AUTOMATIC RELATIVE REGISTRATION OF SPOT5 IMAGERY FOR COLOR MERGING

Leong Keong KWOH and Xiaojing HUANG

Centre for Remote Imaging, Sensing and Processing, National University of Singapore,
SOC1, Level 2, Lower Kent Ridge Road, Singapore 119260
crshxj@nus.edu.sg

KEY WORDS: Multispectral, Registration, Correlation, Multiresolution, Orthorectification, SPOT, Sensor Model

ABSTRACT:

An automatic process for relative registering and merging the panchromatic and multispectral images of SPOT5 is presented. The automated hierarchical local registration process makes use of feature information to select tie points from both images. The tie points are then used to refine the sensor model of the multispectral image by method of least squares solution. Both panchromatic and multispectral images are then orthorectified to a geo-referenced coordinate system using the available one kilometre gridded Digital Elevation Models (GLOBE or SRTM DEM). The merged high resolution colour imagery is then obtained by multiplying each spectral bands with a sharpening factor computed from the intensity values of the orthorectified panchromatic image and the corresponding pixels of the orthorectified multispectral images.

1. INTRODUCTION

The SPOT5 satellite carries two high resolution geometric instruments (HRG) enabling it to map large area of the Earth (60km x 60km). Each HRG instrument on-board the SPOT5 satellite images the ground with the panchromatic (PAN) band (2.5 m or 5 m resolution) pointing slightly forward (0.529°) and the multispectral (XS) bands (10 m resolution) slightly backward (-0.529°). This results in a slight relief displacement in opposite directions for the panchromatic and multispectral images and also a time interval of about 1.5 seconds between the PAN and XS images over a same point on the ground. Therefore there is a need to co-register both images together and to eliminate the relief displacement before they can be merged into a high resolution colour image.

Christophe Latry and Bernard Rouge (Christophe Latry, 2003) used physical sensor models plus massive local image correlation to register the images and eliminate the relief displacement. This technique, however, may not be effective if the images have high amount of cloud cover or water bodies. Kwoh, in an earlier work (Kwoh, 2003), used a method based on coarse orthorectified PAN and XS images with the publicly available 1km gridded GLOBE DEM to eliminate the relief displacement differences due to the different look direction in the Pan and XS image; and then using the simple affine transformation to register the two orthorectified images. The method had assumed that both the Pan and XS image have no relative orientation bias differences, which is logical since the Pan and XS images were taken at a mere 1.5 second apart.

In this paper, we present an improved method where the PAN and XS images are registered relative to one another by refining the physical sensor model of one image relative to the other (instead of the simple affine transformation). The orthorectified images generated with the same DEM will then be well registered to one another and will have no relative relief displacement errors. Any orientation bias differences between the PAN and XS image will also be removed during the sensor model refinement process.

The method is implemented in 4 steps as shown in Figure 1— (1) Automated image matching and tie point selection; (2) Relative SPOT5 sensor model affinement for XS image with respect to PAN image with tie points; (3) Orthorectifying both PAN and XS images to a common georeferenced coordinate system by using one kilometre resolution GLOBE or SRTM Digital Elevation Models (DEM); (4) obtaining high resolution colour image by merging the high resolution (2.5m or 5m) panchromatic image with the lower resolution (10m) multispectral image.

