DIGITAL SURFACE MODELS OF CITY AREAS BY VERY HIGH RESOLUTION SPACE IMAGERY

Jacobsen, K.

University of Hannover, Institute of Photogrammetry and Geoinformation,
Nienburger Str.1, D30167 Hannover
phone +49 511 762 2485, fax +49 511 762 2483, Jacobsen@ipi.uni-hannover.de

ABSTRACT

The geometric resolution of the very high resolution optical space sensors like IKONOS, QuickBird and OrbView-3 do allow the generation of detailed digital surface models also over build up areas. This may be used as base for a 3D-city model. By automatic image matching the height value of the building top do show the form of the houses within the spacing of the height models well, but in most cases the neighboured points on the ground do have some distance from the building caused by viewing shadows. The size of the used sub-matrix for image matching also will cause some rounding of the corners by the fact that the pixels are located partially on top of the buildings and partially on the ground or even the wall.

IKONOS, QuickBird and OrbView-3 stereo pairs have been used for the generation of digital surface models in city areas. By simple theory the height accuracy is linear depending upon the height-to-base-ratio. This is the case for open and flat areas, but not for cities. With a small height-to-base-relation (close to 1:1) only a limited number of points can be determined on the ground, degrading the overall information contents. In city areas it has been shown that a smaller base has some advantages and it is not linear influencing the vertical accuracy - for a small convergence angel of the space images the standard deviation of the x-parallaxes is still better. Of course the mentioned problems are not the same for the very time consuming manual 3D-measurements. But in some cases of very high buildings also problems of the stereo impression are caused by a large convergence angle of the used satellites. Also this is leading to relatively better results for a limited height-to-base-relation.

1. INTRODUCTION

3D-city models are becoming more and more popular; they are required for several purposes. They can be determined by ground survey or remote sensing. Ground survey is very time consuming; that means expensive. In addition usually only the road side can be reached without problems. By these reasons usually the remote sensing will be used. A detailed survey is possible by laser scanning (LIDAR), but laser scanning is also expensive and laser scanners are mainly concentrated to Europe and North America, only very few instruments are in use in other continents. The traditional survey by aerial photogrammetry is a proven technique, but in several countries aerial images are classified. With the now available very high resolution of space images, they can be used for the generation of 3D-city models.

2. METHODS OF IMAGE MATCHING

Automatic image matching tries to find corresponding image points in overlapping images taken from different directions – a stereo model. A sub-matrix of the left image is compared with the sub-image of the right hand image and the optimal match between the shifted and/or transformed sub-matrixes leads to corresponding or conjugate image positions which can be used for the computation of the ground coordinates. For the image matching following methods are used:

1. Feature based matching: image points which can be identified by interest operators are searched in both images and the conjugate positions are determined. Only a limited number of well determined features are available, this is not enough for the generation of a digital elevation model. By this reason, the feature based matching is used as start information for the other methods requiring some approximations.
2. Image correlation: a smaller pattern matrix of one image is compared with a part of a larger search matrix of the other image and the correlation coefficient between both images is computed. The pattern matrix is shifted in the search matrix – this will be done with all possible combinations between both matrixes. The positions with the largest correlation coefficient are identical to the conjugate points. This traditional image correlation by theory is only correct for horizontal objects because tilted ground elements do have a different size in both images. The search matrix should not be too small to avoid problems caused with not accurate start values and it should not be too large to avoid a second maximum – for example if two nearly identical buildings are located in the pattern matrix, for both buildings the correlation coefficient may be the same.
3. Image correlation in epipolar images: with given exterior orientation, the position of a point conjugate to another in the other image must be located on the corresponding epipolar line. So a search of the corresponding position can be reduced to one direction. With this exception, it is identical to the standard image correlation.
4. Vertical line locus: it requires also the exterior orientation. Instead of a search in epipolar lines, matrixes are specified in the object space with a sequence of different height values. These object matrixes are transformed into the images and the correlation coefficient between the corresponding sub-matrixes in both images is computed for every chosen height level. The maximum of the correlation as function of the object height is identical to the object height. This method is very similar to previous, but it has the advantage of fixed object coordinates X and Y, so directly an exact grid of points can be computed, while the image matching in epipolar lines leads only to an approximate grid of points.
5. Least squares matching: Independent from the method of approximate positions, the simple image correlation by theory requires a horizontal object. A tilted surface has a different scale in both used images. This is respected by the least squares matching; it relates the sub-matrix of the one image to the corresponding sub-matrix of the other by affine transformation. In addition also a linear change of the grey values within the sub-matrixes is respected. The least squares image matching by theory is the most accurate method, but it has the disadvantage of a small convergence radius – the approximate relation between both images must be known in advance. So it has to be combined with a method giving approximate conjugate positions.

