COMPARISON OF THE ACCURACY OF DEM FROM SPOT HRS TWO-FOLD STEREO DATA AND HRS/HRG THREE-FOLD STEREO DATA IN BARCELONA TEST SITE

Rupert Müller, Peter Reinartz, Manfred Lehner, Manfred Schroeder
German Aerospace Center (DLR), Remote Sensing Technology Institute
PO Box 1116, D-82230 Wessling, Germany
rupert.mueller@dlr.de

WG I/2

KEY WORDS: Stereoscopic data, image matching, two-fold stereo, three-fold stereo, DEM generation, accuracy analysis

ABSTRACT:

ISPRS and CNES announced the HRS (High Resolution Stereo) Scientific Assessment Program during the ISPRS Commission I Symposium in Denver in November 2002. 9 test areas throughout the world have been selected for this program. For the Barcelona test site not only two-fold stereo HRS data but also HRG data of the same region have been acquired. This leads to the possibility to compare the DEM production and accuracy differences between two-fold and three-fold stereo data sets. For the derivation of the DEMs, the DLR own stereo processing software, developed for the MOMS-2P three line stereo camera is used. The matching process using three images is more robust than with two images due to the possibility to eliminate blunders by matching three stereo pairs. Furthermore the forward intersection of the three image rays allows better constraints for the object space intersection. Due to these reasons the DEM filtering in object space that is necessary for DEMs from two-fold stereo is very much reduced in the case of three-fold stereo data. The derived DEMs are compared with the high precision reference DEM by several methods regarding: height accuracy, location accuracy, blunders, error budget depending on surface properties etc.,.

.

1. INTRODUCTION

Within the HRS Scientific Assessment Program, which gives a user community the possibility to derive digital elevation models (DEM) from along track stereo data from space with the new HRS sensor on SPOT 5 (Gleyzes 2003), the performance of different methods should be tested. This paper deals with the comparison of DEM, derived from HRS two-fold stereo data and HRS/HRG three-fold stereo data. Main emphasis is placed on the use of the very accurately determined absolute and relative position and attitude data delivered along with the image data.

HRS produces image stereo pairs with two optics looking forward and backward (±20 degrees) with respect to the nadir direction. The camera has a spatial resolution of 10 meter across track and along track, but a ground sampling distance of about 5 m along track for obtaining higher resolution of the parallaxes for the DEM generation. The swath of the HRS is 120 km (12000 CCD elements) and one acquisition sequence extends 600 km along track. HRG produces two staggered panchromatic nadir looking images (HMA and HMB) with a resolution of 5 m each, which can be combined to a 2.5 m resolution image.

2. TEST AREAS AND GROUND REFERENCE DATA

The test area is located in Catalonia (Spain) and includes the city of Barcelona, covering dense urban areas as well as mountainous terrain. The data of Catalonia have been acquired on October 15th 2002 with a sun elevation of 39º and no clouds.

The data provided by SPOT IMAGE contain the following parts:

- Two sets of 8 bit panchromatic image data (size 12000 x 12000 pixel = 120 km across x 60 km along track) of Catalonian test area from two viewing directions in TIF format
- Two sets of 8 bit HGR panchromatic nadir looking images (HMA and HMB data) with a ground resolution of 5 m x 5 m of part of the test site (size 12000 x 12000 pixel = 60 km across x 60 km along track)
- XML-files containing all additional information regarding time synchronization, position (DORIS), attitude (ULS: star tracking unit to compute absolute orientation in a celestial frame combined with the attitude orbit control system AOCS), interior orientation (tabulated look angles for each pixel)
- ASCII text files containing information on the delivered data.

The ground reference data are the following:

- 32 color orthoimages (1:5000) with pixel size of 0.5 meter and accuracy better than 1 pixel (1σ)
- DEM with pixel spacing 15.0 meter and orthometric height accuracy of 1.1 meter (1σ)

Both reference data sets (orthoimages and DEM) are provided in UTM zone 31 with the geodetic datum ED50 (European Datum 1950).

