

EVALUATION OF WORLDVIEW-1 STEREO SCENES AND RELATED 3D PRODUCTS

Daniela POLI, Kirsten WOLFF, Armin GRUEN
Swiss Federal Institute of Technology
Institute of Geodesy and Photogrammetry
Wolfgang-Pauli-Strasse 15, 8093 Zurich, Switzerland
E-mail: poli,wolff,agruen@geod.baug.ethz.ch

KEYWORDS: Worldview-1, orientation, matching, DSM, city modelling

ABSTRACT: This paper describes the first investigations carried out at ETH Zurich on the very-high resolution optical satellite sensor WorldView-1. The dataset includes a stereopair acquired in 2007 over Morrison, Colorado, USA. The pair consists of a quasi-nadir scene with mean GSD close to 51cm and an off-nadir scene, with GSD up to 73cm. We show our first results of exploiting Digital Surface Modeling (DSM) and of evaluating the potential of the images for 3D city modeling using the software CyberCity-Modeler, in terms of level of detail. Here, special attention was dedicated to the DSM generation using the software package SAT-PP (Satellite Image Precision Processing) and the accuracy assessment by visual and quantitative analysis with respect to reference data derived from aerial images.

1 INTRODUCTION

Worldview-1 (WV-1) is one of the newest civil very high resolution satellite sensors providing stereo scenes. Launched on 18 September 2007 and operational since November 2007, the satellite operates at an altitude of 496 kilometers along a sun-synchronous orbit, with an inclination of 97.7° and a revisit time of 1.7 days at 1m GSD and 4.6 days at around half meter GSD. Worldview-1 carries a panchromatic imaging system providing in-track stereo images of target areas by rotating along its axis (nominally maximal $\pm 45^\circ$ off-nadir).

The aim of the work was a) to analyze the geometric and radiometric characteristics of the images, b) to do an accurate (subpixel) orientation, c) to exploit digital surface modeling (DSM) and d) to evaluate the potential of the images for 3D city modeling in terms of level of detail. The scenes have been processed using the software package SAT-PP (Satellite Image Precision Processing) and other in-house developed routines for image pre-processing, orientation and DSM generation, and the software CyberCity-Modeler for 3D building model extraction. In this paper we report the first investigations carried out at ETH Zurich on Worldview-1 stereo scenes.

2 DATA DESCRIPTION

2.1 Worldview-1 scenes

The Worldview-1 stereo scenes, kindly provided by Eurimage SpA (Eurimage, 2008), were acquired from the same orbit on 25 November 2007 at around 5:40pm over Morrison (Colorado, USA). One of the two scenes, hereafter called “nadir”, was acquired with mean in-track and cross-track viewing angles of 4.1° and 10.5° respectively, and a mean GSD of 0.51m. The mean viewing angles of the second scene, called “forward”, are 32.6° (in-track) and 11.9° (cross-track), resulting in a mean GSD of 0.73m (Figure 2). The scenes were processed at level 1, that is radiometrically

and at sensor level (all pixels from all detectors are blended into the synthetic array to form a single image), but not geometrically. The dynamic range of the images (11bit) was reduced to 8bit with a linear transformation for later processing with our own software. In addition to the images, Eurimage provided the location and geographic coordinates of 11 ground points, whose distribution is shown in Figure 1.

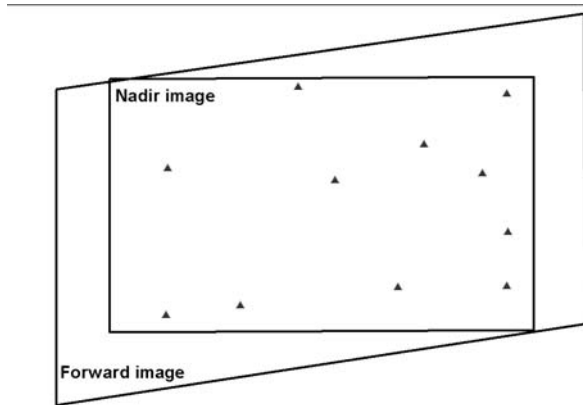


Figure 1. Location of the Worldview-1 stereo scenes and distribution of GCPs.



Figure 2. Zoom on detail from the nadir (top) and forward (bottom) views.