The investigations of the 3D-surface models based on the very high resolution space images were made with the Hannover program DPCOR using least squares image matching. The approximate corresponding image positions are determined by region growing – from very few manual measured conjugate points the neighboured points are determined and from these the next neighboured and so on. So the relation between both images is known in advance based on the relation of the directly neighboured points.

3. PROBLEMS OF IMAGE MATCHING IN BUILD UP AREAS

All methods of image matching for the generation of a digital surface model (DSM) are based on a continuous and differentiable surface. The DSM includes the height of the visible surface; that means, also buildings and vegetation. A DSM is not differentiable because the change of the direction from the facade to the roof top. Also the difference in the view direction of the used images is causing problems. A facade may be shown in one image, but not in the other – so some image pixels do not have conjugate pixels in the other image (figure 1).

fig. 1: same building shown in the images of an IKONOS stereo pair – only in left hand image the south-western facade can be seen

height to base relation: 7.5

For the image matching sub-matrixes are used. The sub-matrixes must have a sufficient size, enabling the matching. If they are too small, random noise may dominate the matching; if they are too large, a building will be deformed into a hill, because parts of the sub-matrixes may show the ground (street level) and parts may show the roof top. Like standard image correlation, the least squares adjustment will lead in the position of the facade to an average height between the ground and the roof top (figure 2).

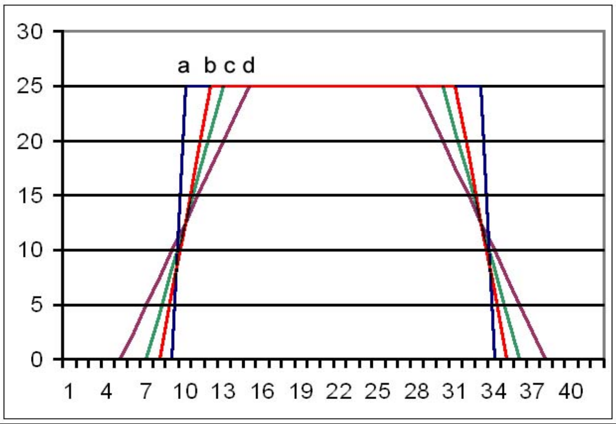


fig. 2: profile through a 25m high building
a: original building
b: sub-matrixes with 4 x 4 pixels
c: sub-matrixes with 6 x 6 pixels
d: sub-matrixes with 10 x 10 pixels

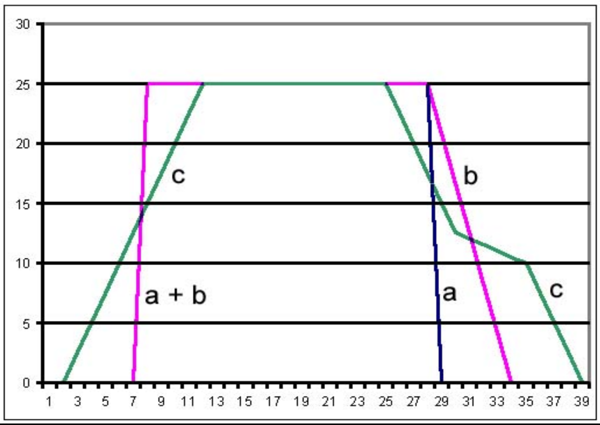


fig. 3: profile through a 25m high building
a: original building
b: shape with nadir 13° angle (view from left)
c: shape caused by nadir 13° angle + sub-matrixes 10 x 10 pixels

Depending upon the size of sub-matrixes used for image matching, the vertical facade will get an inclination (case b, c, d in figure 2). But in any case in a raster representation will also directly lead to an inclined facade because between the last point on the ground and the first point on the roof at least the width of the DEM-spacing will appear (a in figure 2 and figure 3). In addition the view from the projection centre to the object will not be exactly vertical, so at least one side of the building is hiding parts of the ground (e.g. figure 1 left: the right hand side of the tall building is hiding parts of the ground and lower buildings). This will cause, in addition to the effect of the size of the sub-matrixes used for matching, the shape shown as line c in figure 3. A stereo pair is a combination of two images taken from different positions, so the effect shown in figure 3, line c, will happen also for the other image, so the left hand side may get a similar deformation.

The standard deviation of the object height SZ depends upon the height to base relation and the accuracy of the x-parallax.

$$SZ = \frac{h}{b} \bullet Spx$$

Formula 1: standard deviation of object height; Spx [GSD]

For satellite images the flying height *h* is fixed, so the major possibility for changing the height accuracy is to change the base; that means the distance between the both projection centres. The very high resolution satellites do have a free view direction, by this reason the base has to be computed with the collection elevation (nadir angle) and the collection azimuth – it is not just a function of the nadir angle. The standard deviation of the x-parallax is depending upon the image contrast but also the angle of convergence between the corresponding images of a stereo model. The images shown in figure 1 do have just a height to base relation of 7.5 – by simple theory this is causing a poor vertical accuracy, but under these conditions, the corresponding images are quite similar leading to a better image matching (smaller value for Spx). This effect together with formula 1 results in an optimal configuration which is for city areas more in the height to base relation of 3 (see also Börner et al 1997).