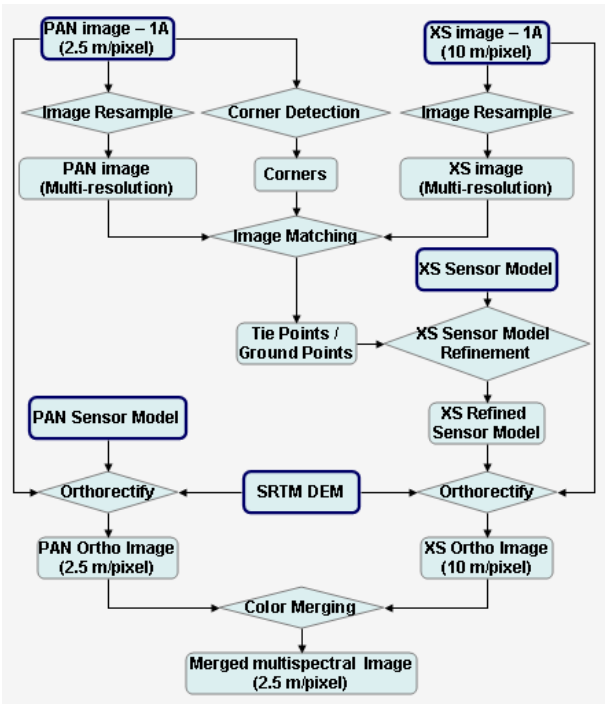

Figure 1. Relative registration and colour merging diagram

2. AUTOMATIC TIE POINT SELECTION AND IMAGE MATCHING

2.1 Selection of Image Patches

In order to find corresponding image tie points for registration, one needs to select and match image features with high information content. Our approach is to automatically search for image patches with good edges and corners. There are many filters (Bergholm, 1987, Iverson, 1995, Rothwell 1995 and Smith 1997) to detect edges and corners. We chose the Canny filter (Canny, 1986) to locate the edges in the image and then Harris (Harris, 1988) corner detector to detect corners from the detected edges. In this way, we reduce unnecessary edges, keep the corners with most feature information, and also increase the processing speed for image matching.

There are a few processing steps in the detailed Canny detector (Canny, 1986): Gaussian smoothing, gradient computation, non-maximal suppression and hysteresis. We used a 11×11 Gaussian window with a standard deviation σ of 2.0 for the Guassian smoothing. A 3×3 Sobel filter is then used to derive the edge strength, i.e. the gradient of image intensity. In performing hysteresis, pixels with edge strength greater than 98.5 percent of all strength or any connected pixel with strength greater than 50 percent are marked as edge pixels. In Harris (Harris, 1988) corner detection, a 5×5 Gaussian window with a standard deviation σ of 1.0 is used as the gradient weight, and 0.04 is chosen as the constant value in the corner/edge response function.

The edges and corners detection are only applied to the high resolution PAN image. Image patches that contain one or more corner points are used as candidates to approximately select the corresponding image patches in the XS image based on the a-priori imaging parameters provided in the image header files.

2.2 Hierarchical local matching

After corner detection, a hierarchical local matching approach is implemented to match the image patches. The PAN and XS images are scaled by various factors to create a few hierarchical layers, with the highest resolution layer having a pixel size equal to that of the PAN image. The matching of two patches at various layers is computed by maximising the normalized cross correlation:

$$C(A,B)=\frac{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}((A_{ij}-\overline{A})\times(B_{ij}-\overline{B}))}{\sqrt{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(A_{ij}-\overline{A})^2\times\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(B_{ij}-\overline{B})^2}} \quad (1)$$

where $C(A, B)$ = the normalized cross correlation between the scaled PAN and XS image patches

N = window size of image patch

A_{ij} and B_{ij} = individual pixel values in the PAN and XS patches, respectively

$\overline{A}$ and $\overline{B}$ = average pixel values in the PAN and XS patches, respectively.

Image matching starts with the lowest resolution layer. The panchromatic patch that contains one or more detected corners is used as the reference. By shifting the multispectral patch around its initial position, the location where the normalised cross correlation of the two image patches is highest within a range of shifting is selected as the matched position. For each higher resolution layer, the matched position derived from the previous lower resolution layer is used as the initial position and matching is performed to improve the accuracy of the matched position. A threshold of 0.85 for the cross correlation coefficient of the highest resolution layer has been set to reject those image patches with poor matching.

The centre coordinates of the matched image patches are used as the respective coordinates of tie point pairs.

3. SENSOR MODEL AND REFINEMENT

3.1 Direct Sensor Model

SPOT5 image is delivered with header file or metadata file, which contains the sensor model imaging parameters such as satellite state vectors, attitudes, pixel look directions etc. Details of the SPOT5 sensor model, including the use of the imaging parameters is described in SPOT Satellite Geometry Handbook, 2002.

