Advances in DSM Generation and Higher Level Information Extraction from High Resolution Optical Stereo Satellite Data

P. Reinartz, J. Tian, H. Arefi, T. Krauss, G. Kuschk, T. Partovi, P. d'Angelo

German Aerospace Center (DLR), Remote Sensing Technology Institute, 82234 Wessling, Germany

Earsel Special Interest Group "3D Remote Sensing & Urban Remote Sensing"

KEY WORDS: Optical stereo data, DSM generation, Image matching, DTM generation, 3D object extraction, 3D change detection

ABSTRACT:

The automatic generation of digital surface models (DSM) of urban areas from high and very high resolution (VHR) stereo data from satellites is still a research issue. Image matching algorithms from computer vision have been introduced and adopted to satellite imagery in recent years. These algorithms do not work using local optimisation like area based matching but try to optimize a global cost function. Analysis shows that matching approaches based on epipolar images like semi-global matching (SGM) and new methods using total generalized variation (TGV) yield the best results. Especially satellites like Worldview-2, GeoEye-1 and Pleiades exhibit very high spatial resolution and geometric quality and can therefore be used to generate DSMs with good properties. If several stereo images from one orbit are available, a combination of DSMs generated by different stereo pairs leads to even better results. Using these DSMs, which already show urban area features in high detail, further higher level information can be extracted and new products can be generated.

The main focus of this contribution is to present and analyse several derived information products which can be generated using the DSMs generated from satellite stereo data together with panchromatic and multispectral images. In many applications the product of interest is not the DSM but the digital terrain model (DTM), which exhibits not the height of objects like buildings or trees but the underlying terrain. In this paper a selected method and example for generating DTM from DSM is presented. A further issue is to extract real 3D objects like buildings using DSM and the derived DTM data. Different methods have been developed and are presented using typical examples in densely build city areas including validation results. At least larger buildings can be automatically extracted with sufficient accuracy; even the roof shape (gable or flat roof) can be extracted and modelled. Since automatic change detection is generally a very difficult topic in image processing, information on height or 3D properties are of advantage when used in the change detection procedure. Therefore, if stereo data sets from different dates are available, automatic 3D change detection can be performed using the corresponding DSMs. Since the change detection results depend very much on the quality of the DSMs they generally have to be improved using the multispectral information. Two methods have been developed and are shown and compared using examples from dense urban and industrial areas. The presented results show that due to the good data quality and resolution of satellite image data and the corresponding DSMs, it has become feasible to derive higher level and detailed geo-information regarding 3D relevant object classes and change detection.

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. First results using satellite data have been achieved in Krauß et al. (2005 and 2008) and d'Angelo et al. (2008). With the improvements developed through the semi-global matching (SGM) algorithm (Hirschmüller, 2008) these results became more and more reliable and robust. But typically some not matchable areas and a small number of outliers is still present, especially 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. An operational processor for Cartosat-1 (IRS-P5) stereo data including automatic outlier detection has been presented in d'Angelo et al. (2009).
Highly agile very high resolution satellite data like WorldView-2, GeoEye-1 and Pleiades data are capable of acquiring multi-view stereo datasets with a ground resolution of 50 cm in a single pass. If several views exist, good results can be achieved through merging the DSMs generated from single image pairs by dense stereo methods (Hirschmüller, 2008,

d'Angelo 2011). Despite its simplicity SGM has also been proven to work very robust for large amounts of satellite and aerial imagery.

Our multi-view matching algorithm leads to a 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_p. 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 neighbourhood 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 favours similar or slightly changing disparities between neighbouring 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, all possible image pairs are matched first. The pairwise image matching allows the detection of occluded regions and avoids merging 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 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, cloud shadows or water areas. These outliers 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 (the main railway station 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°/10°; 0°; +10°/15° together with a low off-nadir (side looking) angle.

Within its lead of a working group of ISPRS Commission I, DLR has setup an open stereo matching benchmark site near Barcelona, where matching algorithms can be evaluated against ground truth data provided by the Institut Cartogràfic de Catalunya. It currently includes a Worldview-1 stereo pair donated by Digital Globe and Cartosat-1 stereo pair provided by Euromap GmbH (Reinartz 2010).

