

# ACCURACY ANALYSIS AND SURFACE MAPPING USING SPOT 5 STEREO DATA

Hannes Raggam  
Joanneum Research, Institute of Digital Image Processing  
Wastiangasse 6, A-8010 Graz, Austria  
[hannes.raggam@joanneum.at](mailto:hannes.raggam@joanneum.at)

## Commission I, WG I/2

**Key Words:** Spot 5, Sensor Model, Accuracy Analysis, Image Matching, Elevation Model

## ABSTRACT

On the Spot 5 satellite stereo data can be acquired simultaneously from the HRS (high resolution stereoscopic) instrument, which comprises 2 cameras looking forward and backward, respectively, at an off-nadir angle of  $\pm 20$  degrees. Enhanced along-track pixel resolution of 5 meters shall further assure a high accuracy with respect to 3D data extraction. In this concern, an HRS study team was installed and specific test sites designated in order to validate the geometric performance of Spot 5 HRS stereo data. This paper summarizes the results which have been achieved from the validation activities of the Institute of Digital Image Processing. The investigations were carried out using a test data set acquired over the city of Barcelona. Study areas showing different topographic characteristics have been investigated. Beside pure HRS stereo data, a glimpse has been further made onto the joint use of the HRS stereo data and a supermode THR image, which is basically a nadir scene with a pixel size of 2.5 meters.

## 1. INTRODUCTION

In order to continue and strengthen the traditional Spot stereoscopic potential, the HRS instrument was realized on Spot 5, allowing the acquisition of in-track stereo data during the same overflight. The HRS instrument comprises two CCD cameras looking forward and backward, respectively, at an off-nadir angle of  $\pm 20$  degrees, allowing stereo data capture at a time interval of about 90 seconds (Gleyzes et al., 2003). The resolution of the stereo data is 5 meters along and 10 meters across-track. Thus, the performance of image matching applied to such data is not degraded due to temporal changes in the scenes.

An HRS assessment program has been established, involving an expert team as well as a number of designated test sites in order to verify mapping accuracies being predicted for Spot 5 HRS stereo data. Following the objectives of this program, this paper is devoted to the analysis of the topographic mapping potential of Spot 5 HRS stereo data. The following issues are specifically addressed:

- Geometric modelling of HRS stereo data, including utilization of given orbital and imaging parameters
- Analysis of a-priori 2-dimensional location accuracy for individual scenes as well as 3-dimensional location accuracy of stereo models using control points
- Optimisation of sensor models utilizing control point measurements and least squares parameter adjustment
- Discussion of the performance of image matching techniques being applied to Spot HRS as well as multi-sensor THR/HRS stereo data
- Extraction of surface models from HRS as well as multi-sensor THR/HRS stereo data
- Accuracy analysis of extracted elevation models through comparison with reference data

For the study the Remote sensing Software package Graz (RSG, Joanneum Research, 2003) was used. This is designed for geometric processing of remote sensing images, including the

aspects of geometric sensor modelling, block adjustment, stereo image matching, and surface model generation. At the Institute, Spot stereo mapping using analogue Spot stereo pairs and an analytical plotter was already an issue in the eighties (Raggam et al., 1989). Also the aspects of multi-sensor stereo data utilization as well as validation of stereo mapping potential (Raggam et al., 1990 and 1996) have been a matter of previous research.

## 2. TEST DATA

The city of Barcelona as well as the areas north(-west) of it were selected as a primary test site for the HRS assessment program. The related test data set was provided for this study and comprises the following image and reference data:

- a Spot HRS stereo pair with a nominal pixel size of 5 meters along- and 10 meters across-track;
- a THR supermode product with a nominal pixel size of 2.5 meters;
- a digital elevation model with a mesh width of 15 meters;
- ortho photo mosaics for 8 detailed study areas with a pixel size of 0.5 meters.

Some 15 high precision control points were measured for each of the detailed study areas from the respective ortho photo mosaics, resulting in an overall pool of 116 points to be used for sensor modelling and accuracy analyses.

