

EVALUATION OF WORLDVIEW-1 STEREO SCENES

Daniela Poli^a, Kirsten Wolff^b, Armin Gruen^b

^aEuropean Commission, JRC, via Fermi 2749, Ispra (VA), Italy - daniela.poli@cec.europa.eu

^bSwiss Federal Institute of Technology, Institute of Geodesy and Photogrammetry
Wolfgang-Pauli-Strasse 15, 8093 Zurich, Switzerland - wolff,agruen@geod.baug.ethz.ch

Commission I, WG I/4

KEY WORDS: Worldview-1, orientation, matching, DSM, analysis

ABSTRACT:

This paper describes the investigations carried out at ETH Zurich on the very-high resolution optical satellite sensor WorldView-1. The dataset includes a stereopair acquired in November 2007 over Morrison, Colorado, USA, consisting of a quasi-nadir scene with mean GSD close to 51cm and an off-nadir scene, with GSD up to 73cm. The radiometry of the images was evaluated through the estimation of the noise level (standard deviation of the digital number) in non-homogeneous areas. For the orientation of the scenes, we used ground control points to estimate the parameters of an affine transformation and improve the orientation related to the Rational Polynomial Coefficients (RPC) provided in the metadata files. The next step included the automatic DSM generation with quality assessment. In all our tests the matching and the final DSM generation was realized with our in-house software package SAT-PP (Satellite Image Precision Processing) which has already been successfully used for other satellite sensors like IKONOS, ALOS/PRISM, Cartosat-1 and airborne linear and frame cameras. For the evaluation of the potential of a WorldView-1 stereo image pair for DSM generation, reference data are fundamental. As no DSM with a suitable resolution was available in this dataset, we generated the reference DSM from 1:15'000 scale infrared aerial images purchased from USGS. The aerial images were acquired in April 2002 with an analogue camera and scanned with a resolution of 14 μ m. No GCPs were available for the area covered by the aerial images. Therefore we defined a sufficient number of well-defined texture points in the oriented stereo pair of WorldView-1 as GCPs and determined their 3D coordinates using the orientation parameter of the satellite stereo pair. These points were used for the image orientation of the aerial images, which leads to a good relative orientation between the two different image sets. By this we investigated the quality of the matching process with WorldView-1 stereo image data mostly independent of the exterior orientation parameters. We will present here our accuracy assessment by visual and quantitative analysis (3D residuals distribution and statistics). The difference between the acquisition times of the two data sets (about 2.5 years) caused changes in the land cover, e.g. in correspondence of vegetation and buildings. In order to avoid the influence of temporal differences in our analysis, we compared the surface models in areas where no significant time depending changes could be expected. In addition, we evaluated the potential of the scenes for 3D building model extraction. The results are presented and discussed.

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 ground sample distance (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 +/- 45° off-nadir).

The aim of the present work is to evaluate a WV-1 stereopair for DSM and 3D information extraction. The study includes the analysis of the radiometric characteristics of the images, an accurate (subpixel) orientation, automatic generation of accurate digital surface modeling (DSM), the extraction of a 3D city model for a small part of the scenes and its evaluation in terms of level of detail.

For image pre-processing, orientation and DSM generation, the scenes have been processed using the software package SAT-PP (Satellite Image Precision Processing), distributed by ETH spin-off company 4DiXplorer AG, and other in-house developed routines, while for 3D building model extraction the software CyberCity-Modeler was used.

In this paper we report and comment the processing steps and results.

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), therefore few months after the launch of the satellite. Table 1 summarizes the main characteristics of the scenes. 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 1). The scenes were processed at level 1B, 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 quality of the images was not overall satisfactory, for example edges were not always detectable, and some radiometric artifacts were present (Figure 2).

In addition to the images, Eurimage provided the rough description of the location and geographic coordinates of 11 ground points (Figure 4), whose distribution is shown in Figure 3.

Table 1. Main characteristics of the WV-1 stereo scenes over Morrison (Colorado, USA).

