Filling Landsat ETM+ SLC-off Gaps Using a Segmentation Model Approach

by Susan Maxwell

The purpose of this article is to present a methodology for filling Landsat Scan Line Corrector (SLC)-off gaps with same-scene spectral data guided by a segmentation model. Failure of the SLC on the Landsat 7 Enhanced Thematic Mapper Plus (ETM+) instrument resulted in a loss of approximately 25 percent of the spectral data. The missing data span across most of the image with scan gaps varying in size from two pixels near the center of the image to 14 pixels along the east and west edges. Even with the scan gaps, the radiometric and geometric qualities of the remaining portions of the image still meet design specifications and therefore contain useful information (see http://landsat7.usgs.gov for additional information).

The U.S. Geological Survey EROS Data Center (EDC) is evaluating several techniques to fill the gaps in SLC-off data to enhance the usability of the imagery (Howard and Lacasse 2004) (*PE&RS*, August 2004). The method presented here uses a segmentation model approach that allows for same-scene spectral data to be used to fill the gaps. The segment model is generated from a complete satellite image with no missing spectral data (e.g., Landsat 5, Landsat 7 SLC-on, SPOT). The model is overlaid on the Landsat SLC-off image, and the missing data within the gaps are then estimated using SLC-off spectral data that intersect the segment boundary. A major advantage of this approach is that the gaps are filled using spectral data derived from the same SLC-off satellite image.

Segmentation

Segmentation is not a new concept to image analysis (Kettig and Landgrebe 1976). Segmentation is basically the process of grouping individual pixels into homogeneous landscape units. There are several approaches to image segmentation (Blaschke et al. 2000), a few of which have become commercially available. The segmentation algorithm used in this study is referred to as the fractal net evolution approach (Baatz and Schäpe 2000) implemented as part of the eCognition[1] object-oriented image analysis software package by Definiens Imaging GmbH (2004).

Usually a segmentation model is developed for the purpose of classifying land cover/land use. In our case, the segmentation model is used to estimate missing spectral data within the SLC-off image gaps. The goal of the model is to reflect, as closely as possible, the landscape characteristics of the SLC-off image. Therefore, the image selected to generate the model should be one that is approximately the same time of year as the SLC-off image acquisition date and within a few years of the acquisition date so that landscape changes are minimized. Once the segment boundaries are created, they are overlaid on the SLC-off image and used to guide the process of estimating missing spectral information in the gaps.

Case Study

An example of using a segmentation model to fill SLC-off gaps is presented here using imagery collected near Sioux Falls, South Dakota. The landscape throughout this region is predominantly agricultural. A Landsat 5 image, acquired on August 9, 2000, was used to generate the segment boundaries. A Landsat ETM+ SLC-on image acquired on August 4, 2001 (i.e., acquired prior to the SLC malfunction), was used to test the fill technique. The images were first georegistered to the same projection. Gaps in the Landsat ETM+ SLC-on image were simulated so that results of the fill method could be compared with actual data. The process is comprised of the following steps (Figure 1):

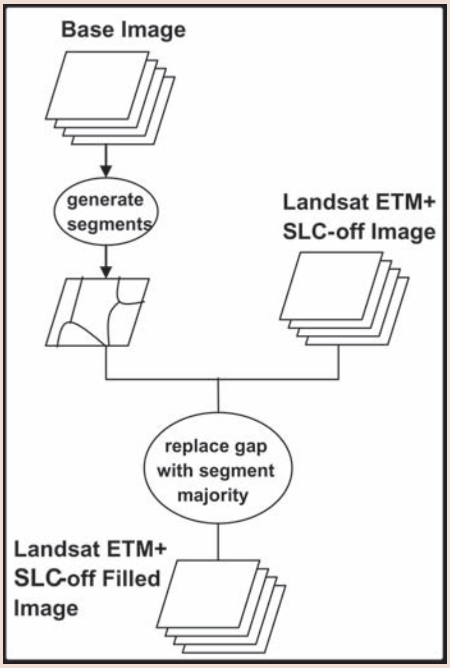

Figure 1. Segmentation model processing flow diagram.

1. Generate the segment boundaries from a complete Landsat image (base image).
2. Overlay the segment boundaries on the SLC-off image and replace the missing data with the majority digital number (DN) value for each segment.