## 3. SENSOR MODELLING

In general, rather detailed and precise information is provided for the sensor model of a Spot 5 image, comprising orbital data as well as attitude parameters of the instrument (Spot Image, 2002). This justifies the feasibility to generate value-added products with high accuracy. Geolocation accuracies in the range of 15 meters in planimetry and 10 meters in height are anticipated with respect to HRS input images without using control points for optimisation (Spot Magazine, 2003).

### 3.1 Single Images

The HRS stereo data as well as the THR supermode scene were included into the geolocation accuracy analysis. For these data only the orbital information was utilized but not the detailed attitude parameters. Instead, constant values were extracted from the header data and initially used for the sensor attitude angles. This results in an a-priori geolocation accuracy of tens to even hundreds of input pixels. Hence, also ground control points had to be used in order to optimise the sensor models of the Spot 5 image data.

The sensor models of these images were optimised using a least squares parameter refinement procedure – an equivalent to photogrammetric bundle adjustment – as implemented in the RSG software. The RMS, minimum and maximum point residuals, resulting after the sensor model optimisation are summarized in Table 1. Sub-pixel location accuracy was achieved for each of the images, being represented by RMS values in the order of 0.7 pixels for the HRS images and of 0.9 pixels RMS for the THR scene.

116 control points		Along	Across	Length
HRS1	RMS	0.72	0.72	1.02
	MIN	-1.58	-1.43	0.01
	MAX	1.59	1.56	2.13
HRS2	RMS	0.79	0.71	1.06
	MIN	-1.83	-1.36	0.05
	MAX	1.96	1.66	2.28
THR	RMS	0.91	0.91	1.29
	MIN	-1.78	-1.96	0.09
	MAX	1.94	1.82	2.51

Table 1: Statistical results of parameter optimisation.

### 3.2 Stereo Models

To evaluate the stereo mapping performance of Spot 5, the standard HRS stereo pair as well as multi-sensor stereo pairs were considered, which comprise the THR supermode scene as well as one of the HRS scenes. For these 3 stereo models the a-priori stereo mapping accuracy was determined. Therefore, ground coordinates are calculated for stereo control points measured in both images of the stereo model using a least squares point intersection algorithm, and 3D point residuals are determined through comparison with given control point coordinates. The RMS, minimum and maximum point residuals being achieved in East, North and Height are summarized in Table 2.

116 control points		East	North	Height	Length
HRS1-HRS2	RMS	6.6	3.9	4.0	8.6
	MIN	-15.2	-11.7	-9.7	1.4
	MAX	12.2	7.7	12.0	16.2
THR-HRS1	RMS	2.4	2.3	8.4	9.0
	MIN	-4.5	-4.6	-18.0	1.6
	MAX	5.7	5.5	16.0	18.4
THR-HRS2	RMS	2.2	2.3	7.9	8.5
	MIN	-4.5	-4.8	-17.2	1.0
	MAX	5.0	5.4	19.3	19.9

Table 2: A-priori stereo mapping accuracy.

For the HRS stereo model with a base-to-height ratio of 0.72 a height accuracy of 4 meters was achieved, while the planimetric

accuracy is about 8 meters and hence worse by a factor of 2. For the multi-sensor models with a base-to-height ratio of 0.36 the achieved height accuracy of some 8 meters is worse by a factor of about 2 in comparison to the HRS model. On the other hand, a planimetric accuracy of less than 2.5 meters in East and North is achieved for both models, which is significantly superior (by more than a factor of 2) in comparison to the HRS model. Hence, planimetric accuracy is improved at the cost of height accuracy for the THR-HRS image pairs. The overall residual length shows an RMS value of about 9 meters is roughly the same in either case.

### 3.3 Image Block

The benefit to merge the HRS scenes with the THR scene to an image triple for 3D data extraction was further investigated. Considering an image block formed by these 3 images, point intersection of homologue control points was performed to evaluate the 3D location accuracy. The overall statistics of resulting point residuals are summarized in Table 3. An RMS height accuracy of 3.4 meters is achieved, i.e. slightly superior to what is achieved from the pure HRS stereo data, while the planimetric accuracy is 2 to 3 meters in East and North, i.e. significantly better than for the HRS image pair and close to the one achieved from THR/HRS multi-sensor image pairs.

