TOWARDS AUTOMATED DEM GENERATION FROM HIGH RESOLUTION STEREO SATELLITE IMAGES

Pablo d'Angelo, Manfred Lehner, Thomas Krauss, Danielle Hoja and Peter Reinartz

German Aerospace Center (DLR), Remote Sensing Technology Institute, D-82234 Wessling, Germany
(Pablo.Angelo, Manfred.Lehner, Thomas.Krauss, Danielle.Hoja, Peter.Reinartz)@dlr.de

Commission IV, WG IV/9

KEY WORDS: spaceborne scanner systems, digital elevation models (DEM), image matching, CARTOSAT-1, orthoimage, accuracy analysis

ABSTRACT:

High resolution stereo satellite imagery is well suited for the creation of digital surface models (DSM). In this paper, a system for highly automated DSM and orthoimage generation based on CARTOSAT-1 imagery is presented. The proposed system processes photometrically corrected level-1 stereo scenes using the rational polynomial coefficients (RPC) universal sensor model. The RPC are derived from orbit and attitude information and have a much lower accuracy than the ground resolution of approximately 2.5 m. Ground control points are used to estimate affine RPC correction. Accurate GCP are not always available, especially for remote areas and large scale reconstruction. In this paper, GCP are automatically derived from lower resolution reference images (Landsat ETM+ Geocover and SRTM DSM). It is worthwhile to note that SRTM has a much higher lateral accuracy than the Landsat ETM+ mosaic, which limits the accuracy of both DSM and orthorectified images. Thus, affine RPC correction parameters are estimated by aligning a preliminary DSM to the SRTM DSM, resulting in significantly improved geolocation of both DSM and orthoimages. Robust stereo matching and outlier removal techniques and prior information such as cloud masks are used during this process. DSM with a grid spacing of 10 m are generated for 9 CARTOSAT-1 scenes in Catalonia. Checks against independent ground truth indicate a lateral error of 3-4 meters and a height accuracy of 4-5 meters. Independently processed scenes align at subpixel level and are well suited for mosaicing.

1. INTRODUCTION

In May 2005 India launched its IRS-P5 satellite with CARTOSAT-1 instrument which is a dual-optics 2-line along-track stereoscopic pushbroom scanner with a stereo angle of 31° and the very interesting resolution of 2.5 m. The operational use of the data is described in (Srivastava et al, 2007). The CARTOSAT-1 high resolution stereo satellite imagery is well suited for the creation of digital surface models (DSM). In this paper, a system for highly automated DSM generation based on CARTOSAT-1 stereo scenes is presented.

CARTOSAT-1 stereo scenes are provided with rational polynomial functions (RPC) sensor model, derived from orbit and attitude information. The RPC have a much lower accuracy than the ground resolution of approximately 2.5 m. Traditionally, subpixel accurate ground control points (GCP) are used in previous studies to estimate bias or affine RPC correction parameters required for high quality geolocation of HRSI images. Such highly accurate GCP are usually derived from a DGPS ground survey or high resolution orthoimages and digital elevation models. For many applications, especially ones demanding near real-time results, such as desaster assessment tasks in remote regions, highly accurate GCP data is often not available. Without accurate GCPs, CARTOSAT-1 scenes are of limited use, since bias or affine RPC correction is required before CARTOSAT-1 scenes can be used for DSM extraction and orthorectification (Lehner et al, 2007).

We propose the use of widely available lower resolution satellite data, such as the Landsat ETM+ and SRTM DSM datasets as a reference for RPC correction. Digital surface models (DSM) are derived from dense stereo matching and forward intersection and subsequent interpolation into a regular grid. Since stereo matching is unreliable in large, homogeneous image areas, such as fields, meadows and water bodies, as well as in complicated terrain with occlusions and shadows, strict consistency checks are used during matching. The first section of the paper describes the process used for DSM generation. The second part evaluates the processor using 9 CARTOSAT-1 stereo pairs.

2. DSM GENERATION

The DSM generation process consists of the following main steps, implemented as part of the DLR XDibias image processing system.

