# Surface Mapping Using Image Triplets: Case Studies and Benefit Assessment in Comparison to Stereo Image Processing

Hannes Raggam

## Abstract

Mapping of the Earth's surface from satellite images is a continuing application in remote sensing, which has been distinctly pushed with the increasing availability of very high-resolution image data. In most of the missions, stereo images can be acquired making 3D mapping from space very attractive, as high accuracy and level of detail become feasible from such data. In this paper, several case studies are presented, which have been applied to stereo pairs acquired by the SPOT sensors as well as the very highresolution Ikonos sensor. As an extension to standard stereo mapping, a 3D mapping approach which makes use of multiple image coverage is presented. For this work, image triplets including multi-sensor data sets have been used to extract 3D topography information and to generate digital surface models. The benefit of this extended approach is comparatively assessed with respect to results achieved from standard stereo mapping.

## Introduction

Driven by the image acquisition capabilities of many of the past and present satellite missions, 3D mapping of the Earth's surface using stereo image data is an ongoing subject of research and applications in remote sensing. The availability of digital stereo data was particularly pushed in the mid-eighties by the first SPOT mission and has meanwhile reached a high level, as stereo coverage can be acquired from space at very high-resolution in the meter and sub-meter range, as provided at present by the Ikonos and the Quickbird sensors. In contrast to the former SPOT stereo data collection from repetitive over-flights and at oblique viewing directions (cross-track stereo), optimized acquisition scenarios are meanwhile used where the data are acquired during a single over-flight at almost the same time through forward and backward viewing with one or two instruments (in-track stereo).

At the Institute of Digital Image Processing, stereo mapping using satellite image pairs has a long tradition. First results were generated from analogue optical SPOT, as well as SAR image pairs through mapping on an analytical plotter (Raggam *et al.*, 1989). Based on this work, the algorithmic baseline to extract 3D information from multi-sensor stereo data was established with respect to digital stereo data utilization (Raggam *et al.*, 1990). Further on, the feasibility to combine image pairs acquired from different instruments was monitored in order to extract 3D information in case that stereo data acquired from the same instrument are not available (Raggam *et al.*, 1996).

In either case, the accuracies which are feasible from stereo mapping are driven by the base-to-height ratio (or the stereo intersection angle) which should be as large as reasonable from a geometric point of view, and by the radiometric image similarity from a radiometric point of view, which is ensured only in case of small base-to-height ratios and in case the images are acquired at more or less the same time and from the same sensor. In order to ease these restrictions, a strategy was taken into consideration, which includes additional images into the traditional stereo mapping approach. These have to fulfil respective requirements in order to be of some benefit, i.e., they should be either geometrically or radiometrically more similar to one or both of the stereo images, than these are with respect to each other. This approach was stimulated by the concept of three-line scanning systems, as realized by the former MOMS instrument (http://www.nz.dlr.de/moms2p/), the future Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM) of the ALOS mission (http://www.eorc.jaxa.jp/ALOS/ about/2prism.htm), or indirectly by the present SPOT 5 system, which can provide like data by acquisition from the HRS as well as the HRG instruments (http://spot5.cnes.fr/gb/ satellite/satellite.htm). Thus, forward, backward, and nadir view of an area are taken almost at the same time.

3D stereo mapping nowadays is an issue for many investigators and in many applications. The aim of this paper is not to give a state-of-the-art of the voluminous work being done in this context, but rather to present recent investigations at the Institute of Digital Image Processing with respect to 3D mapping from space-borne image data sets. Based on the stereo mapping tools yet available, a multi-image based approach was implemented in a generic and flexible manner, allowing to utilize an arbitrary number of images for 3D mapping, which moreover can be acquired from different sensors, provided that a sufficient similarity is inherent to such a multi-image data set. In this paper, a description of this approach is given, and application examples carried out in different case studies are presented. Image triplets were used for 3D mapping investigations for the following test data sets:

• a SPOT 5 test data set acquired over the city and the surroundings of Barcelona, comprising forward and backward

Photogrammetric Engineering & Remote Sensing Vol. 72, No. 5, May 2006, pp. 551–563.

0099-1112/06/7205–0551/\$3.00/0 © 2006 American Society for Photogrammetry and Remote Sensing

Joanneum Research, Institute of Digital Image Processing, Wastiangasse 6, A-8010 Graz, Austria (hannes.raggam@joanneum.at).

acquisitions of the HRS instrument, as well as a nadir supermode HRG image;

- a multi-temporal SPOT 1/2/4 image triplet acquired over a mountainous area in the Austrian Alps; and
- two very high-resolution multi-sensor data sets comprising an Ikonos standard stereo pair and in addition a panchromatic Quickbird image.

