

Automatic Ortho-rectification of ASTER Images by Matching Digital Elevation Models

José A. Gonçalves and André R.S. Marçal

Faculdade de Ciências, Universidade do Porto,
Rua Campo Alegre, 4169-007 Porto, Portugal
jagoncal@fc.up.pt

Abstract. The ortho-rectification of satellite images is normally a time-consuming and expensive task. It can strongly benefit from automatic or semi-automatic procedures in order to avoid field work for ground control survey. This article presents an automated method to automatically georeference satellite images acquired by the ASTER sensor. The method is based on automatic matching of images and Digital Elevation Models (DEM). First a DEM is extracted from the two stereo image bands, which then is matched to the SRTM (Shuttle Radar Topography Mission) DEM in order to correct its position in a map reference system. This allows for the simultaneous correction of image geo-location, which then is followed by the ortho-rectification. The method was applied to an ASTER image from North Portugal and assessed using topographic map data. It was possible to geo-reference images in hilly terrain with positional accuracy better than one pixel.

1 Introduction

Precise registration of satellite images is required for several applications, such as image fusion or change detection by multi-temporal image analysis. Usually images need to be ortho-rectified, using a DEM, and georeferenced to a standard map projection in order to remove relief distortions and allow for the integration with other geo-referenced data in a Geographical Information System (GIS). These procedures frequently require the field survey of ground control points, which in many cases is very expensive or not possible at all, especially in remote areas, where satellite imagery is most important for mapping purposes.

Modern satellite sensors incorporate navigation equipment, such as GPS receivers and attitude sensors that allow for the determination of an approximate image geo-location. This is known to have an uncertainty of the order of 50 meters, in the case of ASTER [1]. However, since the pixel size is a few times smaller (15 meters for the visible and near-infrared sensor), Ground Control Points (GCP) are still required in order to geo-reference images with positional accuracy better than pixel size. Medium scale topographic maps (e.g. scale 1:25,000) are not available for most of the world, so other remote sensing data sources can provide the GCPs, with the advantage of possibly automating their identification.

A possible solution to obtain GCPs is to use global databanks of georeferenced images. A global coverage of Landsat ETM images exists presently – the GeoCover of

the Global Land Coverage Facility – freely available on the internet, with a spatial resolution of 15 meters. However its positional uncertainty is of 40 meters [2], which is not appropriate for ASTER. Another difficulty posed by this solution is that, due to different image characteristics, automatic matching between the two datasets would most likely fail.

Another remote sensing dataset currently available, with a nearly global coverage is a DEM derived from data acquired by the Space Shuttle during the Shuttle Radar Topographic Mission (SRTM), carried out in February 2000. It covers 80 % of the land mass, with a resolution of 3 arc-seconds (90 meters at the equator), a vertical and horizontal accuracy, at 90% confidence, of respectively 7 and 10 meters [3] and it is freely available on the internet for download. An interesting application of the SRTM DEM is to correct the location of other DEM datasets, with positional uncertainty, by means of maximizing the correlation between the two elevation datasets. This is known to be possible with a sub-pixel accuracy of 5 to 10 meters (i.e., better than $1/10^{\text{th}}$ of a pixel) for areas where some relief exists [4].

Since ASTER can provide elevation data from its own images, the idea for the work described in this paper was to use SRTM as ground control to correct the location of the DEM extracted from ASTER images and consequently bring the images to their correct orientation. This methodology would allow for a strong minimization of the field survey efforts to produce reliable ortho-images.

This methodology was implemented, with all software developed for that purpose. The production of the DEM from the ASTER image by stereo matching is not the main objective. It only intends to produce reliable elevation data for the geo-location correction by matching the SRTM DEM. In both matching tasks simple correlation methods were used.

2 Height Data Extraction from ASTER

ASTER has the capability to acquire along-track stereoscopic imagery. The visible and near infrared system, with a ground sampling of 15 m, has two sensors in the band of 0.78 to 0.86 μm (named band 3), one pointing in the nadir direction and the other pointing backwards, with an off-nadir angle of 26° [5]. These bands are referenced as 3N and 3B.

2.1 ASTER Image Geometry

In order to obtain 3D ground coordinates from the stereopair, image geometry must be carefully treated. Optical satellite sensors are linear arrays that acquire images by scanning, being each line a central projection. For each image of level 1A (i.e., images in sensor geometry) a set of points derived from navigation data are made available within the image ancillary data. These are the coordinates of the satellite position (S) and left (L) and right (R) points, for a set of lines, with a spacing of 400 pixels, as shown in figure 1.

