Merging Landsat and SPOT digital data using stochastic simulation with reference images

Júlia Carvalho¹, Jorge Delgado-García² and Amílcar Soares¹

¹ Environmental Group of the Centre for Modelling Petroleum Reservoirs, CMRP/IST
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
Tel.: + 351 218 417 444; Fax: + 351 218 417 389
jcarvalho@ist.utl.pt; asoares@ist.utl.pt

² Dpto. Ingeniería Cartográfica, Geodésica y Fotogrametria, Escuela Politécnica Superior, Univ. de Jaén
Campus de las Lagunillas, Edif. A-3. 23071 Jaén, España
Tel.: + 344 953 212 468; Fax: + 34 953 212 855
jdelgado@ujaen.es

Abstract
There is a wide range of systems providing digital satellite imagery with different spatial and spectral resolutions. But, unfortunately, these resolutions are in most cases opposite; i.e., the high-resolution sensors have low spectral resolution whereas the multispectral sensors have good spectral resolution but bad spatial resolution making their use in detailed applications difficult. The problem is solved using digital image merging procedures. The main objective of these methods is to obtain synthetic images that combine the advantage of the high spatial resolution of one image with the high spectral resolution of another image. Ideally, the method used to merge data sets with high-spatial resolution and high-spectral resolution should not distort the spectral characteristics of the high spectral resolution data. The classical methods of merging procedures (PCA, IHS, HPS, etc.) present several drawbacks. The objective of this paper is to present a geostatistical merging methodology based on direct sequential co-simulation with reference images (Carvalho et al., 2006). With the stochastic simulation one generates a high spatial resolution image with the characteristics of the of the higher spectral resolution image. It is an iterative inverse optimization procedure that tends to reach the matching of an objective function by preserving the spectral characteristics and spatial pattern, as revealed by the variograms, of the higher-spectral resolution images both in terms of descriptive statistics and band correlation coefficients. The method was applied to Landsat TM and SPOT-P images. The results were compared with the original Landsat image and the images provided by classical merging procedures.

Keywords: digital image merging method, geostatistics, stochastic simulation

1 Introduction
Digital images are very frequently used in environmental and cartographic applications. The increasing applications are due to the availability of high quality images for a reasonable price and improved computation power. Nowadays there is a wide range of systems that provide images in digital format, and their interpretation into terrestrial attributes is very dependent on their spatial and spectral resolution. As a result of the demand for higher classification accuracy and the need in enhanced positioning precision there is always a need to improve the spectral and spatial resolution of remotely sensed imagery. Normally, these types of resolution are contradictory: high spatial resolution sensors have a low spectral resolution whereas multispectral sensors have a low spatial resolution.
Digital image-merging procedures are techniques that aim at integrating the multispectral characteristics in a high spatial resolution image, thus producing synthetic images that combine the advantages of both types of images. The main constraint is to preserve the spectral information for tasks like classification of ground cover.

Ideally, the method used to merge datasets with high-spatial resolution and high-spectral resolution should not distort the spectral characteristics of the high spectral resolution data. Not distorting the spectral characteristics is important for calibrating purposes and to ensure that targets that are spectrally separable in the original data are still separable in the merged dataset (Chavez et al., 1991).

Several methods for spatial enhancement of low-resolution imagery combining high and low-resolution data have been proposed. Some widely used ones are: Intensity-Hue-Saturation (IHS) (Chavez et al., 1991), Colour Normalized (CN) (Vrabel, 1996), Principal Components Analysis (PCA) (Pohl, 1998; Chavez et al., 1991) and Brovery transform (Marr, 1982). The classical methods of merging present several drawbacks: a) they do not take into account the information support of the data to merge; b) IHS and PCA are not really merging methods – they consist of the substitution of the high-spectral images with a high-spatial resolution image based on the correlation between both data sets, where the correlation level does not modify the process but influences the final results – and these methods can only be applied to triplets of bands; and, c) they do not provide any additional information about the images (uncertainty, spatial variability, scale of variation, etc.).

