

AUTOMATIC ORIENTATION OF LARGE MULTITEMPORAL SATELLITE IMAGE BLOCKS.

Pablo d'Angelo

German Aerospace Center (DLR), Remote Sensing Technology Institute
D-82234 Wessling, Germany
pablo.angelo@dlr.de

KEY WORDS: Matching, Georeferencing, Digital photogrammetry, bundle adjustment, CARTOSAT-1

ABSTRACT:

One of the main goals of optical 3D mapping satellites, such as SPOT-5, Cartosat-1, ALOS, ZY-3 and similar satellites is the generation of large ortho and digital surface model mosaics. The direct georeferencing performance of the current satellites is worse than the resolution of their data. This leads to geometric errors between adjacent scenes, if they are processed independently and without ground control. These errors can be reduced or eliminated by georeferencing the images with high quality GCPs or with a block triangulation. A complete coverage of larger areas with high resolution satellite imagery usually requires several months or years of acquisition. The images typically contain strong differences due to different seasons, cloud cover and other effects. This paper introduces our automatic image orientation procedure for satellite images, which can deal with these effects and automatically orient large blocks of satellite imagery. We additionally exploit DSMs to minimize the number of ground control points required for orientation of the images. The method evaluated on a block of 1210 Cartosat-1 images covering northern Italy.

1. INTRODUCTION

One of the main goals of optical 3D mapping satellites, such as SPOT-5, Cartosat-1, ALOS, ZY-3 and similar satellites is the generation of large Ortho and Digital surface model mosaics. Unfortunately, the direct georeferencing performance of the current satellites is worse than the resolution of their data. This leads to geometric errors between adjacent scenes, if they are processed independently and without ground control. High quality products should be consistent, and have sub-pixel errors between adjacent scenes. To ensure proper georeferencing, each image can be oriented separately using high quality GCPs, or with a block triangulation that involves neighbouring scenes or strips. To deal with the high data volume of today's satellite systems, a highly automated processing is required for cost effective production of state or continent wide mosaics. This also requires that the image orientation should happen automatically. This includes gathering the data required for adjustment, such as tie points and ground control points.

Very large blocks of aerial imagery are routinely processed using bundle block adjustment (aerial triangulation). While the triangulation process of satellite and aerial imagery is quite similar, except for using different sensor models, satellite data sets are often harder to process. Aerial imagery is usually acquired under controlled conditions in a regular pattern, and within days of each other. For this type of imagery, robust and efficient tie point matching strategies exist, that allow full automation of the processing chain (Wiechert et al., 2012).

Typical satellite imagery blocks are collected across multiple seasons (and sometimes over several years), and under unfavourable conditions, such as significant cloud cover. This results in imagery with large difference in appearance, resulting in matching failures. Additionally, the image quality of some satellites is quite low compared to digital aerial imagery, further

complicating image matching. These factors result in problems when performing automatic tie point matching on blocks with thousands of images. This paper introduces an approach that is successfully used for the automatic orientation Cartosat-1 blocks consisting of more than 1200 images.

2. AUTOMATED ORIENTATION

2.1 Overview

The procedure consists of multiple steps, first, downscaled images are matched using the SIFT operator. For this, a simple but effective approach based on RANSAC is used to verify the pairwise matching. Image pairs with outlier rates of more than 90% can be successfully handled. After outlier removal, the pairwise matches are chained into multi ray matches, and an initial bundle block adjustment is performed. Using the corrected orientation, tiepoints are then refined using local least squares matching on the full resolution images, resulting in tie points with an accuracy of up to 0.1 pixels. Finally, a block adjustment is performed using these high quality tie points.

2.2 Coarse SIFT tie point matching

The first step in the orientation procedure is the estimation of tie points connecting all images in a block. These tie points should be multi ray tie points, connecting all overlapping imagery. For this purpose, we use SIFT feature matching [Lowe 2004]. It is a robust technique for matching of images with different scale, rotation and some local distortion. It is also relatively robust to radiometric changes. While scale and rotation invariance are typically not required for typical satellite imagery, local distortion and radiometric changes are common.