3. PREPROCESSING OF THE ANCILLARY DATA

The delivered SPOT 5 HRS Level 1A product consists of the image data in standard TIF format and the meta data in DIMAP format. The following information is extracted from the XML ancillary file for further processing:

- the ephemeris data containing position and velocity of the satellite measured by the DORIS system every 30 seconds with respect to the ITRF90 (International Terrestrial Reference Frame 1990) system during the data take and at least four times before and after image data acquisition,

- the corrected attitude data with respect to the local orbital coordinate frame measured by gyros and the star tracker unit ULS with 8Hz, the data are already corrected for different effects (Bouillon et al. 2003)
- the look direction table for the 12000 CCD pixel elements expressed within the sensor coordinate frame and
- the data used for time synchronization like line sampling period and scene center time.

According to the „SPOT Satellite Geometry Handbook“ (SPOT IMAGE 2002) Lagrange interpolation of the ephemeris data and linear interpolation of the attitude data are recommended to calculate data sets (position, velocity and attitude) for every scan line. For DEM production the exterior orientation is transformed to a local topocentric system (LTS) with a fundamental point located at the center of the image scene. For orthoimage production the exterior orientation is expressed in the Earth Centered Earth Fixed (ECEF) WGS84 Cartesian frame. The transformed data serve as input for DLR's processing software.

4. ORTHOIMAGE GENERATION AND ACCURACY ANALYSIS

To get an impression of the absolute and relative accuracy of the position and attitude data, and to get an estimation of the necessity to improve the ancillary data by bundle adjustment or other methods, orthoimages are derived using an already derived DEM.

The inputs for the orthoimage production are the interior orientation (extracted from the meta data file), the six parameters of the exterior orientation with respect to an ECEF coordinate frame for each image line (interpolated from the measured sampling points) and the digital elevation model (DEM). In the case of Catalonia the reference DEM described in chapter 3 is used for the orthoimage generation.

The principle of the orthoimage production is based on the intersection of the actual sensor viewing direction (pointing vector) with the DEM applying the rigorous collinearity equation. The orthoimage processor calculates the object space coordinates of the points within the intermediate local topocentric system and then transforms them to the desired map projection of the output image using geodetic datum transformation parameters (Müller et al. 2002,2003). The DEM is internally transformed to the same LTS as the exterior orientation, where an undulation of –18.2 m with respect to the ED50 geodetic datum is taken into account. Bilinear resampling to a 10 x 10 m grid has been performed for the final orthoimage.

After generation of the three orthoimages without any ground control information, a check of the accuracy using 24 of the ground control points has been performed. For the quality assessment the measurements have been carried out in bilinear enlarged orthoimages to achieve sub-pixel accuracy. Table 1 shows the deviation in x (east) and y (north) direction for the orthoimages in comparison to the control points.

Table 1: Mean values and standard deviations for the difference to the orthoimages of 24 ground control points in meter in UTM ED50 coordinate system (Catalonia)
x1, y1 – Coordinates in reference orthoimages
x2. y2 – Coordinates in orthoimage from forward looking
x3. y3 – Coordinates in orthoimage from backward looking
x4. y4 – Coordinates in nadir looking image (HMA)

	x2 – x1	y2 – y1	x3 – x1	y3 – y1	x4 – x1	y4 – y1
MEAN	-9,90	-16,59	-0,36	-11,16	-24,22	-6,22
STDV.	4,64	8,48	5,72	5,23	5,96	5,71

The result shows that even without any ground control, the absolute georeferencing accuracy of the HRS sensor is in the order of one to two pixel, less than 20 meter and standard deviation less than one 1pixel. This is expected, since the values for the absolute pointing accuracy is given by the French colleagues to about 33 meters with 90% accuracy (Bouillon et al. 2003, Airault et al. 2003). Only the x-coordinate in the nadir looking image (x4) shows a slightly higher mean difference than the stereo channels

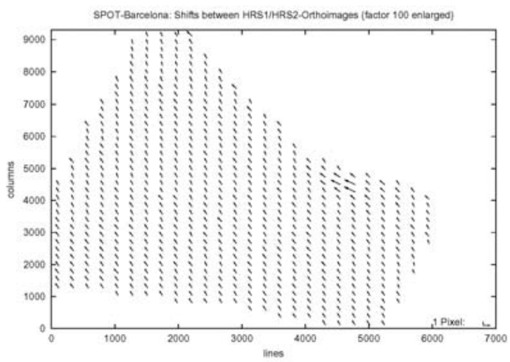

Figure 1: Shifts between the two orthoimages derived from forward and backward looking channels of SPOT HRS (mean values in a regular grid), Catalonia test area