The basic model described in the SPOT Satellite Geometry Handbook is the direct sensor model which relates the pixel coordinates (*samp, line*) to geographical coordinates (*lat, lon*) as follows:

$$\begin{aligned} lon &= F(samp, line, h) \\ lat &= G(samp, line, h) \end{aligned} \quad (2)$$

where F and G are sets of various functions, collectively known as the direct sensor model. The relation of F and G to the various imaging parameters is briefly described below:

For every pixel (*samp*, *line*) in an image, the *samp* coordinate is related to pixel look direction within the sensor, Ψ, while the *line* coordinate is related to time *t.*, The attitude (*yaw, pitch roll*) of the sensor at each line is given in the header file. The satellite sensor position and velocity can be interpolated at t from set of time-tagged satellite state vectors. From the pixel look direction, attitude, position $\vec{R}$ and velocity $\vec{V}$, the intersection of look vector with the topographic surface with altitude h above the earth ellipsoid gives the geographic (*lon, lat*) coordinates.

3.2 Inverse Model

For orthorectification and model refinement by least squares adjustments, it is desirable to relate the pixels coordinates (*line,samp*) in terms of geographical coordinates (*lon,lat,h*).

$$\begin{aligned} samp &= f(lon, lat, h) \\ line &= g(lon, lat, h) \end{aligned} \quad (3)$$

This formulation is commonly referred to as the inverse sensor model. As described in the SPOT Satellite Geometry Handbook, f and g can not be directly derived from the SPOT header information, but is obtained with an iteratively algorithm based on the direct model F and G.

3.3 XS Image Sensor Model refinement

From the PAN and XS image header metadata files, we will be able to establish the respective sensor model independently. To co-register the images, we choose to refine the XS image sensor model relative to the PAN image by introducing 3 parameters, c_{0y}, c_{0p}, c_{0r}, to correct for the biases differences in *yaw*, *pitch* and *roll* respectively. Such that:

$$\begin{aligned} yaw &= yaw_0 + c_{0y} \\ pitch &= pitch_0 + c_{0p} \\ roll &= roll_0 + c_{0r} \end{aligned} \tag{4}$$

where yaw_0, $pitch_0$ and $roll_0$ = initial yaw, pitch and roll value calculated from sensor model of XS image, respectively

yaw, *pitch* and *roll* = refined yaw, pitch and roll, respectively

c_{0y}, c_{0p}, c_{0r} = correction parameters.

The correction parameters, c_{0y}, c_{0p}, c_{0r} can be expanded into higher order polynomials, but for this work, the zeroth order constant term corrections is found to yield sufficient accuracy.

3.4 Height Constrain

In the solution for the model corrections parameters, the geographical coordinates of the tie points were also solved simultaneously. Since the angle subtended between the PAN and XS is only 1.06°, the least squares solution was found to be unstable. This stability problem can be remedied by introducing a height control value to one of the tie points. This height control need not be accurate. We have conveniently chose to extract it from the 1km gridded GLOBE or SRTM DEM (same DEM will be used later for orthorectification of the images).

3.5 Procedures for refinement

By automatic selecting about 20 tie points or more, and one ground point for height control, the refined parameters for yaw, pitch and roll correction can be derived from Equations 3 and 4 by using least mean square method as following:

$$\begin{aligned} \mu_{1i} &= samp_{1i} - f_1(lon_i, lat_i, h_i) \\ \nu_{1i} &= line_{1i} - g_1(lon_i, lat_i, h_i) \\ \mu_{2i} &= samp_{2i} - f_2(lon_i, lat_i, h_i) \\ \nu_{2i} &= line_{2i} - g_2(lon_i, lat_i, h_i) \end{aligned} \tag{5}$$

where 1 = Pan image

2 = Xs image

$i \in [0, n)$

n = total number of tie points, and one ground point

Using the first order Taylor expansion and least-square estimation method, the ground coordinates of tie points and the refined parameters c_{0y}, c_{0p} and c_{0r} can be computed as follows :