As SIFT is a computationally expensive algorithm, when applied to large remote sensing images, we perform the initial

tie point matching on downscaled images. For the experiments in this paper, we used a downscale factor of four. After downscaling, SIFT features are extracted for every image. The standard SIFT parameters were modified to allow feature extraction in areas of low contrast. This is especially important for images with lower dynamic range, such as Cartosat-1 images at high latitudes (such as Iceland). Here the useful gray value range can be as low as 10 gray values. With the default SIFT parameters (Lowe, 2004), no sift features can be extracted in these areas. Lowering the SIFT contrast threshold results in many interest points on images with high contrast. To reduce the number of sift features to a manageable amount; we divide the image into 30x30 grid cells. For each cell, we select 20 interest points with the highest interest score. This limits the maximum number of interest points for each image to 18000.

For matching of “non” agile satellites, that cannot scan sideways, such as Cartosat-1, ZY-3 or ALOS PRISM, we disable the rotation invariance of the SIFT algorithm by ignoring the orientation of the interest point during computation of the feature descriptor. The feature descriptors are then stored on disk for each image. As the feature descriptors can be computed for each image independently, all CPU cores in a computer can be used effectively.

The image metadata or RPC files are used to compute a coarse location for each image. All overlapping image pairs are matched independently. We use the ratio between nearest and second nearest neighbour as matching score and keep all matches with a ratio smaller than 0.75. This high ratio ensures that many matches are found even for radiometrically different images, but also leads to an increased number of blunders.

As bundle adjustment can only tolerate a low rate of blunders, a robust outlier removal is required. We use a simple but effective RANSAC (Fischler and Bolles 1981) based procedure for outlier removal. This procedure first computes the spatial forward intersection for all tie points and their residuals in image space. Then a random point is chosen, and a set of tie points with similar residuals are found. This procedure is repeated 90 times, and the set containing the highest number of points is selected as inliers. The pseudocode for this algorithm is:

```
residuals = fwint of tie points
inliers = {}
max_len = 0
for i in 90 random tie points:
    maybe_inliers = {}
    for j in all tie points:
        if abs(residual_i-residual_j) < t:
            add point to maybe_inliers
    if len(maybe_inliers) > len(inliers)
        inliers = maybe_inliers
```

This procedure assumes that the relative orientation between the two images can be modelled well with a bias correction in image space. This applies to most high resolution satellite sensors where the image strip is cut into roughly quadratic subscenes. The threshold t needs to be chosen, we use 20 pixels. This is a rather conservative choice, for satellites that required only a bias RPC correction, lower values of 2 or 3 pixels also work fine.

Note that the model just has 4 degrees of freedom; these can be estimated using a single corresponding point in the image pair.

This is a major advantage over using image to image similarity transforms such as scale+rotation or affine transformation, which require two or three corresponding points, and thus a larger number of RANSAC iterations. Additionally, our approach is independent of the scene and viewing geometry, and works well for mountainous images captured from significantly different viewpoints. Scale and rotation, or affine transformation cannot model these effects.

These properties allow us to remove large percentages of outliers accurately and efficiently. Using 90 RANSAC trials allows us to find the correct inlier subset with a probability of 99% under the assumption that there are 95% outliers in the matched points. The only drawback of this method is that outliers in epipolar direction of the image pair cannot be detected. Thus a small number of outliers in epipolar direction might still be present after this procedure.

For bundle adjustment it is advantageous to use multi-ray points to reduce the number of unknowns in the block. As the pairwise matches are based on the same SIFT feature points, we can merge all pairwise matches that share the image coordinates in the same image into multi ray tie points.