1. Stereo matching in epipolar geometry
2. Affine RPC correction and alignment to reference DEM
3. Forward intersection and outlier removal
4. Interpolation
5. Orthorectification
6. Quality inspection and manual editing

A CARTOSAT-1 stereo scene consists of a nadir looking image with a along track tilt of -5°, a forward looking image with a along track tilt of 26°. They are named Aft and Fore throughout this paper.

2.1 Stereo Matching

Hierarchical intensity based matching is used for matching the stereo pairs and the reference image. It consists of two major steps, hierarchical matching to derive highly accurate tie points, followed by a region growing step to generate a dense set of tie points.
The initial matching step uses a resolution pyramid (Lehner&Gill, 1992; Kornus et al., 2000) to cope even with large stereo image distortions stemming from carrier movement and terrain. Large local parallaxes can be handled without knowledge of exterior orientation. The selection of pattern windows is based on the Foerstner interest operator which is applied to one of the stereo partners. For selection of search areas in the other stereo partner(s) local affine transformations are estimated based on already available tie points in the neighborhood (normally from a coarser level of the image pyramid). Tie points with an accuracy of one pixel are located via the maximum of the normalized correlation coefficients computed by sliding the pattern area all over the search area. These approximate tie point coordinates are refined to subpixel accuracy by local least squares matching (LSM). The number of points found and their final (subpixel) accuracy achieved depend mainly on image similarity and decrease with increasing stereo angles or time gaps between imaging. The procedure results in a rather sparse set of tie points well suited for introduction into bundle adjustment and as an excellent source of seed points for further densification via region growing.

The second step uses the region growing concept first published by Otto and Chau in the implementation of TU Munich (Heipke et al., 1996). It combines LSM with a strategy for local propagation of initial conditions of LSM. Epipolar stereo images are used in the region growing step, and the propagation strategy is modified to enforce points located on the epipolar lines. Stereo tie points deviating more than 0.5 pixels from the epipolar geometry are removed. A quasi-epipolar stereo pair with epipoles corresponding to the image columns is generated by aligning the columns of the Fore image with the Aft image, using highly accurate matches from the pyramidal matching step.

Various methods for blunder reduction are used for both steps of the matching:

- Threshold for correlation coefficient
- 2-directional matching and threshold on resulting shifts of the coordinates
- Threshold on the deviation from epipolar geometry.

In areas of low contrast the propagation of affine transformation parameters for LSM in region growing leads to high rates of blunders. In order to avoid intrusion into homogeneous image areas (e.g. roof planes without structure) the extracted image chips are subject to (low) thresholds on variance and roundness of the Foerstner interest operator. This and the many occlusions found in densely built-up areas imaged with a large stereo angle create lots of insurmountable barriers for region growing. Thus, for high resolution stereo imagery the massive number of seed points provided by the matching in step one (image pyramid) turns out to be essential for the success of the region growing.

The numbers of tie points found and their subpixel accuracy is highly dependent on the stereo angle. A large stereo angle (large base to height ratio b/h) leads to poorer numbers of tie points and to lower accuracy in LSM via increasing dissimilarity of (correctly) extracted image chips.

2.2 GCP collection and affine RPC correction

Previous studies (Lehner et al., 2007) have shown that the CARTOSAT-1 RPC ground accuracy is in the order of hundred meters. Additionally, forward intersection performance without RPC correction is poor and results in large residuals in image space. The estimation of affine RPC correction parameters requires well distributed GCP with subpixel accuracy. In many application scenarios, such as continent wide reconstruction or crisis support applications, acquiring the required GCP is very tedious or might even be impossible, if a fast response is required.

Gobal and easily available reference datasets are the OnEarth Landsat ETM+ Geocover mosaic and the SRTM elevation data. The accuracy of these datasets is low compared to the high resolution CARTOSAT-1 images. The Landsat ETM+ Geocover mosaic is specified with a lateral error of 50m. The absolute lateral error of SRTM amounts to 7.2m - 12.6m (LE90, depending on the continent), with an absolute height error of 4.7m to 9.8m (Rodriguez et al., 2005).