While the existing operational system allows highly detailed DSM generation, small errors remain, especially for building contours and in shadow areas. This mainly results in noisy building boundaries. First promising results have been achieved by extending the energy functions with additional edge or surface priors that allow better reconstruction in these problematic areas. These energy functions are more complex and harder to optimize with the SGM algorithm. By using a total generalized variation (TGV) regularizer, slanted surfaces can be reconstructed with higher quality, see Fig. 2, than with the fronto-parallel smoothness term used by SGM (Kuschk, 2013a). Convex optimization is used for energy minimization. It has larger computational cost than SGM but can be implemented efficiently on graphical processing units (Kuschk 2013b).

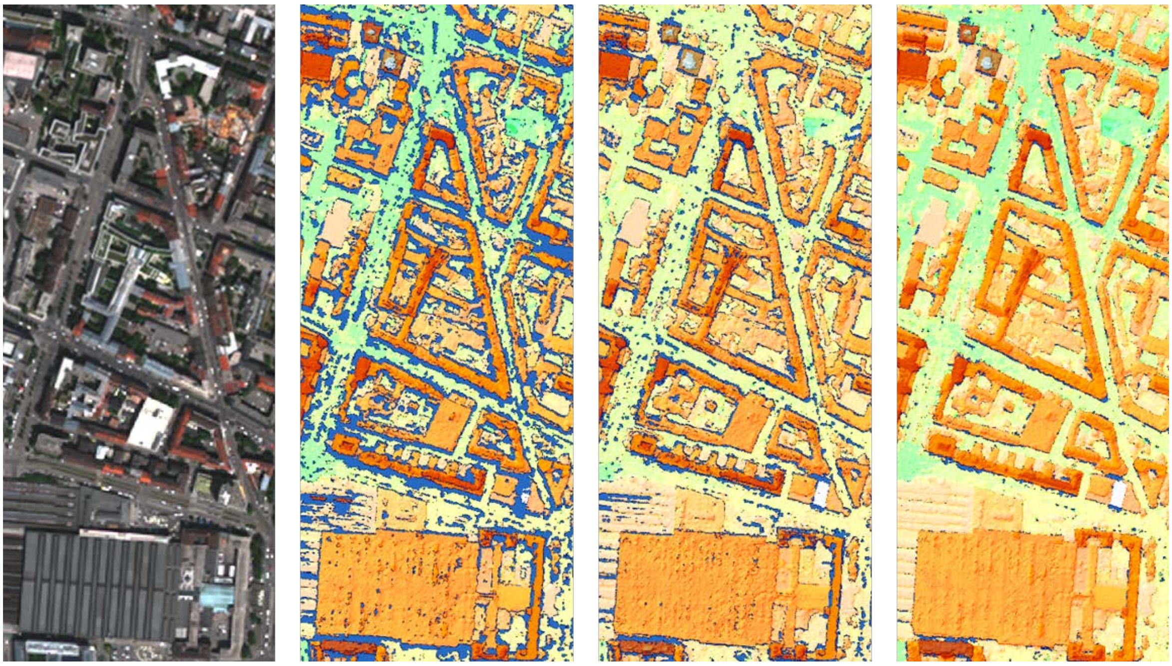

Fig. 1: WorldView-2 image and DSMs generated by different convergence angles and by stereo-pairs and tri-stereo data. (a) WorldView-2 multispectral image; (b) DSM from 24° stereo angle; (c) DSM from 12° stereo angle; (d) DSM from tri-stereo. Blue color indicates unmatched or occluded areas.

