SLC Gap-Filled Products
Phase One Methodology

Pat Scaramuzza, Esad Micijevic, Gyanesh Chander

Introduction:

The Landsat 7 scan-line corrector (SLC), a mechanism designed to correct the undersampling of the primary scan mirror, failed on May 31, 2003. With the SLC now permanently turned off, the ETM+ is losing approximately 22% of the data due to the increased scan gap. This paper describes a local linear histogram matching technique which will be used initially to fill the scan gap with previously acquired Landsat 7 imagery.

Methodology:

Filling the scan gap first requires precise knowledge of what pixels are valid in an image and which are to be filled. To do this a scan gap mask is created for each band that marks existing data as 1 and missing data in the scan gap and fill regions as 0. These scan gap masks are created during L1G product creation by a second pass through the same resampling algorithms used to create the L1G product. They are delivered with the product as 8-bit images having dimensions identical to the corresponding image band files, and are compressed using the gzip utility to avoid expanding the product size.

Once the gaps are located, the linear histogram matching methodology attempts to find a linear transformation between one image and another.

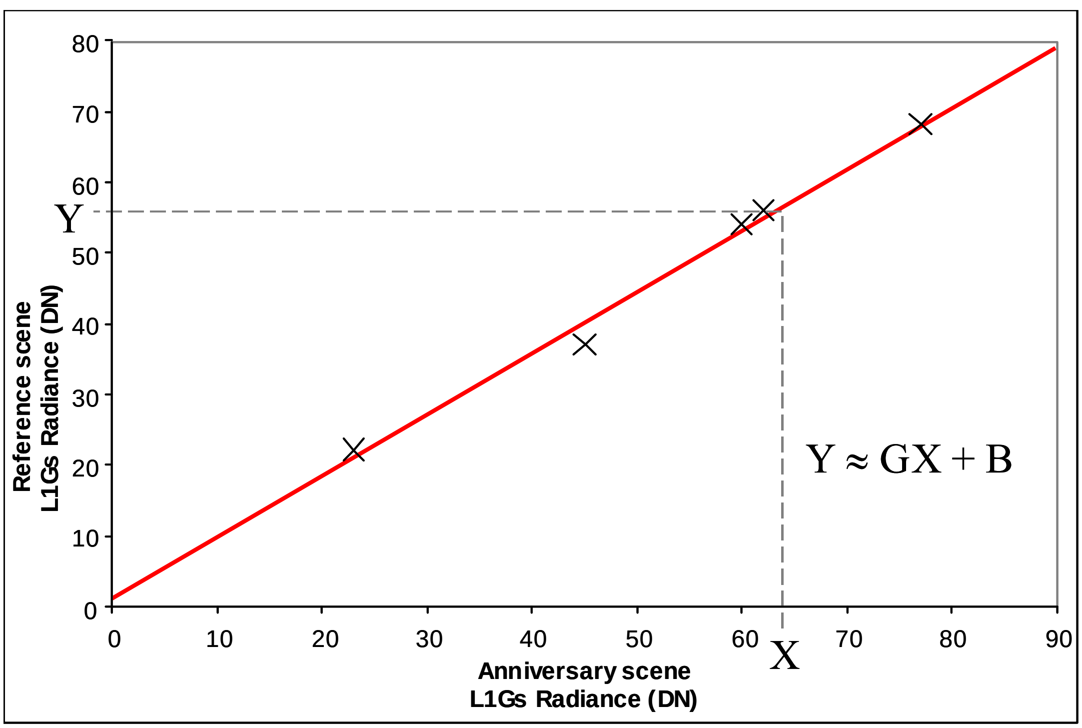


Figure 1: An assumed linear relationship between scenes

In theory, 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 perform a computationally expensive linear fit, the corrective gain and bias can be found using the mean and standard deviation of the data.

With a SLC-On fill scene X and a SLC-Off primary scene Y, we make the assumption that:

$$Y \approx GX + B$$

where	G	=	the gain used to histogram match the fill image to the primary image.
	B	=	the bias used to histogram match the fill image to the primary image.
	X	=	the fill (SLC-On) scene array.
	Y	=	the primary (SLC-Off) scene array.