	Nadir	Forward
Date of acquisition	25 November 2007, 5:40pm (UTC time)	
Number of columns	35170	35170
Number of rows	27892	23708
Mean GSD	0.51m	0.73m
Mean in-track angle	4.1°	32.6°
Mean cross-track angle	10.5°	11.9°
Scan direction	Forward	
Mean sun azimuth	162.1°	161.9°
Mean sun elevation	27.6°	27.6°



Figure 1. Zoom on detail from the nadir (top) and forward (bottom) views.



Figure 2. Artifacts in the images.

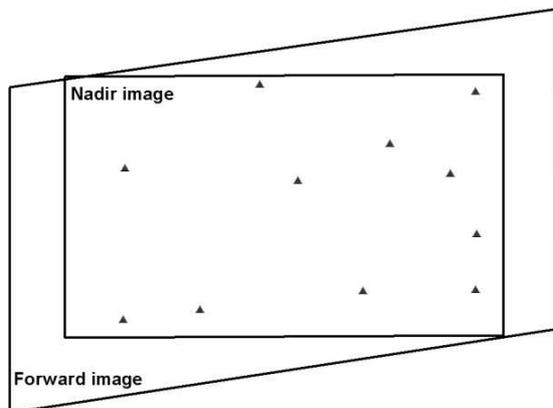


Figure 3. Distribution of GCPs in the stereo scenes.

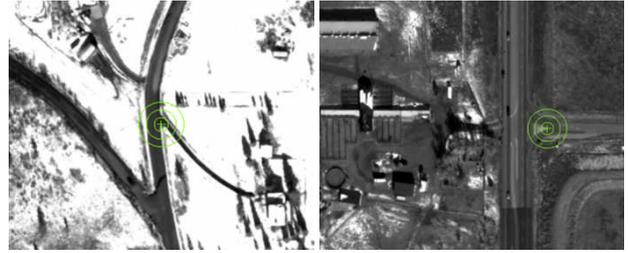


Figure 4. Example of location of GCPs.

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 ground control points (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. RADIOMETRIC PROCESSING

The two WV-1 scenes are processed at level 1B. At this level the images are radiometrically corrected, but we do not have information on the pre-processing applied. The radiometric quality of the images can be evaluated through their noise level in homogenous (e.g. sea and lake surfaces) and non-homogeneous areas (Baltsavias et al., 2001). In our case large homogeneous areas did not exist in the images, therefore we performed our analysis in non-homogeneous areas only. The noise level is estimated as the standard deviation of the Digital Number (DN). In non-homogeneous areas the grey level range is divided in bins, and the standard deviations are assigned to a bin according to the mean grey value of a window that moves in the area corresponding to the bin. The size of the window is small (in our case, 3x3 pixels) in order to get small homogeneous areas between edges. In each bin, the standard deviations are sorted, and the noise is estimated as the mean of the N% smallest standard deviation (in our case, $N=5$). The estimation of noise in each bin is reliable only if the N% sample number is sufficiently large (in our case, >100). Figure 5 shows the noise values for each bin in the original scenes (light grey). We observe that the noise reaches larger values in the forward scene, with respect to the nadir one; in the bins 0-127, 128-255, 1792-1919 and 1920-2047, that contain the majority of pixels, the two scenes present a noise in the range 2.8 – 3.2 grey levels. For noise reduction, an inverse weighted smoothing filter was used, in order to preserve the edges for DSM generation, and then the noise level was re-computed. We observed a significant reduction of the noise level in all bins (Figure 5).

The algorithms above described are implemented in a C++ software available at IGP for the radiometric analysis and improvement of images. The software and algorithm details are described in (Pateraki, 2005).

A great advantage of this software is that the analysis and processing is performed at the original dynamic range of the images (11 bit), so the full radiometric knowledge is preserved. Anyway for the next operations the images have been reduced to 8 bit with a non-linear iterative method, which overcomes the loss of information typical of linear reduction methods (Pateraki, 2005).