4. HEIGHT MODEL BASED ON IKONOS

From the area of the Turkish city Maras an IKONOS stereo model with a height to base relation 7.5, taken from the same orbit, is available. The city is located in a rolling terrain and is densely build up with medium size houses, but also some larger ones.

fig. 4: part of IKONOS image Maras, Turkey

The image matching with the Hannover program DPCOR allows a variation of the sub-matrixes for matching, a variation of the distance of the neighboured points and can set a tolerance limit for the correlation coefficient. With program LISA used for DSM presentation, the height model can be filtered. As tolerance level for the matching, based on intensive experience, a correlation value of 0.6 usually was used. The standard size of the sub-matrixes for matching is 10 x 10 pixels. If directly neighboured pixels are matched, the corresponding sub-matrixes are overlapping by 90% - such a matching result cannot be independent; so usually every third pixel (pixel spacing = 3) is used for matching. This includes the disadvantage of a stronger inclination of facades in the generated DSM (see also figures 2 and 3).

With smaller sub-matrixes the shape of the buildings should be more precise, on the other hand with a smaller number of pixels the noise will be larger (larger standard deviation). To find the pros and cons, sub-matrixes of 10x10, 6x6 and 4x4 have been used with pixel spacing of 3 and 1. In addition in the program LISA the results have been improved by median filter with 3x3 and 7x7 neighboured values.

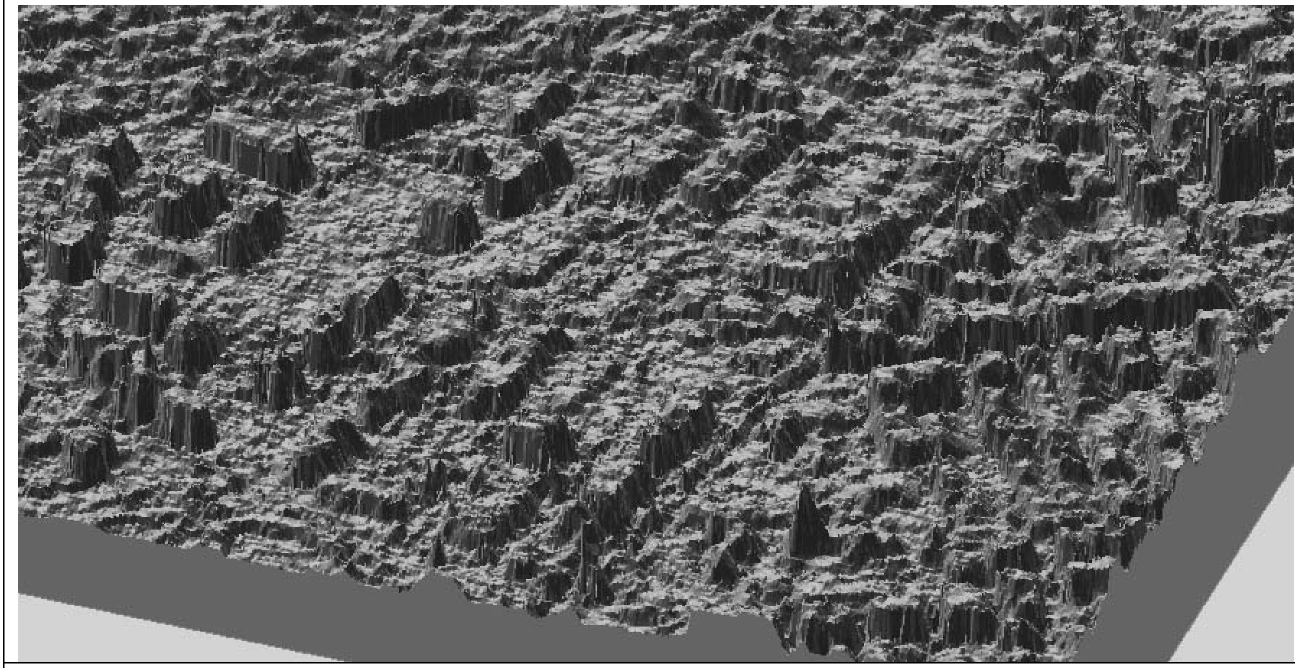


fig. 5: 3D-view to IKONOS DSM (same area like figure 4), matching with sub-matrixes 10x10 pixels, point spacing 3, no filter