We present a new multisensor image merging geostatistical technique – using stochastic simulation with reference images –, based on prior work of Carvalho et al. (2005), to merge low-resolution multispectral images with high-resolution panchromatic images, and compare the results with classical merging methods.

2 Classical digital image merging methods
A variety of different image merging techniques exist in the literature with specific adaptation for particular problems. In order to compare results two classical image merging methods were applied: a) Intensity-Hue-Saturation transformation and b) Colour Normalized.

2.1 IHS transform
The IHS transform is one of the most common methods of merging images. The IHS system offers the advantage that the separate channels outline certain colour properties, namely intensity (I), hue (H), and saturation (S). The intensity describes the total colour brightness and exhibits as the dominant component a strong similarity to the panchromatic image.

This transformation consists of two basic steps. In the first step, RGB colour values for three selected TM multispectral bands are converted to hue, saturation and intensity colour components. Mathematical functions are used to convert RGB values to IHS values. The higher-spatial resolution image is constantly stretched in order to adjust the mean and variance to unit intensity. The second step is the substitution of the stretched panchromatic image for the intensity component of IHS and retransformation to RGB.

2.2 Colour normalised transformation
The colour normalised transformation merges the two spectral and spatial datasets assuming there is a certain spectral overlap between the multispectral bands and the more highly
resolved panchromatic band. This method (Vrabel, 1996) uses a mathematical combination of the
colour image and high-resolution data to merge the higher spatial and higher spectral
resolution images. Each band in the higher spectral image is multiplied by a ratio of the higher
resolution data divided by the sum of the colour bands. The function automatically resamples
the three-colour bands to the high-resolution pixel size using a nearest neighbour, bilinear, or
cubic convolution technique. The output RGB images will have the pixel size of the input
high-resolution data.

3 Geostatistical approach
The proposed procedure is based on the application of the geostatistical simulation technique –
direct sequential simulation with reference images (Carvalho et al., 2006) – to obtain simulated
values of the 10m Landsat TM image from the original 30m Landsat TM values. In this
sequential simulation methodology, any realization of a given variable $z(x_0)$, at location $x_0$, is
drawn from a conditional distribution function built with a reference image, instead from
global cdf like in direct sequential simulation, DSS, (Soares, 2001).

Consider $z^*(x_0)$ and $\sigma_{sk}(x_0)$ the local estimate and simple kriging estimation variance at $x_0$.
Instead of sampling the global cdf, like in DSS, here one draw simulated values $z_i(x_0)$ from
local cdf calculated from local portions of a reference image. In this particular case, centered
at the location $x_0$ of Landsat TM (reference image) local cdf is built, from the set of points
inside a given radius, by using the same approach of DSS.

An additional condition is imposed: the mean value of the 9 pixels of the 10m cosimulated
Landsat TM-SPOT P image (3x3 pixels) must be equal to the 30m Landsat TM original
values.

The simulation process should allow attaining a simulated image that reproduces the statistical
characteristics of the merged images. The simulated image must have the same mean value as
the 30m Landsat TM image and the same variance and variogram as the SPOT PAN image.

This technique allows generating several realizations of the original values with a specific
pixel size, preserving the basic statistical characteristic of the original images and using
information derived from the high-resolution image according to the level of correlation.

Let us consider $TM_i(x)$ as the digital value of the original 30mx30m Landsat TM image for the
band $i$ at the location $x$, $PAN(x)$ the value of the original 10mx10m SPOT-P image at the
location $x$ and $TM_i(x_1)$ the simulated value of Landsat TM image for the band $i$ at the
10mx10m grid (SPOT-P grid) at the position $x_1$.

With the proposed algorithm the simulated $TM_i(x_1)$ image must have the spatial pattern of
SPOT-P image, the histogram of band $i$ of TM but with a variance corrected for the 10mx10m
grid and the same local mean of band $i$ of TM; i.e., the mean of 9 pixels $TM_i(x_1)$ must be
equal to the correspondent value $TM_i(x_0)$.