$$\begin{aligned} \mu_{1i} &\approx samp_{1i} - f_1(lon_{i0}, lat_{i0}, h_{i0}) - \frac{\partial f_1}{\partial lon_i}\cdot \delta lon_i - \frac{\partial f_1}{\partial lat_i}\cdot \delta lat_i \\ &\quad - \frac{\partial f_1}{\partial h_i}\cdot \delta h_i \\ \nu_{1i} &\approx line_{1i} - g_1(lon_{i0}, lat_{i0}, h_{i0}) - \frac{\partial g_1}{\partial lon_i}\cdot \delta lon_i - \frac{\partial g_1}{\partial lat_i}\cdot \delta lat_i \\ &\quad - \frac{\partial g_1}{\partial h_i}\cdot \delta h_i \\ \mu_{2i} &\approx samp_{2i} - f_2(lon_{i0}, lat_{i0}, h_{i0}) - \frac{\partial f_2}{\partial lon_i}\cdot \delta lon_i - \frac{\partial f_2}{\partial lat_i}\cdot \delta lat_i \\ &\quad - \frac{\partial f_2}{\partial h_i}\cdot \delta h_i - \frac{\partial f_2}{\partial c_{0y}}\cdot \delta c_{0y} - \frac{\partial f_2}{\partial c_{0p}}\cdot \delta c_{0p} - \frac{\partial f_2}{\partial c_{0r}}\cdot \delta c_{0r} \\ \nu_{2i} &\approx line_{2i} - g_2(lon_{i0}, lat_{i0}, h_{i0}) - \frac{\partial g_2}{\partial lon_i}\cdot \delta lon_i - \frac{\partial g_2}{\partial lat_i}\cdot \delta lat_i \\ &\quad - \frac{\partial g_2}{\partial h_i}\cdot \delta h_i - \frac{\partial g_2}{\partial c_{0y}}\cdot \delta c_{0y} - \frac{\partial g_2}{\partial c_{op}}\cdot \delta c_{0p} - \cdots - \frac{\partial g_2}{\partial c_{0r}}\cdot \delta c_{0r} \end{aligned} \tag{6}$$

For $i = n-1$, i.e. ground point with height constrain, there is no δh term in equation 6.

The above equation 6 can be grouped as below:

$$E = \sum (\mu_{1i}{}^2 + \nu_{1i}{}^2 + \mu_{2i}{}^2 + \nu_{2i}{}^2) \tag{7}$$

$$V_{(4n)\times 1} + A_{(4n)\times(3n-1+3)} \bullet X_{(3n-1+3)\times 1} = L_{4n} \tag{8}$$

where V stand for $\mu_{1i}, \nu_{1i}, \mu_{2i}$ and ν_{2i},

$$\begin{aligned} L_{4i,0} &= -samp_{1i} + f_1(lon_{i0}, lat_{i0}, h_{i0}) \\ L_{4i+1,0} &= -line_{1i} + g_1(lon_{i0}, lat_{i0}, h_{i0}) \\ L_{4i+2,0} &= -samp_{2i} + f_2(lon_{i0}, lat_{i0}, h_{i0}) \\ L_{4i+3,0} &= -line_{2i} + g_2(lon_{i0}, lat_{i0}, h_{i0}) \end{aligned}$$

$$A_{4i} = \left[\overbrace{0,0,\cdots,0}^{3\times(i-1)}, -\frac{\partial f_1}{\partial lon_i}, -\frac{\partial f_1}{\partial lat_i} - \frac{\partial f_1}{\partial h_i}, \overbrace{0,0,\cdots,0}^{3\times(n-i)-1}, \overbrace{0,0,0}^{3} \right]$$

$$A_{4i+1} = \left[\overbrace{0,0,\cdots,0}^{3\times(i-1)}, -\frac{\partial g_1}{\partial lon_i}, -\frac{\partial g_1}{\partial lat_i}, -\frac{\partial g_1}{\partial h_i}, \overbrace{0,0,\cdots,0}^{3\times(n-i)-1}, \overbrace{0,0,0}^{3} \right]$$

$$A_{4i+2} = \left[\overbrace{0,0,\cdots,0}^{3\times(i-1)}, -\frac{\partial f_2}{\partial lon_i}, -\frac{\partial f_2}{\partial lat_i}, -\frac{\partial f_2}{\partial h_i}, \overbrace{0,0,\cdots,0}^{3\times(n-i)-1}, \overbrace{-\frac{\partial f_2}{\partial c_{0y}}, -\frac{\partial f_2}{\partial c_{0p}}, -\frac{\partial f_2}{\partial c_{0r}}}^{3} \right]$$

$$A_{4i+3} = \left[\overbrace{0,0,\cdots,0}^{3\times(i-1)}, -\frac{\partial g_2}{\partial lon_i}, -\frac{\partial g_2}{\partial lat_i}, -\frac{\partial g_2}{\partial h_i}, \overbrace{0,0,\cdots,0}^{3\times(n-i)-1}, \overbrace{-\frac{\partial g_2}{\partial c_{0y}}, -\frac{\partial g_2}{\partial c_{0p}}, -\frac{\partial g_2}{\partial c_{0r}}}^{3} \right]$$