	East	North	Height	Length
RMS	3.0	2.3	3.4	5.1
MIN	-7.2	-4.7	-9.0	1.1
MAX	8.8	4.9	8.3	12.5

Table 3: A-priori image block mapping accuracy.

## 4. DSM GENERATION

Three of the detailed study areas were selected to apply the DSM generation procedure and to investigate the performance of algorithms as well as the quality of achieved results. Selection was made upon land cover and morphology as follows:

- Rural/hilly area, being partly covered by forests
- Mountainous terrain
- Urban area, represented by the city of Barcelona

Anaglyph presentations of the HRS stereo images of these test areas are shown in Figure 1.

### 4.1 DSM from HRS stereo pair

First, surface models were extracted for selected test areas using the HRS stereo pair. The procedure comprises matching of the stereo images, calculation of ground coordinates from the matching result, and interpolation of a regular surface elevation raster.

For stereo matching, the widely used cross correlation approach was applied. The performance of this image matching approach with respect to these stereo data is summarized in Table 4, which shows the percentage of pixels where no matching was possible. The matching failures in general are caused by homogeneous areas where discrimination of individual pixels is difficult if not impossible (similarity too high), and by major geometric differences (similarity too low).

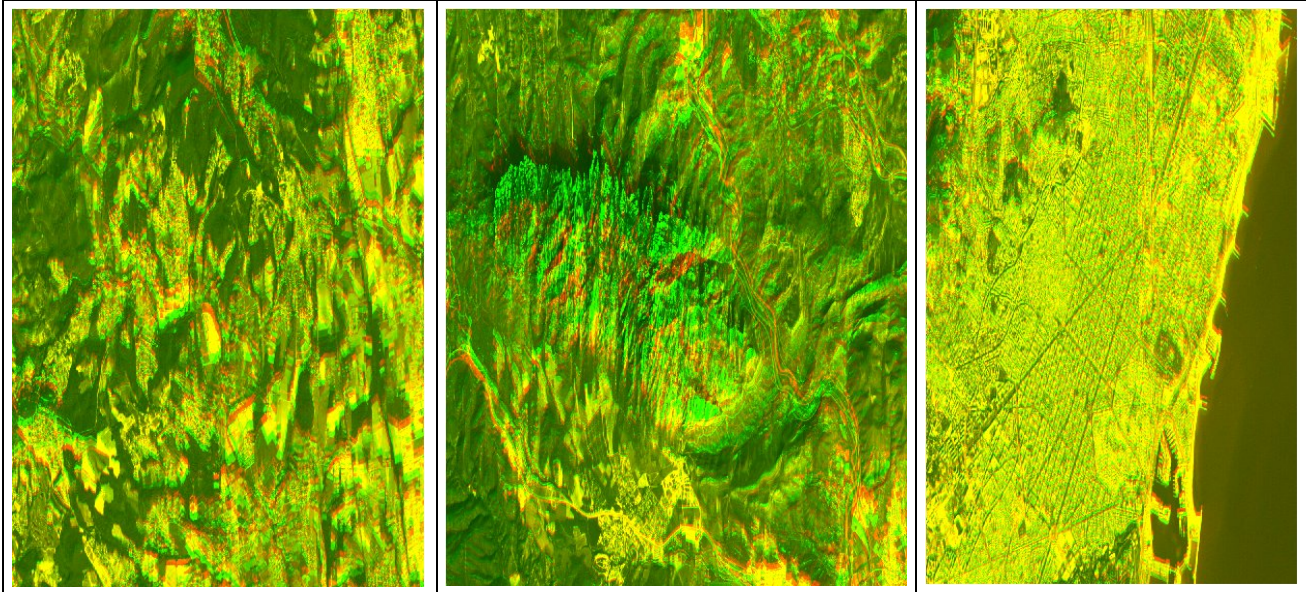


Figure 1: Detailed test areas showing hilly (left), mountainous (mid) and urban (right) terrain in anaglyph presentation (red: forward image, green: backward image).