Using these test data sets, a comprehensive benefit assessment of the multi-image based 3D mapping approach was made in comparison to the standard stereo mapping procedure. In particular, the impact of the multi-image based mapping strategy onto 3D mapping accuracy, matching performance and quality of extracted surface models was comparatively investigated. The performance of the individual processing steps is discussed in comparison to mapping from common stereo pairs.

# **3D Mapping Approach Based on Multiple Images**

Motivated by occasional weaknesses inherent to 3D mapping from stereo images, the stereo mapping procedure being implemented in the Remote Sensing Software package Graz developed at the Institute (RSG; Joanneum Research, 2003) was extended in order to cope with additional, i.e., more than only two images. This extended concept is illustrated in Figure 1, showing the key workflow in order to generate digital surface models from multi-image data sets. The scheme in this figure is based on the assumption of three input images and can be adequately extended to any other number of images to be used, while standard stereo processing is just covered by the left or right processing path being indicated.

A description of the key processing steps as implemented in the RSG software and as necessary to extract 3D surface models is given in the following sections. These are sensor modeling (not included in the workflow), coarse registration of images (optional), image matching, point intersection, and resampling of regular surface model raster.



Figure 1. Work flow of 3D mapping using multiple image data sets.

## **Sensor Modeling**

Sensor modeling refers to the establishment of respective parametric models which enable transformations between image and ground, and the optional optimisation of the sensor models in case they are not accurate enough. This processing step is not included in the workflow shown in Figure 1. For recent satellite missions, rather precise information is provided in order to geo-locate the respective image, i.e., to relate an image point to ground and vice versa. This can be either detailed information on exterior orientation elements like in case of the SPOT sensors, or just by providing so-called rational polynomial coefficients, which is the customary scenario for high resolution Ikonos or Quickbird images.

As it has turned out until now, there is still some accuracy improvement potential in either case. For parametric sensor models such improvement is realized by least squares parameter adjustment procedures, which are equivalent to photogrammetric bundle adjustment and can be applied to either single images, stereo pairs, or multiple image data sets (image blocks). The adjustment relies on the availability of ground control points. Moreover, tie-points can be utilized in order to induce additional constraints. For the Ikonos and Quickbird images which rely on rational polynomial coefficients, an optimised relationship between image and ground is established by means of linear correction (add-on) polynomials for line and column. These are supplied to the given polynomial transformation according to Dial and Grodecki (2002), which then are used to adequately cover the accuracy potential inherent to the Ikonos and Quickbird images.

For multiple images an estimate on the *a priori* 3D mapping accuracy can be determined through a spatial point intersection of control point measurements collected from (a minimum of) two or more images. The point intersection increases in its robustness with increasing point measurements. It is based on a least squares solution with respect to the best fitting ground coordinates. 3D point residuals in East, North, and height may be determined through comparison with given control point coordinates and serve as an indicator for the 3D geo-location accuracy being feasible from the used image data set.

## **Coarse Image Registration**

In general, the images to be used for 3D mapping are displaced with respect to each other, i.e., affected by shift, rotation and occasionally (e.g., in case of multi-sensor images) at different scale. In order to increase the performance of image matching, the images to be processed are in general coarsely registered. Therefore, one image is defined as the reference, while all the others are coarsely transformed with respect to the geometric properties of this reference image. The registration is based on a limited number of tie-points and a linear transformation derived there from. It has to be easily invertible, as the matching results achieved for the registered image have to be related to the original image in the subsequent processing again.

## Image Matching

A key processing step of the 3D mapping procedure is image matching, which in the multi-image scenario, is applied to multiple stereo pairs constituted by the reference image and all the images which have been coarsely registered with respect to it. For matching of the stereo pairs an extended version of the "feature vector" matching method developed at the Institute was used (Paar *et al.*, 1992; Caballo-Perucha, 2004). The components of the "feature vectors" are, in general, represented by various convolution and variance filters or other suchlike features. One essential feature is the cross correlation coefficient, which formerly used to be applied for image matching as such. Hence, cross correlation implicitly is included in this matching approach.

In order to receive feedback on the reliability of matching results, back-matching is used where the stereo pair is switched, and the result of forward matching is used to find the corresponding point in the reference image again. The spatial distance between predefined pixel location and backward matching result (back-matching distance) gives a confidence measure for the matching result.

## Point Intersection and DSM Generation

The matching results achieved for a single or for multiple stereo pairs subsequently are used to determine the corresponding point on ground. This is based on a least squares point intersection procedure as previously described, which requires at least one matching result representing the minimum of two projection lines in space. Multiple matching results define multiple projection lines in space and provide extended options for the determination of the corresponding ground coordinates in comparison to the case of a single stereo pair. In this case, unreliable matching results, which presumably will lead to wrong ground coordinates can be identified under consideration of the back-matching distance and point residuals resulting due to the increased over-determination of the least squares procedure. The unreliable matching results can then be excluded from the intersection, which can still be executed as long as two projection lines (i.e., one acceptable matching result) can be retained. Otherwise, the respective point is completely rejected.

