

# Operational Generation of High Resolution Digital Surface Models from Commercial Tri-Stereo Satellite Data

SEBASTIAN CARL, SAMUEL BÄRISCH, FLORIAN LANG, München  
PABLO D'ANGELO, HOSSEIN AREFI, PETER REINARTZ, Oberpfaffenhofen

## ABSTRACT

High resolution stereo satellite imagery is now well suited for the creation of digital surface models (DSM) in urban areas due to recent developments in data resolution, quality and collection capabilities. A system for highly automated and operational DSM and orthoimage generation based on WorldView-2 imagery is presented using dense matching methodology, with emphasis on the usage of tri-stereo data for the generation of optimized DSMs. Due to constraints given by the three images, which allow six different pair-wise matchings (including left and right matching of each pair), robust results containing only very few outliers can be generated. The proposed system processes level-1 stereo scenes using the rational polynomial coefficients (RPC) universal sensor model. The RPC are derived from orbit and attitude information and exhibit a lower absolute accuracy than the ground resolution of approximately 0.5 m. In order to use the images for orthorectification or DSM generation, a bias or affine RPC correction is required, which can be achieved through a bundle adjustment using further scenes of the covered area from different dates. Furthermore these DSMs can be used to generate higher level products like digital terrain models (DTM), extracted single 3D objects like buildings and for automatic 3D change detection analysis. DLR-IMF and GAF AG developed and implemented within a close co-operation an operational workflow which now provides operational services based on multi source tri-stereo satellite data. The DSM processing is shortly described, some results of generated DSMs are shown and also examples for higher level products are given in the paper.

## 1. INTRODUCTION

Digital surface models (DSM) can be efficiently generated with automatic image matching from optical stereo images. Detailed reconstruction of small structures from very high resolution stereo images requires the use of dense stereo matching algorithms that are not based on correlation of image windows. Very good results with a reasonable runtime have been achieved with the semi-global matching (SGM) algorithm (Hirschmüller, 2008). The results typically contain a small number of outliers, if only a single stereo pair is available. Errors in the DSM normally need to be cleaned by manual editing, which is very costly for large area DSM generation. Classical DSM filtering methods based on the evaluation of local statistics are not suitable for these kinds of outliers, as they affect many well recovered man-made objects such as buildings.

Highly agile very high resolution satellite data like WorldView-2 and Pleiades data are capable of acquiring multi view stereo datasets with a ground resolution of 50 cm in a single pass. Good results can be achieved through merging the DSMs generated from several image pairs by dense stereo methods (Hirschmüller, 2008, d'Angelo 2011). Despite its simplicity SGM has also been proven to work very well for large amounts of satellite and aerial imagery.

Our multi-view matching algorithm uses a 2.5D DSM as final scene representation. We use Census as photo-consistency measure (Zabih et al. 1994), as previous studies have shown its robustness and high performance in real world applications (Hirschmüller, et al. 2009). SGM strives to minimize the data cost and a regularization term. Instead of strong assumptions on the local surface shape, a global energy function  $E$  is minimized for all disparities (local shifts between the stereo pair)  $D$ . SGM performs a semi global optimization by aggregating costs from 16 directions, and finds an image  $D$  which leads to a low energy  $E$ :

$$E(D) = \sum_p \left( C(p, D_p) + \sum_{q \in N_p} \begin{cases} 0, & |D_p - D_q| = 0 \\ p_1, & |D_p - D_q| = 1 \\ p_2, & |D_p - D_q| > 1 \end{cases} \right)$$

The function  $C$  is the photo-consistency/data term between the image pixels for each pixel location  $p$  and possible disparity  $D_p$ . The second term of  $E$  penalizes disparity changes in the neighborhood  $N_p$  at each position  $p$ . The penalty  $p_1$  is added for all disparity changes equal to one pixel. At larger discontinuities (disparity change  $> 1$  pixel), a fixed cost  $p_2$  is added. This cost function favors similar or slightly changing disparities between neighboring pixels, and thus stabilizes the matching in image areas with weak contrast, but also allows large disparity jumps in areas with high contrast. When matching multiple images with SGM, several image pairs are matched first. The pairwise image matching allows the detection of occluded regions and avoids merging most mismatched regions. The individual disparity maps are then fused into a single DSM by using a median filter. For operational processing, a variant of the Semi-Global Matching (SGM) algorithm was developed and thoroughly tested (d'Angelo et al. 2012). Modifications over the standard SGM algorithm include a robust, hierarchical search strategy, which dynamically reduces the search range for flat areas and results in faster computation and denser DSMs. The resulting DSMs can still contain some outliers in problematic regions, such as clouds or water areas. These can be detected using a region based outlier detection algorithm. This works particularly well if three or more images are available, yielding redundant height information.