GCPs are collected by transfer of highly accurate tie points between the CARTOSAT-1 Aft and Fore images to the Landsat reference image and extraction of the corresponding height from SRTM. The matching procedure starts by aligning the CARTOSAT-1 Aft image to the ETM+ reference by using the corner values provided in the CARTOSAT-1 metadata. The first step of the hierarchical matching procedure described in Section 2.1 is applied to obtain tie points between the ETM+ and CARTOSAT-1 Aft scenes. A similar matching could be done by matching the Fore image against the ETM+ image, it would however yield different tie points and thus GCPs for the Aft and Fore image. Since affine RPC correction is performed separately for each image, a good link between the Aft and Fore images is required to ensure good forward intersection behaviour. Thus, highly accurate stereo tie points between the CARTOSAT-1 Aft and Fore images are selected by applying strict thresholds on the bidirectional matching shift (0.1 pixels) and correlation coefficient (0.8). The Aft coordinates of these stereo tie points are then used as interest points and matched against the full resolution ETM+ scene, using the previous Aft vs. ETM+ matching as initial approximation. This yields the geographic position of the stereo tie points. Finally, 3D GCPs for both Aft and Fore scene are obtained by bilinear interpolation of the SRTM DSM. We use a non-interpolated C band SRTM, where holes larger than 2 pixels are still open. This avoids deriving GCP from interpolated heights. Affine RPC correction parameters are estimated both for the Aft and Fore scene.

2.2.1 RPC correction by DSM alignment

After the alignment based on ETM+ and SRTM reference data, forward intersection residuals are significantly improved, but the lateral accuracy is still limited by the ETM+ Geocover reference. To take advantage of the higher accuracy of the SRTM dataset a second RPC correction step is necessary. A 3D point cloud is calculated by forward intersection of a subset of the stereo tie points. The point cloud is aligned to the SRTM DSM. It is assumed that the height z_i of a point P_i located at

(x_i,y_i,z_i) equals the reference DSM height $h_D(x_i,y_i)$ at the corresponding position (x_i,y_i):

$$h_D(x_i, y_i) = z_i \tag{1}$$

A 3D affine transformation is used to align the initial stereo point cloud to the SRTM DSM:

$$\vec{p}_{ti} = A\vec{p}_i \tag{2}$$

where $\vec{p}_i = (x_i \quad y_i \quad z_i \quad 1)^T$ is the original point,
A is a 3x4 matrix,
$\vec{p}_{ti} = (x_{ti} \quad y_{ti} \quad z_{ti})^T$ is the transformed point.

The affine transformation matrix A is estimated using an iterative least mean squares algorithm. Using Eq. (1) and (2), the following observation equation is obtained.

$$v_i = h_D(x_{ti}, y_{ti}) - z_i \tag{3}$$

Since the model is non-linear, the solution is obtained iteratively. An identity transform is used as initial approximation, since the stereo points are not far from the reference. It is likely that the stereo point cloud, and to a smaller extend the DSM contains outliers, which cannot be handled by a standard least mean squares algorithm. After the initial estimation, points with a residual larger than 3 times the standard deviation are removed and a new transformation is estimated. This procedure is repeated until less than 0.3% outliers are detected and the squared sum of the outlier residuals accounts for less than 5% of the squared sum of all residuals.

The estimated affine transformation could be used to align the final DSM to the SRTM reference and thus improve its accuracy. Orthoimages would however still be limited by the ETM+ accuracy. It is desirable to include the correction in the RPC models, too. This is done by aligning the 3D stereo points to SRTM and using them as GCP for a second RPC correction which yields the final affine RPC correction used in all subsequent steps.

2.3 Forward intersection and outlier removal

Forward intersection is done via iterative least squares adjustment using 4 observation equations and derives object space coordinates in Geographic coordinates in WGS84 datum. (Grodecki et al, 2004, Lehner et al, 2007). The residuals in image space are used for a further blunder reduction step. Points with a residual larger than 0.5 pixels are rejected. Of course, only residuals in cross track direction will be effective because wrong row coordinates of tie points are translated into wrong height values if only two stereo partners are available (stereo imaging direction).