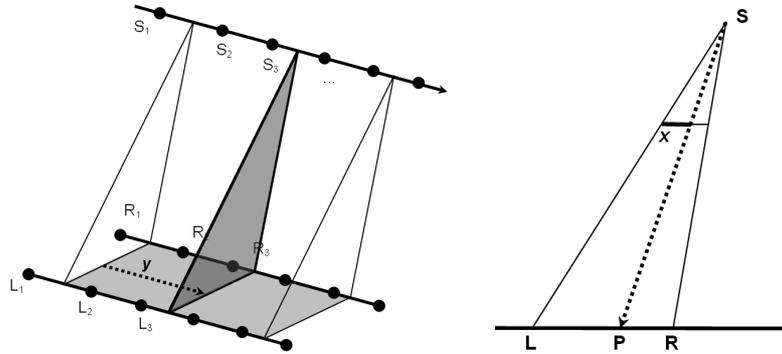


Fig. 1. Image geometry: left image shows the acquisition of line y and right image shows the acquisition plane and the position of pixel, x , along the line

For a given image point P , with image coordinates x and y , in pixel units, it is possible to calculate coordinates of points L , R and S , by interpolation of the given data for the y coordinate. Normally a 3rd degree polynomial is sufficient for that. Then, as shown in figure 1 (right), two unit vectors of directions SL and SR are calculated. The actual pointing direction of SP is calculated from the x coordinate, taking into account that the image is generated by a central projection. In this way it is possible to do an image to object projection by intersecting the pointing direction with the Earth surface, if the terrain height is known [6]. Normally, constant height (h) surfaces are taken as ellipsoids, with semi-axis $a+h$, and $b+h$, where a and b are the semi-axis of the reference ellipsoid, normally WGS84. The intersection procedure is straight forward and can be seen in general terms as a function like:

$$\begin{aligned} \lambda &= F(x, y, h) \\ \varphi &= G(x, y, h) \end{aligned} \quad (1)$$

where λ and φ are the longitude and latitude in the WGS84 ellipsoid. Inversely, similar formulas can be established for the inverse projection, i.e, to calculate image coordinates for a given location (λ, φ, h) . These projections are established for both images, 3N and 3B.

These geometrical procedures are used within this work for two purposes: the first is to approximately register images 3N and 3B. Assuming the terrain coincides with the reference ellipsoid points are projected from image 3N onto height zero and then onto image 3B. A set of points is obtained that can be used to register the images, which in this case was done by a 3rd degree polynomial. The lack of coincidence between the images is essentially due to the relief effect: points are displaced due to the actual terrain height in different manners due to different pointing directions. The matching procedure, described below, determines this parallax effect.

Figure 2 shows one of the full 3N images, from North Portugal, tested and samples of the 3N and 3B (registered to the 3N image space) images. The displacement between images can be seen, as well as a slight difference in shape due to terrain slope.

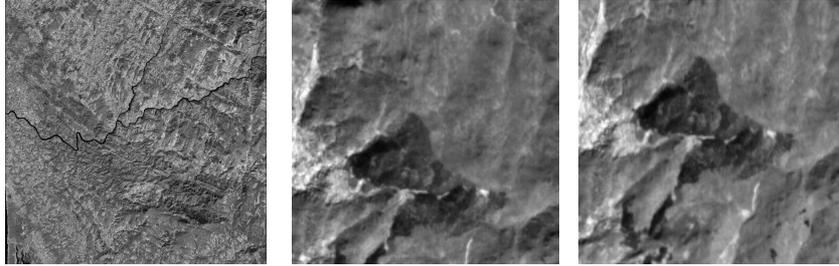


Fig. 2. ASTER image used: left – full 3N image, center – sample of the 3N image (200 by 200 pixels), right – corresponding 3B image (registered to 3N)

Once a conjugate point is obtained, the intersection between lines defined for images 3N and 3B is determined. Since they usually do not exactly intersect, due to errors in image coordinates, the minimum distance point is considered as the intersection, as shown in figure 3.