As an example Fig. 1 shows the results that can be achieved by two to three stereo images and by different convergence angles for a dense urban area in Munich (Hauptbahnhof is located in the bottom of the images). It can be seen, that for urban areas a lower convergence angle is of advantage to avoid occluded areas and tri-stereo data can improve the DSM generation substantially in comparison two just image pairs. The optimal viewing angles for a tri-stereo data are  $-15^\circ/10^\circ$ ;  $0^\circ$ ;  $+10^\circ/15^\circ$  together with a low off-nadir (side looking) angle.

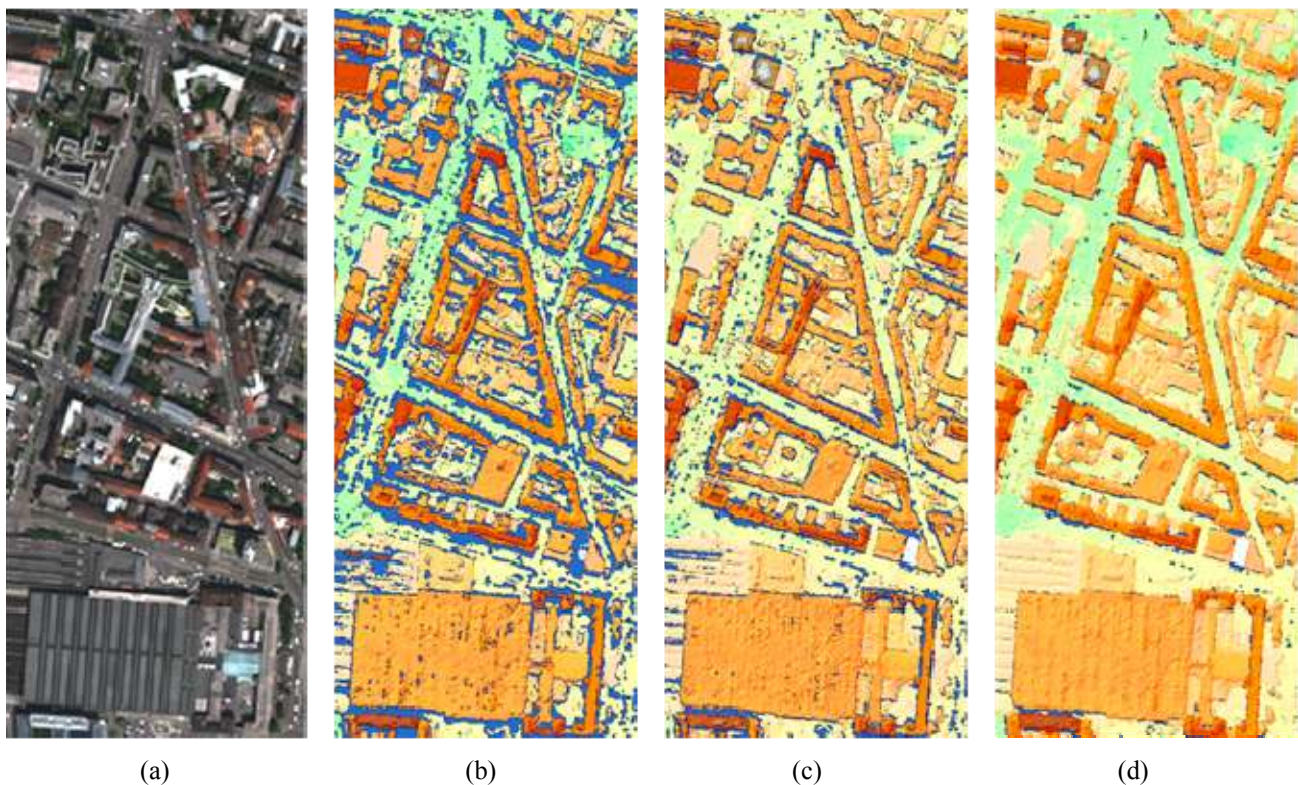


Figure 1: Worldview-2 image and DSMs generated by different convergence angles and by stereo-pairs and tri-stereo data. (a) WorldView-2 multispectral image; (b) DSM from  $24^\circ$  stereo angle; (c) DSM from  $12^\circ$  stereo angle; (d) DSM from tri-stereo. Blue color indicates unmatched or occluded areas.



## 2. RESULTS FROM TRI-STEREO MATCHING AND QUALITY IMPROVEMENTS

We present an example of tri-stereo matching and DSM generation for the city of Tunis (Fig. 2). Besides the generation of the DSM by pair-wise matching of the stereo scenes, two more layers which give information on the local quality of the DSM are generated.