While for the rural and the urban area only 1,45% and 2,83% of the entire pixels were not matched, a failure rate of 10,82% was achieved for the mountainous test area. This is due to the large parallaxes (dissimilarities) between the 2 stereo images, as they are present in case of mountainous terrain and large base-to-height ratio of the data.

Rural area	1,45 %
Mountainous area	10,82 %
Urban area	2,83 %

Table 4: Matching performance for HRS stereo data.

The quality of the surface models resulting from these stereo matching results is checked through comparison with the reference elevation model, i.e. through calculation of height differences. However, the reference elevation model does not represent the a surface model, but rather a ground model, excluding objects like trees or buildings.

The digital surface models resulting for the selected test areas as well as the difference elevation models, which were determined with respect to the given reference DEM, are shown in Figure 2. Moreover, the overall statistics of elevation differences are summarized in Table 5. The following comments can be made:

**Hilly test area:** The elevation errors clearly correlate with ground cover, i.e. positive elevation differences in the order of more than 5 meters are frequently achieved for forested parts of the test area. This also results in a bias in the height differences of 1.6 meters, while the standard deviation is 5.5 meters.

**Mountainous test area:** For this area large height errors are partly achieved over the central mountainous area, which covers an elevation range between 76 and 1206 meters. Local areas with rather extreme elevation errors of up to 300 meters distinctly degrade the standard deviation to 35.8 meters only for this test area.

**Urban area:** For the built-up areas height differences of more than 5 meters are widely achieved, reflecting the heights of the buildings in this area. This leads to a bias of 9.3 meters, while the standard deviation is 8 meters. However, individual buildings can not really be discriminated

Model	Area	Mean	Std.Dev.	Min.	Max.
HRS1-HRS2	2	1.6	5.5	-78.2	48.0
	4	0.3	35.8	-300.4	224.5
	6	9.3	8.0	-23.1	46.3

Table 5: Summary of elevation difference statistics for investigated test cases.

A visual quality check can be made through stereo ortho photos, which are generated from the input stereo images and using these surface models. Location differences of these ortho photos indicate elevation errors in the surface model. A superposition of the HRS ortho photos in red (forward image) and green (backward image) is shown in Figure 3 for the rural and the mountainous test area, respectively. A satisfactory correspondence is achieved for the rural area, while extreme differences result for the mountainous area.

#### 4.2 DSM from THR-HRS stereo pair

The multi-sensor stereo model comprising the THR image in conjunction with the HRS1 stereo image was used in order to investigate the benefit of the THR supermode product with respect to DSM generation. For image matching, the HRS1 stereo image was over-sampled and coarsely registered to the geometry of the THR image.

This approach preserves the high resolution of the THR image. However, a distinctly different level of detail is inherent to the resulting stereo pair. This is shown in Figure 4, where sub-windows of the stereo images are presented for a built-up area and a rural area. It is obvious, that many details being visible in the high-resolution THR image disappear in the over-sampled

HRS1 image. Successful matching cannot be expected for such features.