In general, the point intersection becomes more robust in the multi-image case (provided reasonable multiple matching results are available) and the results of the intersection usage are more precise than in the standard stereo mapping procedure based only on two images. The point intersection leads to an irregular raster of points on ground, known by their East, North, and height coordinates. This raster has to be resampled to a regular raster surface model. Gaps which may still be present in this raster surface model due to unreliable point rejection are interpolated using a versatile interpolation mechanism.

# Case Study 1: SPOT 5 Image Triplet

On the SPOT 5 satellite, stereo data can be acquired simultaneously from the high resolution stereoscopic (HRS) instrument, which comprises two cameras looking forward and backward, respectively, at an off-nadir angle of  $\pm 20$  degrees. The acrosstrack pixel resolution is 10 meters, while an enhanced alongtrack pixel resolution of 5 meters shall assure increased accuracy with respect to 3D data extraction. Independently, a high-resolution geometric (HRG) image may be acquired in a nadir view with a pixel size of 2.5 meters in the "supermode" (see http://spot5.cnes.fr/gb/satellite/satellite.htm).

In this regard, an HRS study team was installed and specific test sites designated in order to validate the geometric performance of SPOT 5 HRS stereo data. The Institute in its role as a member of the HRS study team has elaborated respective results for a test data set over the city of Barcelona, Spain. These results have been basically presented at the 20<sup>th</sup> ISPRS Congress (Istanbul, July 2004), but not published in the conference proceedings (Raggam, 2004). In this section, a revised and extended version of this co-investigator work is summarized. Among the HRS study team, similar work utilizing the HRS image pair, and the HRG image was carried out by Müller *et al.* (2004).

For the investigation, study areas showing different land-cover and morphology have been investigated:

- 1. Rural/hilly area, being partly covered by forests (elevation range 155 to 490 meters);
- 2. Mountainous terrain (elevation range 76 to 1,206 meters); and
- 3. Urban area, represented by the city of Barcelona (elevation range 0 to 212 meters).

The forward and backward HRS, as well as the HRG image are shown in Figure 2, along with an indication concerning the location of the detailed test areas. With respect to the multi-image 3D mapping approach, the HRS scenes were registered to the HRG image, i.e., over-sampled to 2.5 meters. The basic input images, as well as these registered products, are shown in Figure 3 for the selected test sites. As reference data sets, a digital elevation model with a mesh width of 15 meters as well as orthophoto mosaics with a pixel size of 0.5 meters were available.

A comprehensive set of control points was measured from the orthophoto mosaics in order to be used for the



Figure 2. SPOT 5 image data: HRS forward (a), HRG (b), and backward (c). Detailed test areas are indicated in the HRG scene.



Figure 3. Detailed test areas showing rural (top row), mountainous (middle row), and urban (bottom row) terrain characteristics.

TABLE 1.	A Priori Sterec	MAPPING ACCURACY	OF SPOT 5 DATA SET
----------	-----------------	------------------	--------------------

116 GCPs	B/H	East	North	Height	Length
HRS fwd/bwd HRG/HRS fwd HRG/HRS bwd Image triplet	0.72 0.36 0.36	3.3 2.3 2.2 2.2	3.6 2.3 2.3 2.3	4.0 8.4 7.9 3.9	$6.3 \\ 9.0 \\ 8.5 \\ 5.1$

geometric modeling and accuracy assessment of this data set. For the stereo models which can be formed from the HRS image pair and the HRG image, as well as for an image block comprising the image triplet, the *a priori* 3D mapping accuracies were determined from these control points as previously described. The RMS values of the achieved 3D point residuals are summarized in Table 1.

For the HRS stereo model with a base-to-height ratio of 0.72 a planimetric, as well as a height accuracy of 4 meters, were achieved. For the multi-sensor models formed with the HRG image, however, the planimetric accuracy is slightly better due to the higher resolution of the HRG image, but the height accuracy in the order of some 8 meters is worse by a factor of 2. The image triplet combines the benefits of these individual stereo pairs and leads to an optimum planimetric, as well as height accuracy of about 2.5 and 4 meters, respectively.

Image matching as previously described was applied to the HRS image pair on the one hand, and to the HRG and

over-sampled HRS images in any combination on the other. The matching results were then used in the point intersection procedure, in which user-defined thresholds for the back-matching distance, as well as for a maximum allowed point residual serve to reject unreliable matching results. In general, reasonable settings for these thresholds have to be found empirically, based upon the characteristics of the image data. In homogeneous areas, for instance, the backmatching tolerance might be larger than in inhomogeneous and well-textured areas. For the SPOT 5 data sets a backmatching distance of three pixels and a maximum point residual tolerance of one pixel were used. The percentage of rejected points is also an indicator for the matching performance and the amount of matching failures, which may originate from major geometric differences (similarity too low) and from homogeneous areas, where discrimination of individual pixels is difficult if not impossible (similarity too high). The respective percentage rates which were achieved for the SPOT 5 image pairs are summarized in Table 2.