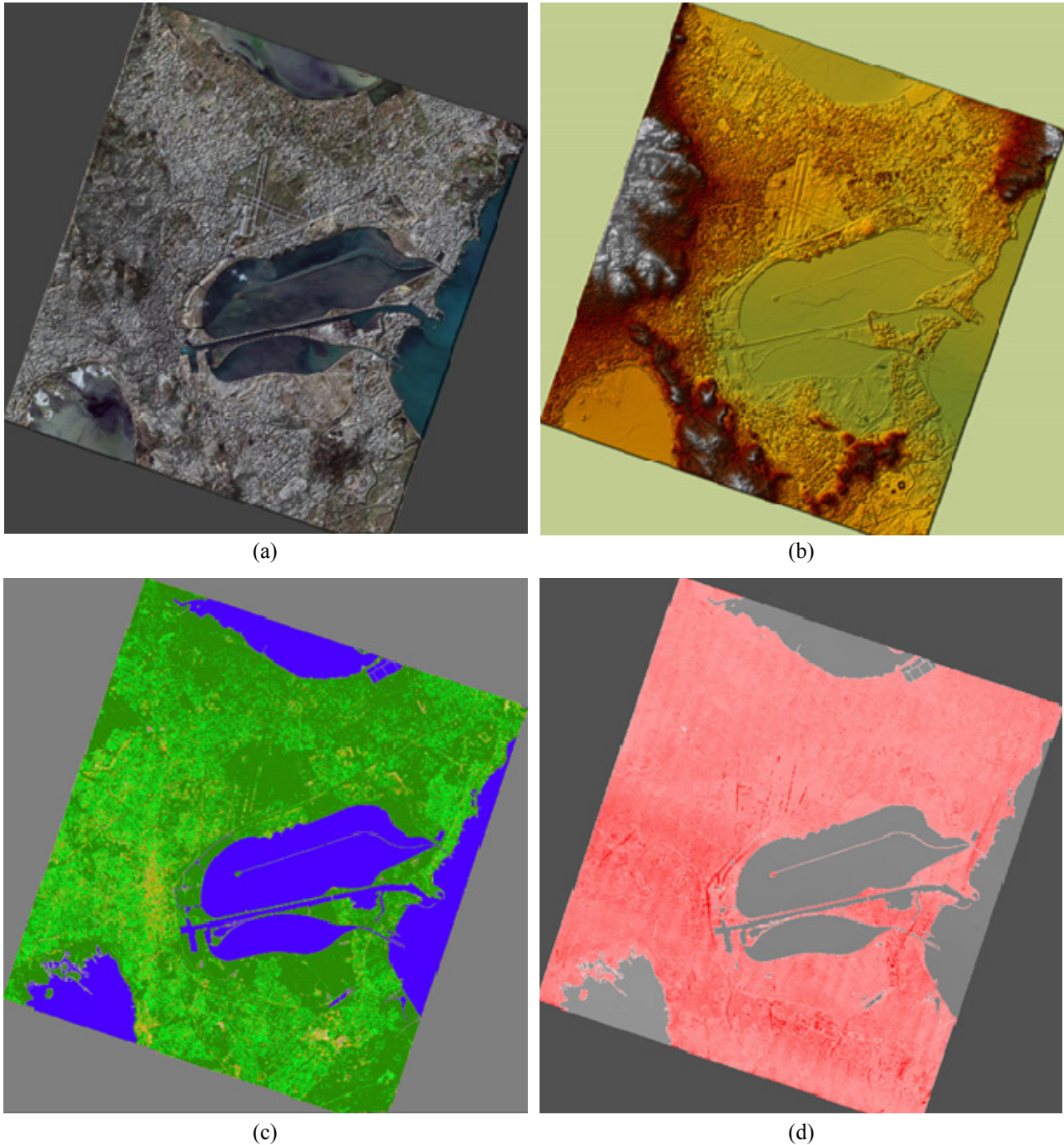


Figure 2: WorldView-2 data of Tunis: (a) multispectral image; (b) generated DSM; (c) number of matched image pairs, dark green:6, light green:5, red:4, yellow:3; (d) local standard deviation.

One layer (Fig. 2c) shows the number of matched pairs (left/right and right/left matching counts each) for each pixel, which implicitly shows partly occluded areas and other places which are difficult to match. It can be seen, that most pixels can be matched five or six times, but some dense

urban areas can be matched only three or four times. A second quality layer (Fig. 2d) shows the standard deviation of the generated heights through matching results for each single pixel. It can be seen that again the dense urban areas show a higher standard deviation and also some linear structures with higher values exist.

Despite the redundant height information, the processed DSM still contains some artifacts, mostly caused by moving objects. Typically, those artifacts manifest themselves by small peaks and pits along the driving direction of moving objects (Fig. 3). Due to the small time gap between the