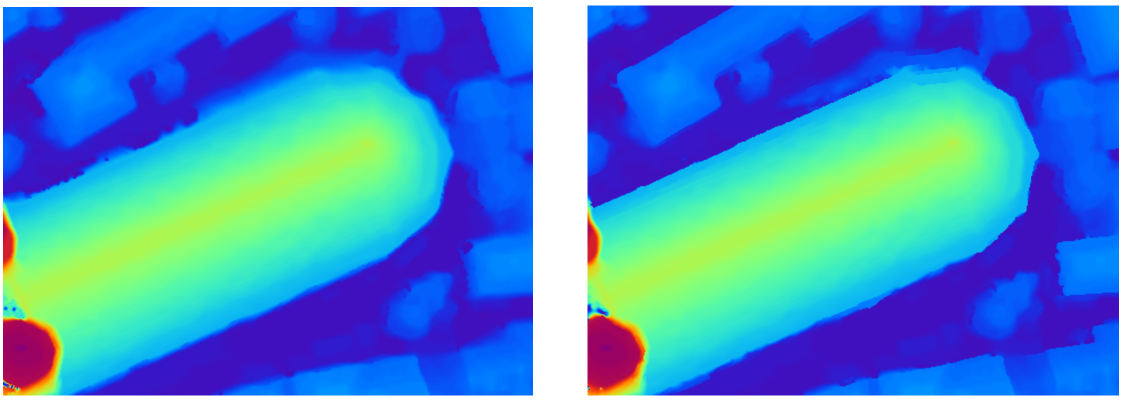

Fig. 2 Improvements of the representation of roof edges through introducing additional edge or surface priors and the total generalized variation regularizer; left SGM result, right: TGV result with priors

PRODUCTS GENERATED FROM DIGITAL SURFACE MODELS FROM SPACE

DTM generation:

A Digital Terrain Model (DTM), as an important post product of DSM, is produced after removing non-terrain regions (like buildings and trees) from the DSM. The presented DTM generation algorithm is in first instance motivated by the "fast hybrid grey-scale reconstruction" algorithm described in Vincent (1993). But here a new hierarchical approach for generating a DTM from DSM data has been developed. Image reconstruction based on geodesic dilation is the core of the algorithm. Morphological reconstruction based on geodesic operation employs two input images; marker and mask. In geodesic dilation the marker image (J) is dilated and the resulting image is forced to remain below the mask image (I).

$$\delta_I^{(1)}(J) = (J \oplus B) \wedge I,$$

The image reconstruction is achieved by applying geodesic dilations until stability is reached.

$$\delta_I^{(n)}(J) = \underbrace{\delta_I^{(1)}(J) \circ \delta_I^{(1)}(J) \circ \cdots \circ \delta_I^{(1)}(J)}_{n \text{ times}}.$$

Most commonly a marker image is generated by subtracting a constant value from the mask image which is the original DSM. To avoid problems caused by an improperly selected offset our solution is to use a sequence of constant offset values to create a sequence of marker images. By subtracting the reconstructed image from the mask image the normalized DSM (nDSM) is obtained. A first classification of terrain and off-terrain points is carried out by binarising the nDSM. Any point (in the nDSM) above zero is collected as an off-terrain point. For further analysis features are determined for each off-terrain region. Here the size of the region and the local height difference along the boundary of the region are utilized as feature descriptors. By subtracting maximum and minimum values in local 3 by 3 windows, moved over the mask image along the region boundary, the local height differences are found. The average local height difference is calculated and used as the second feature. Classification is now using these two features. A new interpolation is calculated using the terrain points as input which basically produces a Digital Terrain Model (Arefi et al. 2011). Fig. 3 shows a result of the generated DTM compared to the DSM.