The segments were derived from the August 9, 2000, Landsat 5 image (Path 29 Row 30). Three different scale parameters (10, 15, and 20) were used to evaluate the effect of segment size on the output product (Figure 2). The scale parameter is an arbitrary value used to limit heterogeneity and segment size (Definiens Imaging GmbH 2004). An SLC-off image was simulated by using a mask applied to the August 4, 2001, Landsat 7 SLC-on image. The segment boundaries generated from the 2000 Landsat 5 image were then overlaid on the

continued on page 1110

continued from page 1109

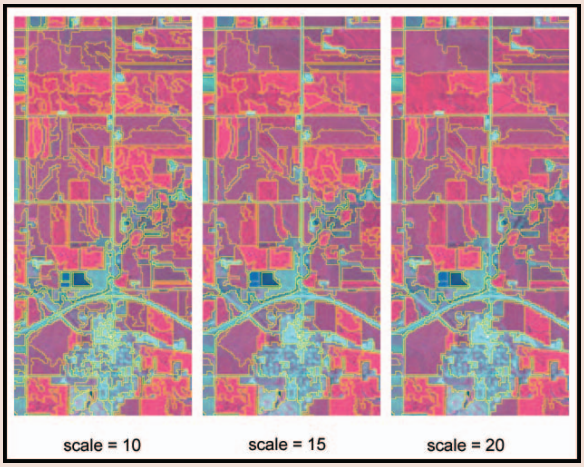

Figure 2. Segment boundaries created from the Landsat 5 image. A comparison of segment boundaries using three different scales is shown. A smaller scale (e.g., 10) creates smaller segment areas.

SLC-off image and the gaps filled with the majority DN value of pixels intersecting each segment.

Results and Discussion

The optimum segment size in our case is one that is small enough so that it represents a meaningful landscape unit while at the same time large enough to ensure that the segments within an SLC-off gap intersect a portion of the image that contains spectral data. The problem is that segments created with a smaller scale value (i.e., scale=10) may not intersect portions of the scene with existing data, especially toward the edges of the image where the larger gaps appear (region 'B' in Figure 3). The larger the segment size, the more likely the boundary will intersect existing data, and missing spectral information can be estimated. The down side to larger segments is that they tend to result in a smoothed appearance due to the larger number of pixels with the same DN value. This is especially apparent in landscape elements with higher spectral texture, such as urban areas. We plan to test alternative methods to maintain landscape texture by using the original image spectral characteristics and the assumption of a normal distribution of pixels within a given segment. The relative histogram position per pixel is expected to remain stable per segment (R. de Kok, Definiens, personal communication). Use of multi-resolution segmentation to fill smaller segments that fall entirely within the gap also is being investigated. For many applications, the appearance is of less importance compared to how the estimated values affect classification performance. A collaborative research effort is currently underway with the U.S. Department of Agriculture National Agricultural Statistics Service (USDA/NASS) to test this method on Landsat imagery used to develop crop maps in the United States.

The major advantage to applying this modeling approach is that same-time spectral data are used to estimate the missing data. A second, complete image is required to develop the segments, yet once the segment boundaries are generated they may be useful for several years, depending on landscape dynamics and purpose of the application. Regions where landscape structures change frequently, such as rapid urban growth or frequent forest clear cutting may require re-segmentation more frequently (e.g., every few years). Conversely, regions that have relatively low rates of change, such as agricultural landscapes, may need re-segmentation less often. In our test case, relatively few fields had changed from the 2000 base image to the 2001 SLC-off image (Figure 4). Analysis of multi-temporal Landsat imagery could be used to determine the rate of change of boundaries in various landscape systems. Further research is planned to:

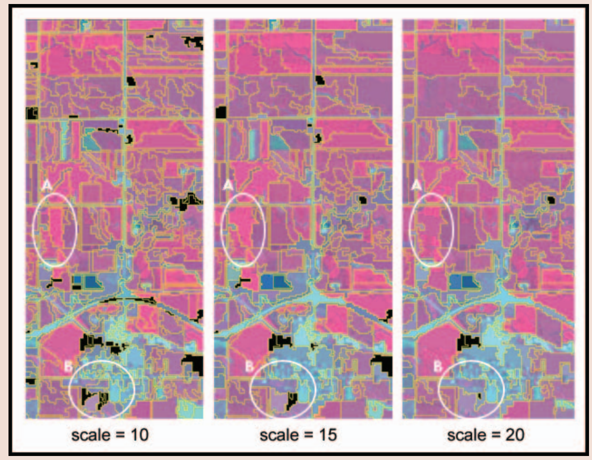

Figure 3. Comparison of the August 4, 2001, Landsat 7 SLC-off image filled with segmentation model approach using three different scales of segment boundaries. Segment boundaries created from the August 9, 2000, Landsat 5 image are overlaid on the SLC-off images. The smallest scale (scale=10) captures landscape variability at a higher resolution (A) yet some of the segments are not filled because they lie entirely within the gap (B). The larger scale (scale=20) results in less segments being missed (B); however, the interpolated values are more noticeable due to the larger number of pixels with the same value (A).

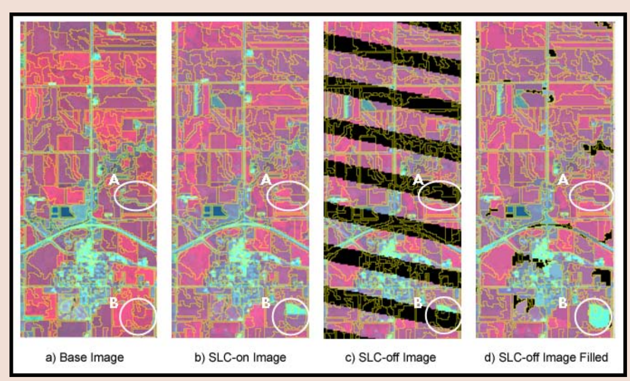

Figure 4. Comparison of a) the August 9, 2000, Landsat 5 base image from which segment boundaries were generated (scale=10), b) the August 4, 2001, Landsat 7 SLC-on image with the 2000 segment boundaries overlaid, c) the August 4, 2001, Landsat 7 SLC-off simulated image, and d) the August 4, 2001, Landsat 7 SLC-off simulated image with gaps filled. Most interpolated values were reasonably close to the SLC-on image values. The circled areas designated as 'A' and 'B' show examples of crop field boundary alterations that resulted in large differences between the interpolated values and the SLC-on values.

- perform tests on a variety of environmental landscapes (e.g., forest, grassland),

- improve the algorithm to maintain texture and to ensure the entire gap is filled,
- study the use of a multi-scale segmentation approach, and
- determine the feasibility of implementing the procedure operationally.

Conclusion

Several options are available to fill gaps in Landsat ETM+ SLC-off satellite imagery. The approach presented in this article utilizes a segmentation model to fill the gaps with coincident spectral data. The approach appears promising; however, further studies are needed to evaluate the results on end-user applications and to determine if it is feasible for the EDC to implement the procedure operationally.

References

Baatz, M. and A. Schäpe, 2000. Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation, In: Strobl, J. and Blaschke, T. (eds.): *Angewandte Geographische Informationsverarbeitung XII*, Wichmann Verlag: 12–23.

Blaschke, T., S. Lang, E. Lorup, J. Stroble, P. Zeil, 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications, In: Cremers, A. and Greve, K. (eds.): *Environmental Information for Planning, Politics, and the Public, Vol. 2*, Metropolis Verlag, Marburg: 555-570.

Definiens Imaging GmbH, 2004, eCognition, Web page: http://www.definiens-imaging.com (accessed August 9, 2004).

Howard, S.M. and J. Lacasse, 2004. Evaluation of gap-filled landsat SLC-off imagery for wildland fire burn severity mapping, *Photogrammetric Engineering and Remote Sensing*, 70:877-880.

Kettig, R.L. and D.A. Landgrebe, 1976. Classification of multispectral image data by extraction and classification of homogeneous objects, *IEEE Transactions on Geoscience Electronics*, 14:19-26.

Author

Susan K. Maxwell is employed by Scientific Applications International Corporation, contractor to the U.S. Geological Survey, EROS Data Center, Sioux Falls, SD 57198 (maxwell@usgs.gov). Work was performed under U.S. Geological Survey contract 03CRCN0001.

(Endnotes)

[1] Any use of trade, product, or firm name is for descriptive purposes only and does not imply endorsement by the U.S. Government.