acquisition of the first four channels and the last four channels of the WorldView-2 satellite, a change detection using the principal component analysis can be applied on every single orthorectified image. The resulting change detection layers are used as an indicator to speed up a manual editing step, which aims to identify, mark and interpolate the affected areas using the inverse distance weighting interpolation. Due to the impossibility to match water bodies and the target to achieve an aesthetic and realistic DSM, the water bodies are identified on the nadir-near image using a region-growing technique. The heights of the identified areas are set to one level. To guarantee an aesthetic and realistic DSM, the heights of rivers are manually stepped down, with regards to the surrounding topography. As a last finishing step to complete the DSM, a manual quality control step is indispensable to identify remaining artifacts.

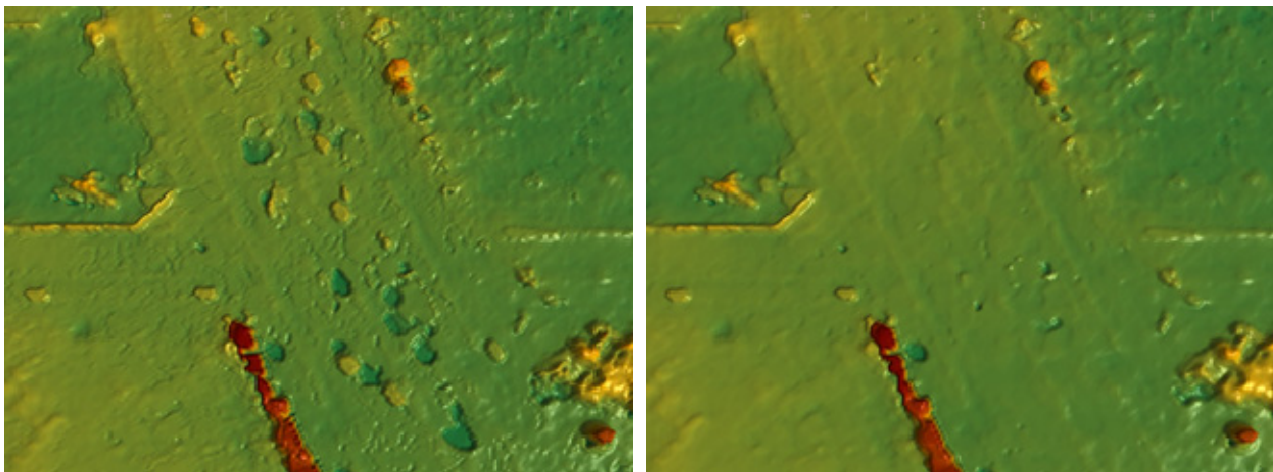


Figure 3: generated DSM of Tunis before and after post-processing: left: DSM with visible artifacts on a road; right: interpolated and median (3x3 window) filtered DSM.