2.2 Reference data

For the DSM quality evaluation, reference data at suitable resolution was generated. To this purpose we used 1:15'000 scale infrared aerial images from USGS. The photos were acquired in April 2002 with an analogue camera ($f=153\text{mm}$) and scanned with a resolution of $14\mu\text{m}$. After orienting the images with LPS (Leica Photogrammetric Suite), the DSM was generated with the software package SAT-PP with 1m grid spacing. No GCPs were available for the area covered by the aerial images. Therefore we chose well-defined GCPs in the oriented stereo pair of WorldView-1 and used them for the orientation of the aerial images. For our investigations of the matching process for DSM generation a good relative orientation between the two stereo models is sufficient. The resulting RMS error for the triangulation of the aerial images is 0.4 pixels, using the same a priori standard deviations for the GCPs coordinates of 2m in X- and Y-direction and 1.5m in Z-direction.

3 GEOREFERENCING, DSM GENERATION AND QUALITY ANALYSIS

The geometric processing of WV-1 scenes for 3D information extraction was carried out with our program package SAT-PP (Satellite Image Precision Processing). Details of the underlying algorithms can be found in (Zhang, 2005) and (Zhang and Gruen, 2006) and examples of using different kinds of high resolution satellite image data can be found in (Poli et al., 2004) for SPOT-5, (Baltsavias et al., 2006) for IKONOS and (Wolff and Gruen, 2007) for ALOS/PRISM. The main advantage of SAT-PP for image matching and DSM generation is that it matches very densely with three different match-point primitives (regular image grid points, interest points and edgels).

As orientation parameters for the two WV-1 scenes, we used the RPCs available in the metadata files, followed by a 6 parameter affine transformation, estimated with 9 GCPs. In general the description of the ground points was not always sufficient for accurate measurement and two GCPs were deleted. As RMSEs we got around 0.3m in

planimetry and height, but we still observed small y-parallaxes in some areas. However, this might have only a local influence in some cases. To improve the conditions for feature extraction and matching we preprocessed the images using a Wallis filter. As grid space we chose 4 times the GSD, which leads to 2m. For the matching process an initial DSM is needed. The required quality of the initial DSM depends on the GSD, the topography, the texture and the image quality. It can be defined just by a plane or by a given DSM like the SRTM. If no initial DSM is available, SAT-PP has the possibility to measure seed points manually in a stereomodel which will be used to define the initial DSM. As experiences show, one can expect blunders caused by the topography, image acquisition geometry, poor texture and y-parallaxes. Most of these blunders can be detected visually e.g. in a shaded model of the DSM (see Figure 3 - top left). To avoid such blunders additional seed points can be measured to improve the initial DSM locally (see Figure 3 – top right).

