'Photogrammetric Week '95', D. Fritsch & D. Hobbie, Eds., Wichmann Verlag, Heidelberg, 1995



# Automatic DTM Generation using MOMS-02/D2 Image Data

FRANZ SCHNEIDER AND MICHAEL HAHN, Stuttgart

**ABSTRACT**

The Digital Terrain Model (DTM) is one of the most elementary and important products of processing stereo image data. The 3D shape of the imaged scene is the basis for topographic mapping, orthophoto generation, geocoding multispectral image data, etc. An automatic procedure for generating DTM's inclusive orthophotos from three-line imagery is developed during the last years. The main features of the algorithm are feature based matching of points and edges extracted in all three channels, consistency checks in image and object space using the known orientation of the image strips, finite element modelling for surface representation and a coarse-to-fine directed processing strategy which controls the overall processing steps. In the paper this procedure is descriped in detail. Processing scenes of Andes orbit 115 and Australia orbit 75B of MOMS-02/D2 mission shows that the procedure is successful even in terrain as well as in low texture scenes. Matching the three panchromatic stereo channels is done fast and reliable. The experimentally found height accuracy of 3D points determined by error propagation is about 10 - 15 m. This accuracy level is confirmed for the Australia scene by independent check measurements using DGPS.

## 1. INTRODUCTION

The MOMS-02 camera has some very unique features for mapping the Earth's surface from space. It offers image data capture in different modes of one of the three categories: (a) along-track three-line stereo, (b) multispectral imaging with 4 channels and (c) combinations of stereo and multispectral imaging.

The stereo module of MOMS aims at high quality 3D reconstruction of the imaged scene. To achieve this goal prerequisites are a geometric calibration of the camera, the recording of image data with good contrast (enabled by properly chosen gain settings) and, what is most important, with high geometric resolution. If comparing with existing digital remote sensing systems the improved MOMS performance becomes obvious. The pixel sizes of Landsat Thematic Mapper and SPOT (panchromatic) are 30 m and 10 m, respectively. Significantly higher is the resolution of the panchromatic MOMS nadir channel with a pixel size of 4.5 m.

The other important property of stereo MOMS is the along-track stereo capability. This means, that in the stereo mode a scene is recorded line by line with the forward (FW), the nadir (N) and the backward (BW) looking channels quasi simultaneously (with time offset of 2 x 20 seconds). The common way to get stereo with existing Earth observation systems is off-track scanning, in which a scene is imaged again by the sensor with a time lag of days or even weeks.

The generation of DTM's is a main objective of the MOMS stereo concept (Ackermann et al., 1989). The philosophy of automatic mensuration processes is simple: just measure a lot of object points which densly cover the imaged terrain surface. Surface shape is reconstructed by finite element filtering. This process applies robust statistics and data reduction (of several object points to 1 DTM grid point) to obtain a DTM accuracy comparable to that received by selective (manual) DTM sampling procedures (Hahn, 1989). A first proposal is presented by Hahn and Schneider (1991). The process applies feature based matching in object space, uses the given orientation and relies on coarse-to-fine processing including image pyramids and multiresolution representation of the DTM. With some modifications the procedure is now implemented in software. The concept of the approach is discussed in the next section.

More details of the scientific goals of the MOMS stereo segment can be found in Ackermann (1993). First results on photogrammetric processing of MOMS-02/D2 data are reported in Fritsch (1995).

## 2. TRINOCULAR STEREO SOLUTION FOR DTM RECONSTRUCTION

During the MOMS-02/D2 mission 26. April - 6. Mai 1993 stereo image data have been recorded of selected regions by capturing strips of 37 km swath width. These strips can be very long, for example, hundreds or thousands of km. Therefore, a DTM reconstruction procedure dealing with those huge image data sets should operate without any need of human interaction. This was an ultimate goal in developing the MOMS DTM module.
An overal characterization of the matching strategy is plotted in figure 1.

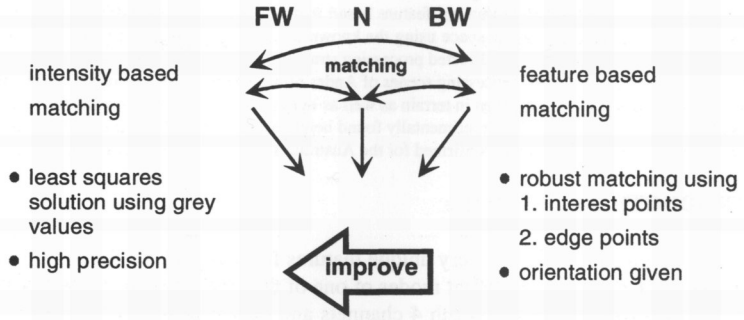


Figure 1: Trinocular stereo solution for DTM reconstruction

Considering matching of image pairs just indicates a condition of the three-line image approach: points transfered by matching the channels BW-to-N and N-to-FW have to be consistent with matching BW-to-FW. This condition can be applied in object space as well as in image space by using given orientation of the images.
The feature based module extracts interest points and strong edge points as shown in figure 2.