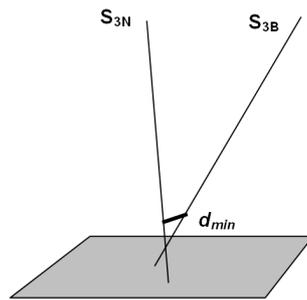


Fig. 3. Space intersection of pointing directions for a 3N and 3B conjugate point

This procedure is applied to the set of conjugate points that results from the matching process, in order to generate a DEM. The matching and the DEM generation procedures are described in the following sections.

2.2 Matching Between ASTER Image Bands 3N and 3B

The objective of this task is to have as many tie points, or conjugate points, as possible for the generation of a reasonable DEM. The nadir view ASTERs band 3N is considered as reference image and the backward view, ASTERs band 3B, as input image.

Initially the reference image is used to produce a Tie Point Suitability Index (TPSI) image. The purpose of a TPSI is to identify the best locations on an image to search for tie points. These locations should have distinct features in order to maximise the chances of a correct selection on the image matching process. For a given target matrix of size t (for example 3×3 for $t=3$) a TPSI value is attributed to each pixel of the image, thus producing a TPSI image, where the highest values should correspond to

the most promising locations to search for tie points. Four TPSI were used: Basic, Composed, Ratio and Prewitt. A detailed description of these TPSI is available at [7].

The reference image is divided in small 6 by 6 pixel sections, which for an ASTER scene of 4200 by 4100 pixels results in over 450 000 sections. For each section, the pixel with highest TPSI value is selected as candidate tie point.

The tie point selection process is based on image matching by normalized two-dimensional cross-correlation in the spatial domain. A target matrix T (of size t , usually small) is established in the reference image and a search window S (of size s , larger than t) is examined in the input image. The convolution between T and all sub-window of S (of size t) is performed, resulting in a set of correlation coefficients (r), from -1 to 1. The best match will be the pixel of highest correlation in the search window. The MATLAB implementation of the normalized two-dimensional cross-correlation function was used in this work [8], with a target matrix of 5x5 and a search window of 9x61 pixels. This choice for the search window size was done as the main displacement between images 3N and 3B are along one direction and are mainly due to the surface relief. Once a set of conjugate points from a strip of sections is made available, a further refinement is made using sub-pixel matching algorithm based on MATLAB, which is effective if the region around the central point does not have zero standard deviation and if they are not poorly correlated. [8].

For each pixel selected as candidate tie point in the reference image, the matching process will provide a conjugate pair in the input image. This will be the location in the search window where the convolution between the target window and the search sub-window is maximum. However, this will not necessarily be a suitable match, as the presence of clouds, noise, or other similar areas within the search window, might result on the selection of the wrong location. A criterion to reject doubtful matches was used. If all correlation coefficients between the target window and all sub-windows in the search window are below 0.75, the point is rejected. Using this criterion, a total of 15% of the points initially made available (460,000) were rejected at this stage.

2.3 Generation of a DEM

Three-dimensional coordinates were determined for the matched points by the intersection procedure described before. In each intersection process a distance between the two lines is determined, which can also be used as a further criterion for rejecting points, since a large displacement between the lines indicates a poor matching in x direction. The criterion used was to reject points with a separation larger than 15 meters (1 pixel). Only very few were rejected at this stage.

The DEM was generated from these valid points, by a TIN (Triangulated Irregular Network) and linear interpolation, with a spacing of 90 meters, closer to the resolution of the SRTM DEM (3 arc-seconds). Figure 4 shows a portion of the extracted DEM, for a region of 5 by 5 km². Patches of this size were extracted, since the matching to the SRTM DEM will be done independently for different parts of the image area. Occasionally blunders appeared, but in small number. A DEM filtering procedure, for example by eliminating points where slope is too large, could be used to remove them.

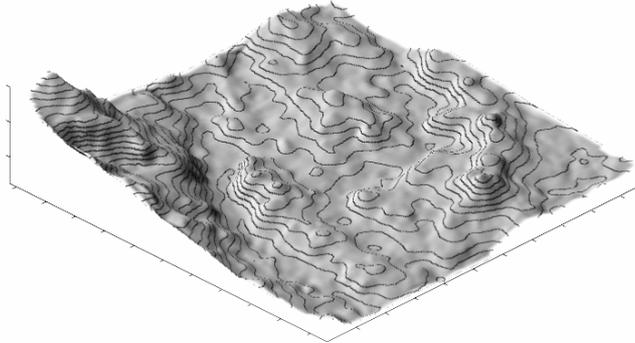


Fig. 4. Perspective view of the DEM (shaded relief and contours) generated from matched points, for an area of 5 km by 5 km

3 Improvement of Image Orientation

The DEM extracted from ASTER is not correctly geo-located because it was calculated from the initial image orientation parameters, known to have a positional uncertainty of the order of 50 meters. The next step is to determine the shift to apply to this DEM in order to match it to the SRTM DEM. The following sections describe this process in general and its application in the present case.