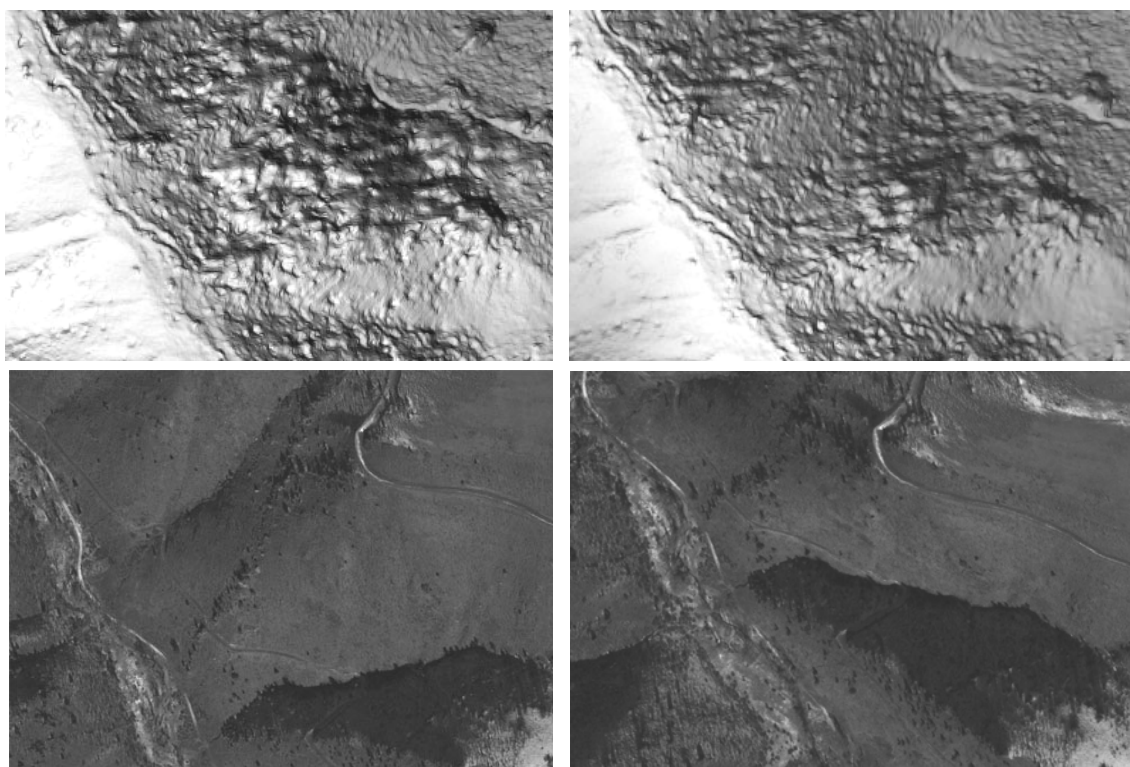


Figure 3. Top left: Shaded model of a critical area with blunders (grid space 2m). Top right: Improvement by measuring additional seed points Bottom: original stereo images of the same area (left: nadir, right: forward).

For quality evaluation, we compared the WV-1 DSM with the reference aerial image-based DSM, described in 2.2. As the time difference between the aerial images and the WV-1 scenes is around 5.5 years, we focused the analysis on a 1.1km by 1.0km open area, called Robinson Hill Road, with separated buildings and single trees (see Figure 4 – bottom left), and deleted some evident new buildings to reduce the influence of temporal changes. Figure 4 – top shows the reference and the WV-1 DSM over the test area. From a first visual analysis we can see that the houses and bigger streets are visible in both DSMs.

The quantitative analysis was carried out using the software package LS3D (Least Squares 3D Surface Matching), which was developed in our group (Gruen and Akca, 2005). It allows to evaluate the potential of the WV-1 data and the matching algorithm only, without the influence of the possibly different absolute orientations of the two

datasets. In fact with LS3D we first co-registered the reference and the WV-1 DSMs with three shift parameters to remove possible offsets between the two datasets, and we observed a significant shift in planimetry (0.70m in X and 1.57m in Y direction). In a second step the Euclidian distances between the two DSMs were computed, and the 3D RMSE was split into X, Y, Z components. The values are summarized in Table 1. The RMSE of the Euclidean distances is 0.71m, which is equivalent to the GSD of the forward scene. This is in the range of what we expected for this kind of project. The main component of the residual error is in Z direction. Figure 4 – bottom right visualizes the distribution of the 3D residuals (Z-component of Euclidean distances), after co-registration. Local blunders are probably due to temporal differences, low texture and the y-parallaxes. We are currently continuing with the evaluation.