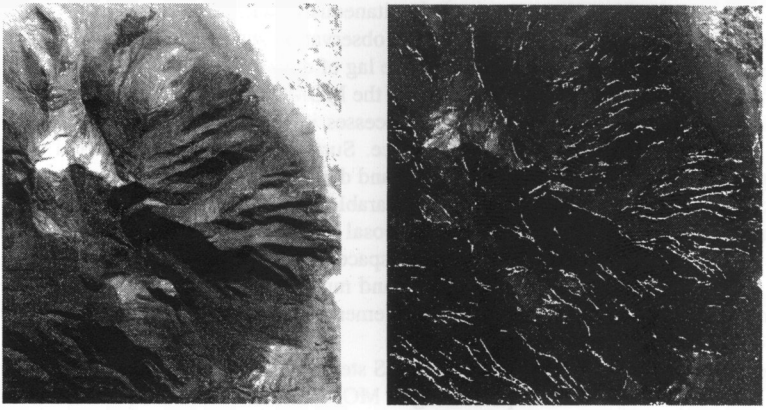

Figure 2: Section of the Andes scene (orbit 115).
The left image displays interest points, the right image interest points and edge points.

Algorithmically the feature based matching module consists of following steps:

start : Generate image pyramids for all channels and
start matching on a coarse resolution channel

proceed from coarse-to-fine:
extract interest points and edge points
match:
- search for corresponding features in BW - N - FW channels
(epipolar search windows and error check)
- transfer to object space, determine 3-D coordinates
(direct spatial intersection and error check)

surface modelling: recover surface shape
(finite element approximation)

goto next pyramid level

The geometric relationship between object space and image space of N, BW and FW channels used for determining 3D point coordinates is given by the following formulas (Ebner, Kornus, Ohlhof, 1992).

$$\underline{x} = k\,\underline{M}^T\underline{D}^T[\underline{X} - (\underline{X}_0 + \Delta\underline{X}_0)] \quad \text{with} \quad \Delta\underline{X}_0 = \underline{D}\,\Delta\underline{s} \tag{1}$$

**M**: rotation matrix, describes the relative latitude of the FW- and BW- looking channel to the nadir channel.

**D**: rotation matrix, describes the relation between object space and image space of the nadir looking channel

k: scale factor

**X:** 3D object coordinate triple

$\mathbf{X_0}$**:** coordinates of the projection center at time of point $t_i$

$\Delta$s: displacement of the FW- and Bw- looking channels to the nadir channel in the camera coordinate system

$\Delta X_0$ is a displacement of the projection center of forward (or backward) looking channels relative to the nadir channel. With $M^TD^T$ the spatial vector from a projection center to an object point is rotated. Finally scaling with factor *k* defines the image coordinates in the corresponding channel. With this equation the epipolar search window in image space is controlled as well as a direct formulation for spatial intersection of point triples is derived.
The second module in figure 1 is intensity or area based least squares matching. The matching process starts at the location of matched feature points with the intention of improving the matching precision or equally the height of 3D points. Furthermore, this module is also embedded in coarse-to-fine

processing. But this option is only applied if not a sufficient number of point triples is found by the feature based process.
In the experimental investigations scenes of Andes (orbit 115) and Australia (orbit 75B) are processed. For these experiments the high resolution nadir channel is resampled to the resolution of the backward and forward looking channels. Theoretical investigations into feature point matching exploiting the high resolution channel are reported in Schneider and Hahn (1992, 1994). Experiments show that a certain problem is the changing location of feature points: points may move in an arbitrary direction in the order of 3-5 m if image resolution changes from a pixel size of 4.5 m to 13.5 m.

## 3. EXPERIMENTAL INVESTIGATIONS

In the experiments we want to focus on two aspects. First is the reliability of DTM data capture and second is theoretical and empirical precision of the derived DTM's.

