DEM GENERATION FROM VERY HIGH RESOLUTION STEREO DATA IN URBAN AREAS

Thomas Krauß[1], Peter Reinartz[1], Manfred Lehner[1], Manfred Schroeder[1], Uwe Stilla[2]
[1]German Aerospace Center (DLR), Remote Sensing Technology Institute
PO Box 1116, 82230 Wessling, Germany
thomas.krauss@dlr.de
[2]Photogrammetry and Remote Sensing, Technische Universitaet Muenchen
Arcisstrasse 21, 80333 Muenchen, Germany
stilla@bv.tum.de

KEY WORDS: VHR data, Stereo Data, DEM generation, image matching, object extraction

ABSTRACT:

In this article an attempt for improving an existing method creating DEMs fully automatically from stereoscopic image pairs is presented. Results applying this method to very high resolution (VHR) satellite imagery is shown. The herein discussed “column” algorithm is illustrated in brief and compared to a conventional algorithm frequently used in generating DEMs from satellite line scanner data. The presented algorithm is producing DEMs with an improved matching of steep objects as can be found in urban satellite imagery but introduces also some blunders in the generated DEMs.

1. INTRODUCTION

Actual very high resolution (VHR) satellite image data with resolution of about one meter offer the possibility to see many details in urban areas. Objects like houses, trees and cars can be interpreted easily with visual methods, while the automatic interpretation of this kind of space borne data is still a very busy field of research.

Especially, an automatic DEM generation from along track stereo pairs of VHR optical data from space, like IKONOS or Quick Bird images, offer new challenges for developing techniques for the automatic interpretation of homologous image structures and further for object extraction.

So, at present a primary target in city modelling is the quest for algorithms creating fully automatically good and reliable digital elevation models from relatively easy to acquire high resolution stereo satellite imagery [1, 2].

The image pairs acquired from two viewing directions of about 40 degrees difference in along track direction with a resolution of about one meter show a lot of variations in object representations. Certain features are only visible in one image, while they are hidden in the other.

Many attempts are made therefore in creating algorithms resembling the simple task of stereoscopic vision which is a easy walk through for human eyes. See for more details the article of Scharstein and Szeliski [3] and the very good associated web site.

Conventional techniques of image matching like the one described in [4] have to be modified to be able to reconstruct vertical walls, streets, trees and other objects, in order to obtain a surface/object model of a city which is close to reality. If it's possible to create automatically such high resolution digital elevation models (DEMs) the cost for good three dimesional representations of urban areas will drop dramatically and many new applications will arise.

In this article first steps of a new attempt for creating digital elevation models from VHR stereoscopic images is shown and discussed. The resulting DEMs from an urban scene created by a conventional algorithm [4] and the herein described new “column” algorithm are compared and discussed.

2. DESCRIPTION OF THE “COLUMN” METHOD

The presented method referred herein as “column” algorithm for extracting DEMs is based on epipolar geometry. Those images are e.g. corrected images from satellite scanner data as delivered from IKONOS or the test data sets in [3].

In this geometry there are line-for-line correspondences of the two images of a stereo pair of images like the used test data set “Athens” as shown in figure 1. (Flight direction from left to right, scan lines runs up-down, the columns from left to right.)