The forward intersected points still contain a small amount of blunders due to matching errors in regions with sparse texture. To eliminate gross outliers, a reference check against the SRTM DSM is performed. All points whose height deviates more than 3 times the height error of the SRTM are rejected. The SRTM height error map was found to be a good approximation of the true height error (Rodriguez et al. 2005), and is used to dynamically adjust the height difference threshold. Typical thresholds are 24 m in flat areas, and 75 m in mountainous areas.

2.4 DSM interpolation

Result of matching and forward intersection is a set of 3D points representing the Earth surface (including f.e.. tree tops) acquired by the stereo images. To ease further applications, the irregular point cloud is transferred to a regularly spaced grid with a spacing of 10 m. If multiple points fall into the same grid cell, their heights are averaged to form a new point. The points are connected by Delauney triangulation into a triangulated irregular network (TIN). Finally, the triangles are superimposed on the regularly spaced grid of the resulting DSM. For each triangle the plane defined by the three vertices is calculated. To each pixel inside the triangle the height value interpolated on this plane is assigned (Hoja et al., 2005).

2.5 Orthorectification

Orthoimages with user defined datum and projection are created by orthorectification of the Aft image with the generated DSM and the affine corrected RPC.

3. EVALUATION

The DSM creation process described above is evaluated with 9 CARTOSAT-1 scenes of Catalonia. Scene Cat was part of the Cartosat scientific assessment programme, while the remaining 8 scenes have been provided by Euromap. The Landsat ETM+ Geocover mosaic and the SRTM C band DSM have been used as sources for GCP collection. Two reference DTM with a GSD of 15 m and 10 orthoimages with a resolution of 0.5 m have been provided by the Institut Cartographic de Catalunya (ICC) and are only used as ground truth during the evaluation. The location of the scenes and ground truth data is shown in Figure 1. The scenes are mostly cloudless. Scene 117/207 contains two large clouds in the upper left corner, covering most of the overlap between with scene 116/207. The scenes located near the coastline contain both flat areas along the coast as well as the Montseny mountain range with peaks of over 1600 meters. As shown in Table 1, the scenes were aquired early in the year, leading to large shadows in the mountainous areas.

Abbreviation for the paper for aft and fore scenes	Imaging date
Cat-A/F	01 Feb. 2006
115/207-A/F	16 Feb. 2008
115/208-A/F	16 Feb. 2008
116/207-A/F	05 March 2008
116/208-A/F	05 March 2008
116/209-A/F	05 March 2008
117/207-A/F	25 Jan. 2008
117/208-A/F	25 Jan. 2008
117/209-A/F	25 Jan. 2008

Table 1: CARTOSAT-1 scenes evaluated in this paper

3.1 Matching

The hierarchical matching procedure described earlier yields a large number of good matching points, see Table 5 for the amount of mass tie points extracted from each scene.

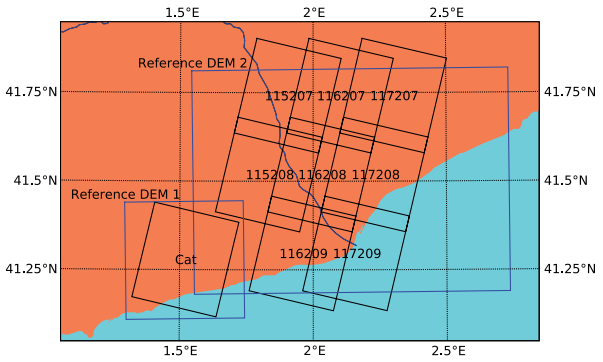


Figure 1: Geographic layout of evaluated CARTOSAT-1 scenes and ICC reference data used for verification.