As there still might be many multi ray tie points inside the block, we need to further thin out the multi ray points. This thinning should keep the overall connectivity to ensure that the block does not break into multiple sub-blocks and also enforce a strong block layout with many multi-ray points. A grid based thinning is used for this purpose. It keeps all connections between images and greedily selects the points with the highest number of rays for each bin and image pair. The following algorithm is used:

```
for i in all_images:
    for g in grid bins:
        for j in connected images:
            for point in grid bin g of i and j:
                keep point with highest number
                of rays.
```

This procedure results in multi ray tie points with a density controlled by the number of grid bins in each image.

2.3 Coarse Bundle adjustment

The coarse SIFT based tie points are used to perform an approximate orientation of the input imagery, which will later be used for the tie point refinement. This is especially important for Cartosat-1 imagery, where we have occasionally observed scenes with a direct geo referencing error of up to 1.7 km.

For a coarse orientation it is sufficient to estimate a zero order RPC bias correction parameters for each scene. This typically allows orientation with a relative accuracy of several pixels. Using a higher order RPC correction could potentially lead to slightly better orientations for some satellites, but our coarse tie point accuracy is also limited and in the range of 2 pixels only. The lower number of parameters in bundle adjustment also decreases the risk of convergence to a bad solution and increases the robustness to gross outliers in epipolar direction that might still be present in the coarse multi ray tie points. More details on the used bundle block adjustment are given in Section 3.

2.4 Tie point refinement

For a high quality adjustment, tie points with sub pixel accuracy are required. A refinement of the SIFT tie points is therefore required, which we perform using local least squares matching (Ackermann, 1983). We first compute the object positions of all SIFT based tie points before thinning and then project these into the still unmatched images. This provides approximate point positions inside all images with an accuracy of a few pixels. For each point, a single master image is chosen and the point is matched to all possible images. As the convergence radius for local least squares matching is limited and only in the order of one pixel, we use a pyramidal approach to successfully match pixels with a larger initial error. Strict thresholds on correlation coefficient and bidirectional matching result in high quality results with almost no outliers. To reduce the number of points to a manageable level for the block adjustment, we perform thinning of the multi ray points as described in Sec. 2.2.

2.5 Final block adjustment

Using these high quality tie points, we perform a bundle block adjustment using the RPC correction order required by the particular satellite. This provides the final orientation of the images used in further processing.

3. BUNDLE BLOCK ADJUSTMENT

The bundle adjustment procedure follows the approach given by [Grodecki 2003], and estimates corrections in image space. As most of the high resolution imagery is delivered with RPC coefficients we use these as sensor model. We use the following model, where r, c are the row and column image coordinates predicted by the RPC model, and r' and c' are the corrected image coordinates:

$$\begin{aligned} r' &= a_0 + a_1 r + a_2 c \\ c' &= b_0 + b_1 r + b_2 c \end{aligned} \quad (1)$$

For a coarse correction, a bias correction with parameters a_0, b_0 is sufficient. Depending on the satellite, an affine correction using 6 parameters a and b . Some satellites might even require an extension of Eq. (1) with quadratic terms.

A bundle block adjustment can be used to estimate correction parameters for more than one image at a time. We follow the RPC bundle block adjustment of (Grodecki 2003), which is based the following observation equations:

$$\begin{aligned} \Delta r_{ji} &= r_j(\phi_i, \theta_i, h_i) - r_{ji}^* \\ \Delta c_{ji} &= c_j(\phi_i, \theta_i, h_i) - c_{ji}^* \end{aligned} \quad (2)$$

where ϕ_i, θ_i, h_i are the object coordinates of point i , r_j and c_j are the object coordinates to image coordinates projection functions for row and column, depending on the RPC coefficients and the correction parameters a and b . r_{ji}^* and c_{ji}^* are the measured image coordinates in image j .