$$X = \left[\overbrace{\delta lon_0, \delta lat_0, \delta h_0, \cdots, \delta lon_{n-1}, \delta lat_{n-1}}^{3\times n-1}, \overbrace{\delta c_{0y}, \delta c_{0p}, \delta c_{0r}}^{3} \right]^T$$

$$X_0 = \left[\overbrace{lon_0, lat_0, h_0, \cdots, lon_{n-1}, lat_{n-1}}^{3\times n-1}, \overbrace{c_{0y}, c_{0p}, c_{0r}}^{3} \right]^T \quad (9)$$

$$X_{new} = X_0 + X \quad (10)$$

In equation 9, there are no derivation terms to δh_{n-1}, which is the height of ground point.

Since the inverse f_1, g_1, f_2 and g_2 are set of functions, some of which are iterative functions, their partial derivatives in A cannot be directly derived. They are instead analytically computed.

With the initial h_i of tie points to 0, lon_i, lat_i of both tie points and ground point calculated from equation 2, and 3 coefficients to 0, X can be solved with equation 8. Then the initial values are replaced with the new values X_{new} with equation 10. The least square estimation is repeated until E is stabilised to a minimal.

4. IMAGE ORTHORECTIFICATION

The original sensor model of PAN and the refined sensor model of XS images are used to orthorectified PAN and XS images respectively. As mentioned earlier, the look angles between the PAN and XS Image for SPOT 5 HRG instrument is 1.06°. This is equivalent to a B/H ratio of less than 0.02. If the error in the DEM used is 100m in height, the relief displacement error between the PAN and XS images will be about 2m, less than a PAN pixel size. This insensitivity to DEM accuracy suggests that a coarse orthorectification with the 1km gridded GLOBE or SRTM will be accurate enough for removing the relative displacement error to sub pixel accuracies.

5. COLOR MERGING

Merging of the high resolution colour imagery is performed by multiplying each spectral band with a sharpening factor computed from the intensity values of the panchromatic image and the corresponding pixels of the multispectral images. The equation for merging is described below:

$$I_{i_{merge}} = \frac{I_{pan}}{I_{xs}} \cdot I_{i_{xs}} \quad (11)$$

where i = the band index of SPOT5 imagery

I_{pan} and I_{xs} = intensity of panchromatic and multispectral imagery, respectively

$I_{i_{xs}}$ and $I_{i_{merge}}$ = grey value of each band in multispectral and merged imagery, respectively

The factor (I_{pan} / I_{xs}) is referred to as the sharpening factor.

Since the spectral range of the PAN image overlaps only the RED and GREEN bands of the XS image, we compute the $I_{i_{xs}}$ by taking averaging of pixel values of only the RED and GREEN bands. The non-overlapping bands (NIR band and SWIR band) can either be sharpened with the same sharpening factor or merely resampled to the resolution of the panchromatic imagery. The pseudo-natural colour SPOT 5 imagery can also be merged with the panchromatic imagery in the same way.

6. TEST IMAGERY AND RESULTS

To test the method developed, we chose a site in Bandung, Indonesia, where there are sullen volcanoes and the terrain height variation from a few tens to more than 3000 meters. A pair of 2.5m PAN and 10m XS SPOT 5 images taken from the same instrument were selected over the area. Figure 2 and Figure 4 show the overview of level 1A images of PAN and XS respectively. Parts of the full resolution images are shown in Figure 3 and 5.

The edges of PAN imagery detected by Canny detector are shown in Figure 6 in grey, while the corners detected by Harris corner detector are shown in black in the same figure. Total 40 image patches have been matched and the tie points selected automatically for the sensor model refinement. The rms residual after least squares solution of the refined XS sensor model is 0.7 pixel (PAN image pixel size).

Figure 7 shows the overview of orthorectified PAN image after applying its sensor model. And Figure 8 shows the overview of orthorectified XS image after applying its refined sensor model. The final merged 2.5m pixel pseudo-natural color imagery (merged from both orthorectified PAN and XS images) is show in Figure 9 in overview and Figure 10 in full resolution.