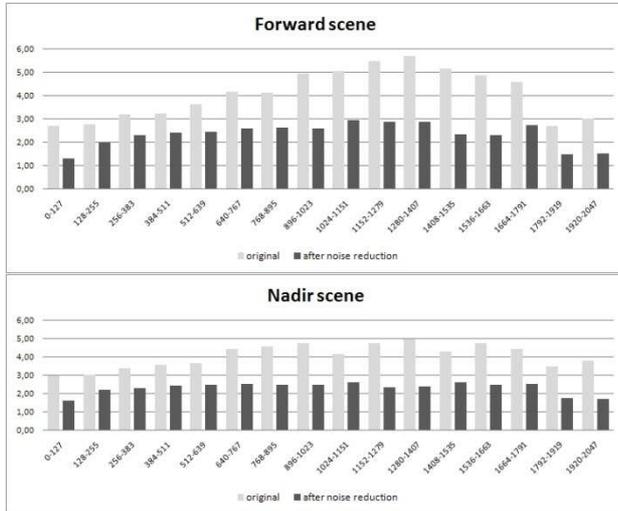


Figure 5. Noise levels in forward and nadir scenes for each bin, before (light grey) and after (dark grey) noise reduction.

4. GEOREFERENCING, DSM GENERATION AND QUALITY ANALYSIS

The geometric processing of WV-1 scenes for 3D information extraction was carried out with the program package SAT-PP. 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, for example, in (Poli et al., 2004) for SPOT-5/HRS, (Baltsavias et al., 2006) for IKONOS and (Wolff and Gruen, 2007) for ALOS/PRISM.

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 isolated areas. However, this might have only a local influence in some cases.

The next step was the image matching and DSM generation. The main advantage of SAT-PP is that it matches very densely with three different match-point primitives (regular image grid points, interest points and edgels). To improve the conditions for feature extraction and matching we preprocessed the images using a Wallis filter.

A first DSM was generated with a simple initial surface model and a grid space equal to 4 times the GSD, which leads to 2m. As expected, we visually detected local blunders in a shaded model of the DSM, caused by e.g. topography, image acquisition geometry, poor texture or y-parallaxes (see Figure 7 - top left). To avoid such blunders, SAT-PP has the possibility

to measure seed points in a stereomodel manually which improves the initial DSM locally in such problematic areas. (see Figure 7 – top right).

For quality evaluation, we compared the WV-1 DSM with the reference aerial image-based DSM, described in 0. 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 6), and deleted some evident new buildings to reduce the influence of temporal changes.



Figure 6. Original WV-1 nadir scene of the test area.

Figure 8 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; in this operation 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. The histogram of the results in Figure 9 shows the normal distribution of the residuals and that we do not have any systematic errors. The 3D RMSE was split into X, Y, Z components, whose statistics are summarized in Table 2. 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. In the Table we see that the main component of the residual error is in Z direction. Figure 9 visualizes the distribution of the 3D residuals (Z-component of Euclidean distances), after co-registration.

We verified that local blunders are due to temporal differences, low texture (snow cover) and the y-parallaxes.