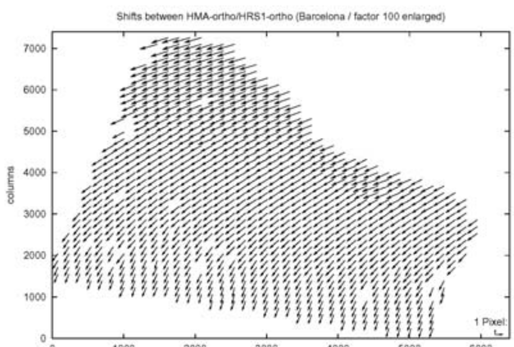

Figure 2: Shifts between the two orthoimages derived from forward and nadir looking channels of SPOT HRS and HMA (mean values in a regular grid), Catalonia test area

An interesting behavior shows up when looking at the comparison of the three orthoimages in detail. The automatic matching of the orthoimages reveals that the difference vectors (each a mean value in a 200 x 200 pixel squared area) show a very homogeneous behavior, mean length is about 15 meter (Fig. 1) – the shift shows up predominantly in flight direction. This means that by using one very good and exact ground control point, the absolute accuracy of the orthoimage can be improved and the images can be used as matched correctly (see e.g. Nonin et al. 2003). Only the nadir looking channel (HMA) shows a different behavior, since here the matching differences (arrows in Fig. 2) show variable shift, which depends on the position in the CCD array.

Remark: This systematic behavior, which shows mainly a constant shift between the two images, is a result of good relative orientation for the single images, but an absolute pointing change between the forward and backward data acquisitions (~90 sec. time difference). By using the values of table 1 the corrections of the angular changes are 0.0013º for HRS1 and 0.0008º for HRS2 for the Catalonia test site. These values are in line with the accuracy specifications of the data provider. The reason for these residual rotations are probably due to the uncertainty of the initial attitude values, but should be discussed with the data producers for further investigations. The measured residual orientation values have been applied for corrections to the attitude values, which leads to nearly accurate matched orthoimages. Similar results are found for test site Bavaria (Reinartz 2004).

5. DEM PRODUCTION FROM TWO RAY STEREO DATA

The first matching of the two images is performed purely in image space with DLR software. Details on this software are described in Lehner et al. 1992. It relies on a 7-step image resolution pyramid and applies intensity matching in two forms: normalized correlation coefficient for pixel accuracy and subsequent local least squares matching (LLSQM) for refinement to sub pixel accuracy (for mass points 0.1 to 0.3 pixel standard deviation, depending on the radiometric quality of the imagery). First interest points are generated with a Förstner operator and the homologous points are searched for in the other image. Only points with high correlation and quality figure are selected as tie points for bundle adjustment (see chapter 7) and a less stringent criterion is valid for the usage as seed points for the subsequent Otto-Chau region growing procedure for dense matching (Heipke et al 1996). This local least squares matching starts with template matrixes of 13 x 13 pixels around the seed points with a constant step in each direction (here three pixel). For cross checking a backward match is performed for all points found. From the differences of the image coordinates a standard deviation of about 0.14 pixel is found. Points showing differences larger than 0.5 pixel in the backward matching are eliminated.

Having the mass points from the matching process as well as the exterior and interior orientation of the camera system, the object space coordinates can be calculated using forward intersection. This is done by least squares adjustment for the intersection of the image rays. Intersections with weak geometry (threshold determined using intersection constraints of high quality homologous points) are rejected.

The irregular distribution of points in object space after the forward intersection has to be regularized into a equidistant grid of about 15 x 15 meter pixel size. The interpolation process is performed by a moving plane algorithm (Linder 1999). The resulting DEM, which are surface models, are compared to the reference DEM, which are terrain models. Therefore a distinct difference is expected e.g. in forest areas.