It must be noted, that for the multi-image based mapping approach the HRG image is the reference image, and the matching results achieved from the HRG/HRS forward and the HRG/HRS backward image pairs, respectively, are merged in order to provide a bundle of three projection lines. Table 2 shows, that some 20 or even more percent of unreliable matching results are inherent to these individual image pairs for either of the detailed test areas. When merging these matching results into an image triplet, this rejection rates

TABLE 2. PERCENTAGE OF REJECTED PIXELS FOR SPOT 5 IMAGE DATA

	F	Rejected Pixels (%	5)
Involved Images	Area 1	Area 2	Area 3
HRS fwd/bwd	4.67	12.86	18.03
HRS over-sampled	15.55	21.90	20.53
HRG/HRS fwd	22.27	25.02	26.88
HRG/HRS bwd	18.23	20.67	26.69
Image triplet	6.75	7.92	13.71

drop significantly down to 6.75, 7.92, and 13.71 percent for the three test areas. This is due to the oppositional effect of these two image pairs. Critical matching areas in the pair containing the forward image are less problematic in the pair containing the backward image and vice versa (e.g., back- versus fore-slopes). Frequently, at least one useful matching result is then achieved for such areas when these data are merged in the multi-image point intersection.

The table also shows the unreliability for the original as well as the over-sampled HRS image pair. Due to the larger base-to-height ratio and the increased geometric dissimilarity, most of these percentage rates are also significantly worse than in case of the image triplet, with the exception of Area 1, where the rejection rate for the original HRS pair is even smaller (4.67 percent) and the matching performance is fairly good due to the moderate terrain. However, it has to be considered that due to the coarser resolution less dense information still is available in this case. This is illustrated in Figure 4, which shows surface models where the rejected areas are not filled by interpolation ("raw" DSM). The extension of these areas is distinctly larger in case of the HRS stereo DSM than in case of the DSM resulting from the image triplet.

Such gaps use to be filled by proper interpolation methods in order to achieve full height coverage. The DSMs resulting in this way from the original HRS stereo pair, as well as from the image triplet, were further analysed with respect to the reference DEM. Statistical values of resulting height differences are summarized in Table 3, while illustrations of the key results are collected in Plate 1 for the three detailed test areas as follows:

- DSM derived from the image triplet;
- a stack of ortho images comprising HRS forward and backward as well as the HRG image, which have been generated using the DSM derived from the image triplet. The correspondence of these ortho images is an indicator for the quality of the DSM; and
- color-coded height difference models showing the differences between the DSMs derived from the HRS pair and from the image triplet, respectively, and the available reference elevation model.

The reference elevation model does not represent the surface, but is rather a ground model, where objects like trees or buildings are excluded. This becomes well visible in the height difference models shown in Plate 1. For the rural area, there is a clear correlation between forested areas and height differences larger than 5 meters, which basically reflect the heights of the trees (yellow and red areas in difference model versus dark areas in ortho image stack). A similar effect can be seen for the urban test area, where the height differences in a similar way correspond to the built-up areas and result in a distinct bias of 9.1 meters in the HRS stereo and 8.2 meters in the image triplet approach.

The DSMs derived from the image triplet are obviously superior to those derived from the HRS stereo pair. This is documented by the height difference models, where a



Table 3. Summary of Elevation Difference Statistics for Investigated Test Cases  $% \left( {{\Gamma _{\mathrm{S}}} \right)^{2}} \right)$ 

	Area 1	(rural)	Are (mount	ea 2 ainous)	Are (Ur	ea 3 ban)
	Mean	Std.D.	Mean	Std.D.	Mean	Std.D.
HRS fwd/bwd HRG/HRS fwd HRG/HRS bwd Image triplet	1.5 3.6 0.1 1.7	4.9 5.7 5.4 4.8	$4.4 \\ 5.0 \\ -0.7 \\ 1.8$	24.1 7.1 8.0 6.5	9.1 10.2 6.3 8.2	8.2 9.2 7.9 8.3

significant reduction of blunders can be seen, specifically for the mountainous test area, where the standard deviation of the height errors drops down from about 24 meters to 6.5 meters. By means of the urban area it is further shown, that a higher level of detail can be preserved through the over-sampling of the HRS images with respect to the HRG image. The road network of the city of Barcelona is much



more detailed in the DSM and the difference model resulting from the image triplet.