If the ground coordinates $\phi_i^*, \theta_i^*, h_i^*$ for some points are known (Ground control point, GCP), for example by GPS measurement or from matching against a reference image, additional observations are added:

$$\begin{aligned} \Delta \phi_i &= \phi_i - \phi_i^* \\ \Delta \theta_i &= \theta_i - \theta_i^* \\ \Delta h_i &= h_i - h_i^* \end{aligned} \quad (3)$$

Prior information about the rpc correction terms can be added, if available:

$$\begin{aligned} \Delta a_j &= a_j - a_j^* \\ \Delta b_j &= b_j - b_j^* \end{aligned} \quad (4)$$

An extension to (Grodecki, 2003) is the DEM height difference observation term, similar to the approach by (Strunz, 1993):

$$\Delta h_i = h_i - D(\phi_i, \theta_i)^* \quad (5)$$

where $D(\phi_i, \theta_i)^*$ is the bilinearly interpolated height in the reference DEM at position ϕ_i, θ_i .

For satellites requiring an affine RPC correction, well distributed GCPs are essential for high quality results. The 3D alignment in the bundle block adjustment using Eq. (5) works well in areas where significant terrain is available throughout the whole scene. If an absolute accuracy of approximately 10m CE90 is sufficient, the SRTM C Band DEM is a suitable reference. In practice, this method works very well for images with enough terrain (Utenthaler et al., 2011), and the whole orientation process can be fully automated. Scenes that contain no usable relief for DSM alignment require the use of GCPs, or reference images. Problems might also occur for hilly scenes that contain large flat areas, as the 3D matching does not provide lateral constraints in flat areas. This is similar to using an uneven distribution of ground control points.

Using the observations equations (1)-(5), and corresponding weights W , based on prior knowledge, an iterative least squares estimation is used to estimate tie point object coordinates and the RPC correction parameters. In classical bundle block adjustments, only a small number of well distributed tie points are required for each image. If only DEM observations are used, many well distributed tie points are required. These tie points should capture the main relief in the scene to add horizontal constraints to the adjustment. This requires the use of computationally efficient estimation algorithms, as millions of unknowns need to be estimated for larger blocks.

As the SIFT based multi ray tie point typically still contain some outliers in epipolar direction, we use robust estimation using the Huber m-estimator (Huber, 1981). The adjustment program is based on the g2o graph optimization framework (Kümmerle 2011), and can process large blocks with thousands of images and millions of tie points.

4. EVALUATION

The matching and bundle adjustment algorithm is evaluated on a large Cartosat-1 mosaic, consisting of 1210 images, and covering complex areas, including highly mountainous areas as well as flat agricultural areas with changing land use.

4.1 Description of the North Italy test area

The evaluation is performed using a block of 605 Cartosat-1 stereo pairs covering the northern part of Italy. Each stereo pair

covers an area of approximately 28 by 30 km. An overview of the test area and the locations of the scenes are shown in Fig. 1.

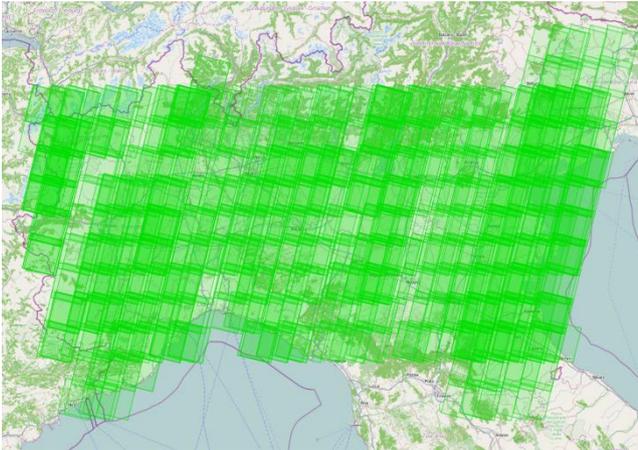


Figure 1. Overview of the Cartosat-1 North Italy block. Darker colors indicate a higher overlap of scenes.

Multiple stereo pairs are available for many areas, indicated by the dark green colour. The scenes are mostly cloud free, but some scenes contain a cloud cover of more than 20%. They were acquired in the years 2008-2012, and cover all seasons. This complicates tie point matching, as only a few points can be found between winter scenes with snow cover and summer scenes without snow cover. This caused matching failures for some scenes from overlapping strips, especially if they primarily contain agricultural fields and little man made structure. These areas are mostly located in the central part of the block.