3.1 DEM Matching

The DEM matching is applied to small portions of two DEMs, a DEM to adjust (in our case the ASTER DEM) and a reference DEM (the SRTM DEM). Figure 5 shows the profiles of SRTM and one of the extracted DEMs. A slight shift can be seen.

The first DEM is shifted in longitude and latitude directions until finding the maximum correlation. Since images of elevation data are very similar, the correlation coefficient is normally high, very close to 1. Anyway, the slight shift required to align both DEMs can be detected with a very good degree of accuracy. Figure 6 shows an example plot of the correlation coefficient (between the ASTER DEM and the SRTM DEM) as a function of the shifts applied to the ASTER DEM. The initial correlation (shift zero) was more than 98% but, for a shift of 39 m in east direction and 8 m to north, a maximum correlation of more than 99% could be obtained. Tests with SRTM DEM and DEMs derived from topographic maps revealed an accuracy of 5 to 10 meters in the shift determined [4].

Since different shifts may be needed for different parts of the image area, the matching was applied to small portions, with a dimension of 5 by 5 km². In fact the DEM to adjust does not need to be a grid but can be only a subset of the points (λ , φ , h_{ASTER}) extracted from the matching. For those locations (λ , φ), heights are interpolated from the SRTM DEM; then the correlation coefficient between both height datasets is calculated. Small shifts in longitude and latitude directions, within a small range (± 100 m), are tested in order to search for the maximum correlation. Once the maximum correlation is obtained, a vertical adjustment between surfaces is also done by calculating the mean of height differences.

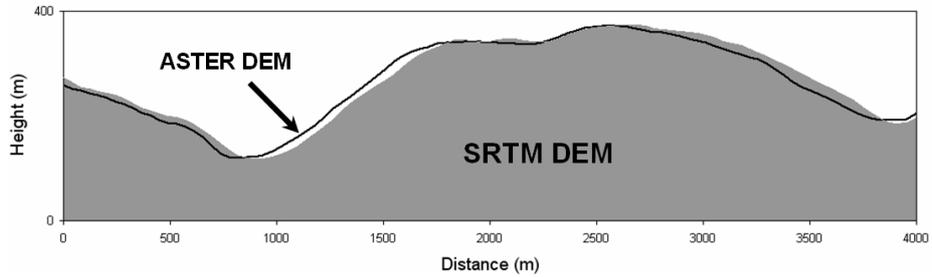


Fig. 5. Example of two DEMs: SRTM (filled in grey) and ASTER (black line), representing the same relief shape but with slightly different locations

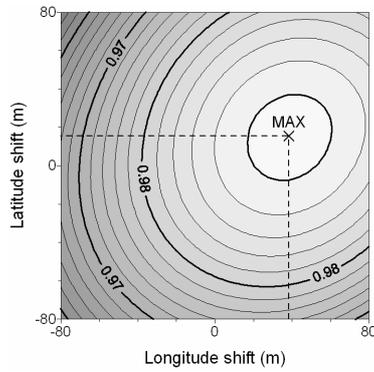


Fig. 6. Plot of the correlation coefficient as function of the shift applied to an ASTER DEM in order to align it to the SRTM DEM

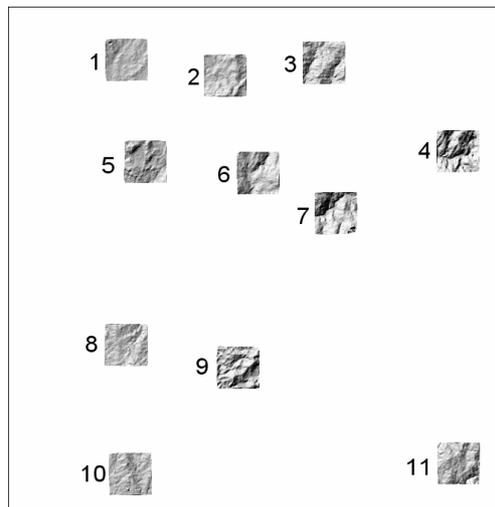


Fig. 7. Eleven DEM patches of 5 by 5 Km² extracted from the ASTER DEM. The boundary corresponds to the 3N image

Figure 7 shows the 11 DEM patches extracted, which were matched to the SRTM DEM. The important fact for this method to work is that relief exists in the area, even if moderate, with slopes in different directions.