In short the simulated $TM_i(x_1)$ must satisfy the following:
1. For any image radiance (DN): $\text{prob}\{TM_i(x) < \text{DN}\} = \text{prob}\{TM_i(x) < \text{DN}\}$; where
$TM_i(x)$ is the corrected TM variable for the variance of a 10mx10m grid;
2. $\gamma_{PAN}(h) = \gamma_{TM_i}(h)$, where $\gamma_{PAN}(h)$ and $\gamma_{TM_i}(h)$ are the variograms of the original
SPOT-P and simulated Landsat TM merged image, respectively;
3. Conditioning of the simulated images to the local means: the mean of the pixels grouped according to the 3x3 pixels scheme must be equal to the 30m Landsat TM original image values.

The idea of the proposed algorithm is to use the CoDSS with reference images to generate $T_M^*(x)$ in a 10m grid using SPOT-P(x) as secondary information. The spatial correlation between primary and secondary variables, Landsat and SPOT-P images (after upscaling to the 30mx30m grid), cannot be considered homogeneous and representative of the entire image. Hence a local model of co-regionalization is applied using the Markov-type approximation (Pereira et al., 2000). This means that local correlation coefficients between the two images are calculated (inside local windows), and adopted as the co-regionalization model of the two variables in the cosimulation procedure. To meet condition 3, i.e. the local means of the original TM, the proposed simulation methodology uses a gradual correction for local means:

Suppose $T_M(x_0)$ the value of Landsat at $x_0$. Each one of the nine values $T_M^*(x_j)$, $j=1,9$, included in $x_0$ is corrected as soon as they are simulated:

$$T_M^* = T_M^*(x_j) - m_t + T_M(x_0), \quad j = 1,N \text{ with } N<=9$$

where:

- $m_t$ - mean of simulated values $T_M^*(x_j)$ up to step N.

With this gradual correction one guarantee that after all 9 points of $x_0$ are simulated the mean $m_t=T_M(x_0)$.

The geostatistical image merging method can be summarized in the following steps:

1. Calculation of the basic statistics, correlation matrix and variograms of the several images (bands) that take part in the merging process. The calculation is applied to the Landsat TM bands and SPOT-P image.

2. Generation of a cosimulated image for each band. This image is generated using the direct cosimulation method, with each of the Landsat image’s bands as primary information, and the high-resolution image as secondary information and the local correlation coefficient between Landsat TM and SPOT-P (defined in a 150x150 m window). During the simulation process a correction for local means is performed in a 3x3 pixels of the simulated image, rectifying those values to the same mean value as the 30m Landsat TM image.

**4 Example data**

To show the capabilities of the proposed method, the merging procedures presented in section 2 were applied to a test area. The selected area covers a 2400mx2400m area in the Jaén province (South of Spain), with several kinds of land use (urban, olive trees, riverside vegetation, roads, etc).

The dataset used for this application comprises a portion of the following images (Figure 1):

- Landsat-TM images. Scene: 20034/95. Date: 08/26/1995. Image size: 80x80 pixels. GSD=30m (TM6 band has not been considered).
Figure 1 SPOT-P (left) and Landsat TM4 (right) images with 1% linear expansion.

The basic statistics of the different bands are presented in Table 1.

<table>
<thead>
<tr>
<th>Bands</th>
<th>Mean</th>
<th>Variance</th>
<th>Std. Dev.</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM1</td>
<td>99.10</td>
<td>224.68</td>
<td>15.00</td>
<td>97</td>
<td>66</td>
<td>166</td>
</tr>
<tr>
<td>TM2</td>
<td>51.65</td>
<td>103.33</td>
<td>10.17</td>
<td>50</td>
<td>26</td>
<td>97</td>
</tr>
<tr>
<td>TM3</td>
<td>67.37</td>
<td>214.87</td>
<td>14.66</td>
<td>66</td>
<td>28</td>
<td>131</td>
</tr>
<tr>
<td>TM4</td>
<td>74.99</td>
<td>206.18</td>
<td>14.37</td>
<td>74</td>
<td>34</td>
<td>133</td>
</tr>
<tr>
<td>TM5</td>
<td>123.56</td>
<td>762.55</td>
<td>27.62</td>
<td>122</td>
<td>45</td>
<td>226</td>
</tr>
<tr>
<td>TM7</td>
<td>66.67</td>
<td>281.00</td>
<td>16.77</td>
<td>66</td>
<td>22</td>
<td>139</td>
</tr>
<tr>
<td>PANa</td>
<td>140.58</td>
<td>662.78</td>
<td>25.74</td>
<td>137</td>
<td>63</td>
<td>254</td>
</tr>
</tbody>
</table>

