then stitched the border values to the background areas. Since we knew the exact original pixel values, it was possible to check if the results got better or worse. Error estimates were made by calculating the mean absolute difference between the values of pixels that existed in both the matched fill scene and the SLC-Off image. The result of statistical error estimates is given in table 1, and the percentage improvement in table 2.

We have applied this algorithm separately to every image band of Landsat, but chose to show the result of Band 4, since it had higher contrast and wider histogram, which made the effect of this method more observable.
Table 1 Numerical error estimates of stripe-stitched images according to ground-truth data (DN values) in three cases: Sole (striped image of LLHM and AWLHM), First (implementation of method on first outer border of stripes), Second (implementation of method on second outer border of stripes)

<table>
<thead>
<tr>
<th></th>
<th>Sole</th>
<th>First</th>
<th>Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWLHM</td>
<td>1.5740</td>
<td>2.3069</td>
<td>1.3812</td>
</tr>
<tr>
<td>LLHM</td>
<td>2.4181</td>
<td>2.0208</td>
<td>1.8081</td>
</tr>
</tbody>
</table>

Table 2 Same as table 1, but for different bands in percentage

<table>
<thead>
<tr>
<th></th>
<th>First stage</th>
<th>Second stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWLHM</td>
<td>9.9</td>
<td>2.6</td>
</tr>
<tr>
<td>LLHM</td>
<td>22.4</td>
<td>11.8</td>
</tr>
<tr>
<td>AWLHM</td>
<td>12.4</td>
<td>3.3</td>
</tr>
<tr>
<td>LLHM</td>
<td>24.1</td>
<td>10.8</td>
</tr>
<tr>
<td>AWLHM</td>
<td>12.8</td>
<td>4.2</td>
</tr>
<tr>
<td>LLHM</td>
<td>23.0</td>
<td>9.5</td>
</tr>
<tr>
<td>AWLHM</td>
<td>14.4</td>
<td>4.6</td>
</tr>
<tr>
<td>LLHM</td>
<td>28.2</td>
<td>12.3</td>
</tr>
<tr>
<td>AWLHM</td>
<td>12.3</td>
<td>3.8</td>
</tr>
<tr>
<td>LLHM</td>
<td>30.1</td>
<td>13.1</td>
</tr>
<tr>
<td>AWLHM</td>
<td>11.8</td>
<td>3.6</td>
</tr>
<tr>
<td>LLHM</td>
<td>28.3</td>
<td>11.9</td>
</tr>
</tbody>
</table>

Figure 5 shows the effects of applying this algorithm to the border pixels. The added values have resulted in magnified intensity. By comparing it to the Band 4 image, it is obvious that this method affected the complex urban areas more than the homogeneous rural areas.

The quantitative validation results are given in tables 1 and 2 and shown in figure 6, in which the horizontal and vertical axes show the original and estimated values, respectively. From this figure, it is clear that, by using of this method, data points became less scattered (or the precision was improved) and more aligned with the identity line (or the estimated pixel values became closer to the real values) in general.

Various factors may cause different degrees of errors in gap-filled SLC-Off images. To see how this method functions regardless of magnitude of error before or after the gap-filling process, we calculated relative errors in percentage for the different bands (from the 1st to the 6th) of the same Landsat image. The results are given in table 2.

Figure 5 In band 4, the difference of pixel values at the stripe borders before and after the application of the algorithm

4 Conclusions

In this paper, we addressed a striping problem caused by a previous effort to compensate for SLC-Off defect, and proposed a simple improvement method applicable to the output data of LLHM and AWLHM algorithms. This method deals with pixels at the border line between the striped area and its adjacent area, and develops a stitch line between these areas by assuming similarity of neighboring pixels. Data from band 4 were selected for validation tests, because the image had a higher contrast. Such tests and error evaluations, using mean absolute difference, showed both quantitative and qualitative improvements by this method. We can see how the precision was improved through this algorithm and how the estimated pixel values became closer to the real values. However, this method also introduced a smoothing effect at the stripe borders, due to the fact that we used mean estimator. However, this effect is negligible, because there was a selection process before
the estimation. By using the ground-truth pixels, we assured the reliability of this method.

In the future, it may be of interest to apply this method to study data from all the bands, or to use images for the same area obtained by other ancillary sensors. Also, it may be worthwhile to look into some appropriate nonlinear central estimators, both for estimating the center values of the adjacent pixels, and for error estimation and comparison.

References