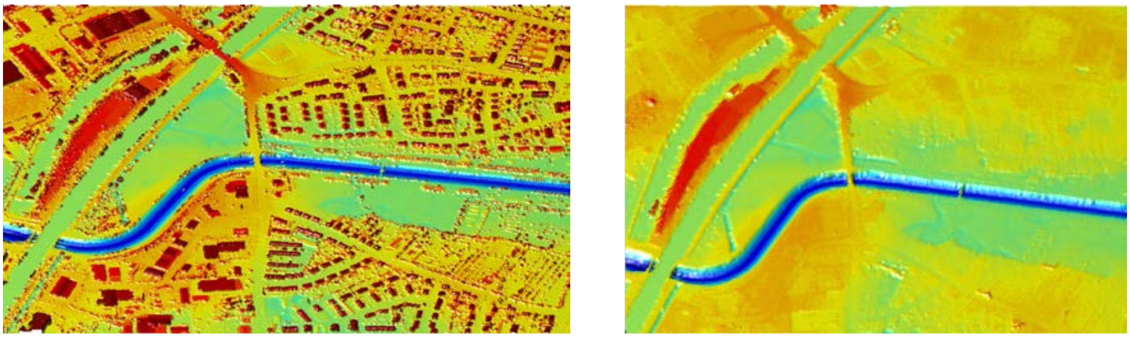

Fig 3: DTM generation by geodesic dilation: left DSM; right DTM

Building reconstruction

Extracting buildings from DSM data generated from space borne stereo data is a brand new topic in photogrammetry. Just a few different methods have been developed for satellite data to achieve a good representation of single buildings by using DSM, DTM and panchromatic data (e.g. Sirmacek et al. 2012, Arefi et al. 2012).The developed method shown here aims at simplifying the 3D reconstruction of the building blocks by decomposing the overall model into several smaller ones corresponding to each building part. The model enables representing the parametric roof shapes such as gable and hip models. Accordingly, the automatic 3D building reconstruction algorithm comprises the following major steps:

- Ridge-based decomposition of building parts
- Projection-based reconstruction of parametric roofs
- Prismatic model generation related to flat roof segments
- Merge parametric and prismatic models and refine the corner nodes

The process begins with feature extraction followed by ridge line reconstruction using orthoimage and height information. According to each ridge line, a projection-based algorithm is employed to transfer the 3D points into 2D space by orthogonal projection of the corresponding pixels of each building part onto a 2D plane that is defined based on the orientation of the ridge line. Based on the type of the roofs, a predefined 2D model is fitted to the data, and in the next step, the 2D model is extended to 3D by analysing the third dimension of the points. A final model regarding the parametric roof structures of the building block is defined by merging all the individual models and employing some post-processing refinements regarding the coinciding nodes and corners to shape an approximated building. Additionally prismatic models with flat roofs are provided regarding the remaining objects which are not containing ridge lines. Finally, all parametric and prismatic models are merged to form a final 3D model of the building. The results are very promising for larger buildings, since the approximated heights of ridge and eve lines could be approximated to about the pixel size (0.5 m) for WorldView-2 data over Munich city center (Fig. 4). More details can be found in Arefi et al. (2013). 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.

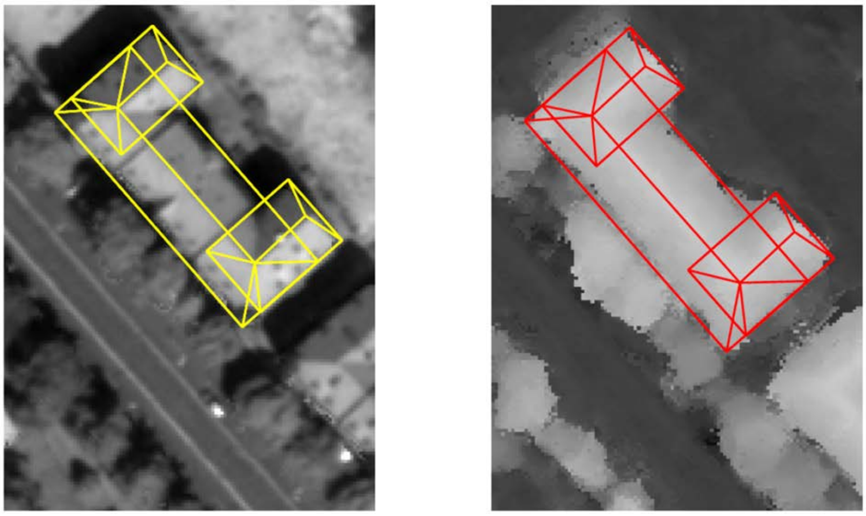

Fig. 4: WorldView-2 image (left) and generated DSM (right) with generated model as overlay

In Fig. 5 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; Std: 0.5 m; Eaves height: Mean -0.45 m; Std: 0.4 m.