*a PAN image is considered resampled to 30m pixel size.*

It is very important to bear in mind that the correlation coefficient between the Landsat TM visible and SPOT panchromatic bands is quite high (around 0.83), but this value decreases considerably (to about 0.70) for the Landsat infrared bands. This difference between the visible and non-visible channels of multispectral imagery is always verified.

When simulating the Landsat image at a 10mx10m spatial resolution, Landsat data is treated as the primary variable and the SPOT-P data as the secondary variable. With the CoDSS algorithm the simulated image is conditioned to reproduce the histogram of the Landsat image (corrected for the variance) and the variogram of SPOT, which means that we will obtain an image that has the spectral characteristics of the Landsat and the spatial distribution of the panchromatic SPOT.

5 Experimental results

The semivariograms and histograms of the panchromatic SPOT image and, as an example, the Landsat TM4 are shown in Figure 2.
For all the Landsat bands and the SPOT image the variograms are omnidirectional, and we have fitted to the sample values exponential models. The sills meet the sample variances. The range of the variograms is 480m for TM1 and TM4, 420m for TM2, 450m for TM3, and 660m for TM5 and TM7, for the panchromatic SPOT image the variogram presents a range of 330m.

Local correlations were computed to account for local differences in Landsat-SPOT data correlation (Figure 3). To compute the local correlations, a window size of 150mx150m, which is half of the variogram range, was considered the most appropriate. Local correlations range from 0 to 0.981, 0.988, 0.985, 0.954, 0.976 and 0.973 for Landsat TM1, TM2, TM3, TM4, TM5 and TM7, respectively.

The simulations were computed thus producing the hybrid Landsat 10mx10m images. The final simulated images were checked for a correct visual appearance. Inherent to the procedure, the merged images reproduce the global histograms of the Landsat bands and the variogram of the SPOT-P (Figure 4), and respect the value of the low-resolution data at their locations when they are subjected to a 30mx30m up scaling.
6 Comparison between geostatistical method and classical approaches

To demonstrate the potential of the proposed methodology, the images presented in section 4 were used in the application of the classical methods of image merging (presented in section 2) and of the proposed geostatistical approach. First of all, we can evaluate the visual appearance. In Figure 5 and Figure 6 the results for TM4 are presented. These would be more evident with colour composites images of the merged bands.

Figure 5 Variogram and histogram of the TM4 simulated in a 10mx10m grid.

Figure 6 Landsat TM4 (upper-left); simulated TM4 (upper-right); IHS 4 (lower-left); CN 4 (lower-right). Linear expansion 1%.
The main characteristic of the images obtained from the classical methods is that they have a close spatial resemblance with the SPOT-P. This could make the photo interpretation easier, but is an illusory advantage, once the spatial features are all reproduced in the original SPOT-P. Furthermore, these images have a final aspect of softly coloured SPOT-P images, in which the colour tones have been obtained from the Landsat TM ones. This is a clear drawback, as a thematic classification can hardly be done with those digital values.

The geostatistically merged image is more similar to the Landsat TM original images, but with better spatial feature details (derived from SPOT-P). For example, several linear features (roads, river, etc.) that are difficult to distinguish in the Landsat images are visible in these merged images. But more importantly, colours of the different features have been preserved by this method. This allows for the thematic classification of the simulated images. An unsupervised classification (K-means) was performed for five spectral classes (Figure 7).

![Figure 7 Classification of Landsat (left); classification of geostatistically merged image (right).](image)

The improvement in the detail of the classified image is obvious, while preserving the proportions of the classes (see Table 2).