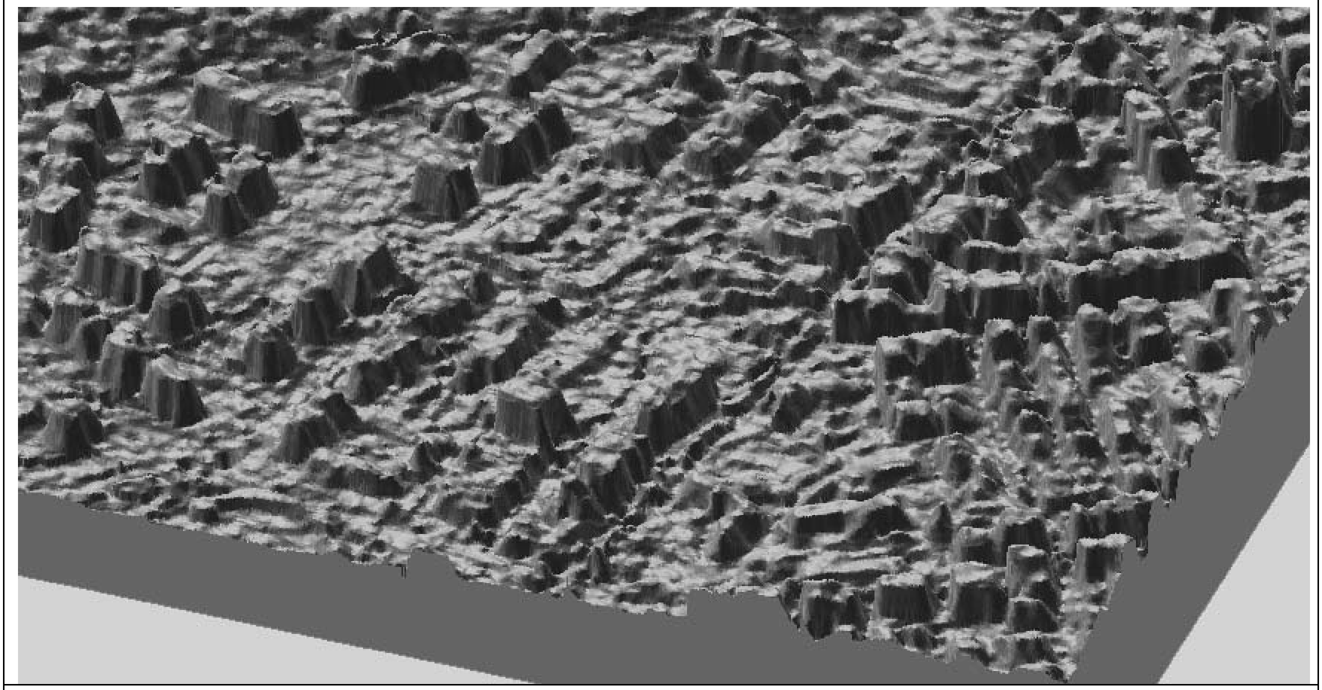


fig. 6: 3D-view to IKONOS DSM (same area like before), matching with sub-matrixes 10x10 pixels, point spacing 1, median filter 7x7

The difference of the matching results shown in figures 5 and 6 are mainly based on the median filtering. The more dense point spacing in figure 6 does not show a clear advantage. The large buildings clearly can be separated and identified in the 3D-shading models while the smaller buildings in the upper left corner are more shown like hills.