The ortho-image stacks serve as another quality analysis tool, as the individual images should fit acceptably over each other in case of a sufficiently accurate DSM. Wherever this is not the case, an indication on height errors of the underlying DSM is given. The amount of displacement correlates with the amount of height error. This is partly visible in the ortho image stack of the mountainous test area, where such displacements are visible in the upper left part, specifically at slopes exposed to north and being affected by shadow.

## Case Study 2: Multi-temporal SPOT Image Triplet

Another study was carried out for a multi-temporal triplet of SPOT images for an alpine Austrian area. These were acquired

in the summer season of the years 1987, 1992, and 1997 at incidence angles of  $-0.6^{\circ}$ ,  $-22.3^{\circ}$ , and  $20.3^{\circ}$ , and are addressed by "*spot87*", "*spot92*" and "*spot97*" in the following figure. Subsets of the images are shown in Figure 5 along with close-ups, which highlight that the data are affected by a significant temporal change of the agricultural areas over the years. Moreover, the scenes are partly affected by cloud and haze cover in the mountainous areas.

From a geometric point of view, the image pair composed from the left looking *spot92* and the right looking *spot97* image (resulting in a base-to-height ratio of 0.74) would be the best choice for 3D stereo mapping. However, the geometric differences in the mountainous areas are rather large and the *spot97* scene is severely covered by clouds. On the other hand, the similarity of e.g., *spot87* and *spot92* is significantly higher, but the geometric disposition is also worse by a factor of about 2 (B/H ratio of 0.38).



 TABLE 4.
 A Priori
 Stereo
 Mapping
 Accuracy
 for
 Multi-temporal

 SPOT
 Data
 SPOT
 Data
 SPOT
 <t

28 GCPs	B/H	East	North	Height	Length
spot87/spot92 spot87/spot97 spot92/spot97 Image Triplet	0.38 0.36 0.74	4.4 4.2 4.3 3.5	$4.1 \\ 4.4 \\ 4.7 \\ 4.2$	16.3 13.7 8.4 8.5	17.4 15.0 10.5 10.1

The 3D mapping potential was analyzed for each of the individual stereo pairs to be formed from this data set and for the image triplet as well. The *a priori* 3D mapping accuracies for these data sets have been determined from a set of 28 control points (measured in precise ortho photos). The RMS values resulting from the 3D control point residuals are summarized in Table 4 and confirm the above statements. While for the *spot92/spot97* stereo pair, a height accuracy of 8.4 meters is predicted, a value of 16.3 meters is achieved for the *spot87/spot92* image pair only. The height accuracy achieved for the image triplet is adequate to the one of the *spot92/spot97* stereo pair.

TABLE 5. PERCENTAGE OF REJECTED PIXELS FOR MULTI-TEMPORAL SPOT DATA

Involved Images	Rejected Pixels (%)
spot87/spot92	13.88
spot87/spot97	21.66
spot92/spot97	29.58
Image Triplet	8.76

The images *spot92* and *spot97* have been coarsely registered to the *spot87* nadir scene; matching was performed for each of the stereo pairs, and spatial point intersection was applied in order to determine ground coordinates from the matching results. Back-matching and point residual thresholds of 1.5 and 1 pixel, respectively, were used to reject unreliable/erroneous matching results. In the multiimage based approach the matching results achieved from the *spot87/spot92* and the *spot87/spot97* pairs were merged. The matching and hence the 3D mapping performance is documented in Table 5, which gives the percentage rates of rejected points for the various mapping scenarios. For the



Figure 5. Subsets of the SPOT images (top row) and close-ups (bottom row) illustrating radiometric differences due to temporal change.

left/right looking SPOT stereo pair *spot92/spot97*, almost 30 percent of the points were rejected due to geometric as well as radiometric dissimilarity. This on the one hand improves for the other stereo pairs, for which however a worse height accuracy was predicted on the other hand. Again, these problems are distinctly reduced in the image triplet scenario, where only some 9 percent of matching results are rejected, and where a high height accuracy was predicted as well. This is due to the utilization of three projection lines, which are achieved from the three images and which span large base-to-height ratios as well. In case that no (reliable) matching is achieved from one stereo pair (e.g., for the cloud-affected areas in the *spot87/spot97* pair), a result may still be achieved from the other pair.

"Raw" as well as final DSMs (without/with interpolation of rejected areas) were generated for these stereo and image triplet mapping scenarios. The accuracy of the final DSMs was further analyzed by means of height differences with respect to a reference elevation model. Plate 2 shows the "raw" DSMs, as well as the height difference models of the final DSMs as achieved from the *spot92/spot97* standard SPOT stereo disposition and from the image triplet, respectively. The mean and standard deviation values of the height differences are summarized in Table 6. The amount of point rejection, as well as its effect onto the quality of the final

TABLE 6. STATISTICS OF HEIGHT DIFFERENCES

Model	Mean	Std. Dev.
spot87/spot92 spot87/spot97 spot92/spot97 Image triplet	2.6 -3.7 -9.0 3.9	23.6 77.0 85.8 17.8

DSM, is clearly visible in the illustrations of Plate 2. Due to the lower rejection rate, the quality of the DSM generated from the image triplet is clearly superior.