The area covered show besides the city of Barcelona, the very steep mountainous area of Montserrat as well as the moderate mountains of Tibidabo and others. The comparison of the DEM is therefore performed in different areas: cities, open areas and forest areas, which are masked using classification results of the orthoimages. Fig. 3 shows the derived SPOT-DEM calculated by using two ray intersection.

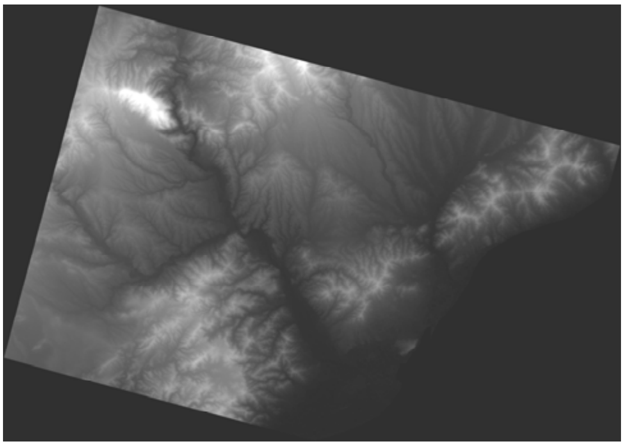

Figure 3: SPOT-DEM of Barcelona and surroundings

First the “best” homologous points for two-fold imagery as projected to object space, are investigated. The result is very close to the result achieved in Bavaria (Reinartz 2004) and shows again a very good absolute accuracy without using any ground control information (table 2).

Table 2: Comparison of height for high quality homologous points in SPOT-DEM derived from two ray intersection and the reference DEM of 67 x 67 km² with 1.1 m accuracy

Mean Height Difference [m]	Std. Dev. [m]	# Points
8.8	3.4	101858

6. DEM PRODUCTION FROM THREE RAY STEREO DATA

For Catalonia test site the images of four cameras are available, the off-nadir looking HRS1/2 and the nadir looking HMA/B (two 5 meter resolution bands). This offers the possibility to derive DEM from the stereo channels HRS1 and HRS2 (called two ray intersection) and additionally to take into account the nadir looking bands (called three ray intersection). For the investigation only the band HMA was included for DEM generation (no interpolation to 2.5m resolution of the HMA / HMB Supermode image was performed). The overlap region can be seen in figure 4.

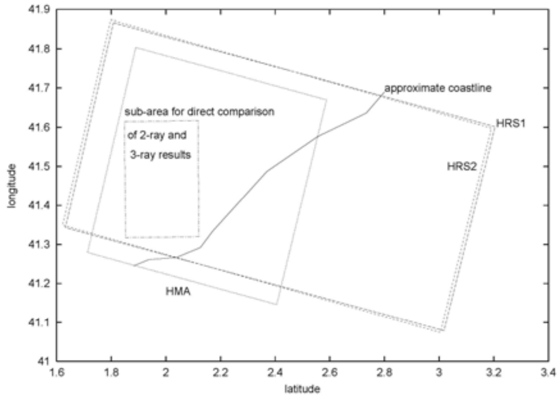

Figure 4: Location of HRS/HMA scenes for the Catalonia test site

Matching was done in several steps in order to generate 3-ray tie points with sub-pixel accuracy. The first step is the generation of seed points with DLR matching software as mentioned in chapter 5. A resolution pyramid of 7 levels (factor 2 reduction in resolution from level to level) was used for a triple of HRS1, HRS2 and a HMA version which was resampled to HRS1/2 resolution (reduction by a factor of 2 in across-track direction). This first matching – via maximum of normalized correlation coefficient and subsequent refinement to sub-pixel accuracy with local least squares matching (LLSQM) – generated about 190000 3-ray tie points between HRS1/2 and reformatted HMA imagery. On the highest resolution level (original image resolution of HRS1/2) a LLSQM window size of 29 rows (along-track, 5 m ground pixel size) and 17 columns (across-track, 10 m ground pixel size) has been used to compensate for the different resolutions in row and column directions (5 m and 10 m, respectively). Next, to get 3-ray tie points in original resolution (5m x 5m), the HMA column coordinates of the tie points from the first step are changed to original resolution and then put into a new LLSQM matching with images of differing resolutions (which is easily possible via LLSQM because of the inbuilt estimation of an affine transformation – together with some interpolation scheme: bilinear interpolation of grey values is used by the DLR matching software to get HMA-resolution chips out of HRS imagery). After careful sub-selection about 20000 tie points were transferred to the next step of LLSQM with Otto-Chau region growing.