3.2 RPC correction

Matching and thus GCP collection between CARTOSAT-1 and ETM+ scenes is hindered by the large time and resolution differences. Most scenes of the ETM+ mosaic have been captured between 1999 and 2001, resulting in large differences between CARTOSAT-1 and ETM+ scenes. Due to the large differences in appearance, relatively loose thresholds on correlation (0.7) and bidirectional matching coordinate shift (0.5 pixels in the ETM+ image) have been used during GCP collection. A sufficient number of well distributed GCP is found for all 9 scenes. An iterative outlier removal procedure is applied during the affine RPC correction estimation. Between 856 and 93 GCP are used for RPC correction. The image space residuals of the GCP are quite large, with standard deviation between 1.7 and 2 pixels, mostly due to the large resolution difference between CARTOSAT-1 and ETM+. No systematic error is visible in the residuals.

The C-SAP demonstrated that subpixel residuals can be achieved when a few high quality and well distributed GCP are available (Lehner et al., 2007). From the above results, it is expected that the Landsat ETM+ mosaic is not a suitable base for deriving GCP with the accuracy required for CARTSAT-1. The DSM based RPC refinement described in section 2.2.1 is thus used to further reduce the error. To estimate the true accuracy of the two RPC correction approaches, 68 checkpoints have been measured in the ICC orthoimages and the stereo partner Cat-A. The height of each checkpoint is derived from the ICC DTM. These measurements have been automatically transformed into Cat-A/F tie points via least squares matching. 6 window sizes from 17 to 27 have been used in LSM in order to get statistical values for the accuracy. Forward intersection of these stereo tie points results in object space positions. The lateral and height differences between stereo points and checkpoints are given in Table 2. It is obvious that the correction based on GCP derived from ETM+ and SRTM leads to a high shift in location and height. Considering the 15 m resolution of ETM+, a mean difference of 12.5 m is still a good result and indicates subpixel accuracy of the ETM+ Geocover mosaic in the studied area. After aligning the stereo points to SRTM and re-estimation of the affine RPC correction, the lateral displacement reduces to 3.5 m. This is a very good result, especially when considering the 90 m grid spacing of the SRTM. Figure 2 shows the lateral shifts of all checkpoints.

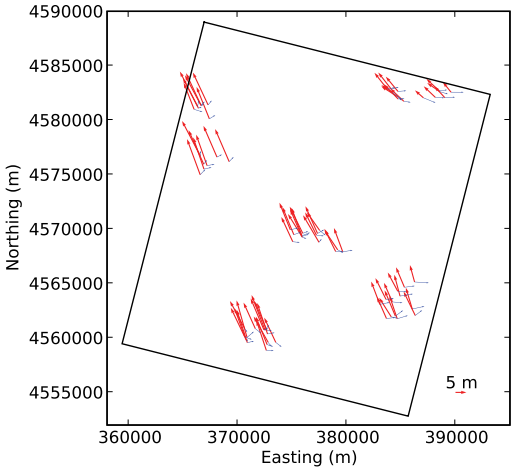


Figure 2: Lateral error of Cat scene for the two affine RPC correction methods, measured using independent checkpoints. Red, thick arrows: ETM+ and SRTM GCP. Blue, thin arrows: Align to SRTM. The arrow lengths are scaled by a factor 200.

RPC correction reference	Lateral difference (m)		Height difference (m)	
	Mean	σ	Mean	σ
ETM+, SRTM	12.51	3.25	1.20	2.40
Align to SRTM	3.48	1.10	0.30	1.47

Table 2: Accuracy of the two RPC correction procedures, measured using well distributed, independent checkpoints.

3.3 Forward intersection and outlier removal

Forward intersection of the mass stereo tie points performs well; no points are discarded with a relatively strict threshold of 0.5 pixels on the image space residuals. A few gross blunders with mismatches along the epipoles remain and are rejected due to their deviation from the SRTM DSM.