7. CONCLUSION

To conclude, an automated relative image registration method based on feature information (edges and corners) and refinement of XS sensor model relative to the PAN sensor model is described. The GLOBE or SRTM DEM used in orthorectification was sufficient to eliminate the relative relief displacement to better than one PAN pixel. The PAN and XS orthorectified images generated can then be merged to produce a high resolution multispectral image, that has the high resolution of PAN and colour information of XS. The method has been tested on a few SPOT5 PAN and XS image pairs of various incidence angles and various relief situations. The results were consistent.

We expect the automatic matching method to have problems in imagery where there are too much clouds or water. In such situations, the program may not be able to find enough image patches with good edges and corners content. An option has been provided to manually provide the corresponding image registration points and these need only be approximately

located. The hierarchical image matching, sensor model refining and colour merging can still be applied in the same way.

References:

Bergholm, F., 1987. Edge forcusing. *IEEE Transactions on pattern analysis and machine intelligence*, vol. PAMI-9, no. 6, pp. 726-741.

Canny, J., 1986. A computational Approach to edge detection. *IEEE trans on pattern analysis and machine intelligence*, vol. PAMI-8, no. 6, pp. 679-696.

Christophe Latry, Bernard Rouge, 2003. Supper resolution: quincunx sampling and fusion processing. *IEEE International Geoscience and Remote Sensing Symposium, CDROM*.

Harris, C., and Stephens, M., August 1988. A combined corner and edge detector. *Proc. 4th Alvey vision conf., Manchester*, pp. 189-192.

Iverson L.A. and Steven, W. Z., 1995. Logical/linear operators for image curves. *IEEE Transactions on pattern analysis and machine intelligence*, vol. 17, no. 10, pp. 982-996

Kwoh, L.K. and Huang X., 2003. Automatic Image Registration and Color Merging for SPOT5 Imagery. *IEEE International Geoscience and Remote Sensing Symposium, CDROM*.

Nalwa, V.S. and Binford, T.O., 1986. Logical/linear operators for image curves. *IEEE Transactions on pattern analysis and machine intelligence*, PAMI-8, no. 6, pp. 699-714

Rothwell, C.A., Mundy, J.L., Hoffman, W. and Nguyen, V.D., 1995. Driving vision by topology. *International Symposium on Computer Vision*, pp. 395-400.

Smith, S. M., and Brady, J. M., 1997. SUSAN -- A new approach to low level image processing. *International journal of computer vision*, 23(1), pp. 45-78.

SPOT Satellite Geometry Handbook, 2002., S-NT-73-12-SI.