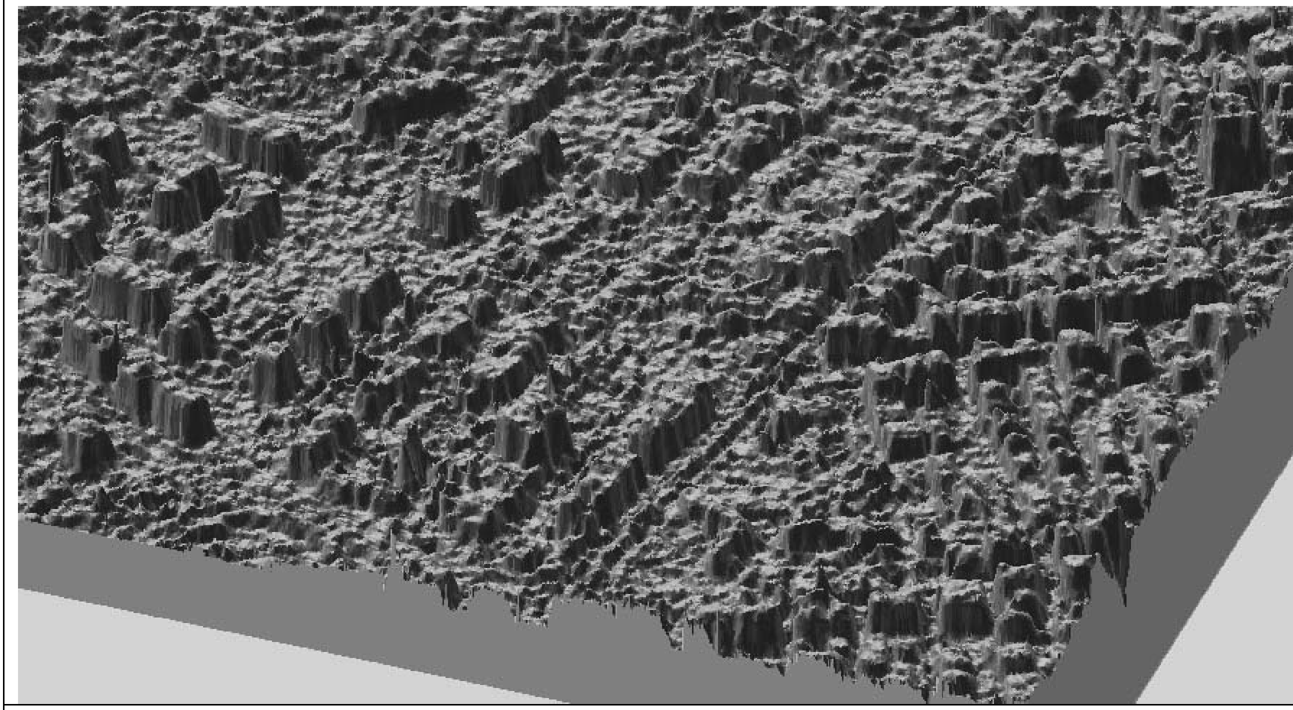


fig. 7: 3D-view to IKONOS DSM (same area like figure 4), matching with sub-matrixes 6x6 pixels, point spacing 1, median filter 7x7

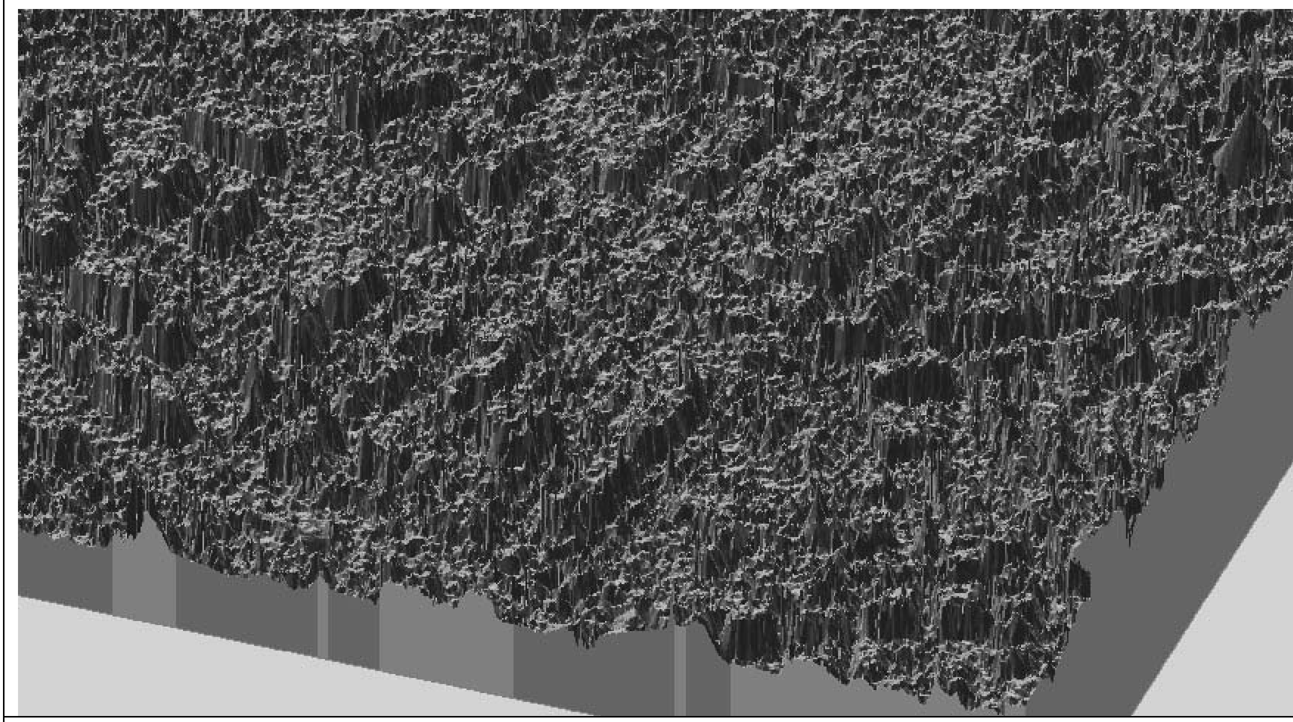


fig. 8: 3D-view to IKONOS DSM (same area like figure 4), matching with sub-matrixes 4x4 pixels, point spacing 3, no filter