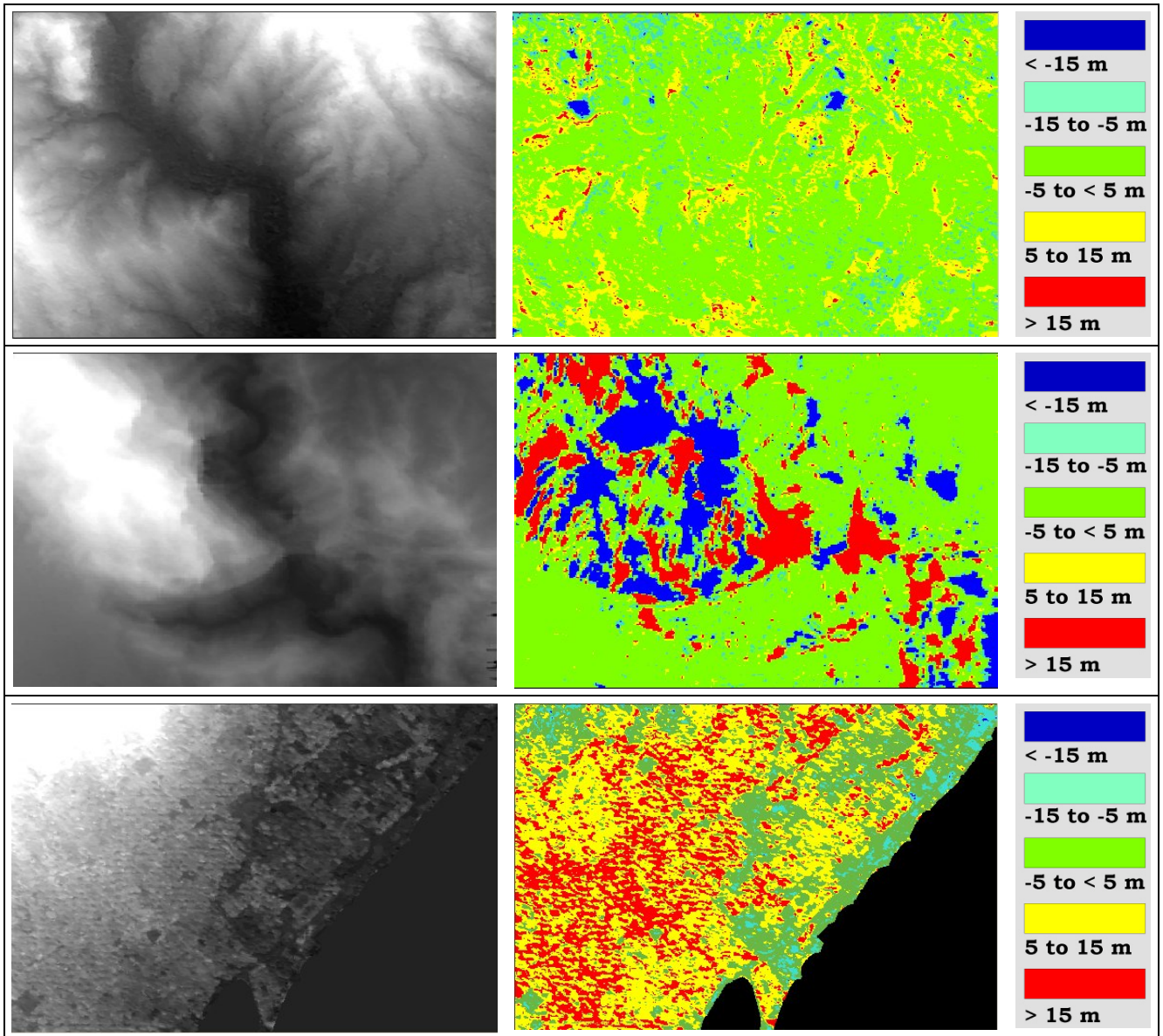


Figure 2: Digital surface models (left) and difference DEMs (right) achieved for rural (top), mountainous (mid) and urban (bottom) test area from HRS stereo data.