### 3. HIGHER LEVEL PRODUCTS

**3.1 DTM generation:** The generation of digital terrain models consists of two major steps: generation of digital surface models (DSM) from stereo scenes and hierarchical image filtering for DTM generation. In the second step, the DSM pixels are classified into ground and non-ground regions using an algorithm motivated from gray-scale image reconstruction to suppress unwanted elevated pixels. Non-ground regions, e.g. 3D objects, as well as outliers are iteratively separated from the ground regions. The DTM generation algorithm is in first instance motivated by the “fast hybrid gray-scale reconstruction” algorithm described in Vincent (1993). The goal of the morphological reconstruction procedure is to suppress unwanted regions with high intensity while preserving the intensity in the regions of interest, which are seeded with “markers” in order to initialize the algorithm. Image segmentation based on gray-scale reconstruction is at the core of the algorithm for filtering non-ground regions from the elevation image (Arefi et al. 2011). Fig. 4 shows a result of the generated DTM compared to the DSM.



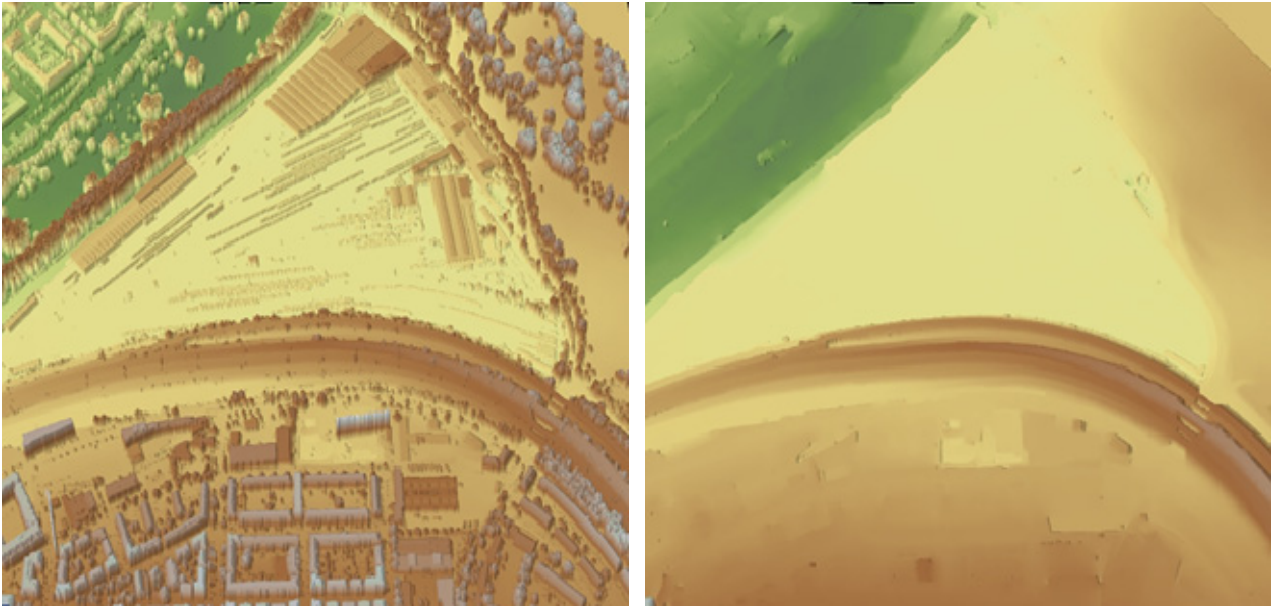


Figure 4: DTM generation: left DSM; right DTM.