The matching results based on smaller sub-matrixes do not show some advantages of clearer object details against the matching with 10x10 pixels, but they are noisier. The results based on sub-matrixes of 4x4 pixels are not very useful. In general, the small base length and the imaging from the same orbit resulted in very high values for the correlation coefficients and a good success rate; between 95% and 97% of the possible points have been matched. For 64% up to 82% of the points, the correlation coefficient exceeded the value 0.95 (figure 13). The smaller percentage belongs to the matching of 4x4 pixel sized sub-matrixes, but also here only 18% of the points have had correlation coefficients below 0.90.

A combination of two IKONOS images taken in June and October over the area of Zonguldak city failed in automatic image matching. The sun elevation changed from 62° to 41°, causing quite different shadows. So finally not more than 30% of the area could be covered by height information and the correlation coefficients have been poor.

5. HEIGHT MODEL BASED ON QUICKBIRD

From the area of Phoenix, Arizona some QuickBird images have been used for the analysis of automatic matching in the city area. The first combination of two images have been taken with 10 days difference in time, this has not caused problems with changed shadows. The build up area was matched successfully (figure 9). Difficult was the height determination on large parking places at shopping centres – the cars have moved and so the matching failed (figure 10).

fig. 9: overlay of matched points to QuickBird image

matched points = white

dark parts = matching failed

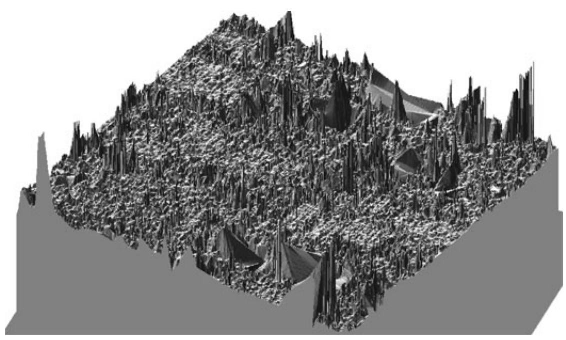
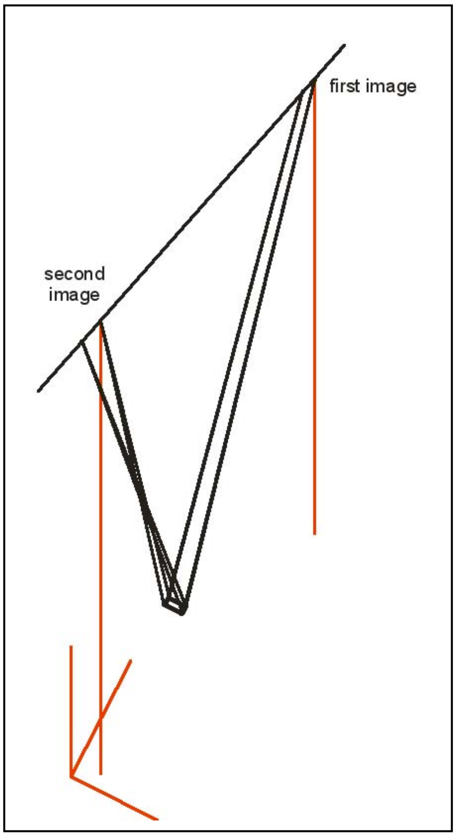


fig. 10: shaded DSM generated by QuickBird image matching

Dark image parts, moved cars and repeated elements, especially on roof tops, caused in some areas errors in matching (peaks in figure10). The caused height errors have been eliminated automatically with the Hannover program RASCOR (Jacobsen, Passini 2001). After this, the result was satisfying. The dominating small individual houses with trees and bushes did not lead to clear shapes of buildings.

6. HEIGHT MODEL BASED ON ORBVIEW-3

An OrbView-3 stereo model taken over the city area of Zonguldak was investigated. OrbView-3 has different imaging characteristics like IKONOS and QuickBird. Opposite to these both, OrbView-3 has no transfer delay and integration (TDI) sensor, corresponding to an electronic forward motion compensation. Without TDI a pixel size of 1m corresponds to an imaging time of 0.14ms and this is not enough for a satisfying imaging quality. OrbView-3 is enlarging the time by staggered CCD-



lines – a combination of 2 CCD-lines shifted ½ a pixel against each other. So the projected pixel size is 2m on the ground, but neighboured pixels are overlapping 50%, that means, the distance between neighboured ground pixels centres, the ground sampling distance (GSD), is 1m. In addition OrbView-3 has only a sampling rate of 2500 lines/sec for the combination of the two staggered lines, but the footprint speed of OrbView-3 is 7.1km/sec. This has to be compensated by a permanent rotation of the satellite during imaging with a slow down factor of 1.4 (1.4 times the distance in the orbit is used for imaging like required for an imaging with same view direction in relation to the orbit).

fig. 11: imaging geometry of OrbView-3 in Zonguldak, height to base relation = 1.4

Caused by the difference between pixel size on the ground and the GSD, the image quality is a little less than for IKONOS and QuickBird (figure 14). In addition the quite different view direction is generating problems (figure 12). The view from north shows the northern facade and this mainly in the shadow, while the view from south shows the southern facade in the sun light.