The SRTM-C Band DSM, version 2, unfilled and edited by NGA is used as primary reference. The 5m Euro-Maps 2D (Euromap 2012) mosaic and SRTM-C Band DSM were used to extract checkpoints. Local least square matching has been used to extract 96472 checkpoints from the Euro-Maps 2D mosaic. The Euro-Maps 2D mosaic has a nominal horizontal accuracy of 15 m CE90. The horizontal accuracy of SRTM over Europe is 10 m CE90.

4.2 Tie point generation

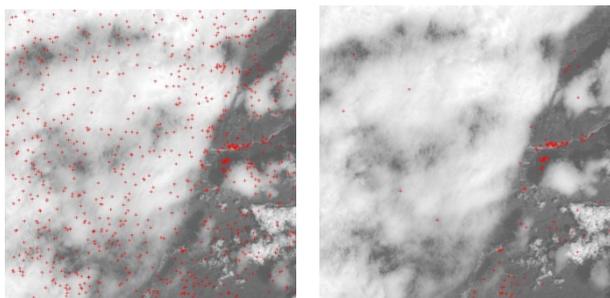


Figure 2. Pairwise SIFT matching and outlier removal. Red crosses indicate matched points. Left image: matched points in a single stereo pair. Note that many points on the clouds were matched. Right image: Remaining points after RANSAC based outlier removal. Almost all mismatches in the clouds have been removed.

The coarse SIFT matching described in Sec. 2.2 resulted in 6,838,585 multi ray SIFT tie points, with an average number of 2.5 rays per tie point. After thinning 659,977 SIFT tie points, with an average number of 3.8 rays per point remained. The block was fully connected with tie points.

Figure 2 provides an example of the tie point locations by pairwise SIFT matching before and after RANSAC based outlier removal. Note that this is an extreme case, and SIFT matching does produce much better results for images without cloud cover. We have chosen this image to show the performance of our pairwise outlier removal algorithm. For automatic processing, it is important that the processor can properly handle scenes with large cloud or water coverage.

The multi ray SIFT points are used for the initial bundle block adjustment, as described in Sec. 2.3. During the adjustment, 10% of the points were removed as outliers. A RPC bias correction is estimated for all scenes. The tie point RMSE was between 0.8 and 3.4 pixels for each scene. Figure 3 shows the multi ray tie points for image shown in Figure 2 before and after the block adjustment. It can be seen that the remaining blunders in the clouds have been detected by the block adjustment and were removed.

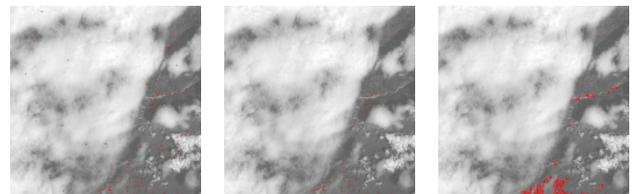


Figure 3. Left: SIFT points in after multi ray points generation. Middle: points after coarse bundle block adjustment. Right: Tie points after refinement with local least squares matching.

After coarse orientation, the fine matching procedure described in Sec. 2.4 was applied to the SIFT tie points and additional points matched between the Aft and Fore images of the stereo pair using a standard pyramidal area based matching using cross correlation. The resulting points are shown in Fig. 3. The multi-ray LSM tie points were thinned using a 10x10 grid, resulting in 103451 tie points with an average number of 7 rays per point. The high number of points is required when using the DEM based geo referencing, Eq. 5.

4.3 Image Orientation

In many projects, highly accurate GCPs are not easily available. It is thus interesting to evaluate whether the freely available SRTM DSM can be used as sole or major ground control. We have performed multiple image orientation experiments:

- Individual orientation of each stereo pair, using SRTM as only ground control
- Bundle block adjustment using tiepoints and GCPs placed in scenes at the corners of the block.
- Bundle block adjustment using tiepoints and SRTM DSM.