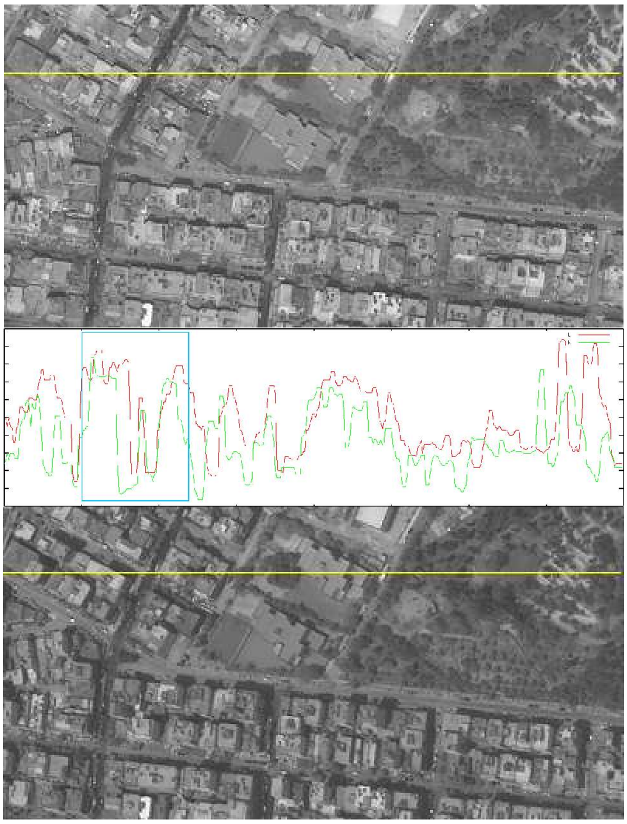


Figure 1 a, b, c. Test data set “Athens” with gray value profile along one scan column. a) left stereo image (top), b) gray value profile along yellow marked corresponding columns (center), c) right stereo image (bottom), see fig. 2 for blue marked detail

The gray value profile (fig. 1b) shows some of the problematic features like occlusion of a shadowed street (see detail in fig. 2), vanishing walls or remarkable differences in radiance between the two images of the stereo pair all over the column.

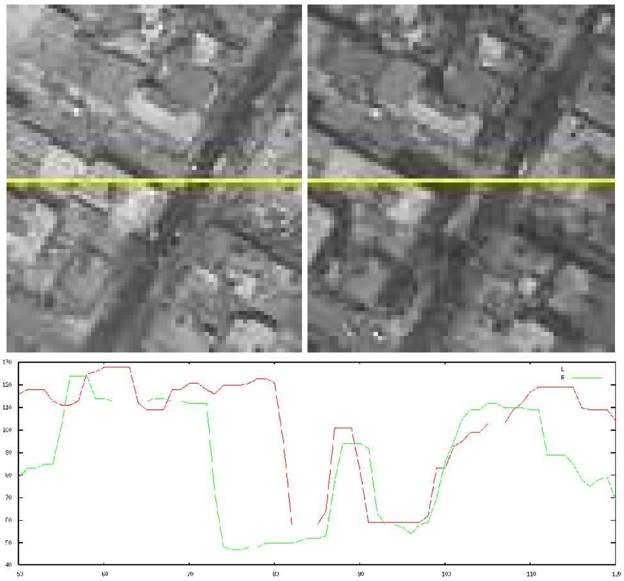


Figure 2: Detail of scene from fig. 1: Left house, narrow and wide shadow, sidewalk, street and right sidewalk and house can be identificated (profile median filtered with size 5).

The following processing steps are carried out for each column of the pair of stereo images:

- Extract the corresponding columns l_i and r_i from the two images;
- Repeat the following two steps for every position i in the column and displacements d ranging over ±distance with left-to-right and right-to-left image columns respectively:
  - Radiometric correction: Attempt to fit the actual patch of size *fit-size f* with the correlated patch in the other image shifted by *displacement d* to obtain local brightness and contrast parameters h_1 and h_2 and hence a radiometric corrected left image column $l_i' = h_1 + h_2 \cdot l_i$;
  - Image matching: Correlate the two patches with a size of *window-size w* yielding a correlation distance $D_{i,d} = \sum_{j=0}^{w} \text{abs}(\, l_j' - r_j \,)$;
- Retrieve the best correlation $D_{i,d}$ and thus the displacement (parallax) d for every pixel i in this column by following a path of minimum "potential energy" across the column;
- do a linear interpolation of all points which could not be correlated;
- generate an orthoimage based on the parallaxes and the two stereo images.

3. ALGORITHM IN DETAIL

Since epipolar geometry is given every scan column in the first (stereo left) image corresponds to the same scan column in the second (stereo right) image. Thus the problem of generating a DEM can be reduced at first to the task of finding correlations of the gray value profiles of these corresponding scan columns. Later in a refinement of the process the influence of other nearby columns can be taken into account.

To find correlations between the two gray value profiles small patches of size *window-size* have to be compared. A distance $D = \sum \text{abs}(\, l_j' - r_j \,)$ over all pixels j of the patch of *window-size w* is taken as measure. Since there are differences in the brightness and contrast of the two images of the stereo pair the two patches are corrected for their radiometry beforehand.