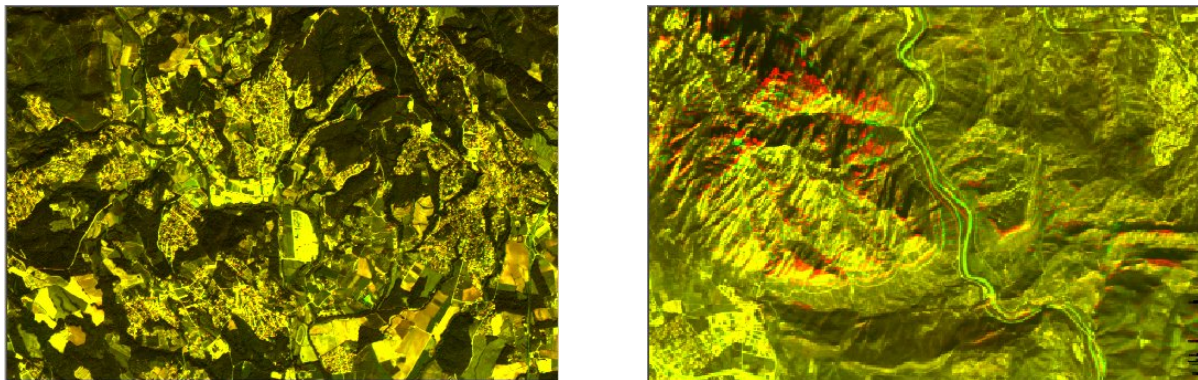


Figure 3: Stereo ortho photos for rural (left) and mountainous (right) area.

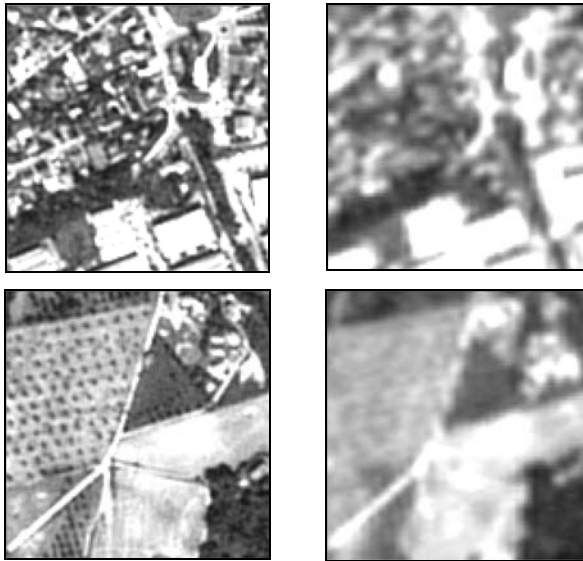


Figure 4: Detail of multi-sensor stereo data (left: THR, right: HRS) for built-up area (top) and rural area (bottom).

This is confirmed in Table 6, which shows the performance of image matching for the selected test areas. Now, more than 10% of the pixels were not matched for all of the three test areas. In a relative sense, however, the matching performance of the mountainous area was not as drastically degraded as for the rural or the urban test site. This is due to the fact that the stereo images now are more similar even in the mountainous areas and image matching is facilitated, although on the other hand the stereo condition was significantly degraded by a factor of 2.

Rural area	10,07 %
Mountainous area	12,25 %
Urban area	13,78 %

Table 6: Matching performance for THR-HRS stereo pair.

The digital surface models which have been generated from these stereo data are shown in Figure 5 together with the difference DEMs, which were determined with respect to the given reference DEM. Statistical parameters like mean, standard deviation, minimum and maximum of these elevation differences are summarized in Table 7. The following conclusions can be made:

**Rural area:** The extension of unreliable areas, specifically represented by larger negative height differences, was significantly reduced, leading to an increase of the bias to 4 meters, which may realistically be caused by the forest areas (yellow and red areas in difference model).

**Mountainous area:** The maximum elevation errors are drastically reduced, although large elevation errors of some 150 meters are still locally present. The standard deviation is reduced to about 9 meters, while the bias was increased to 5 meters. Although not really clear, this could again be due to vegetation and forests, the surface of which should have been tentatively reconstructed.

**Urban area:** The surface model clearly shows the road network of the city of Barcelona. Elevation differences in the built-up areas are typically in yellow, i.e. in the order of 5 to 15 meters and correspond well to the potential height of

buildings. Hence, also a bias of 11.4 meters is achieved for this test area.

For each of these test areas the standard deviation corresponds well to the 8.4 meters RMS error which has been achieved in the a-priori analysis based on control points (see Table 2). For visual quality control again stereo ortho photos were generated. These are shown in Figure 6 in an anaglyph presentation for the mountainous and the urban test area. Again, a significant improvement can be immediately notified for the mountainous area, although major elevation errors still cause geocoding errors, which are well visible in the stereo ortho photos overlay.