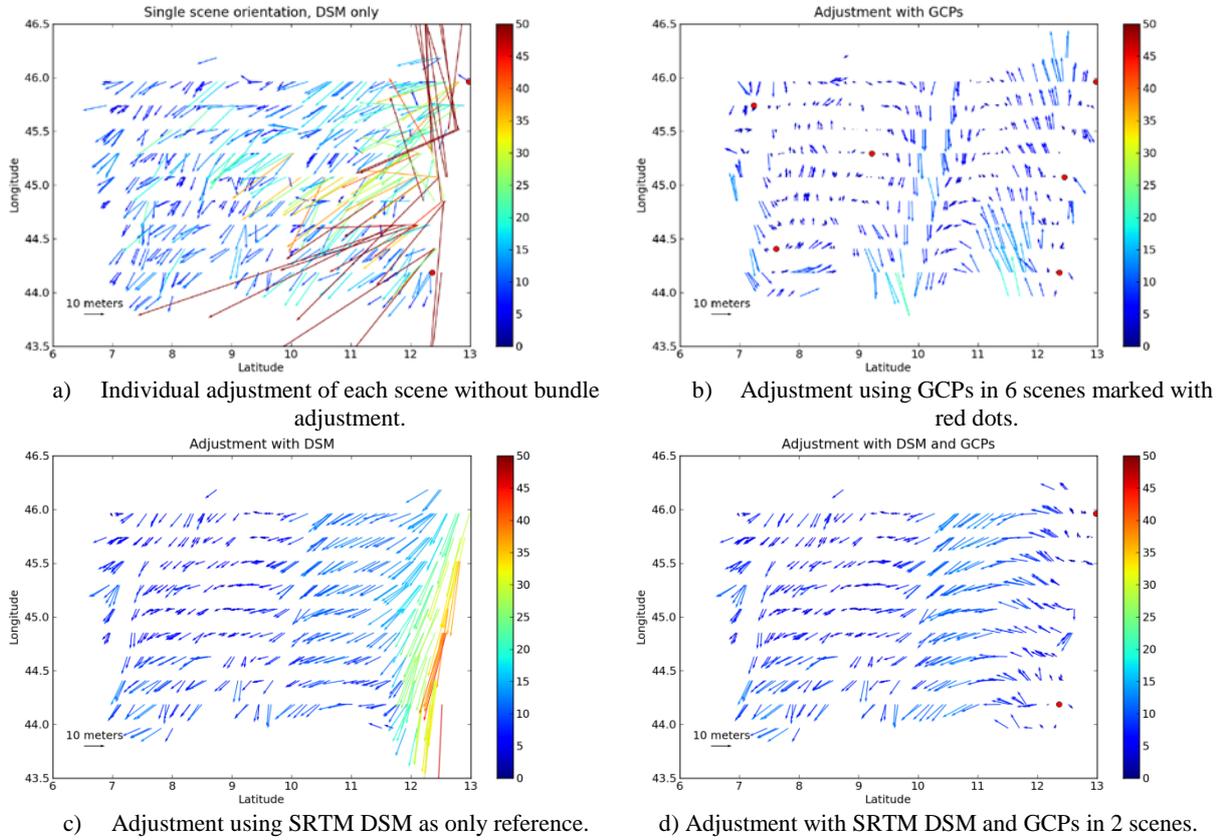


Figure 4. Checkpoint residuals after adjustment. Each arrow indicates the median checkpoint reprojection error in m of a single scene.

d) Bundle block adjustment using tiepoints, SRTM and GCPs in extended, flat areas.

The resulting median checkpoint reprojection errors of each scene after the adjustment are shown in Fig. 4. Corresponding statistics are shown in Table. 1.

Adjustment type	RMSE in m	Mean in m	Std in m	Max in m
a) Single, DSM	36.8	19.5	48	610.0
b) BA GCP	4.7	5.5	3.7	22.0
c) BA DSM	9.3	10.9	7.3	46.2
d) BA GCP+DSM	5.65	7.5	2.8	13.9

Table 1. Statistics on checkpoint errors for the whole block.