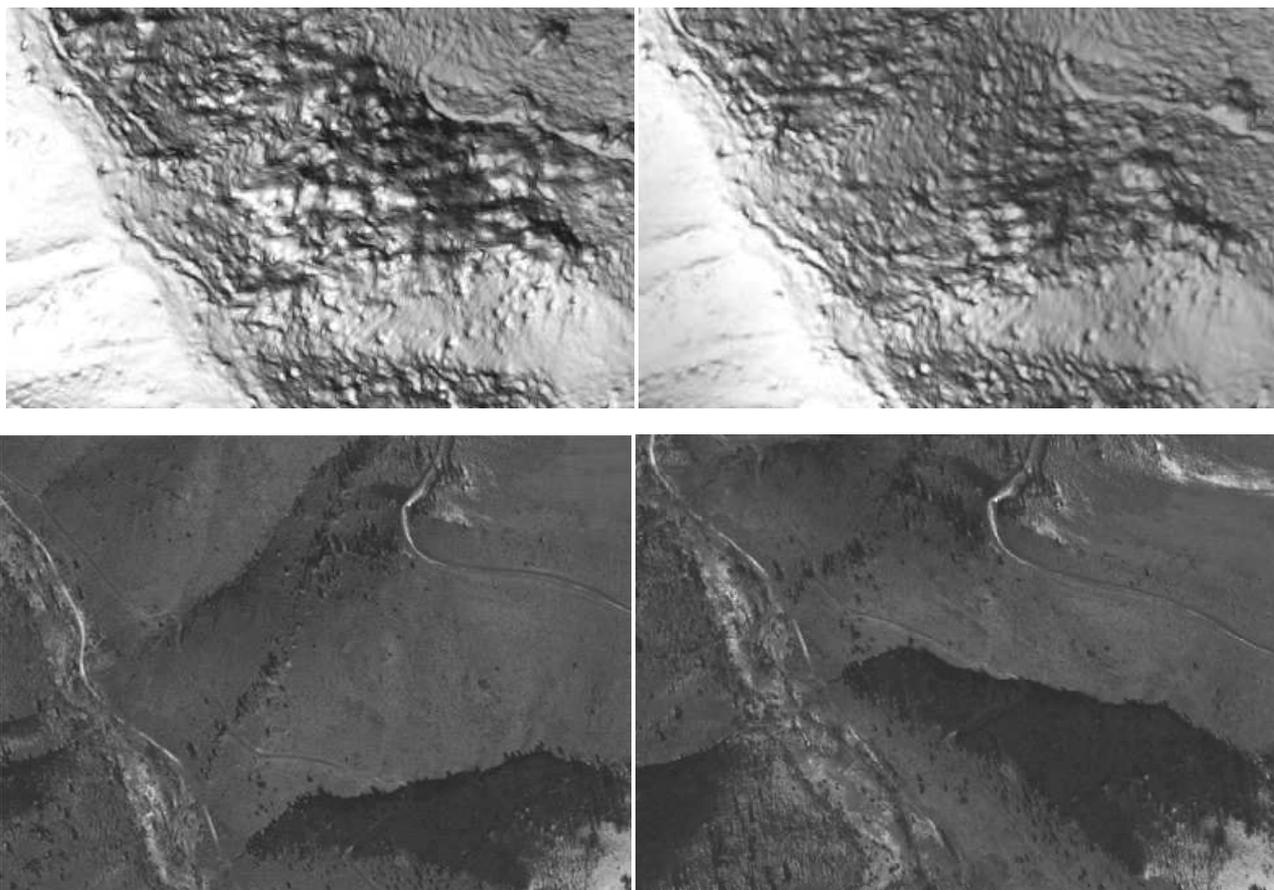


Figure 7. Top left: Shaded model of a critical area with blunders (grid space 2m). Top right: Improvement by measuring additional seed points Bottom: original stereo images of the same area (left: nadir, right: forward).