The gain and bias can be calculated using the standard deviation and mean of arrays X and Y:

$$G = \frac{\sigma_Y}{\sigma_X}$$

$$B = \overline{Y} - G\overline{X}$$

where	σ_X	=	the standard deviation of data in fill image X.
	σ_Y	=	the standard deviation of data in primary image Y.
	$\overline{X}$	=	the mean of the fill (SLC-On) scene array.
	$\overline{Y}$	=	the mean of the primary (SLC-Off) scene array.

This transformation can be applied to the entire fill scene, giving a global linear histogram match. However, while the global histogram match performs well over scenes with invariate terrain such as deserts and rocky areas, it creates visible errors in scenes with transient objects such as clouds. To improve performance in cloudy imagery, high saturated pixels are excluded before calculating the corrective gain and bias.

But even with these appropriate exclusions, the global histogram match is not optimal. For greater precision and a product which is visibly better looking, the corrective gains and biases can be calculated in a moving window around each pixel in the scene. This is the basis of the localized linear histogram match (LLHM).

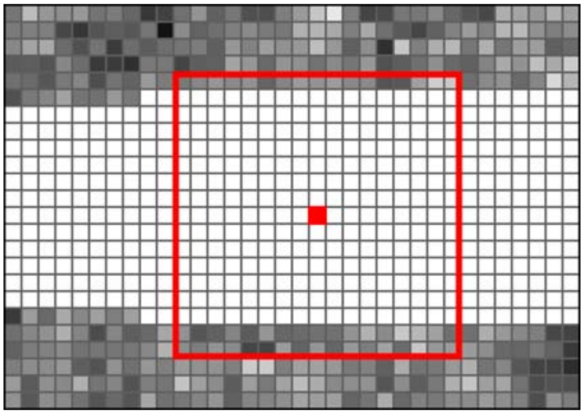


Figure 2: Illustration of localized linear histogram matching window in the scan gap.

In the 30 meter bands (bands 1-5 and 7), the SLC gap is a maximum of 14 pixels wide. A 17 pixel wide window was chosen as the minimum desirable window width. In this 17 pixel square window, coincident data is collected from both the fill scene and the primary scene. Appropriate exclusions are made, and the corrective gain and bias are calculated and applied to the center pixel value from the fill scene. The resulting value is used to fill the missing pixel in the primary image.

Results:

Two test SLC-Off scenes, along with two compatible fill scenes, were selected for analysis. Local linear histogram matching was used to fill the gaps in the SLC-Off data for each scene. Error estimates were made by calculating the mean absolute difference between pixels that existed in both the matched fill scene and the SLC-Off image. These error calculations were made on 30-meter data for every test scene, but due to processing constraints matched images for the thermal and pan bands were only created for one test scene each.

	Path 36, Row 37 -- Mesa, AZ (Arid/Urban scene, no clouds)	Path 39, Row 28 -- SW Montana (Cloudy/Mountainous scene, 30% clouds)
Band	Error (L1 DN)	Error (L1 DN)
1	2.89	5.12
2	3.27	5.14
3	5.17	6.09
4	2.43	4.96
5	5.13	6.17
7	5.09	5.79
6L	-	3.37
6H	-	5.97
8	3.25	-

Table 1: Mean absolute difference error estimates of LLHM test scenes.

Visually, most land types are restored adequately in the gap-filled images. Some striping appears around land cover changes, such as cropland and coastlines, and around transient objects such as clouds. There are also some artifacts around small, sharp edges in otherwise homogenous regions, where the histogram of the data is very narrow.