Case a) simulates independent processing of each stereo pair. Scenes in the mountainous and hilly areas are corrected well by the DSM constraint, whereas scenes over the flat coastal areas in the east show large errors, up to several hundred meters, cf. Table. 1. In this area, the flat DSM does not provide any horizontal constraints, allowing the scenes to drift in an uncontrolled way. Additionally, neighbouring scenes do show larger discrepancies, leading to coregistration errors between overlapping areas.

Case b) shows the result when GCPs in 6 scenes at the corners and center of the block are used, and a full block adjustment is performed. Here, the errors are better distributed across the

block, but there still exists a systematic, wavy pattern. Some scenes still show a high location error, a maximum error of 22m. This could probably be fixed by using more GCPs. However, additional GCPs are often not available, or hard to acquire for extended areas.

Case c) shows the result of bundle adjustment with SRTM as only reference. Compared to case a), no large jumps between adjacent scenes are visible, indicating a good coregistration. Except for the flat, coastal area in the east, the errors seem reasonable. As expected, the DSM constraint does not constrain the horizontal position in flat areas, resulting in large shifts, but the effect is less pronounced as in case a). A systematic shift between the bundle block and the checkpoints is visible, its -7 m across track and -5 m along track. As the reference datasets are specified with a CE90 of 10 m and 15 m respectively, this systematic shift can be explained by the location difference of SRTM and the Euro-Maps 2D mosaic used for automatic checkpoint extraction.

To reduce the errors in block, we used GCPs from the Euro-Maps 2D mosaic in 2 stereo pairs on the flat side of the block, as shown in case d). The GCPs stabilize the horizontal location of the critical coastal area. Fig. 4 d) shows that the systematic errors are strongly reduced. Especially, scenes on the flat coastal area are now properly stabilized. Most shifts are now below 10m. The RMSE is 5.7 m, and the mean across and along track shifts are -6.6 m and -3.4 m. Note that case b) provides a better RMSE value. This is because both GCPs and checkpoints have been extracted from the Euro-Maps 2D mosaic. The results for case a), c) and d) are additionally influenced by the shift between the SRTM DSM and the Euro-Maps 2D mosaic,

leading to a larger RMSE value. However, the results of d) seem to be more consistent, as visible in Fig. 4 d) and the lower standard deviation in Table 1. Additionally, it has the lowest maximum error of 13.9 m. This shows that it is feasible to orient large stereo blocks based on the DSM constraint. Extended flat areas can be stabilized with GCPs in a few scenes.

4.4 DSM generation

Each stereo pair is matched using Semi Global Matching (SGM) [d'Angelo 2010], resulting in a dense disparity map. After bundle adjustment, a DSM is generated for each stereo pair by forward intersection of the disparity map and reprojection into the target coordinate system. These pairwise DSMs are merged by taking the median at each output grid cell. The resulting DSM still contains holes in occluded areas and regions where the matching failed or outliers were removed. As most applications require a DSM without holes, a hole filling procedure is required. Small holes in the DSM are filled using inverse distance weighted interpolation; while large holes, often caused by clouds or large water bodies, are filled with SRTM data using the delta surface fill method by (Grohman et al., 2006). Fig. 6 provides an overview and details of the produced DSM.