**3.2 Building reconstruction:** A further step after generating a DSM and a DTM is to derive single building objects from these 2.5D models. To achieve this, several steps are necessary before being able to extract ridge lines and building boundaries. Several different methods have been developed to achieve a good representation of single buildings by using DSM, DTM and panchromatic data (Sirmacek et al. 2012, Arefi et al. 2012). The automatic 3D building reconstruction algorithm shown here comprises the following four major steps, details can be found in Arefi et al. (2013):

1. Ridge-based decomposition of building parts
2. Projection-based reconstruction of parametric roofs
3. Approximation of the polygons relating to flat roof segments
4. Merge parametric and prismatic models

Fig.5 shows the generated roof for a larger building in a WorldView-2 scene of Munich city. Although the contour of the building is not correctly represented through the DSM, the model can be generated quite well through using also the information from the panchromatic image.

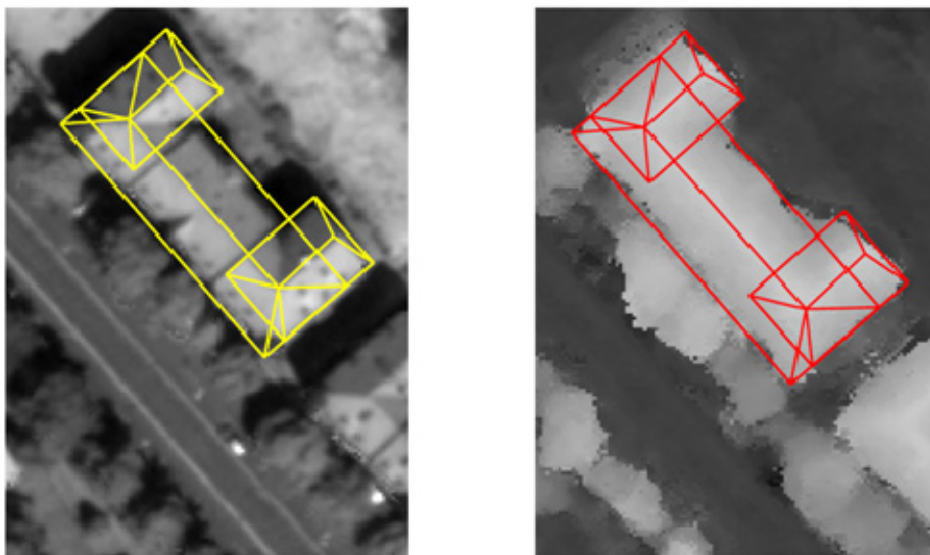


Figure 5: WorldView-2 image (left) and generated DSM (right) with generated model as overlay.

In Fig. 6 the resulting CAD models for eight buildings are displayed. The results are very promising since the overall accuracies in comparison to data from the Munich authorities are: Ridgeline: Mean -0.05 m; StdDev: 0.5 m; Eaves height: Mean -0.45 m; StdDev: 0.4 m.

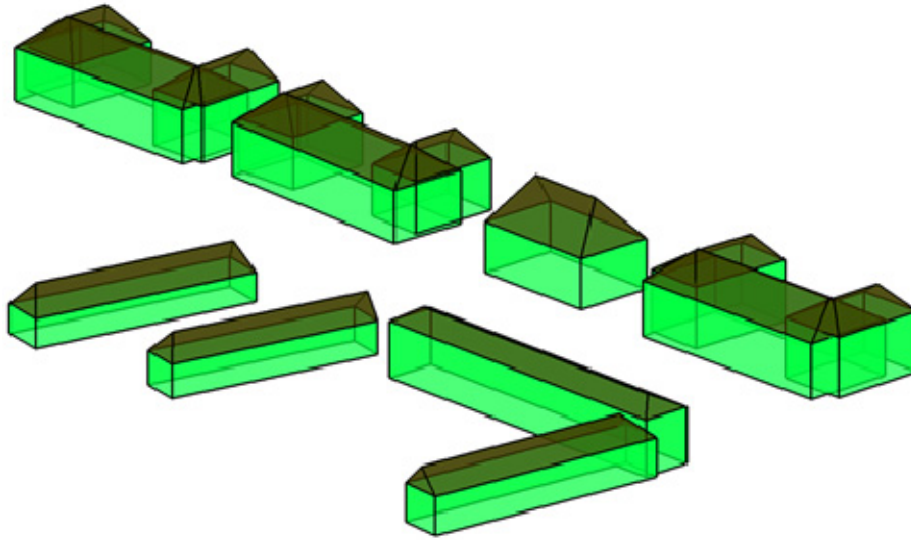


Figure 6: CAD models for 8 buildings, derived from WorldView-2 data.