To exploit the checking possibility available for three stereo pairs three matching steps with the software for region growing (see chapter 4) have been performed using a grid spacing of 3x3 for the growing:

1. Matching between HMA and HRS1 (original resolutions) using seed points from DLR matching software (result: 9720000 points)
2. Matching between HMA and HRS2 for all HMA points resulting from step (1.) and combination to 3-ray points (result: 9404000 points)
3. Check by matching between HRS1/HRS2 for all pairs found in steps (1-2), (4% could not be matched)
4. Sub-selection by a threshold of 0.5 pixel between HRS2 coordinates from steps (2.) and (3.) (5% did not match the threshold condition); for the 91% of accepted points the shifts of the HRS2 coordinates have a mean of 0 pixel and standard deviations of 0.17 (along-track, 5m resolution) and 0.09 pixel (across-track, 10m resolution) respectively

The whole matching process ended with a number of 8555000 3-ray points.

7. COMPARISON OF DEM FROM TWO AND THREE RAY STEREO DATA

Table 3 shows the comparison of the mass points from the Otto-Chau region growing dense matching algorithm with the reference DEM after forward intersection and rejection of points with weak intersection geometry (the threshold for rejection is determined by the intersection geometry of high quality homologous points). Image matching with three images is supposed to provide better control mechanism for the homologous points by the improved check via a third stereo image discussed above. The three ray matching process was based on HMA 5 x 5 m resolution imagery, which results in about twice as much points for region growing. Slightly better results of 0.6m less in standard deviation are obtained from three ray forward intersection. Also minimum and maximum values are reduced significantly. In all two cases more than 99.96 % of the matching points are fitting the reference DEM better than 50m after adjusting the mean height.

Table 3: Comparison of mass points derived from two (HRS1, HRS2) and three ray (HRS1, HRS2, HMA) intersection with the reference DEM.

	Intersection of two rays	Intersection of three rays
Amount of Points	1418965	2662143
Points < 20m	1407636 (99.20 %)	2652764 (99.65 %)
Points < 50m	1418414 (99.96 %)	2662076 (~100 %)
Standard deviation [m]	5.76	5.16
Mean height difference [m]	9.54	9.89
Min. height difference [m]	-151.5	-61.7
Max. height difference [m]	148.2	105.3

After regularization of the SPOT DEM into a equidistant grid by a moving plane algorithm (Linder 1999) (15m x 15m pixel size), a comparison with respect to different classes (forest, settlement and open areas) is performed. Table 4 shows that the mean height differences and the standard deviation are best for open areas, because the reference DEM is a digital terrain model compared to the surface model of the SPOT DEM. Slightly better results of about 0.5 m reduced standard deviation are obtained for the case of three ray forward intersection.

Table 4: Comparison of a regularized SPOT-DEM derived from two (HRS1, HRS2) and three (HRS1, HRS2, HMA) ray intersection with the reference DEM for three different classes (forest, open areas and settlements)

Intersection of two rays

	Forest	Open Areas	Cities
Mean height difference [m]	10.82	9.84	10.69
Standard deviation [m]	7.24	4.73	5.02

Min. height difference [m]	-79.8	-81.0	-75.8
Max. height difference [m]	79.0	90.9	59.6

Intersection of three rays

	Forest	Open Areas	Cities
Mean height difference [m]	11.57	9.82	10.74
Standard deviation [m]	7.16	4.22	4.47
Min. height difference [m]	-122.0	-75.0	-106.1
Max. height difference [m]	66.4	52.8	61.2

As shown in the accuracy analysis of the orthoimages the channel HMA differs about 24 m in one direction with respect to ground control points and keeping in mind the more complex shift patterns in figure 2. This can be the reason for the only slightly better results in the case of using three images for the DEM production.