According to the numerical values given in Table 6, the image triplet approach leads to the best correspondence with a standard height deviation of 17.8 meters. While the *spot87/spot92* stereo pair shows a slightly worse but still fairly good correspondence, the stereo pairs containing the cloud-affected *spot97* image are already worse by a factor of 4 to 5. Further, a bias of about 3 and 4 meters is achieved for the *spot87/spot92* stereo pair and the image triplet scenario. This presumably is due to vegetation and buildings, which are partly included in the DSMs (although not really precise) but are not included in the reference elevation model.

## **Case Study 3: Ikonos and Quickbird Image Triplet**

Recent activities at the Institute are related to the utilization of Ikonos stereo pairs and 3D mapping procedures in order to extract vegetation heights, in particular forest stand heights. This is a key objective in the European Union (EU) 5<sup>th</sup> Framework Project FIREGUARD (Monitoring Forests at the Management Unit Level for Forest Fire Prevention and Control) where the feasibility to use a remote sensing-based technology for the determination of tree heights, and subsequently timber volume and biomass, shall be demonstrated (Stelzl *et al.*, 2005).

In this context, problems inherent to such very highresolution image data are highlighted for two Mediterranean test sites of the FIREGUARD project, located in Greece and Portugal, respectively, and the option to overcome such problems using a multi-image based mapping approach is investigated. In particular, the effects of image dissimilarity and of local cloud cover (frequently a problem in remote sensing)



Figure 6. Registered subset of high-resolution data set for the Greece test site (top row) and the Portugal test site (bottom row) comprising Ikonos forward (a), Ikonos backward (b), and Quickbird panchromatic (c) images.

are treated. In addition to the standard Ikonos stereo acquisitions, which are acquired by a forward and a backward view of the Ikonos instrument, a panchromatic Quickbird image was acquired and could be included into the 3D mapping investigations. Subsets of these, already coarsely registered image data sets are shown in Figure 6 for both the Greece and the Portugal test sites. For image matching, the Ikonos forward image was selected as the reference scene for the Greece test site, while the panchromatic Quickbird image served as reference for the Portugal test site. The image acquisition data and parameters of these data sets are summarized in Table 7.

The Greece subset covers a very steep ravine, which presents itself extremely dissimilar in the Ikonos stereo pair from a geometric as well as the radiometric point of view. According to the image acquisition parameters, the panchromatic Quickbird image provides a compromise in between, and thus a higher similarity with respect to the Ikonos forward image, particularly along the ravine. For the Portugal subset, the Ikonos stereo images are affected by clouds as

TABLE 7. IKONOS AND QUICKBIRD IMAGE ACQUISITION PARAMETERS

Test Area	Image	Acquisition Date	Azimuth	Elevation
Greece	Ikonos forward	19 June 2003	61.96°	73.36°
	Ikonos backward	19 June 2003	$160.05^{\circ}$	66.69°
	Quickbird	05 June 2003	228.63°	89.20°
Portugal	Ikonos forward	04 July 2003	1.10°	67.33°
0	Ikonos backward	04 July 2003	226.11°	80.09°
	Quickbird	14 June 2003	307.87°	$87.09^{\circ}$

well as the shadow of the clouds, while the panchromatic Quickbird image is free of such disturbances.

3D mapping from such high-resolution images provides a high level of accuracy and detail. However, also the demand on ground control point quality is increased, as GCPs have to be precise at an adequate level. For these test data sets, for instance, control points have been acquired by GPS field measurements, and have been used to determine the linear add-on polynomials to be appended to the RPCs, leading to an "optimized" RPC mapping model. Also, image matching is more ambitious for such data. Verification of achieved results is more difficult, because reference information which is necessary for this purpose, e.g., reference height models, is frequently not available with the required accuracy and level of detail. Additionally, processing times increase severely due to the large data volumes associated to such image data. The processing of these subsets with respect to 3D mapping and extraction of surface models is described in the following sections.

#### **Greece Test Site**

The *a priori* mapping accuracies which have been determined for this test site for the various stereo models, as well as for the image triplet, are summarized in Table 8. Determination of the respective point residuals is based on the optimized RPC model. Height accuracies of about one meter are feasible from the Ikonos stereo data, which were collected at a base-to-height ratio of 0.53. A slightly better accuracy is predicted for the image triplet.