**3.3 3D change detection:** Automatic change detection from satellite image data is a difficult task since many changes are due to seasonal or illumination changes, which are not the real changes of interest. If two DSMs of different dates are available, these data can be additionally used to improve change detection capabilities. Several change detection methods using a combination of height changes from DSMs, optimized region boundaries and spectral change from panchromatic images have been developed. Firstly, DSMs with 1 m GSD are generated with SGM from WorldView-2 or other VHR stereo data as described in the introduction. After orthorectification of the near-nadir images the initial regions are obtained from the segmentation results of these orthorectified images of two dates (Tian et al. 2013).

The change analysis regions are generated through a combination of these segmentation results. A rule based region merge approach is provided to reach better consistency between the homogeneous regions of the two dates. Several height and spectral features are then extracted at region level. Additionally to the above mentioned methods, DSMs of the two dates are used in a newly developed combined change detection procedure. Fig. 7 shows an example for an industrial area. The change probability maps (middle column) already give a good indication about possible changed objects, while the rightmost column shows the changes on a per building level. Region or building level results lead to better performance in comparison to pixel level results (Tian et al. 2013).

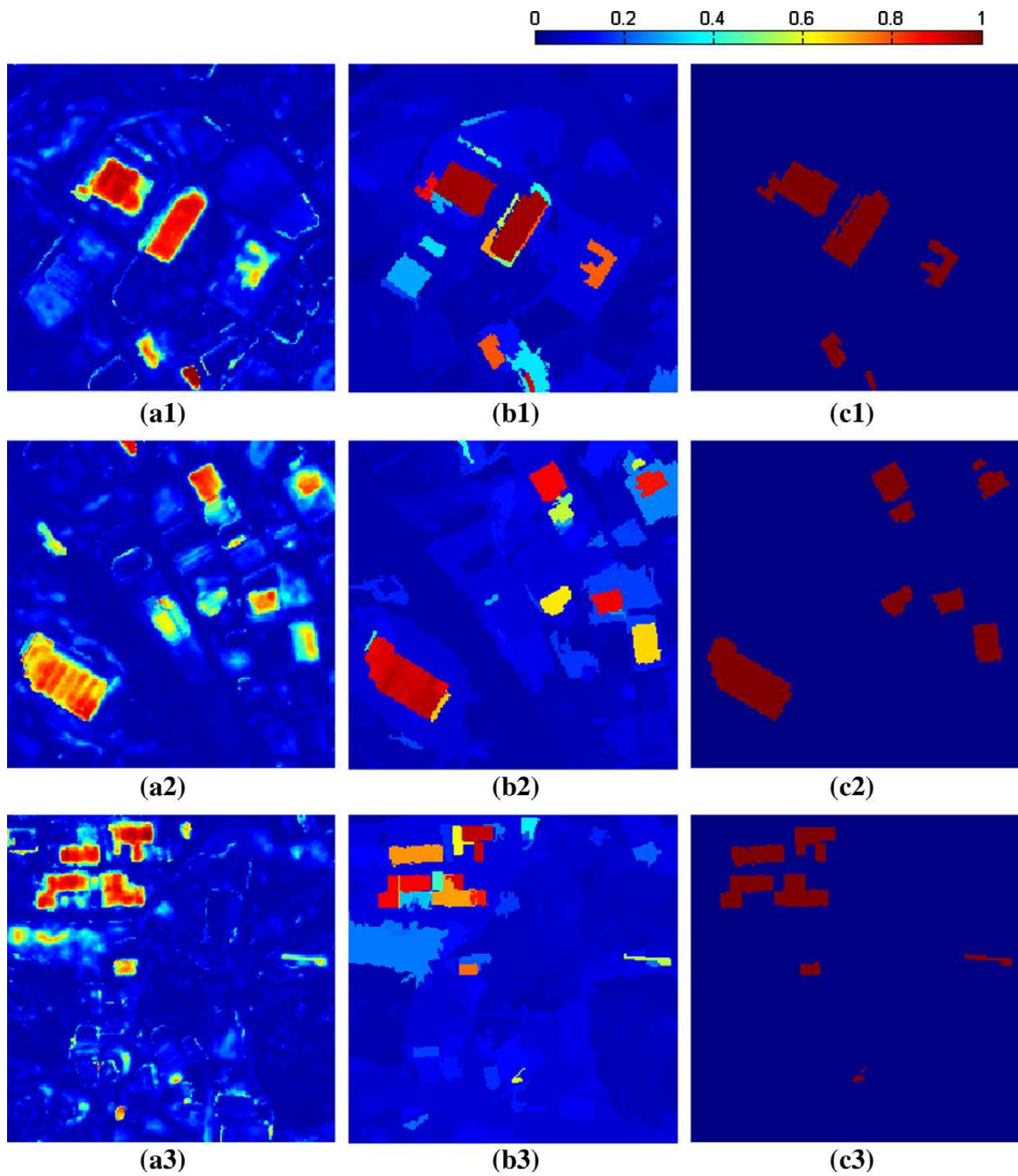


Figure 7: Change detection result for an industrial area: (a1–a3): pixel-based change probability maps; (b1–b3): region-based change probability maps; (c1–c3): change detection masks; Blue: Change probability = 0; Red: Change probability = 1.

#### 4. CONCLUSION

The experiments and results shown in this paper describe the potential of the new generation of spaceborne stereo data. It becomes obvious that the quality of DSM generated from tri-stereo data with low convergence angles together with dense matching methodologies results in a new level of details that can be reached by optical images from satellites. The methods are now developed to a point that fully automatic processing is possible for commercial applications. The described quality assured DSM based on tri-stereo satellite data from several sources is now operational available and commercialized through GAF AG. Especially for areas where no airborne data (optical or Lidar) are available, spaceborne data can give very detailed 3D information. With new developments in high level processing the results are not limited to DSM generation but also DTM, single building reconstruction and 3D change detection can be successfully performed with these data.