<table>
<thead>
<tr>
<th>Class</th>
<th>Classified pixels</th>
<th>% of classified pixels</th>
<th>Classified pixels</th>
<th>% of classified pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>1515</td>
<td>23.7</td>
<td>13454</td>
<td>23.4</td>
</tr>
<tr>
<td>Class 2</td>
<td>1836</td>
<td>28.7</td>
<td>16960</td>
<td>29.4</td>
</tr>
<tr>
<td>Class 3</td>
<td>1662</td>
<td>26.0</td>
<td>14964</td>
<td>26.0</td>
</tr>
<tr>
<td>Class 4</td>
<td>1018</td>
<td>15.9</td>
<td>8986</td>
<td>15.6</td>
</tr>
<tr>
<td>Class 5</td>
<td>369</td>
<td>5.8</td>
<td>3236</td>
<td>5.6</td>
</tr>
</tbody>
</table>

Also interesting is the comparison of the statistical characteristics in Table 3. Here we can see that some basic statistics of Landsat TM are most closely honoured by the geostatistical
merging procedure: the merged bands have an almost equal mean and variance. We emphasize that this failure of the traditional methods is due to the necessary transformation that is applied previously to the merging process.

The geostatistical merging procedure reproduces the spatial pattern of SPOT-P as they are revealed by the variograms.

<table>
<thead>
<tr>
<th>Bands</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TM3</td>
<td>67.37</td>
<td>14.66</td>
<td>28</td>
<td>131</td>
</tr>
<tr>
<td>TM4</td>
<td>74.99</td>
<td>14.37</td>
<td>34</td>
<td>133</td>
</tr>
<tr>
<td>TM5</td>
<td>123.56</td>
<td>27.62</td>
<td>45</td>
<td>226</td>
</tr>
<tr>
<td>Geostatistical method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geo3</td>
<td>66.90</td>
<td>15.50</td>
<td>23</td>
<td>146</td>
</tr>
<tr>
<td>Geo4</td>
<td>74.49</td>
<td>15.22</td>
<td>24</td>
<td>146</td>
</tr>
<tr>
<td>Geo5</td>
<td>123.07</td>
<td>29.05</td>
<td>30</td>
<td>241</td>
</tr>
<tr>
<td>IHS method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IHS3</td>
<td>75.14</td>
<td>45.83</td>
<td>0</td>
<td>245</td>
</tr>
<tr>
<td>IHS4</td>
<td>65.41</td>
<td>36.38</td>
<td>0</td>
<td>202</td>
</tr>
<tr>
<td>IHS5</td>
<td>108.27</td>
<td>60.09</td>
<td>0</td>
<td>255</td>
</tr>
<tr>
<td>CN method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN3</td>
<td>42.89</td>
<td>9.42</td>
<td>17</td>
<td>91</td>
</tr>
<tr>
<td>CN4</td>
<td>39.42</td>
<td>7.19</td>
<td>20</td>
<td>79</td>
</tr>
<tr>
<td>CN5</td>
<td>69.43</td>
<td>12.20</td>
<td>27</td>
<td>123</td>
</tr>
</tbody>
</table>

The IHS and CN merging methods reduce the mean values (reaching a half of the original values for the CN method). The IHS method increases the variance (up to three times); giving final values higher than the corresponding bands that are merged. In contrast, the CN method produces a decrease of the variance in opposition to the reduction in pixel size.

Another important aspect concerns the global correlation coefficients between the different images (bands) used in the merging process. The quality of the spatially enhanced images can also be measured, for each band, by the correlation coefficient between the pixel values of the SPOT-P and the corresponding values of the spatially enhanced images. Column 1 of Table 4 shows the correlation coefficient between SPOT-P and the real Landsat TM image for each band, and columns 2, 3 and 4 the equivalent statistics for the geostatistical procedure, IHS and CN, respectively.