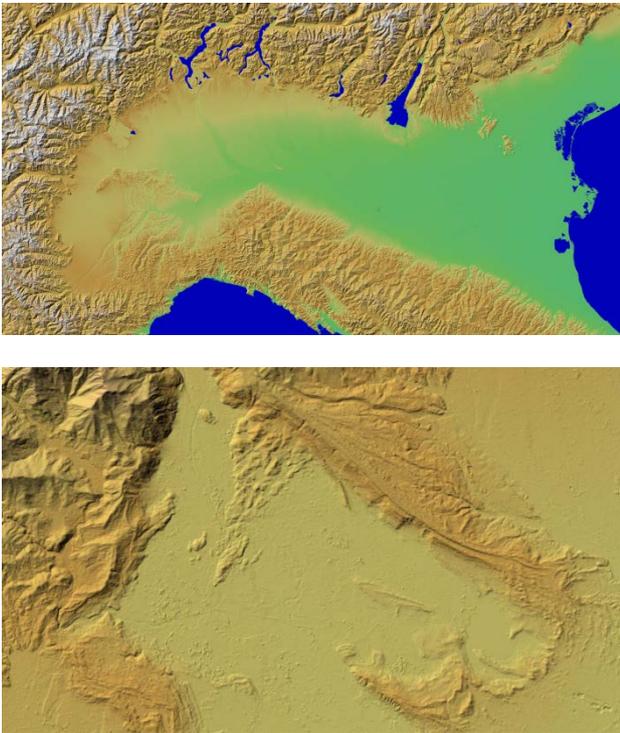


Figure 5. Upper: Full North Italy DSM. Lower: DSM detail

5. CONCLUSIONS

The main contribution of this paper is the presentation of a robust workflow for automatic orientation of satellite imagery. This includes automatic tie point measurement that can deal with large radiometric and geometric changes in images and the resulting high number of outliers during image matching. A fully automated DSM based adjustment procedure is used to minimize or completely avoid the use of GCPs derived from reference imagery or GPS measurements. This is important for state or continent wide DSM generation, where highly accurate GCPs are hard to acquire. The procedure is part of the Catena

processing system (Krauss et al., 2013) and is used by Euromap GmbH for wide area DSM generation from Cartosat-1 imagery (Euromap GmbH, 2013).

6. ACKNOWLEDGMENTS

The authors would like to thank Euromap GmbH for providing the Cartosat-1 stereo scenes and the corresponding Euro-Maps 2D mosaic.

7. REFERENCES

- Ackermann, F., 1983 High Precision Digital Image Correlation. In: *Proceeding of 39th Photogrammetry Week*, University of Stuttgart, pp. 231-243
- Euromaps GmbH, 2009. Euro-maps 2D. http://www.euromap.de/products/prod_005.html.
- Euromaps GmbH, 2012. Euro-maps 3D. http://www.euromap.de/products/prod_008.html.
- Fishler, M.A. and Boles, R.C. 1981 Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. Assoc. Comp. Mach.*, Vol 24, No 6, pp. 381-395
- Grodecki, J. and Dial, G., 2003. Block adjustment of high resolution satellite images described by rational functions. *Photogrammetric Engineering and Remote Sensing* 69(1), pp. 59–70.
- Grohman G, Kroenung G and Strebeck J, 2006, Filling SRTM voids: The delta surface fill method. *Photogrammetric Engineering and Remote Sensing* 72 (3), pp. 213-216.
- Huber, P.J., 1981. *Robust Statistics*. John Wiley & Sons, New York
- Krauss, T., d'Angelo, P., Schneider, M., Gstaiger, V., 2013. The Fully Automatic Optical Processing System CATENA at DLR. In: *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-1/W1, pp 177-181. ISPRS Hannover Workshop
- Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K. and Burgard, W., 2011. g2o: A General Framework for Graph Optimization. In *IEEE International Conference on Robotics and Automation (ICRA)*
- Lowe, D.G., 2004 Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60, 2, pp. 91-110.
- Strunz, G., 1993. *Bildorientierung und Objektrekonstruktion mit Punkten, Linien und Flächen*. Deutsche Geodätische Kommission. Vol. C Number 408.
- Utenthaler, A., d'Angelo, P., Reinartz, P., Hass, T., Carl, S. and Barner, F., 2011. A Concept for a Standardized DSM Product Automatically Derived from IRS-P5 Cartosat-1. In: *Geospatial World Forum*.
- Wiechert, A., Gruber, M., and Karner, K., 2012. ULTRAMAP: THE ALL IN ONE PHOTOGAMMETRIC SOLUTION, In

Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XXXIX-B3, pp. 183-186