Model	Area	Mean	Std.Dev.	Min.	Max.
THR-HRS1	2	4.0	6.9	-31.0	44.4
	4	5.0	8.8	-172.0	156.0
	6	11.4	9.9	-29.0	61.4

Table 7: Summary of elevation difference statistics for investigated test cases.

## 5. SUMMARY AND OUTLOOK

A Spot 5 image data set acquired over the city of Barcelona was used to investigate the accuracy of 3D data being extracted stereoscopically. Stereo modelling using high quality control points has shown a height accuracy of some 4 meters for the HRS stereo pair, while the planimetric accuracy was worse by a factor of 2. When using a multi-sensor THR/HRS image pair, the planimetric accuracy can be improved to less than 3 meters, but the height accuracy is degraded by a factor of 2. Surface models were extracted from HRS image pairs as well as from a THR/HRS image pair for different type of terrain. However, a comprehensive and thorough quality analysis is hardly possible for vegetated and built-up areas, because only a ground model but no surface reference data are available. Future work will focus on the utilization of an image triple comprised by the HRS stereo images as well as the THR scene. This promises a significant upgrade of achievable accuracies in the order of a few meters in planimetry as well as height for any type of terrain.

## REFERENCES

- J.P. Gleyzes, A. Meygret, C. Fratter, C. Panem, S. Baillarin, C. Valorge, 2003. SPOT 5: System Overview and Image Ground Segment. *IEEE International Geoscience and Remote Sensing Symposium*, published on CD-ROM.
- JOANNEUM RESEARCH, 2003. RSG in Erdas Imagine. *Software Documentation, Version 2.1*, RSG-Release 4.60.
- H. Raggam, M. F. Buchroithner and R. Mansberger, 1989. Relief Mapping Using Non-Photographic Spaceborne Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 44, No. 1, pp. 21-36.
- H. Raggam and A. Almer, 1990. Mathematical Aspects of Multi-Sensor Stereo Mapping. *IEEE International Geoscience and Remote Sensing Symposium*, Vol. III, pp. 1963-1966.
- H. Raggam and A. Almer, 1996. Assessment of the Potential of JERS-1 for Relief Mapping Using Optical and SAR Data. *International Archives of Photogrammetry and Remote Sensing*, Vol. 31, Part B4, Commission IV, pp. 671-676.
- Spot Image, 2002. SPOT Satellite Geometry Handbook. Ref. S-NT-73-12-SI, Edition 1, Revision 0.