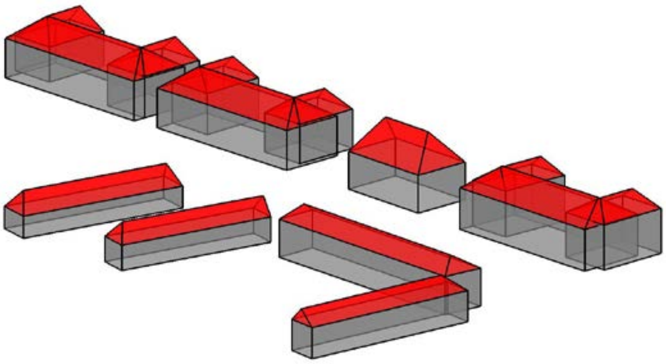

Fig. 5: CAD models for 8 buildings, derived from WorldView-2 data.

Another approach for building extraction is based on a model-driven strategy (Partovi et al. 2013). For reaching reliable results, refined orthorectified panchromatic images are introduced into the process as additional data. The idea of this method is based on ridge line extraction and analysing height values in direction of and perpendicular to the ridgeline direction. After applying pre-processing to the orthorectified data, some feature descriptors are extracted from the DSM, to improve the automatic ridge line detection. Applying RANSAC a line is fitted to each group of ridge points. Finally these ridge lines are refined by matching them or closing gaps. In order to select the type of roof model the heights of point in extension of the ridge line and height differences perpendicular to the ridge line are analysed. After roof model selection, building edge information is extracted from canny edge detection and parameters derived from the roof parts. Then the best model is fitted to the extracted façades of the roofs based on the detected type of model. Each roof is modelled independently and final 3D buildings are reconstructed by merging the roof models with the corresponding walls.

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. The major challenge for optical image based change detection is that only the changes related to the reflectance values and/or local textural changes can be detected. Without the height information from a DSM these changes in vertical direction, such as building height or forest growth are easily overlooked. Such information can play an important role in different applications such as disaster assessment (e.g. after an earthquake) and urban area construction and/or destruction monitoring. DSM assisted change detection are further of interest for city monitoring and disaster damage evaluation (Tian et al. 2011).

Limited to the quality of DSMs generated from satellite stereo imagery, it is hard to reach precise change detection result using only the DSMs and a DSM subtraction. Therefore, depending on the quality of the DSM, the availability of multispectral channels and the requirements of the change detection task, several approaches for change detection of buildings and forest areas by fusing of DSM and optical image information have been developed.

The initial 3D change detection approach is built on refinement of the height difference map. The refinement can be processed based on only the DSM, or assisted by the extracted land cover information from multi-spectral images.

A two-step region based change detection procedure is developed in fusing the height changes from DSM, optimized region boundaries and spectral changes from panchromatic images. In the first step segmentation on orthorectified panchromatic images is performed to obtain initial regions. Regions from two dates (Date1 and Date2) are combined to get an initial segmentation map. To correct the over-segmentation resulting from the region combination, a region merging strategy is proposed to reach a reasonable segmentation level. In the second step, several height and spectral features were extracted at region level to extract the relevant changes (Tian et al. 2013).

Another DSM assisted change detection approach is built on decision fusion. The general term of decision fusion consists of fusion of different change indicators. These change indicators can be extracted directly from the images, but can also be a change detection result from selected methods. The aim of fusing decisions of separate indicators is to increase the overall performance. A decision fusion based change detection method has been developed by the joint use of height changes and Kullback-Leibler divergence similarity measure between the original images. The Dempster-Shafer fusion theory is adopted to combine these two change indicators to improve the accuracy. In addition, vegetation and shadow classifications are used as no-building change indicators for refining the change detection results. In the end, an object based building extraction method based on shape features is performed (Tian et al. 2014).

Fig. 6 shows an example of the first method 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.

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 DSMs generated from stereo data with low convergence angles and especially with multiple-views, together with dense matching methodologies, result in a new level of details that can be reached by optical images from satellites. 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 derivation, single building reconstruction and 3D change detection can be successfully performed with these data.