Table 3 shows the number of accepted points and their mean height difference to the ICC reference DTM. When comparing the differences with the checkpoint evaluation in section 3.2, the larger height errors and standard deviations are noticeable. This is caused by the suboptimal conditions for image matching, such as the low sun angle and the mostly mountainous terrain with vegetation. Especially in the Montseny mountain range located in the upper right of the block, very large black shadows with a diameter of several km can be found, leading to large interpolation facets and resulting in a rather coarse DSM with large facets. Most of scene 117/209 is covered by the ocean and the city of Barcelona with large build up areas. When comparing the generated surface model with the bare earth DTM provided by ICC, a negative height difference, as well as a larger standard deviation is expected in such areas. The negative mean height difference observed for all scenes is a good sign and shows that the CARTOSAT DSM is located above the ICC DTM.

Scene	Number accepted points (Mio.)	Height difference to ICC DTM	
		Mean (m)	σ (m)
Cat	21.5	-0.80	3.41
115/207	23.9	-1.74	4.14
115/208	24.1	-2.14	4.64
116/207	27.9	-3.02	4.84
116/208	26.3	-2.82	5.40
116/209	15.8	-3.70	5.43
117/207	18.0	-3.26	4.98
117/208	21.2	-3.31	5.59
117/209	7.5	-4.83	6.42

Table 3: Accepted object points statistics

3.4 DSM and Orthoimage creation

After the outlier removal step, a DSM on a 10 meter grid is interpolated for all scenes. Figure 4 shows the difference between the ICC DTM and the generated DSM for the CAT scene. The consistency of the produced Aft/Fore orthoimages is assessed by image matching. Highly accurate tie points between each ortho pair have been established using LSM. The mean shift varies between 0.28 and 0.30 pixels, the corresponding standard deviation between 0.15 and 0.18 pixels. Figure 3 shows that the remaining subpixel shift between the Aft and Fore orthoimages is slightly regular.

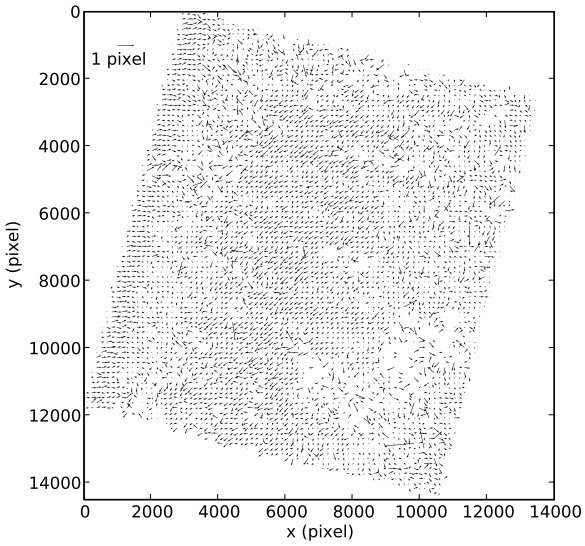


Figure 3: Shifts (averaged on a 150x150 pixel grid) between orthorectified Aft and Fore image of scene 116/208

3.5 Mosaicing

After each scene has been processed independently, DSM and Aft orthoimages of scenes (115-117)/(207-209) are combined into orthoimage and DSM mosaics. No further radiometric or geometric adjustment has been applied. To assess the consistency of the resulting mosaic, shifts between the neighbouring orthoimages are computed by LSM matching and shown in Table 4. This is a good result, considering that no bundle block adjustment was used, and all scenes were processed independently. The Ortho and DSM mosaics are shown in Figure 5.

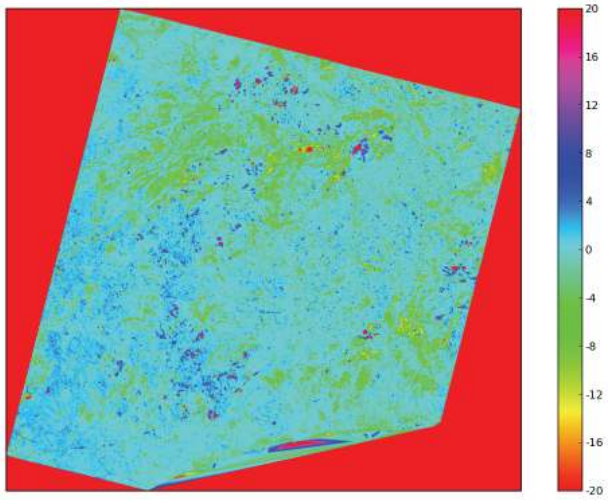


Figure 4: Height difference (meter) obtained by subtracting the CARTOSAT-1 DSM from the ICC DTM.