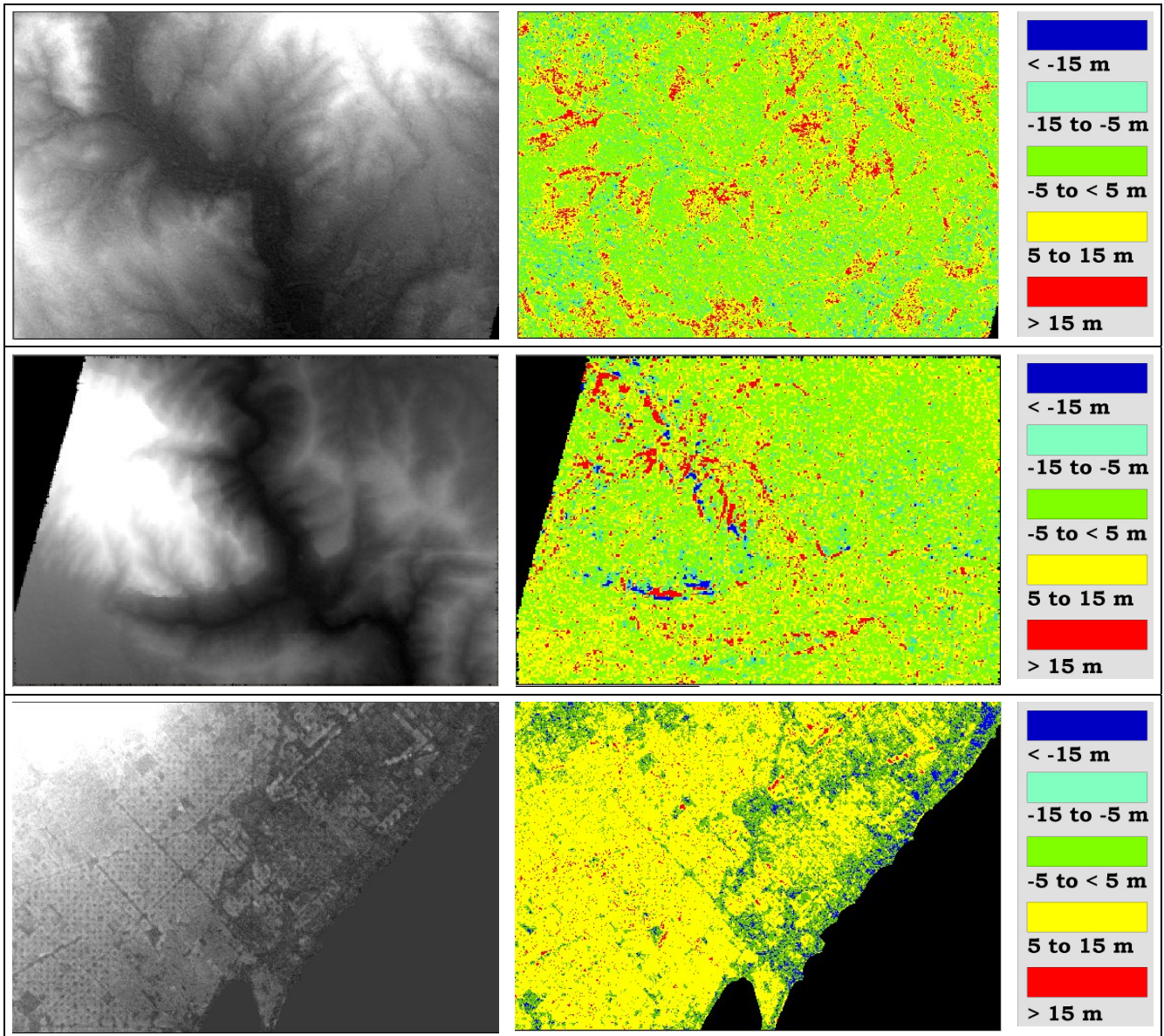


Figure 5: Digital surface models (left) and difference DEMs (right) achieved for rural (top), mountainous (mid) and urban (bottom) test area from multi-sensor THR-HRS stereo data.

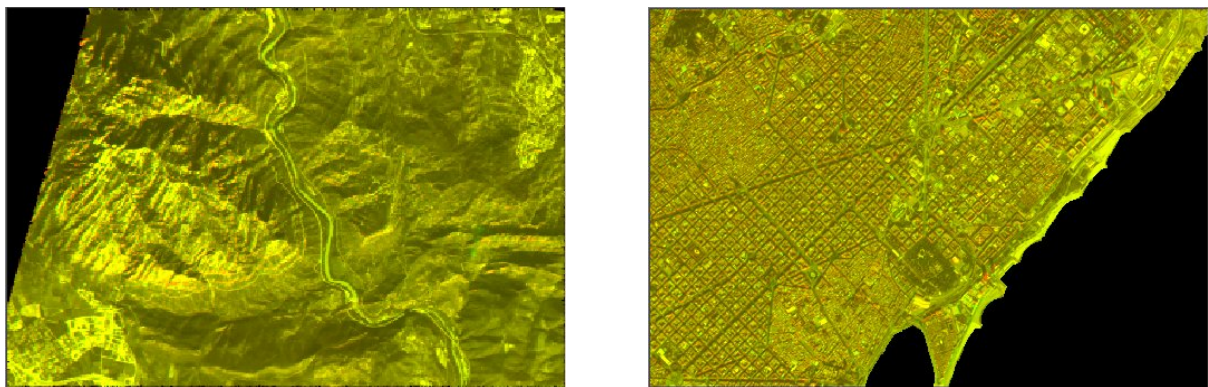


Figure 6: Stereo ortho photos for mountainous (left) and urban (right) area