ACKNOWLEDGEMENTS

The authors would like to thank European Space Imaging (EUSI) for providing the WorldView-2 stereo data of Munich, Germany, and GAF AG for providing the Cartosat-1 data of Istanbul, Turkey, for scientific purposes.

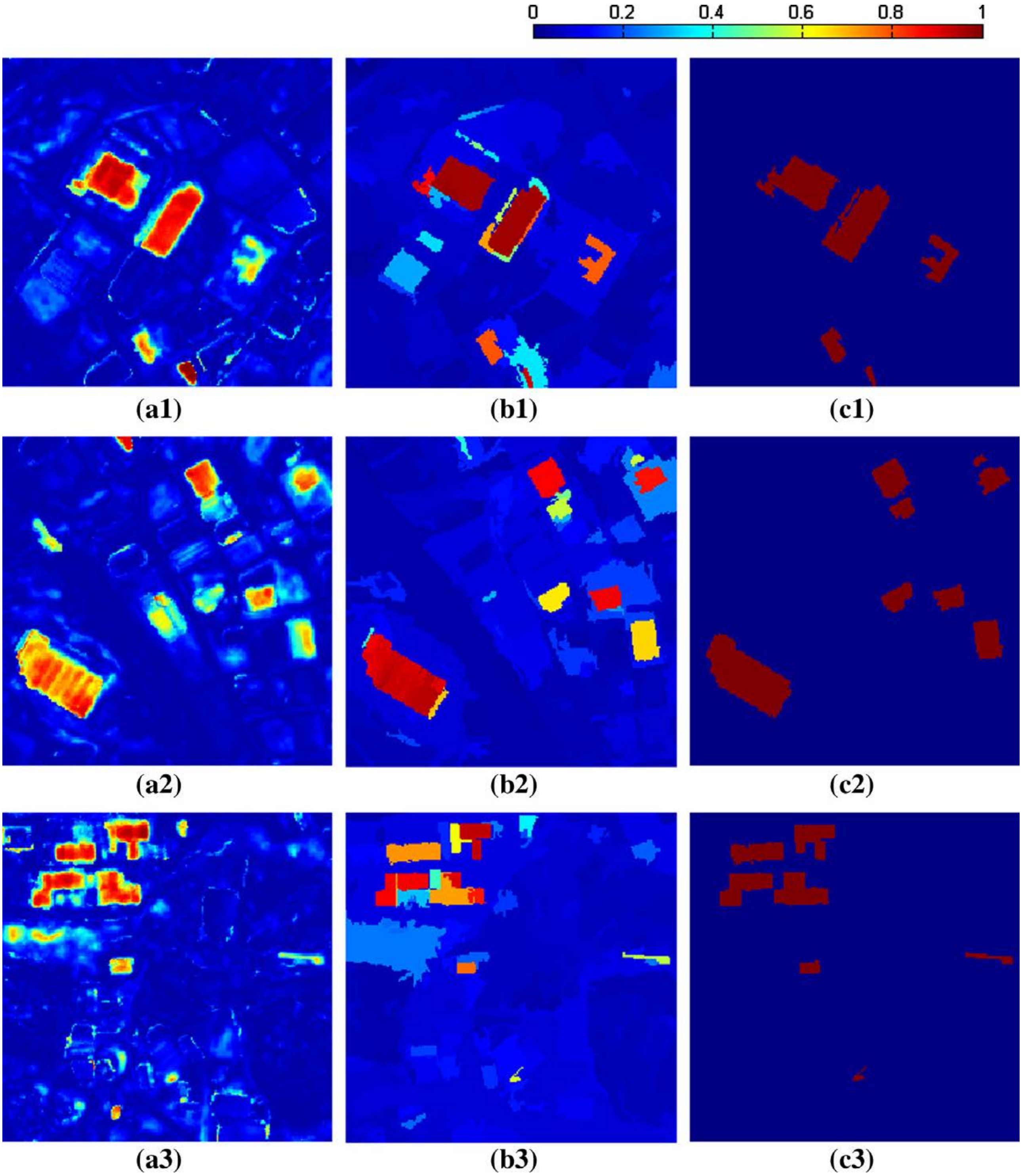

Fig. 6: Cartosat-1 data, Istanbul: Change detection result for an industrial area: (a1–a3): pixel-based change probability maps; (b1–b3): region-based change probability maps; (c1–c3): change detection masks; Blue: Change probability = 0; Red: Change probability = 1