In the following, image matching was applied to the Ikonos stereo pair and the Ikonos forward and panchromatic

TABLE 8. A PRIORI STEREO MAPPING ACCURACY (GREECE TEST SITE)

32 GCPs	B/H	East	North	Height	Length
Ikonos fwd/bwd Ikonos fwd/Quickbird Ikonos bwd/Quickbird Image triplet	0.53 0.30 0.40	$0.4 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2$	0.3 0.2 0.2 0.2	0.9 1.3 0.9 0.8	$1.0 \\ 1.3 \\ 0.9 \\ 0.8$

Quickbird image pair. The results of the point intersection procedure in terms of the percentage of rejected points is given in Table 9 for these stereo pairs, as well as for the image triplet, which results from merging the matching results of the two stereo pairs. Here, a threshold of one pixel was used for the back-matching distance as well as for the maximum allowed point residual. While more than 20 percent of the points are rejected for both stereo pairs, this rate reduces to about 11 percent for the multi-image based mapping approach based on the image triplet. Graphically, this effect is shown in Plate 3, which illustrates the "raw" surface models as generated from these point intersection results, as well as final DSMs achieved after interpolation of rejected areas. Here, a red/green overlay of the two "raw" DSMs which were generated from the stereo pairs is shown. Areas, which have been rejected in one image pair (green areas), have been frequently successfully matched and retained in the other and vice versa (red areas). In the point intersection procedure for the image triplet an implicit merge of the stereo matching results is performed, and more or less only the black areas of this overlay remain rejected.



Plate 3. Surface models generated from Ikonos stereo pair and image triplet (Greece test site): (a) "Raw" Ikonos stereo DSM, (b) Final Ikonos stereo DSM, (c) Overlay of two stereo DSMs, and (d) Final image triplet DSM.

TABLE 9. PERCENTAGE OF REJECTED PIXELS FOR GREECE TEST SITE

Involved Images Priorited Divel	(0/)
Involveu inlages Rejecteu rixer	5 (%)
Ikonos fwd/bwd23.02Ikonos fwd/Quickbird20.87Image triplet11.08	

No detailed reference information was available to verify the results of the 3D mapping procedures. However, a visual comparison of the DSM extracted from the Ikonos stereo pair and the one generated by the multi-image mapping approach leads to the conclusion that the latter result is superior. It shows more detail and a more realistic shape of the ravine. This is also justified by ortho image stacks which show a better correspondence when using the image triplet-based DSM, although there is still a significant difference between the individual ortho-images anyway due to the different occlusion and shadow effects. Therefore, no illustration is included.

## **Portugal Test Site**

Using the collected GCPs and the optimized RPC model, an *a priori* 3D mapping accuracy of 1.4 meters in height was achieved for this test site, while the image triplet led to a predicted RMS height error of 1.3 meters (see Table 10). The results of the 3D mapping procedures applied to the Ikonos stereo pair as well as to the image triplet are shown in Plate 4 and Table 11, respectively, in an adequate manner



TABLE 10. A PRIORI STEREO MAPPING ACCURACY (PORTUGAL TEST SITE)

13 GCPs	B/H	East	North	Height	Length
Ikonos fwd/bwd Quickbird/Ikonos fwd Quickbird/Ikonos bwd Image triplet	0.53 0.37 0.17	0.7 0.6 0.7 0.6	0.9 0.7 0.7 0.8	1.4 1.4 3.7 1.3	1.8 1.7 3.9 1.6

TABLE 11. PERCENTAGE OF REJECTED PIXELS FOR PORTUGAL TEST S	ble 11.	<ol> <li>PERCENTAGE</li> </ol>	OF REJECTED	PIXELS FOR	Portugal	Test	SITE
---	---------	--------------------------------	-------------	------------	----------	------	------

Involved Images	Rejected Pixels (%)
Ikonos fwd/bwd	25.34
Ikonos over-sampled	47.60
Quickbird/Ikonos fwd	35.01
Quickbird/Ikonos bwd	23.60
Image triplet	12.60

to the results achieved for the Greece test site. Here, the panchromatic Quickbird image was used as the reference image for the multi-image based mapping approach.

As mentioned before, the sub-area which was investigated is covered by clouds. Therefore, the Ikonos stereo pair does not result in a useful DSM, because extended areas either cannot be matched or are rejected during point intersection. When using a threshold of one pixel for the back-matching distance, as well as for the maximum allowed point residual, the rate of rejected areas is about 25 percent for the original stereo pair and 47.60 percent for the over-sampled stereo pair. Interpolation of these areas then does not result in a reasonable surface model. Also, the individual stereo pairs utilizing the Quickbird image show very high rejection rates (35.01 and 23.60 percent). However, when merging the matching results in the multi-image based point intersection, the rejection rate again is distinctly reduced to 12.60 percent. This effect is also shown by the red/green overlay of the stereo-based DSMs, where a cloud-covered area in one DSM is not present in the other and vice versa (green versus red areas). Thus, only limited unreliable height values remain in the DSM derived by means of the multi-image mapping approach. Appropriate filter operations might be applied to further reduce these erroneous areas.