Path/Row	Lateral shift (m)	
	mean	σ
(115,116) / 207	1.66	1.46
(116,117) / 207	0.92	0.71
(115,116) / 208	1.12	0.64
(115,116) / 209	3.12	1.31

Table 4: Lateral shift in overlapping areas of neighbouring orthoimages. Except for the overlap between scene 115/209 and 116/209, subpixel shifts are found.

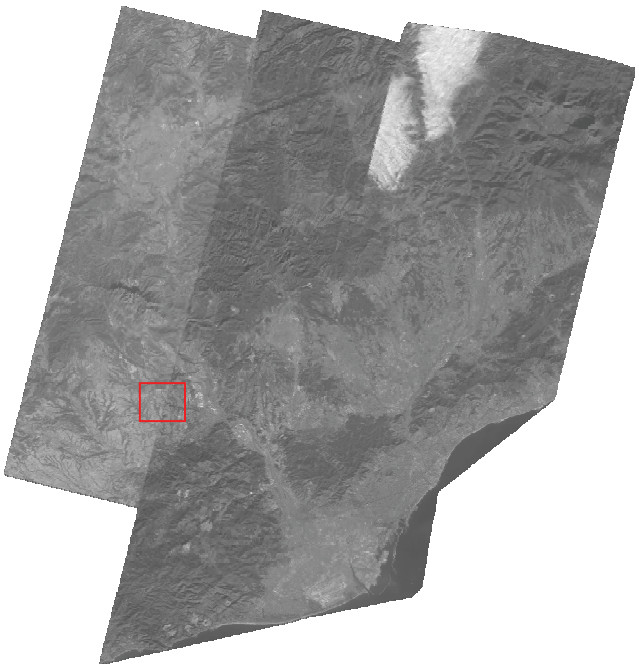
4. CONCLUSIONS

A highly automated DSM processor for CARTOSAT-1 scenes is presented. Key features include the automatic collection of reference data from Landsat ETM+ Geocover mosaic and the SRTM DSM, a novel affine RPC correction estimation approach which leads to ortho and DSM products with much higher lateral accuracy than the ETM+ Geocover mosaic. Comparison with reference data in Catalonia indicates a location accuracy of 3-4 meter, and a height accuracy of 3-4 m in terrain with good pattern matching characteristics. In mountainous terrain and with rather low sun elevations the normal problems of too steep slopes and large shadow areas without enough matching targets lead to coarse reconstruction and local artefacts in the DSM.

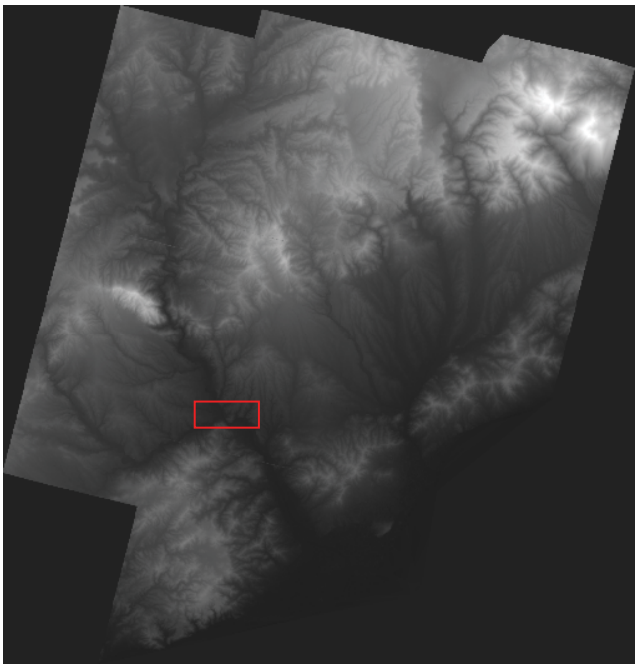
The ability to produce quality DSM and ortho products without manually measured, highly accurate ground control points is especially valuable for emergency applications, mapping of remote areas and large scale DSM production. Seamless mosaics can be formed, even if all scenes were processed independently. Further work includes development of comfortable manual quality inspection and editing tools, as well as improved computational efficiency and better handling of occluded and shadowed areas in the stereo matching module.