<table>
<thead>
<tr>
<th>PAN</th>
<th>PAN</th>
<th>PAN</th>
<th>PAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAN</td>
<td>PAN</td>
<td>PAN</td>
<td>PAN</td>
</tr>
<tr>
<td>TM1</td>
<td>0.83</td>
<td>Geostat1</td>
<td>0.80</td>
</tr>
<tr>
<td>TM2</td>
<td>0.83</td>
<td>Geostat2</td>
<td>0.82</td>
</tr>
<tr>
<td>TM3</td>
<td>0.82</td>
<td>Geostat3</td>
<td>0.80</td>
</tr>
<tr>
<td>TM4</td>
<td>0.74</td>
<td>Geostat4</td>
<td>0.72</td>
</tr>
<tr>
<td>TM5</td>
<td>0.72</td>
<td>Geostat5</td>
<td>0.71</td>
</tr>
<tr>
<td>TM7</td>
<td>0.70</td>
<td>Geostat7</td>
<td>0.71</td>
</tr>
</tbody>
</table>
IHS and CN methods produce a very significant increase in the correlation coefficients between the merged bands and the panchromatic one. These coefficients that are around 0.82 (for visible bands) and 0.73 (for infrared bands) in the original images, as mentioned before in section 4, reach values higher than 0.98 for the IHS merging method and 0.91 for the CN method, both methods presenting very similar values of correlation for visible and infrared bands. On the other hand, the proposed geostatistical method presents very similar correlation values to the original Landsat image, and reproduces the differences in the correlation values for the visible and non-visible channels.

The conservation of the correlation coefficients is produced at both global and local levels. Local correlation coefficients can be calculated inside a 150mx150m moving window. In Figure 8, TM4 vs. SPOT-P local correlation coefficients distributions are shown.

In the original TM/SPOT-P minimum correlation values are around -0.61. This value is related to the presence of riverside vegetation (label E in Figure 8), which produces large reflectance values in TM4 and small on the visible (panchromatic) bands (see Table 5). This behaviour is preserved only in the geostatistical method that has a minimum correlation coefficient of -0.41, while the other methods always produce positive correlation coefficients.

7 Results and discussion

This study proposes geostatistical multi-sensor image merging, based on direct sequential simulation with reference images (Carvalho et al., 2006) and on local correlogrational models. It shows that this algorithm produces images that, unlike those from other classical merging procedures, preserve the spectral characteristics of the higher-spectral resolution images.

Visual and statistical evaluation of the merged images indicates that IHS and CN change the DN of the images, which means that the spectral features are distorted, thus making impossible
a coherent spectral classification of these merged images. In the other hand, the geostatistical approach produces images that allow for an improved classification, when comparing to the classification of the original Landsat images. Moreover, it preserves the proportions of the spectral classes.

The visual aspect of the geostatistically merged images is different from that of the images obtained with classical methods (these images produce a relevant spatial resolution improvement that makes their interpretation easier), but reveals pertinent spatial features of SPOT-P, honouring the variogram and the statistics of each band.

The geostatistical method takes into account the global and local correlation coefficients between the images in the integration, and those coefficients are preserved in the merged image. This is important when working with non-visible spectral bands, which are poorly correlated with higher spatial resolution images that are usually panchromatic.

Multiscale image merging is usually a trade-off between the spectral information extracted from multispectral images and the spatial information extracted from the high spatial resolution images. Classical merging images have a rich spatial quality (same as SPOT-P) but a poor spectral quality, which makes them transformed images suitable only for visual interpretation, and useless for thematic classification, since the spectral characteristics are distorted. The geostatistical method produces an image with improved spatial resolution (compared with the original Landsat) and, in addition, it preserves the radiometric characteristics of the original high-spectral resolution image.

Most classical methods are not considered really merging methods but substitution methods. They consist of simple substitution of the high-spectral images with a high-spatial resolution image based on the correlation coefficient between the two data sets. The geostatistical method can be considered a true integration of the multisensor data, producing an image that can be upscaled back to the spatial resolution of the lower spatial resolution image with exactly the same radiometric characteristics. This is a swift procedure that permits processing large sets of data.

References