### 3.1 Reliability of DTM data capture

Reliability is adressed by considering area coverage and local point density of matched points. The processed scenes cover an area of 17.6 x 17.6 square km. The point distribution with respect to the pyramidal multigrid DTM representation is shown in figure 3. The number of successfully matched points within the individual grid cells varies to a certain extend which just reflects that local regions always differ with respect to the number of textural features suitable for matching. Nevertheless, only a very small number of grid cells includes just some few points. This is true for coarse as well as for fine resolution levels of the hierarchical process. In consequence, it should be guaranteed that the DTM estimates derived on each pyramid level have a homogeneous quality. This again helps to keep the search window sizes for finding corresponding points at the next refinement level relatively small. At the bottom level 0 of the pyramid the DTM grid width is reduced to 22 x 22 pixels with the intention of deriving a high resolution DTM in the final stage of the process. Consequently the local measurement redundancy (expressed by the number of object points per grid cell) is definitely smaller than at the levels above. Comparing the averaged point density between Andes scene (14.7 points per grid cell) and Australia scene (7.9 points per grid cell) shows a difference with a factor of 2. The reason for this discrepancy is a considerable difference in the contrast of both scenes with a factor 6. The contrast of the Andes imagery is quite good, but for the Australian imagery it is rather low (cf. table 1).

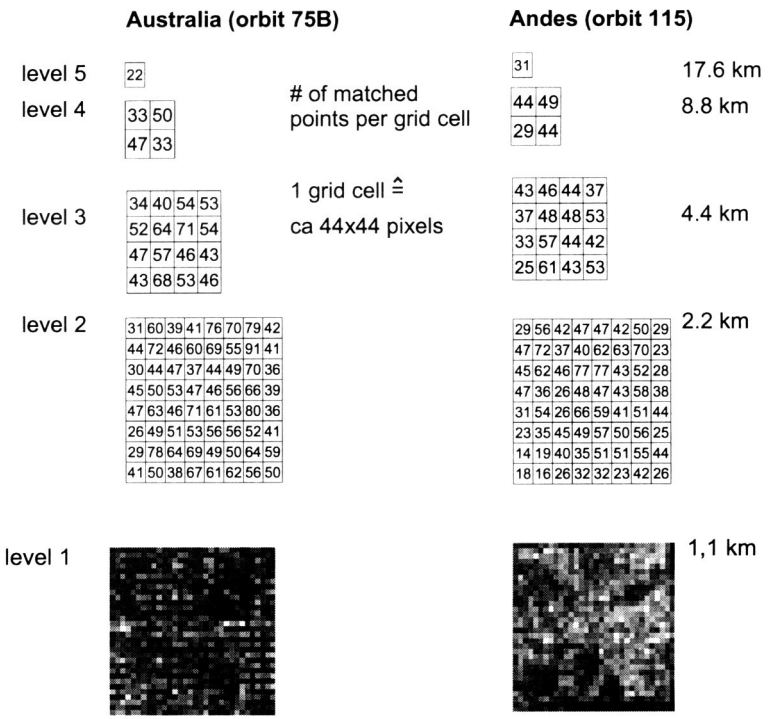


Figure 3: Point density in grid cells. For level 5 - 2 numerical values, for level1 an image displays the numer of matched points per grid cell (white is high, black is low density). the size of a grid cell in image space is about 44 x 44 pixels.

| | Australia (orbit 75B) | | Andes (orbit 115) | |
|---|---|---|---|---|
| channel | mean g | strength $\sigma_g$ | mean g | strength $\sigma_g$ |
| 5 (N) | 81 | 6.0 | 115 | 30.9 |
| 6 (FW) | 77 | 6.1 | 121 | 39.9 |
| 7 (BW) | 77 | 5.4 | 123 | 37.3 |

Table 1: Mean grey level and gray level strength in the Australia scene and the Andes scene

### 3.2 Accuracy

For assessing the accuracy of the photogrammetric processing chain (point transfer, aero triangulation, DTM reconstruction) check points have been measured for the Australia scene by Differential GPS (DGPS).
With respect to DTM accuracy the theoretical precision estimates of single 3D points found by errror propagation is considered using the aerotriangulation results provided by TU Munich. Secondly, the height precision of the DTM is assessed by calculating the rms difference between the DGPS measured heights of check points and the corresponding heights obtained by DTM interpolation.

| | tie points | | check points |
|---|---|---|---|
| measured by | intensity based matching | feature based matching | Operator, manually |
| $\sigma_0$ [pixels] | 0,34 | 0,78 | 0,5 |
| $\sigma_{XY}$, $\sigma_Z$ [m] | 4 - 6, 9 - 12 | 8 - 11, 19 - 22 | 5 - 7, 10 - 14 |

Table 2: Theoretical accuracy for tie points and check points

The theoretical precision for tie points and check points is listed in table 2. The sample size is about 6000 points measured by feature based matching. About 90% of them have been successfully matched by intensity based least squares matching. The 3D point precision of about 5 m in X, Y and 10 m in Z obtained by intensity based matching gives a first idea of what can be expected from a corresponding DTM. The measurement accuracy of the check points is not as good as expected a priori. Even though the check points have been measured thoroughly at a Digital Photogrammetric Workstation in stereo mode, due to identification problems a manual measurement accuracy of only 0.5 pixels has been obtained. The height differences between DGPS check point measurements and DTM are determined and an rms-value is calculated with a result of 15.7 m. The biggest difference observed was 35 m which indicates a quite reasonable DTM quality in this first ground truth experiment.