The a little lower image quality together with the quite different view direction is causing problems for the image matching. The used region growing method partially has had problems in stepping over the difference in the image details. It was not possible like with the IKONOS stereo model, having a height to base relation of 7.5, to use a matching with sub-matrixes of 4x4 or 6x6 pixels. Reverse, the sub-matrixes have been enlarged to 12x12 up to 16x16 pixels.

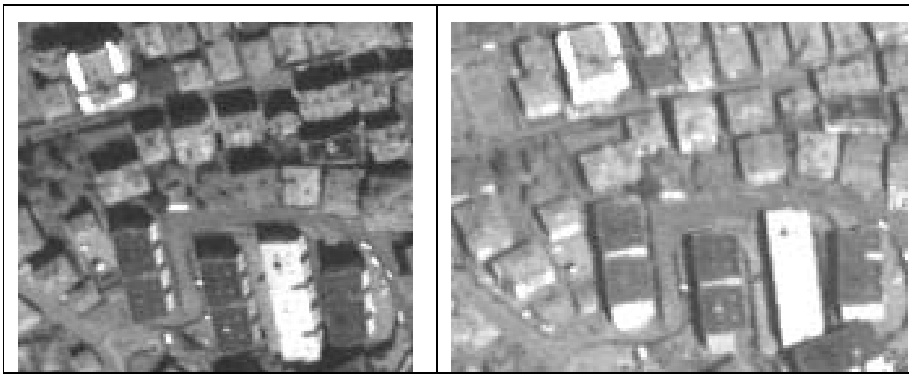


fig. 12: corresponding OrbView-3 sub-images

left: view from north

right: view from south

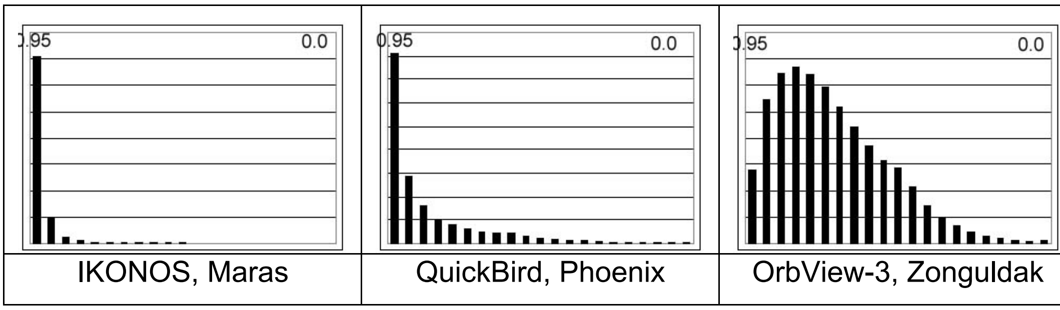
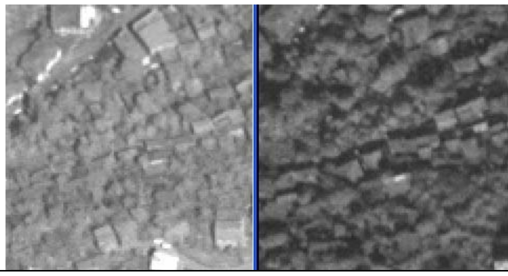
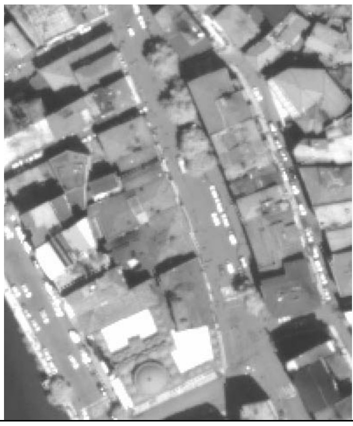
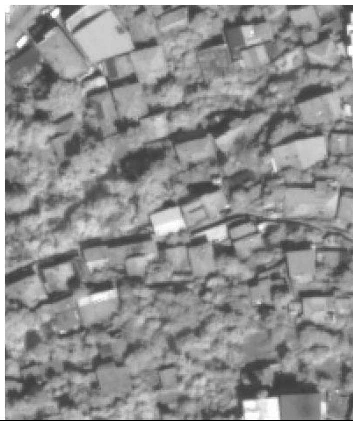
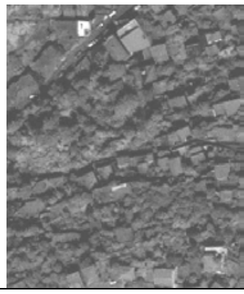