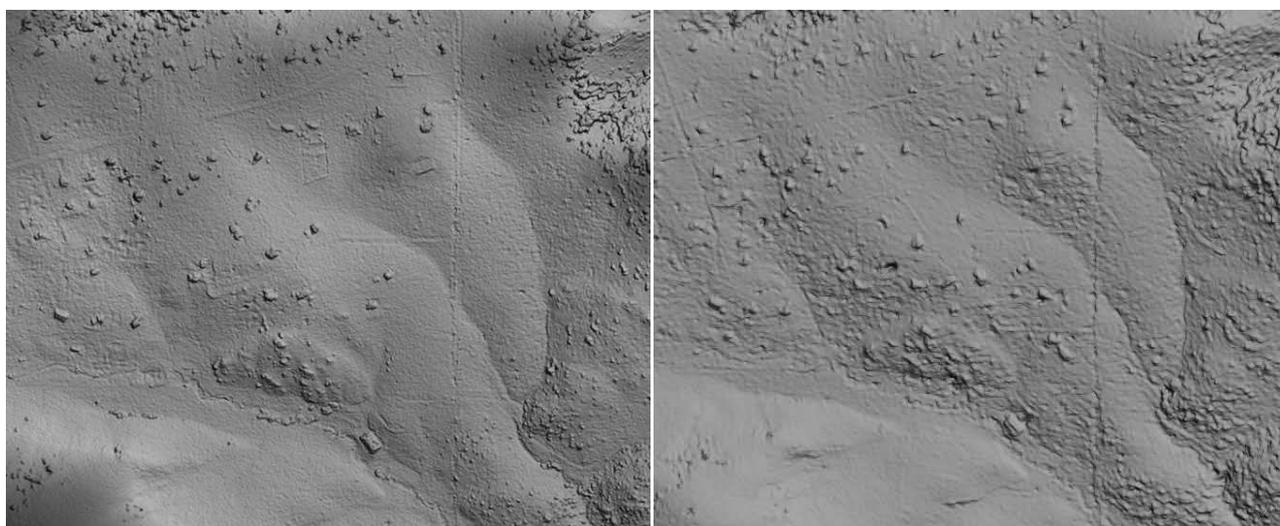


Figure 8. Left: shaded model of the reference DSM (1.1km by 1.0km, grid spacing 1m). Right: shaded model of the WV-1 DSM (grid spacing 2m).

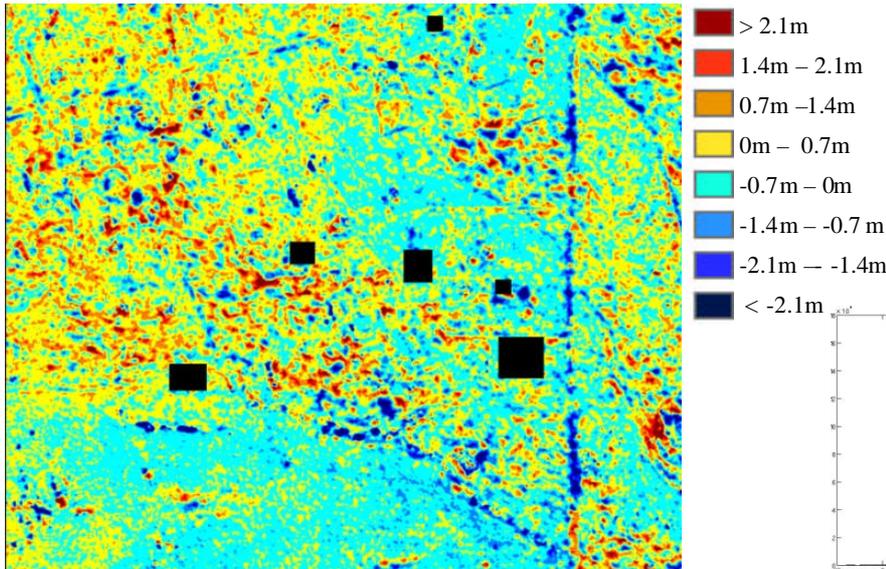


Figure 9. Residuals of the Z-component of the Euclidean distance between reference and WV-1 DSMs. Areas with new buildings, deleted and not considered in the analysis, are shown as black boxes.