Coarse-to fine processing with respect to the DTM is illustrated in figure 4.

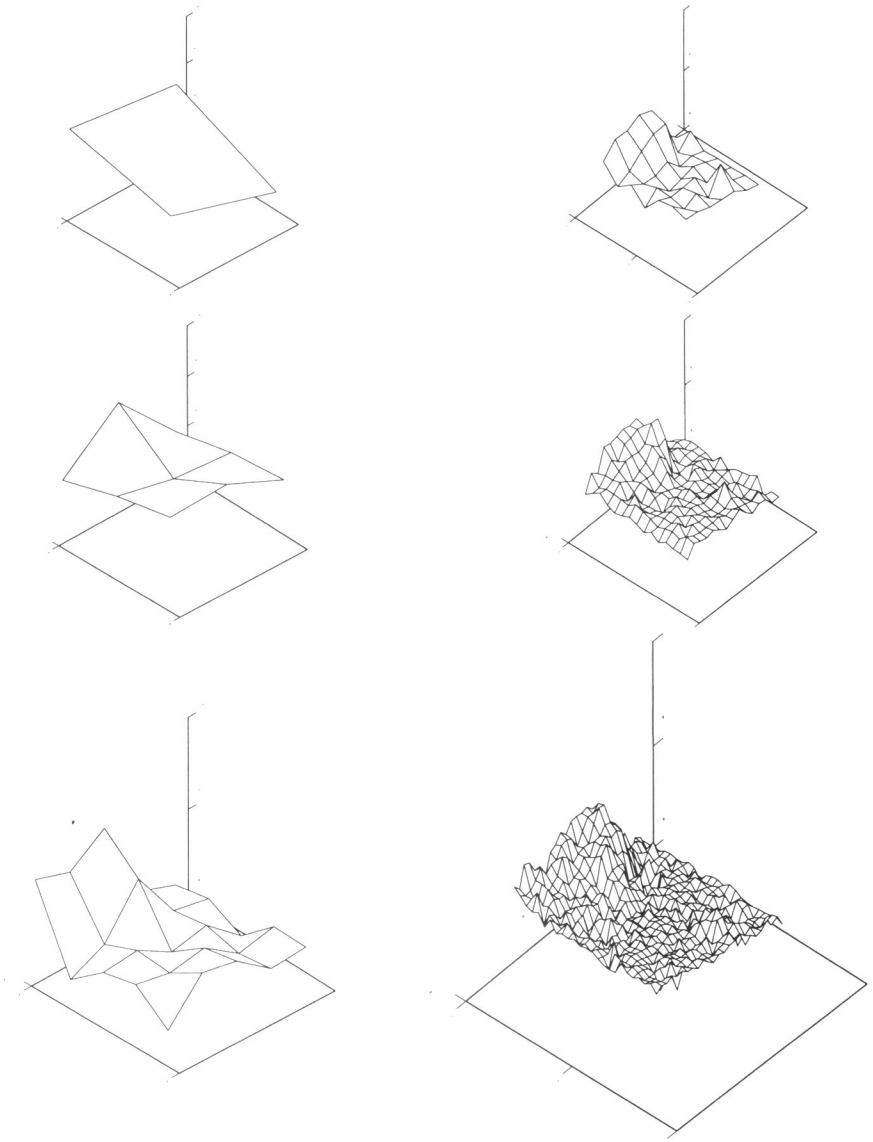

Figure 4: DTM pyramid for Andes scene. The size of the DTM is 17,6 km * 17,6 km. From level to level the grid width decreases linealy by a factor 2.

'Photogrammetric Week '95', D. Fritsch & D. Hobbie, Eds., Wichmann Verlag, Heidelberg, 1995



The finally reconstructed DTM is plotted in figure 5. In this figure a perspective view of the Andes DTM is generated using the orthoimage as an overlay.