In this DSM also, the shape of the forested areas becomes clearly visible. This confirms the feasibility to extract vegetation heights from high-resolution stereo images with reasonable accuracy. Still, ongoing work is related to detailed accuracy assessment of the results achieved so far, using in-situ tree height reference measurements.

## **Conclusions and Further Work**

A 3D mapping approach which makes use of multiple images was implemented and comprehensively tested for different test scenarios and in comparison to the standard stereo mapping approach. The modeling of such multiple image data sets is based on an image block scenario equivalent to photogrammetric block adjustment. Image matching is applied to multiple stereo pairs which are based upon the selection and definition of a dedicated reference image. The multiple matching results are jointly used in a spatial point intersection procedure in order to determine the corresponding ground coordinates.

The anticipated benefits of this approach could be confirmed in the various tests which have been carried out. In particular, the following beneficiary features of this approach can be stated:

- Image matching is facilitated through the utilization of image pairs, which are more similar from a geometric point of view due to a smaller base-to-height ratio, than standard stereo image acquisitions might be.
- Despite matching of images with a comparably weak stereo disposition, geometric robustness may be retained, as the multiple matching results are merged in the point intersection procedure and the multiple projection lines implicitly use to span the larger base-to-height disposition inherent to a standard stereo acquisition.
- The multi-image based approach provides extended options for spatial point intersection, as the least squares solution is applied to multiple projection lines and hence benefits from a higher over-determination, which (besides the back-matching distance) allows the identification of matching failures and displaced projection lines resulting there from.

Ongoing work is related to the utilization of the multiimage based approach to a multiple airborne image data set acquired by the digital UltraCam- $D^{TM}$  camera in order to generate vegetation height models over forested areas and to determine the height of trees. These data are acquired with about 90 percent overlap and a resolution in the decimeter range. As matching in forested areas is ambitious due to the different displacement of trees in consecutive images, stereo pairs with small base-to-height ratio can be merged to multisensor data sets for this data acquisition scenario in order to benefit from the strengths of the multi-image based 3D mapping approach. First results of this work are very promising and will be published in Hirschmugl *et al.* (2005).

## Acknowledgments

The SPOT 5 image data set was provided by Spot Image<sup>TM</sup> within the HRS assessment program. The high-resolution Ikonos and Quickbird images for the Mediterranean test areas were acquired within the EU 5<sup>th</sup> Framework Project FIREGUARD. Valuable support and stimulating discussions have been contributed by K. Gutjahr and M. Franke from the Institute of Digital Image Processing. These inputs are thankfully acknowledged.

## References

- Dial, G., and J. Grodecki, 2002. Block adjustment with rational polynomial camera models, *Proceeding of ASCM-ASPRS Annual Conventions*, April, Washington, D.C.
- Hirschmugl, M., M. Ofner, M. Schardt, and H. Raggam, In review. Single tree detection in very high resolution remote sensing data, *Remote Sensing of Environment*.
- Joanneum Research, 2003. RSG in Erdas Imagine, Software Documentation, Version 2.1, RSG-Release 4.60.
- Müller, R., P. Reinartz, M. Lehner, and M. Schröder, 2004. Comparison of the accuracy of DEM from SPOT HRS two-fold stereo data and HRS/HRG three-fold stereo data in Barcelona test site, *International Archives of Photogrammetry and Remote Sensing*, XX<sup>th</sup> ISPRS Congress, WG I/2, Istanbul, 12–23 July, unpaginated DVD.
- Raggam, H., M.F. Buchroithner, and R. Mansberger, 1989. Relief Mapping Using Non-Photographic Spaceborne Imagery, *ISPRS Journal of Photogrammetry and Remote Sensing*, 44(1): 21–36.
- Raggam, H., and A. Almer, 1990. Mathematical Aspects of Multi-Sensor Stereo Mapping, *IEEE International Geoscience and Remote Sensing Symposium*, Vol. III, pp. 1963–1966.
- Raggam, H., 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.
- Raggam, H., 2004. Accuracy Analysis and Surface Mapping Using Spot 5 Stereo Data, International Archives of Photogrammetry

and Remote Sensing, XX<sup>th</sup> ISPRS Congress, WG I/2, Istanbul, 12–23 July, unpaginated DVD.

Stelzl, H., A. Wimmer, H. Raggam, O. Sacherer, A. Almer, and C. Banninger, 2004. Monitoring Mediterranean Wildlands with Very High Spatial Resolution Satellite Imagery for Fire Prevention and Control, *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Anchorage, 20–24 September, unpaginated CD-ROM.