This is achieved by a least squares fit with $l_i' = h_1 + h_2 \cdot l_i$ which maps the grey values l_i of one of the two columns to the radiometric corrected values l_i'.

Least squares adjustment yielding contrast differences of a factor 5 or more – thus h_2 larger than 5 or lower than 1/5 – are marked as "out of range" (*OOR*). In this case the correlation is not carried out assuming there was no least square fit for the radiometry possible and therefore as well no matching will be possible. The factor 5 has been retrieved experimentally by comparing the quality of the results in a series of carried out matches with factors in a range from 1 to 20.

Based on these least square results a correlation based on a one dimensional pattern matching of an area of *window-size w* for all cells i of the left image is done with all possible displacements d in the range of *-distance* until *+distance* in the right image and vice versa. These arrays can be shown as 3D correlation mountains with the distance $D_{i,d}$ as height (fig. 3):

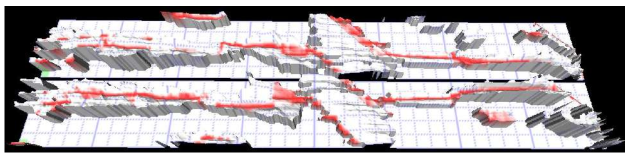


Figure 3. Correlation results for f=50 and w=50; top: left-to-right correlation; bottom: right-to-left-correlation; scan lines (rows) i running right, displacement d running up; lowest 10 % of $D_{i,d}$ in each every row marked red.

The flat areas depict the OOR areas which doesn't match the least squares criterion from above. The lowest parts of the *D*-mountains resemble the best correlations (marked red for better perceptibility).

Good correlations show up as steep valleys surrounded by OOR regions (like the ones in the right third of the image). The retrieved parallax is in this case the displacement of the bottom of the valley. Low correlations yield flat undulating areas with no sharp global minima. As seen in figure 3 a correlation left-to-right results in a mostly opposite displacement as the corresponding right-to-left correlation as has been expected.

The two resulting distance arrays $D_{i,d}$ (one for left-to-right and one for right-to-left matching) are summarized (fig. 4). To retrieve the parallaxes a path requiring a minimum of energy (taking the distance $D_{i,d}$ as a measure for potential energy) connecting all rows from left to right is extracted (red).

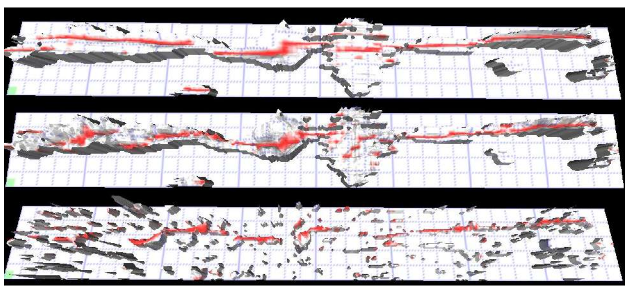


Figure 4 a, b, c. Sum of left-to-right and right-to-left correlations with parameters w=50, f=50 (top); w=10, f=50 (middle); w=50, f=10 (bottom).

For areas obscured in one or the other image the two correlations left-to-right and right-to-left are not identical. Such areas are often also marked as OOR by the least squares radiometry matching or result in low correlations.

Calculating the correlations and extracting the parallaxes for all columns of an image yields the digital surface model of the stereoscopic image pair as shown in figure 7.

Applying the previously described algorithm to the test data set “Athens” (fig. 1) yields following results at given parameters *window-size w* and *fit-size f* (fig. 5):

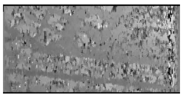
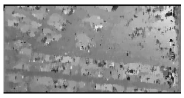
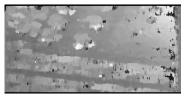
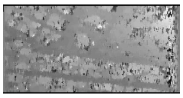
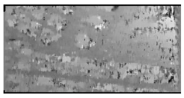
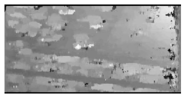
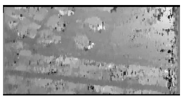
w	*f* = 8	*f* = 16	*f* = 32
8			
16			
32			

Figure 5. Results applying the “column” algorithm to test dataset “Athens”

4. DISCUSSION

As depicted in figure 4b lower window-sizes *w* lead to more rapid parallax changes (shorter red lines) and thus to more details; smaller fit-sizes *f* result in greater OOR regions (fig. 4c).