#### 5. ACKNOWLEDGEMENTS

We would like to thank European Space Imaging (EUSI) for providing Worldview-2 data for scientific purposes.

#### 6. REFERENCES

- Arefi, H., Reinartz, P. (2013): Building reconstruction using DSM and orthorectified images. *Remote Sensing* Vol. 5, pp. 1681-1703.
- Arefi, H., Reinartz, P. (2012): Multi-Level Building Reconstruction for Automatic Enhancement of High Resolution DSMS. In: *Proc. XXII ISPRS Congress 2012, XXXIX-B2*, pp. 11-16.
- Arefi, H., d'Angelo, P., Mayer, H., Reinartz, P. (2011): Iterative approach for efficient digital terrain model production from CARTOSAT-1 stereo images. *Journal of Applied Remote Sensing*, vol. 5, pp. 1-19.
- D'Angelo, P., Kuschik, G. (2012): Dense multi-view stereo from satellite imagery. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 6944-6947.
- D'Angelo, P., Reinartz, P. (2011): Semi-global matching results on the ISPRS stereo matching benchmark. In: *ISPRS Hannover Workshop, Int. Archives Photogrammetry, Remote Sensing and Spatial Information Sciences, Hannover, Germany, 2011*, vol. XXXVIII-4/W19.
- Hirschmüller, H. (2008): Stereo processing by semi-global matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 328-341.
- Hirschmüller H., Scharstein, D. (2009): Evaluation of stereo matching costs on image with radiometric differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 9, pp. 1582-1599.
- Sirmacek, B., Taubenböck, H., Reinartz, P., Ehlers, M. (2012): Evaluation of automatically generated 3-D city models based on six different DSMs from airborne and space-borne sensors. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 1, pp. 59-70.



- Tian, J., Reinartz, P., d'Angelo, P., Ehlers, M. (2013): Region-based automatic building and forest change detection on Cartosat-1 stereo imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 79, pp. 226-239.
- Vincent, L. (1993): Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *IEEE Trans. Image Process*, vol. 2, pp. 176-201.
- Zabih, R., Woodfill, J. (1994): Non-parametric local transforms for computing visual correspondence. In: *ECCV '94: Proceedings of the Third European Conference-Volume II on Computer Vision*, London, UK, 1994, Springer-Verlag, pp. 151-158.