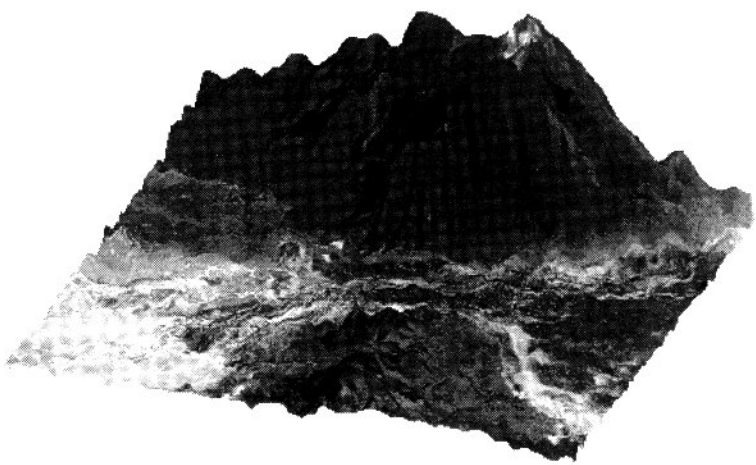

Figure 5: Perspective view of the Andes DTM (overlay: orthoimage)

## 4. FINAL COMMENTS AND OUTLOOK

In this paper the concept and realization of a DTM reconstruction module using MOMS stereo data is described. The experimental minvestigations with image scenes of Australia (orbit 75B) and Andes (orbit 115) focus on reliability and accuracy of DTM generation and show that the algorithm works quite well even in low texture imagery and in mountainous terrain. The DTM accuracy obtained for the Australia scene is in the range of 10 m (theoretical precision) to 15 m (rms value from check points). For the accuracy evaluation ground truth was provided by DGPS measurements of 65 check points.

At the present state of programming the algorithm has received a preoperational stage. Some aspects like the exclusion of problem areas (for example clouds) from matching is not taken into account up to now by the algorithm. Also the performance has to be improved further (cf. table 3).

| Used computer: SPARC 10/41 | | |
|---|---|---|
| Time needed for | project: Andes | project: Australia |
| ▸ matching 1 point: | 0.2 s | 0.3 s |
| ▸ matching & DTM generation 1 DTM grid point: | 3.6 s | 5.8 s |
| ▸ capturing a DTM of 17.6x17.6 km² (grid width: 275 m): | 6.45 h | 8.8 h |

The computing time on a Indigo 2, R 4400 (200 MHz) is 5 times shorter than on SPARC 10/41

Table 3: Performance considerations

## 5. REFERENCES

Ackermann, F., Bodechtel, J., Dorrer, E., Ebner, H., Kaufmann, H., Koch, B., Lanzl, F.,Seige, P., Winkenbach, H., Zilger, J. (1989): " MOMS-02/D2 - Wissenschaftsplan". 85 pages, DLR, Oberpfaffenhofen.

Ackermann, F. (1993): "Das MOMS-02-Stereosegment - Ein hochgenaues System der digitalen Photogrammetrie". Geo-Informations-Systeme, Wichmann Verlag, Vol. 6, No. 1, pp. 16-22.

Ebner, H., Kornus, W., Ohlhof, T. (1992): " Simulation Study on Point Determination for the MOMS-02/D2 Space Project using an Extended Functional Model". Int. Arch.PhRS (29) B4, IV, pp 458-464.

Fritsch D. (1995): "Ableitung digitaler Geländemodelle aus MOMS-02/D2 Bilddaten - erste Ergebnisse". Geo-Informations-Systeme, Wichmann Verlag, Vol. 8, No. 2, pp. 13-20.

Hahn, M., (1989): "Automatic Measurement of Digital Elevation Models with ImageMatching Techniques". Proceedings of the 42. Photogrammetric Week, pp. 141 - 151, Stuttgart.

(1991): "Feature Based Surface Reconstruction - A Hierarchical Approach Developed for MOMS-02 Imagery". Proceedings of the International Geosience and Remote Sensing Symposium, Vol. III, pp. 1747-1750, Helsinki.

Hahn, M., Schneider, F. (1994): "Point transfer in Moms Images of Different Geometric Resolution - A Theoretical Case Study". FIG XX. International Congress, Comm 3, 8 pages, Melbourne.

Schneider, F., Hahn, M. (1992): "Matching Images of Different Geometric Scale". International Archives of Photogrammetry and Remote Sensing, vol. 29, B3, pp. 295-302, Washington.

Schneider, F., Hahn, M. (1992): "Matching Images of Different Geometric Scale". International Archives of Photogrammetry and Remote Sensing, vol. 29, B3, pp. 295-302, Washington.