Calculating the whole DEM with small window sizes *w* produce more details but also more blunders as can be seen in figure 5. Larger window sizes on the other hand smooth out details in column direction. The same applies in a greater extend for the fit size *f*. Although it seems not to make much sense, small *f* combinded with large *w* also work tolerably well.

Comparing this test data set with parameters $w = 16$ and $f = 16$ with the “conventional” algorithm developed for the MOMS three line scanner [4] yields figure 6. This existing matching algorithm is based on image pyramids, finding interest points with the Förstner operator [5] matching these to subpixel range and applying a region growing algorithm.

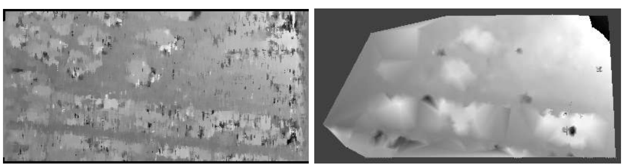


Figure 6 a, b. column vs. conventional algorithm “Athens”

As can be seen the “column” algorithm in contrast to the conventional algorithm produces much more details but also more blunders. The DEM produced by the conventional algorithm is much smoother and sharp details can not be reproduced satisfyingly. (The missing borders are due to a lack of good interest points in this regions.) Some of the blunders introduced by the “column” algorithm can be reduced e.g. by median filtering (see fig. 7).

5. SUMMARY

The new presented “column” algorithm yields more detailed results in dense 3D scenes but also more blunders than the conventional algorithm. In column direction a blur depending on increasing window size as well as increasing fit size can be found. Higher values of window- or fit-size produce less blunder but more blur. Application of a two step median filter along horizontal and vertical directions can reduce blunders significantly.

Further research will concentrate on a better algorithm for blunder reduction and introduce feature based optimizations. Also all results have to be compared with reliable high resolution DEMs e.g. from laser data.

REFERENCES

[1] Toutin T., “DTM Generation from Ikonos In-Track Stereo Images Using a 3D Physical Model”, Photogrammetric Engineering & Remote Sensing 70(6): 695-702

[2] Eisenbeiss H., Baltsavias E., Pateraki M., Zhang L., “Potential of IKONOS and QUICKBIRD imagery for accurate 3D point positioning, orthoimage and DSM generation”, Proc. 20th ISPRS Congress. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35, Part B

[3] Scharstein D., Szeliski R., “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms”, IJCV 47(1/2/3):7-42, http://cat.middlebury.edu/stereo

[4] Lehner M., Gill R.S., “Semi-automatic derivation of digital elevation models from stereoscopic 3-line scanner data”, ISPRS, Vol. 29, Part B4 68-75

[5] Förstner W., Gülch E., “A fast operator for detection and precise location of distinct points, corners and centers of circular features”, ISPRS Intercommission Workshop, Interlaken, June 1987

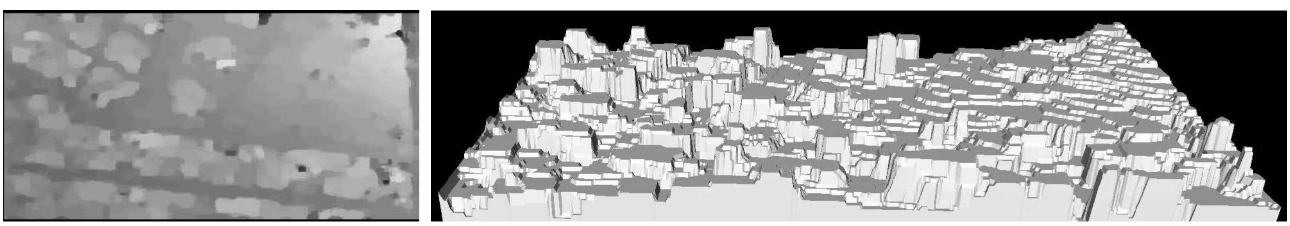


Figure 7. DEM from fig. 6a smoothed by hoizontally and vertically median filter of size 10 (left) and resulting DEM in 3D (right)