For the purpose of this test the squares were manually chosen. In a fully automatic approach they might be selected randomly, with the criterion of not being flat. A mean slope of at least 5 % would be necessary.

This procedure was done for a total of 11 locations distributed over the image area, all with some relief. For each area a shift in longitude, latitude and height ($\Delta\lambda$, $\Delta\varphi$, Δh) was determined. Table 1 summarizes the results obtained: the shifts found, in meters, and the maximum correlation coefficients. The mean and standard deviations of the shifts were also calculated.

Table 1. Shifts obtained on the 11 location for longitude, latitude and height (expressed in meters), maximum correlation and statistics of shifts

Location	$\Delta\lambda$ (m)	$\Delta\varphi$ (m)	ΔH (m)	Max. correlation
1	35	15	5	0.9924
2	36	18	5	0.9955
3	27	9	7	0.9953
4	40	16	8	0.9973
5	34	8	9	0.9974
6	33	4	7	0.9982
7	38	15	-1	0.9909
8	33	7	5	0.9859
9	36	14	8	0.9971
10	35	8	10	0.9957
11	35	8	12	0.9940
Mean	35	11	6	
Std. dev.	3.4	4.6	3.3	

These results agree with the expected geo-location accuracy of the orientation data derived from sensor navigation equipment, which was expected to be better than 50 m. The shifts are nearly constant (standard deviations smaller than 5 m) because they result from very small errors in attitude angles, which are expected to have a similar pointing effect throughout the image.

They have dimensions of a few pixels and are more significant in longitude direction. Ortho-images generated without any further corrections would not have a positional accuracy required for normal operations, such as the integration with other images. Image orientation improvement is achieved by applying this shift to ground coordinates obtained in the intersection procedure or, for a single image, being added to equation 1.

3.2 Assessment of Image Orientation

In order to assess the positional quality of this procedure, a set of independent check points (ICP), measured on topographic maps of scale 1:25,000, were used. A set of 12 well defined points were identified on the images and on the maps. The intersection

procedure was applied to the image coordinates in order to determine ground coordinates; then the shift previously determined was applied; finally these coordinates were compared to the ones measured on the maps. Table 2 shows the geographic coordinates of the check points and the respective errors. The root mean square errors were also calculated.

Table 2. Coordinates of the ICPs obtained by the intersection plus correction and differences (Δ) to the map coordinates (expressed in meters)

ICP	λ (°)	ϕ (°)	H (m)	$\Delta\lambda$ (m)	$\Delta\phi$ (m)	ΔH (m)
A	-8.414862	41.072676	70.0	10	-5	0
B	-8.277938	40.987405	366.9	10	7	-17
C	-8.564163	41.241107	122.9	10	12	48
D	-8.554773	40.953458	205.6	9	5	9
E	-8.308710	41.201958	202.4	7	19	28
F	-8.077801	41.093137	105.3	11	-21	0
G	-7.990382	41.270764	431.8	19	-2	-7
H	-7.899100	41.129816	101.4	-3	-8	4
I	-7.934493	40.932641	780.7	-15	5	-1
J	-8.289022	41.088286	73.9	5	16	1
K	-8.017600	41.105492	101.3	17	23	4
L	-8.207203	41.237425	280.9	10	10	34
RMS				11.4	13.0	19.8