ACKNOWLEDGEMENTS

EUROMAP provided most of the CARTOSAT-1 scenes used for this analysis, with the exception of the Cat scene, which was provided by ISRO/SAC as part of the CARTOSAT-1 Scientific Assessment Program. Acknowledgements also go to the Institut

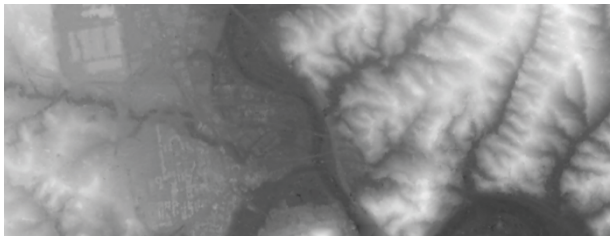


Orthoimage mosaic



DSM mosaic

(c) Detail of orthoimage mosaic, showing the consistency of the independently processed scenes 115/208 and 116/208.



Detail of CARTOSAT-1 DSM

Corresponding SRTM C Band DSM

Figure 5: Resulting orthoimage and DSM mosaics.

Cartographic de Catalunya for the delivery of adequate ground truth for Catalonia.

REFERENCES

Grodecki, J., Dial, G., Lutes, J., 2004: *Mathematical Model for 3D feature extraction from multiple satellite images described by RPCs*, ASPRS Annual Conf. Proc., Denver, Colorado, USA

Hoja, D., Reinartz, P., Lehner, M., 2005: *DSM Generation from High Resolution Satellite Imagery Using Additional Information Contained in Existing DSM*, Proc. of the ISPRS Workshop 2005 High Resolution Imaging for Geospatial Information, Hanover, Germany

Kornus W., Lehner M., Schroeder, M., 2000: *Geometric inflight calibration by block adjustment using MOMS-2P 3-line-imagery of three intersecting stereo-strips*, SFPT (Société Francaise de Photogrammétrie et Télédétection) , Bulletin Nr. 159, pp. 42-54

Lehner M., Gill, R.S., 1992: *Semi-Automatic Derivation of Digital Elevation Models from Stereoscopic 3-Line Scanner Data*, IAPRS, Vol. 29, part B4, Commission IV, pp. 68-75, Washington, USA

Lehner, M., Müller, Rupert, Reinartz, P., Schroeder, M., 2007: *Stereo evaluation of CARTOSAT-1 data for French and Catalonian test sites*, Proc. of the ISPRS Workshop 2007 High Resolution Earth Imaging for Geospatial Information, Hanover, Germany, May 29 – June 1

Rodriguez, E., Morris, C.S., Belz, J.E., Chapin, E.C., Martin, J.M., Daffer, W., Hensley, S., 2005: *An assessment of the SRTM topographic products*, Technical Report JPL D-31639, Jet Propulsion Laboratory, Pasadena, California, 143 pp.

Srivastava P.K., Srinivasan T.P., Gupta Amit, Singh Sanjay, Nain J.S., Amitabh, Prakash S., Kartikeyan B., Gopala Krishna B., 2007: *Recent Advances in CARTOSAT-1 Data Processing*, Proc. of the ISPRS Workshop 2007 High Resolution Earth Imaging for Geospatial Information, Hanover, Germany, May 29 – June 1