8. CONCLUSION

It could be shown that a stereoscopic evaluation of SPOT-HRS data, only using ancillary data delivered by the image provider, leads to an absolute accuracy of terrain heights in the order of about 9 meter (mean height error), with standard deviations of about 3 meter for high quality single points (two ray stereo data) and 5 to 6 meter standard deviation for mass points of the two and three ray stereo data in comparison to the reference DEM. The relative and absolute accuracy for overall comparison of the interpolated DEM (surface model) with the reference DEM (terrain model) of course depends on land use classes and terrain steepness, since image matching algorithm depend on these features. For open areas a mean height error of about 10 meter and a standard deviation of about 6 meter is achieved, whereas a slightly better result (0.5 meter reduced standard deviation) is obtained for three ray stereo data. The expected improvement of three ray stereo data evaluation, which offers better control mechanism for the image matching procedure, probably is compensated by the lower performance of the meta data (interior orientation) of the HMA nadir looking channel.
The absolute accuracy can be improved by using ground control points to reach a mean height difference of about 1 meter (Reinartz 2004). The standard deviations can probably be improved by DEM filtering, although not very large improvements are expected.
Orthoimages can be derived to an absolute location accuracy of 1 to 2 pixels (10 to 20 meter) without ground control, which is in line with the performance specification. The shifts between the orthoimages of the HRS1/2 stereo channels are highly constant throughout the images, offering corrections with few ground control points.

9. REFERENCES

S. Airault, P. Gigord, L. Cunin, E. Breton, A. Bouillon, R. Gachet, M. Bernard, "Reference3d Location Performance Review and Prospects", Proceedings of the ISPRS Workshop High Resolution Mapping from Space, Hanover, Oct. 6.-8. 2003, IPI, Universität Hanover, on CD-ROM

A. Bouillon, E. Breton, F. De Lussy, R, Gachet, "SPOT5 Geometric Image Quality", IGARSS 2003, Toulouse, 21.-25. July 2003, IEEE International Geoscience and Remote Sensing Symposium

J.P. Gleyzes, A. Meygret, C. Fratter, C. Panem, S. Baillarin, C. Valorge : "SPOT5 : System overview and image groud segment", IGARSS 2003, Toulouse, 21.-25. July 2003, IEEE International Geoscience and Remote Sensing Symposium

C. Heipke, W. Kornus, A. Pfannenstein, "The evaluation of MEOSS airborne 3-line scanner imagery – processing chain and results", Photogrammetric Engineering and Remote Sensing, Vol. 62, No. 3, pp. 293-299, 1996

P. Nonin, S. Piccard, "Performance analysis of DEM automatic extraction from SPOT5 sensors", IGARSS 2003, Toulouse, 21.-25. July 2003, IEEE International Geoscience and Remote Sensing Symposium

M. Lehner, R.S. Gill, "Semi-automatic derivation of digital elevation models from stereoscopic 3-line scanner data", IAPRS, Vol. 29, Part B4 Washington, USA pp. 68-75, 1992

W. Linder, "Geo-Informationssyteme", Springer-Verlag Berlin, Heidelberg, Germany, ISBN 3-540-65276-0, 1999

Ru. Müller, M. Lehner, Ra. Müller, P. Reinartz, M. Schroeder, B. Vollmer, "A Program for Direct Georeferencing of Airborne and Spaceborne Line Scanner Images", ISPRS, Comission I Mid-Term Symposium "Integrated Remote Sensing at the Global, Regional and Local Scale", Denver, CO USA, 10-15 November 2002

Ru. Müller, G. Palubinskas, P. Reinartz, M. Schroeder, V. Amann, R. Stätter, "From Airborne Digital Raw Data to Image Maps" PFG Photogrammetrie, Fernerkundung, Geoinformation, Germany, 4, 4, (2003), S. 317-326

P. Reinartz, M. Lehner, Ru. Müller, M. Schroeder, "Accuracy for DEM and Orthoimages derived drom SPOT HRS Stereo Data without using GCP", Proceedings of the ISPRS XXth Congress, Istanbul 12-23 July 2004

SPOT IMAGE, "SPOT Satellite Geometry Handbook", S-NT-73-12-SI, Edition 1, Rev. 0, Toulouse, France 2002