fig. 13: frequency distribution of correlation coefficients

As it can be seen in figure 13, the distribution of the correlation coefficients is quite different for the different sensors. IKONOS shows the best correlation coefficients with the majority above 0.95. For QuickBird it is a little below, but in other areas it has been shown, that it is usually on the same level like IKONOS. For QuickBird-3 the correlation was quite more difficult. Of course, partially this is caused by the large base with height to base relation, but also by the lower image quality (figure 14).

		fig. 14: conjugate sub-images from areas with problems of matching OrbView-3 1m GSD
		same area in QuickBird image with 0.62m GSD
		same area in IKONOS image with 1m GSD

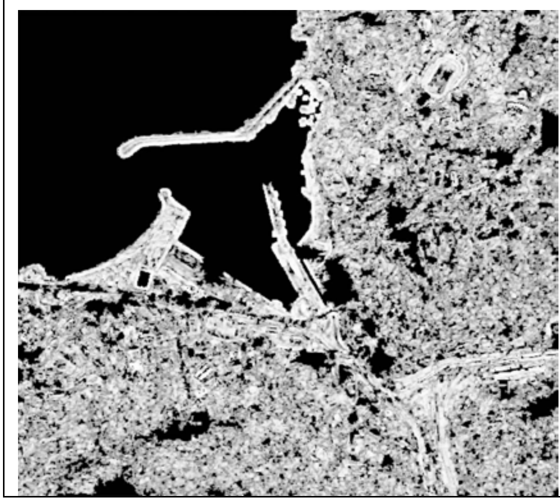


fig. 15: quality image of matching of OrbView-3

grey value 255 = correlation coefficient = 1.0

grey value 127 = correlation coefficient = 0.5

grey value 0 = correlation coefficient = < 0.5

Figure 15 shows the areas with problems in matching the OrbView-3 images. Of course the water surface cannot be matched and in the other areas there was no contrast or the conjugate sub-matrixes were too different (figure 14). Partially even for the human operator it was difficult to find conjugate positions. Finally only 53% of the area could be matched with a sub-matrix size of 12x12 pixels. Approximately 25% of the area is covered by water, so for 22% of the area the correlation coefficient was below the chosen limit of 0.5. In addition 11% of the points have been matched with correlation coefficients between 0.5 and 0.6, so with the same tolerance limit of 0.6, used for IKONOS and QuickBird, 33% had not been accepted.

7. ACCURACY

The building heights in Maras have been determined also by the length of the shadows. This was leading to an accuracy of the building height based on a matching with sub-matrixes having10x10 pixels of 1.7m corresponding to a standard deviation of the x-parallax (Spx) of 0.22 pixels. This was confirmed by the standard deviation of the y-parallaxes with 0.25 pixels for 10x10 pixels, 0.32 pixels for the sub-matrixes 6x6 and 0.44 pixels for sub-matrixes 4x4 pixels. That means the higher details of the smaller sub-matrixes are leading to a lower accuracy.

The QuickBird height model was checked against a reference DEM not including the buildings and the vegetation; nevertheless as RMSZ 4.8m have been reached corresponding to 0.8 pixels Spx. If the influence of the buildings and the vegetation is respected, the accuracy is in the same level like for IKONOS.

For OrbView-3 no precise reference height model is available and the buildings in the city area are a lot higher as in Phoenix. By this reason an accuracy analysis has not been made.

CONCLUSION

Digital surface models can be generated in city areas with very high resolution space images. Of course a manual plotting of the buildings leads to the highest accuracy and detailed results, but this is very time consuming. In build up areas the matching of space images with a small base (large value of height to base relation) has advantages against the matching with a large base, which is causing a loss of matched points by reason of differences in the images taken from different direction and is not leading to a good shape of the buildings because of the large view shadows. So the often used height to base relation of 1.6 or even 1.4 has advantages for the open areas, but for the build up areas a larger value (smaller base) should be preferred. By theory with a smaller size of the sub-matrixes used for matching, the building shape should be clearer, but together with the lower accuracy and the required filtering it has not a real advantage. The point spacing of just 1 pixel for matching leads to a slightly better shape, but this extends the computing time.

REFERENCES

Börner, A., Hilliger, E., Reulke, R. 1997: The Optimization of the Stereo Angle of CCD-Line-Sensors with end to end Simulation Systems, „Sensors and Mapping from Space“, University of Hannover 1997

Jacobsen, K., Passini, R. 2001: Filtering of Digital Elevation Models, GIS 2001, Vancouver