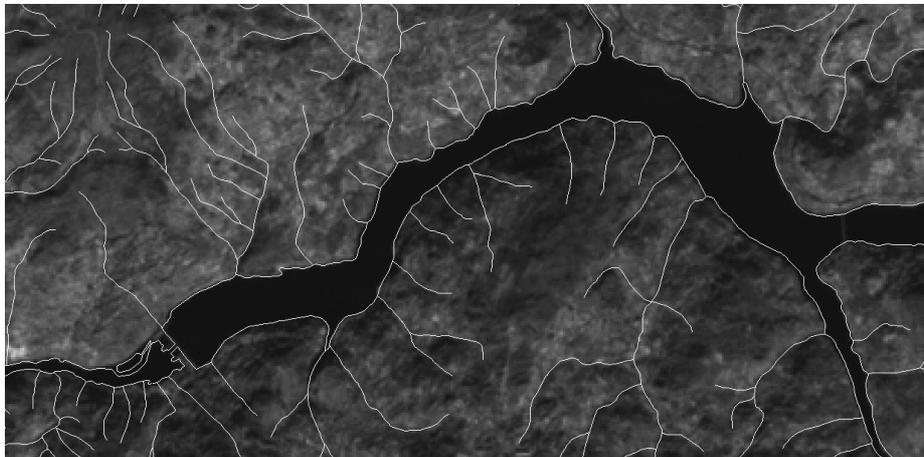


Fig. 8. Ortho-image with superimposed vector data (hydrographic network)

The RMSEs of planar coordinates are smaller than 1 pixel (15 meters) indicating that the image orientation was corrected to a sub-pixel accuracy. It should also be noted that the ICP coordinates also contribute to the errors: they were measured on maps, and are expected to have an accuracy of 10 meters.

The next step is to produce the ortho-rectified images. This is done by projecting 3D coordinates from a DEM (the SRTM DEM in our case) onto the image space, using the inverse of equation 1, and resampling grey scale values.

Figure 8 shows a portion of one ortho-image created (6 by 3 Km²). A vector layer of rivers and streams (digital map data, from topographic maps of scale 1:25,000) was overlaid, showing a very good coincidence .

In the absence of any data for a positional accuracy of the output ortho-images the 3N and 3B images can be overlaid. If relief effects were correctly remove they should now have a very good coincidence.

The test described was done with an image of North Portugal, acquired in 2001. Another image, also from North Portugal but from a different date and different path, was also tested. The method worked equally well, yielding a positional correction of the same order (< 50 m) and a very good coincidence with the map data.

4 Conclusions

In general, precise orientation of satellite images requires the use of GCPs, which can be obtained from maps and other image datasets. This is a strong limitation because a large user intervention is required. Moreover in many cases field surveys may be necessary, which significantly increases the cost of remote sensing projects.

A method was described to overcome these facts for ASTER imagery, based on the extraction of a DEM and matching to a global DEM. Although the initial ASTER image orientation is very good, the method improved the image location to a sub-pixel positional accuracy.

An important fact for this method to work is that relief exists in the area, even if moderate, with slopes in different directions. An empirical conclusion from the DEM matching tests is that a mean slope of 5% is needed.

The method is of particular interest for mountainous areas, where the availability of reliable ground control by standard procedures is very difficult. It avoids field work and can be largely automated.

Acknowledgments

This work was done with the support of Centro de Investigação em Ciências Geo-Espaciais, Faculdade de Ciências da Universidade do Porto, financed by the Portuguese National Science Foundation (Fundação para a Ciência e a Tecnologia - FCT).

References

1. Abrams, M., Ramachandran, B.: ASTER: user handbook – version 2. JPL (2002)
2. GLCF: Internet Web pages of the Global Land Cover Facility (2006), <http://glcf.umiacs.umd.edu/index.shtml>
3. Rodríguez, E., Morris, C., Belz, J.: A Global Assessment of the SRTM Performance. *Photogrammetric Engineering and Remote Sensing* 72(3), 249–268 (2006)

4. Gonçalves, J.: Orientation of SPOT stereopairs by means of matching a relative DEM and the SRTM DEM. In: EuroCOW 2006. International Calibration and Orientation Workshop (2006), <http://www.isprs.org/commission1/euroCOW06/>
5. Yamaguchi, Y., Kahle, A.B., Pniel, M., Tsu, H., Kawakami, T.: Overview of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *IEEE Transactions on Geoscience and Remote Sensing* 36(4), 1062–1071 (1998)
6. Spotimage Spot Satellite Geometry Handbook (document S-NT-73-12-SI) (2002)
7. Marçal, A.R.S: Automatic geometric correction of hyperspectral satellite images from CHRIS/PROBA. *LNCS*, vol. 4522, pp. 553–561. Springer, Heidelberg (2007)
8. MATHWORKS: Using Matlab, Version 6.5. The MathWorks, Inc. Natick. MA (2002)