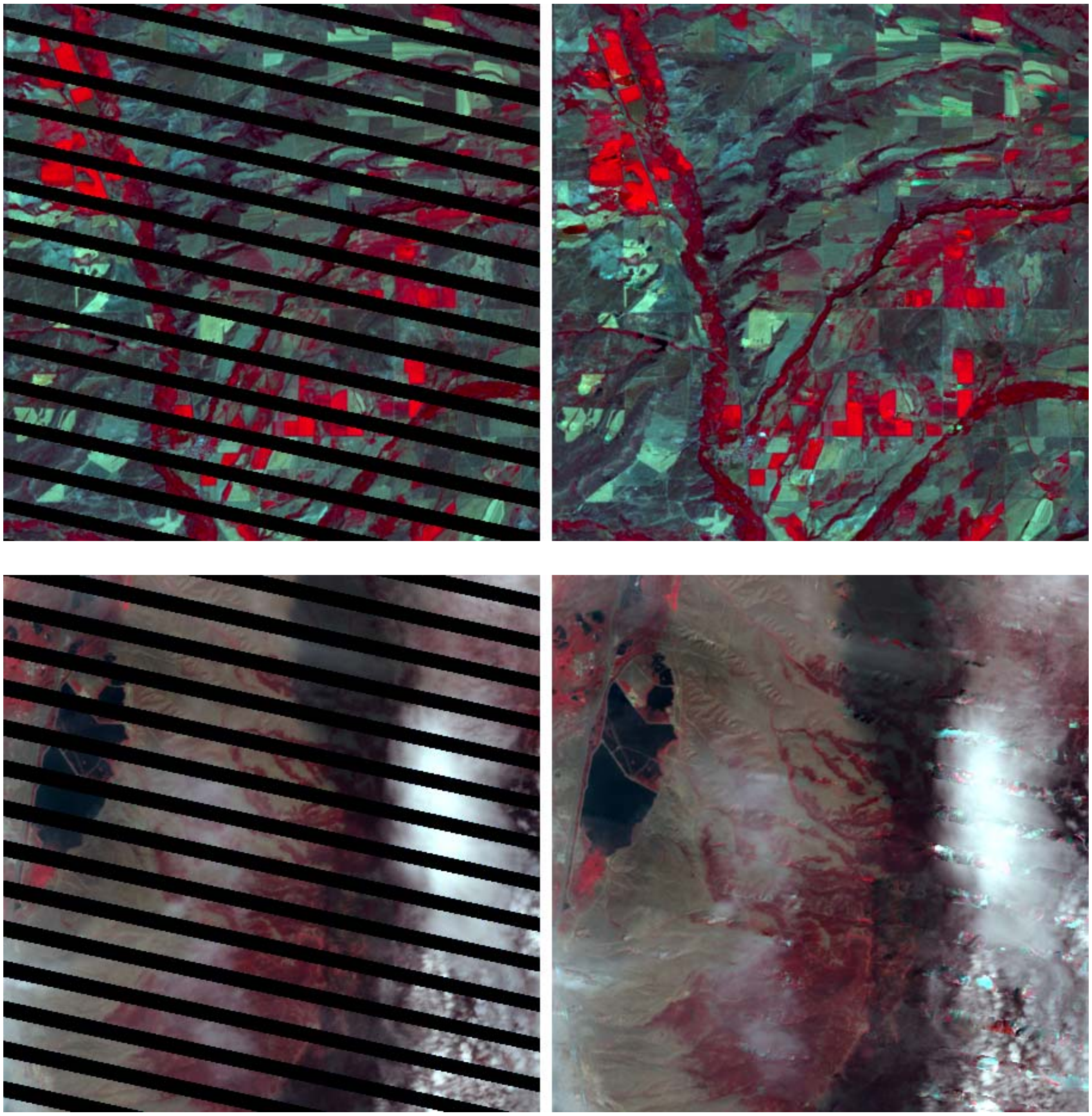
Gap-filled product examples:

Both test scenes are available to the public in their entirety, including primary scenes, fill scenes, and completed gap-filled product images. Links to the data are on the Landsat 7 Gap-filled product information page on the web at http://landsat7.usgs.gov/slc_enhancements/gapfilled1.php .

SLC-Off Local Linear Histogram Match

Figure 3: Examples of LLHM results, using bands 3, 2, 1

SLC-Off scene is path 36 row 37, acquired 6/24/2003
LLHM scene is the same image, filled with matched data acquired 7/7/2002
Both the primary and fill scenes are cloud-free.



SLC-Off Local Linear Histogram Match

Figure 3: Examples of LLHM results, using bands 4, 3, 2

SLC-Off scene is path 39 row 28, acquired 8/16/2003
LLHM scene is the same image, filled with matched data acquired 8/13/2002
The primary scene contains 30% clouds, the fill scene is cloud-free.