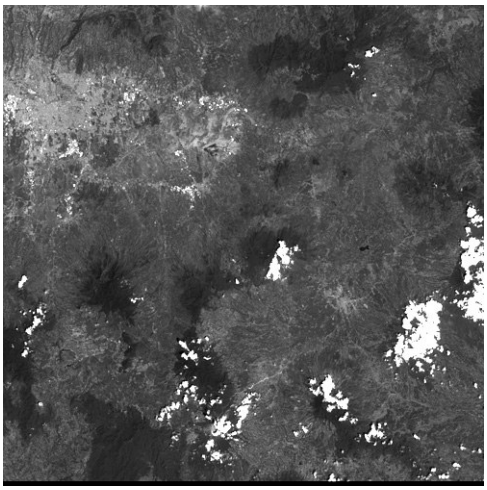

Figure 2. Overview of SPOT5 1A THR image
© CNES 2003

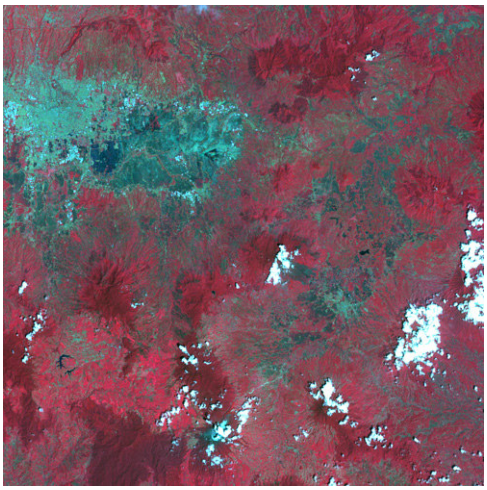

Figure 4. Overview of SPOT5 1A XS image
© CNES 2003

Figure 3. Part of SPOT5 1A THR image in full resolution 2.5m/pixel © CNES 2003

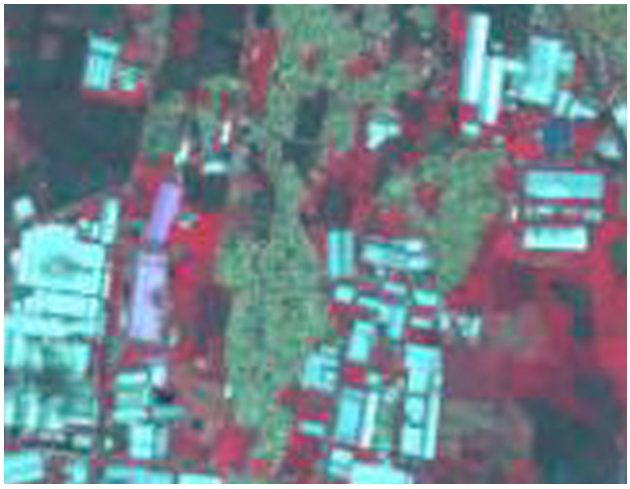

Figure 5. Part of SPOT5 1A multispectral image in full resolution 10m/pixel © CNES 2003

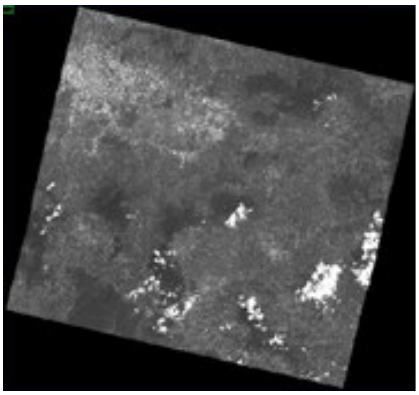

Figure 7. Overview of ortho-rectified PAN image

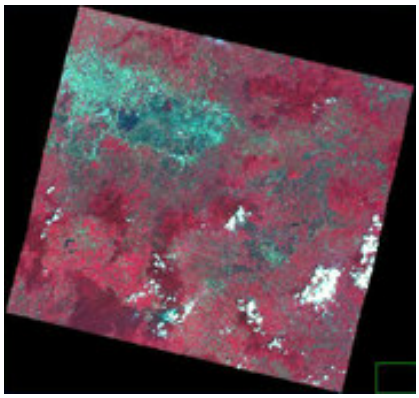

Figure 8. Overview of ortho-rectified XS image with refined sensor model

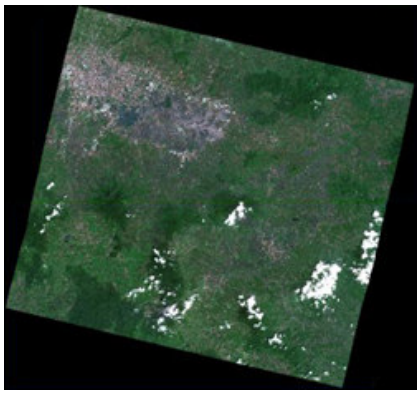

Figure 9. Overview of merged pseudo-natural colour image

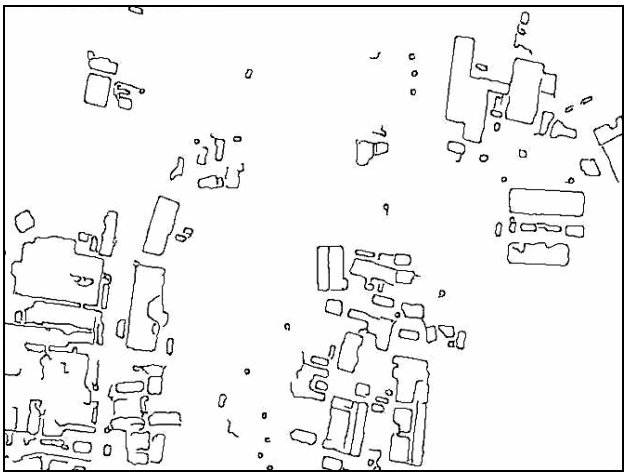

Figure 6. The edges and corners of PAN imagery

Figure 10. Part of merged pseudo-natural colour image in full resolution 2.5m/pixel