REFERENCES

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

Arefi, H., Reinartz, P. (2012): Multi-Level Building Reconstruction for Automatic Enhancement of High Resolution DSMS, in Proc. XXII ISPRS Congress 2012, Int. Arch. Photogramm. Remote Sens. and Spatial Inf. Sci., Vol. XXXIX-B2, pp. 11-16

Arefi, H., Reinartz, P. (2013): Building Reconstruction using DSM and Orthorectified Images, Remote Sensing, Vol. 5, pp. 1681-1703

D'Angelo, P., Lehner, M., Krauss, T., Hoja, D., Reinartz, P. (2008): Towards Automated DEM Generation from High Resolution Stereo Satellite Images, Int. Arch. Photogramm. Remote Sens. and Spatial Inf. Sci., Vol. XXXVII. Part B4, pp 1137-1142

D'Angelo, P., Schwind, P., Krauss, T., Barner, F., Reinartz, P. (2009): Automated DSM Based Georeferencing of Cartosat-1 Stereo Scenes, ISPRS Hannover Workshop: High-Resolution Earth Imaging for Geospatial Information, Int. Arch. Photogramm. Remote Sens. and Spatial Inf. Sci. Vol. XXXVIII-1-4-7/W5, pp. 1-6

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, pp 1-6

D'Angelo, P., Kuschk, G. (2012): Dense multi-view stereo from satellite imagery, IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE Xplore, pp. 6944 – 6947

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

Krauß, T., Reinartz, P., Lehner, M., Schroeder, M., Stilla, U. (2005): DEM Generation from Very High Resolution Stereo Satellite Data in Urban Areas Using Dynamic Programming, ISPRS Hannover Workshop on "High-Resolution Earth Imaging for Geospatial Information", Hannover, Germany, pp. 1-6

Krauß, T., Lehner, M., Reinartz, P. (2008): Generation of Coarse 3D Models of Urban Areas from High Resolution Stereo Satellite Images, Int. Arch. Photogramm. Remote Sens. and Spatial Inf. Sci., Vol. XXXVII. Part B1, pp 1091-1098

Kuschk, G., Cremers, D. (2013a): Fast and Accurate Large-scale Stereo Reconstruction using Variational Methods, IEEE International Conference on Computer Vision Workshops, IEEE Xplore, pp. 700-707

Kuschk, G., d'Angelo, P. (2013): 3D Reconstruction Chain - From Images to 3D City Model. SMPR Conference 2013, Int. Arch. Photogramm. Remote Sens. and Spatial Inf. Sci., Vol. XL-1/W3, 247-251

Partovi, T., Arefi, H., Krauß, T., and Reinartz, P. (2013): Automatic Model Selection for 3D Reconstruction of Buildings from Satellite Imagary, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-1/W3, pp. 315-320

Reinartz, P., d'Angelo, P., Krauß, T., Poli, D., Jacobsen, K., Buyuksalih, G. (2010): Benchmarking and Quality Analysis of DEM Generated from High and Very High Resolution Optical Stereo Satellite Data, Int. Arch. Photogramm. Remote Sens. and Spatial Inf. Sci. Vol. XXXVIII part 1, pp. 1-6

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., Chaabouni-Chouayakh, H., Reinartz, P. (2011): 3D Building Change Detection from High Resolution Spaceborne Stereo Imagery, International Workshop on Multi-Platform Multi-Sensor Remote Sensing and Mapping, pp. 1-4

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

Tian, J., Cui, S., Reinartz, P. (2014): Building change detection based on satellite stereo imagery and digital surface models. IEEE Transactions on Geoscience and Remote Sensing, 52 (1), pp. 406-417

Vincent, L., 1993. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. IEEE Trans. Image Process. Vol. 2(2), pp. 176–201

Zabih, R. and 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, Springer-Verlag, London, UK, pp. 151–158