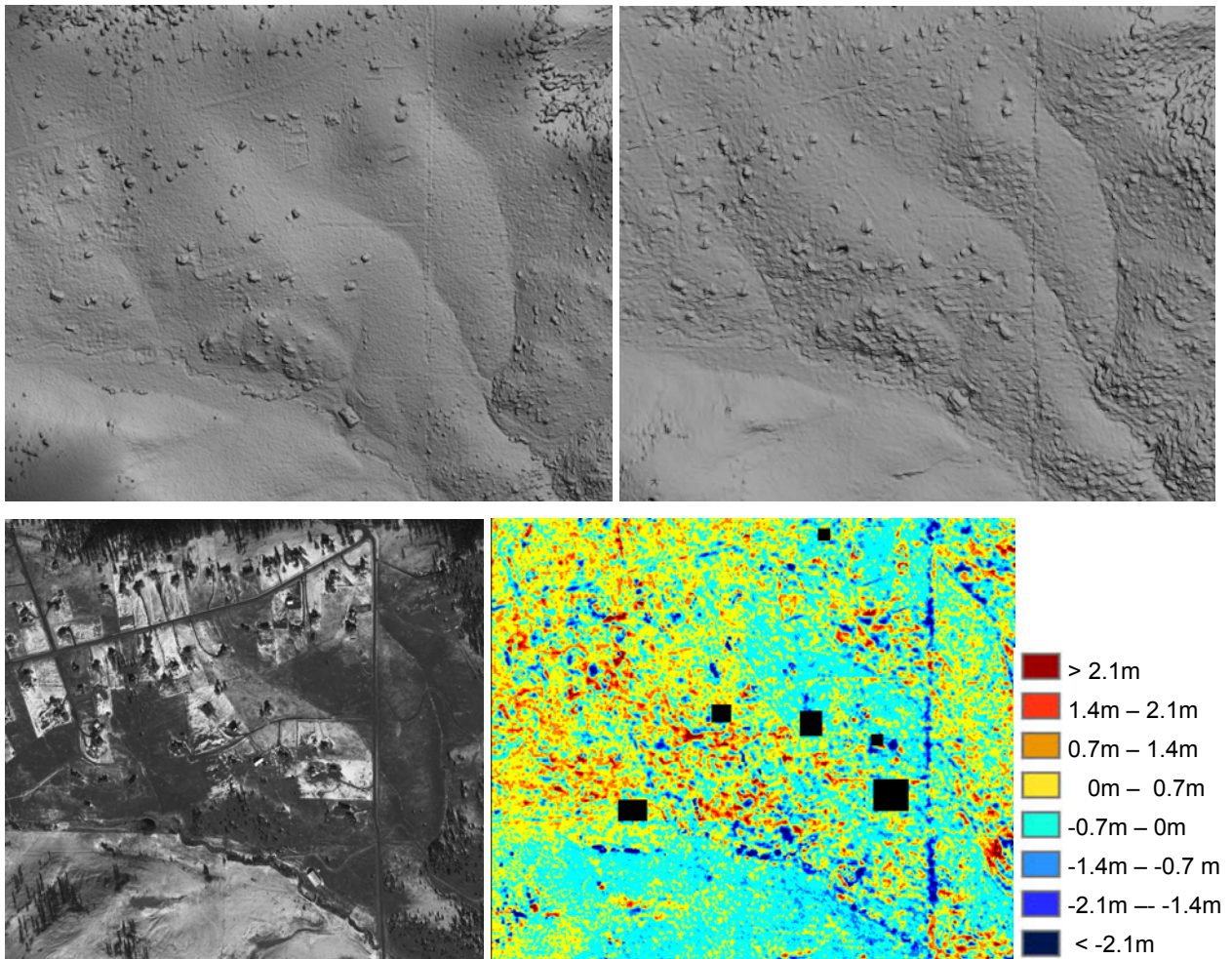


Figure 4. Top left: shaded model of the reference DSM (1.1km by 1.0km, grid spacing 1m). Top right: shaded model of the WV-1 DSM (grid spacing 2m). Bottom left: original WV-1 nadir scene of the test area. Bottom right: Residuals of the Z-component of the Euclidean distance between reference and WV-1 DSMs. Areas with new buildings, deleted and not considered in the analysis, are shown as black boxes.

Table 1. Statistical values of the residuals between reference DSM and WV-1 DSM after co-registration and shift parameters. The number of points is 1074729.

	RMSE (m)	Mean (m)	Min (m)	Max (m)	Shift (m)
Euclidian distance	0.71	0.02	-9.57	9.44	
X component	0.20	-0.01	-5.21	5.60	-0.70
Y component	0.21	-0.02	-5.34	6.06	-1.57
Z component	0.64	0.02	-8.60	9.22	0.10

4 3D CITY MODELING

For the generation of 3D city models the CyberCity-Modeler™ (CCM) was used. The software allows the generation, editing, management and visualization of 3D city models from aerial images, laser scanner data and satellite images in a semi-automatic mode (Gruen and Wang, 2001). In case of Worldview-1 images, the relevant roof points were measured 3-dimensionally on the digital station LPS, following specific rules given by the modelling strategy, then imported as a point cloud in CC-Modeler module and automatically fitted with roof faces. The building walls result from the vertical projection of the roof eaves onto the Digital Terrain Model, previously measured. The geometry of the 3D models was checked in the module CC-Edit and, when necessary, improved (e.g. right angles, parallel lines, planar faces, correction of overlays and gaps). We modeled buildings located in both residential and industrial areas. In Worldview-1 scenes the internal roof edges on the residential houses can be distinguished, therefore the buildings can be modeled with their main roof structures (Figure 5). It is still not possible to measure roof superstructures, like small dormers. Geometric building parameters as heights, volumes, areas and façade inclinations are automatically calculated.