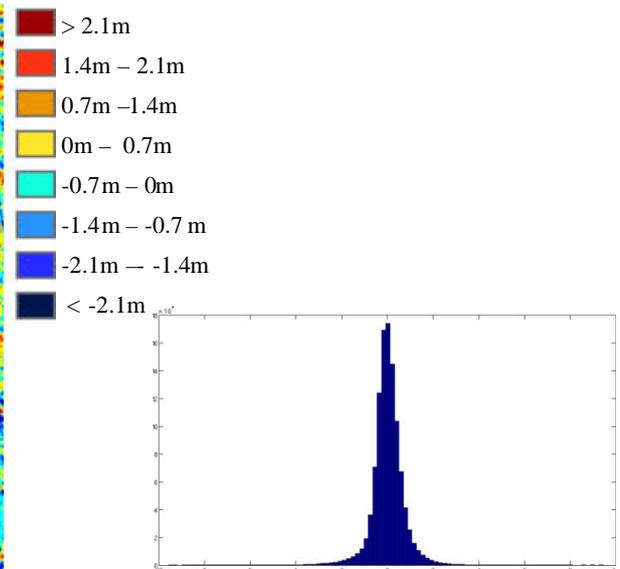


Figure 10. Histogram of the residuals of the WV-1 DSM for the test area.

Table 2. 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

5. 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 could be distinguished, therefore the buildings could be modeled with their main roof structures (Figure 11). It was still not possible to measure roof superstructures, like small dormers. Geometric building parameters as heights, volumes, areas and façade inclinations could be automatically calculated.

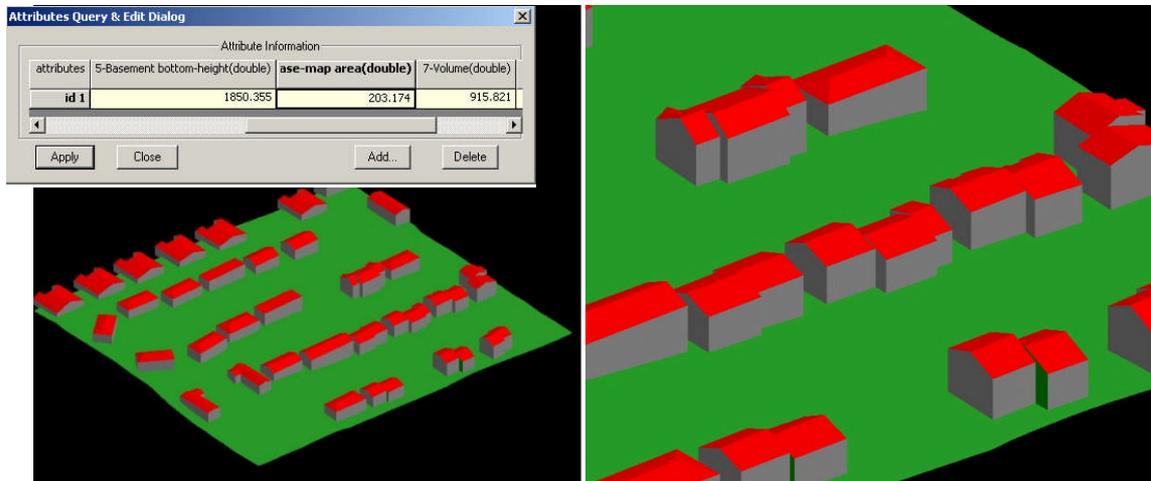


Figure 11. 3D models of residential houses with main roof structure (visualization in CC-Edit) and with automatically derived geometric attributes.

6. CONCLUSIONS

The paper reports the 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). The first part of the processing included the radiometric quality evaluation and image orientation. Then image matching was applied for DSM generation. Even with non-perfect conditions of the test area, our 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 too, 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.

AKNOWLEDGEMENTS

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

REFERENCES

- Baltsavias, E., P., Pateraki, M., Zhang L., 2001. Radiometric and geometric evaluation of IKONOS Geo images and their use for 3D building modeling. Proceedings of ISPRS Workshop "High Resoluition Mapping from Space,, Hannover, Germany. On CDROM.
- 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.
- M. Pateraki, 2005. Adaptive Multi-Image Matching for DSM Generation from Airborne Linear Array CCD Data. Diss., Technische Wissenschaften ETH Zurich, 2005, IGP Mitteilung N. 86.
- 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.