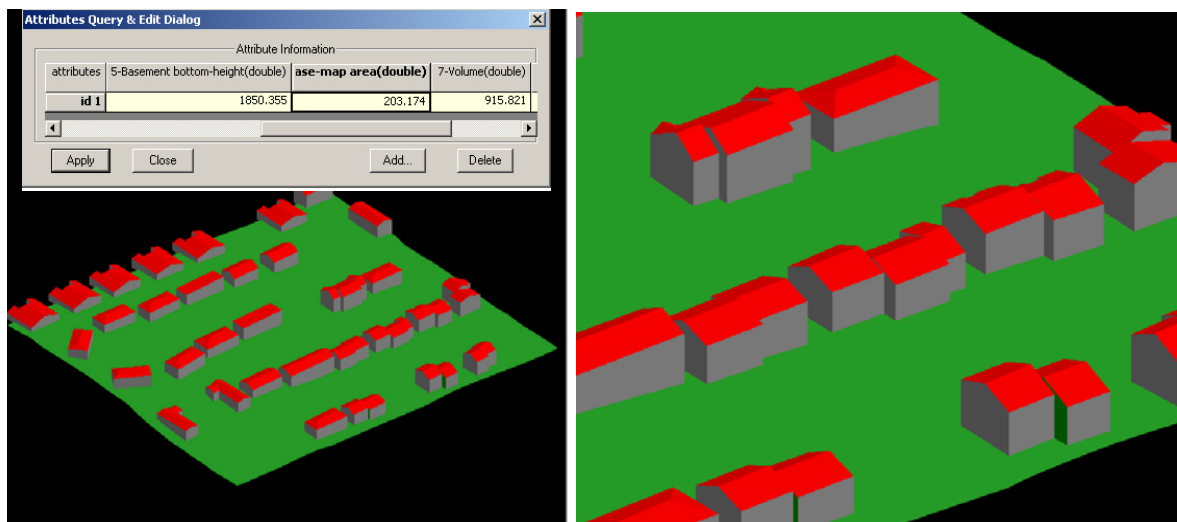


Figure 5. 3D models of residential houses with main roof structure (visualization in CC-Edit) and with automatically derived geometric attributes.

5 CONCLUSIONS

The paper reports the first results obtained with the evaluation of Worldview-1 stereo scenes for 3D information extraction. The test has been carried out on a stereo pair, consisting of a quasi-nadir and forward (30°) scene, acquired in November 2007 over Morrison (Colorado, USA).

Even with non-perfect conditions of the test area, our first experiences with DSM generation from WV-1 show generally satisfying results. The RMSE for the heights of an open area with separated houses and single trees is around 0.71m. The sensor has a good potential for 3D modeling of buildings, not only as simple block models but also with their main roof structures. Our future work will include the analysis of the dependency of the matching results from criteria like the geometry of the image acquisition, the local topography, local texture and others in combination with the small GSD. We also would like to extend the reference area and use other stereo models for more complete investigations.

6 ACKNOWLEDGEMENTS

The authors would like to thank Eurimage SpA for providing the data used for the work presented in this paper.

7 REFERENCES

- Baltsavias, E., Zhang, L. Eisenbeiss H, 2006. "DSM Generation and Interior Orientation Determination of IKONOS Images Using a Testfield in Switzerland.", *Photogrammetrie, Fernerkundung, Geoinformation*, (1):41-54.
- Eurimage, 2008. <http://www.eurimage.com/> (last visit October 2008)
- Gruen A., Akca. D., 2007. "Least Squares 3D Surface and Curve Matching", *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(3):151-174.
- Gruen, A., X. Wang, 2001. News from CyberCity Modeler, In E. Baltsavias, A. Gruen, L. Van Gool (Eds.) *Automatic Extraction of Man-made Objects from Aerial and Space Images (III)*, Balkema, Lisse, pp. 93-101.
- Poli, D., Zhang, L., Gruen, A., 2004. „SPOT-5/HRS stereo images orientation and automated DSM generation.“ *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(B1): 421-432.
- Wolff K., Gruen, A., 2007. "DSM Generation from early ALOS/PRISM data using SAT-PP". *High-Resolution Earth Imaging for Geospatial Information, Proceedings of ISPRS Hanover Workshop*, May 29- June 1 (on CD-ROM).
- Zhang, L., 2005 "Automatic Digital Surface Model (DSM) Generation from Linear Array Images", Ph. D. Dissertation, Report No. 88, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland.
- Zhang, L., Gruen, A, 2006. "Multi-image matching for DSM generation from IKONOS imagery", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 60, No. 3, pp.195-211.