AUTOMATIC DEM GENERATION AND 3D CHANGE DETECTION FROM SATELLITE IMAGERY

Thomas Krauß, Pablo d'Angelo, Jiaojiao Tian, Peter Reinartz

DLR – German Aerospace Center, 82234 Oberpfaffenhofen, thomas.krauss@dlr.de

ABSTRACT

In this paper we present a method for fully automatic generation of digital surface models (DSM) from very high resolution (VHR) satellite imagery and the consecutively automatic change detection from the derived 3D information.

Common change detection methods are normally only based on spectral change detection. These methods will fail for e.g. comparing summer and winter scenes with the latter covered with snow. Introducing the digital elevation model into the change detection process will allow for a more detailled object modeling and also for the possibility to detect more sophisticated changes like volume estimation of mining activities.

Here we present the involved methods for the generation of the high resolution surface models, the fusion and classification and finally the automatic 3D change detection. The methods are applied to some VHR stereo test data sets and the results are evaluated for quality and usefulness for automatic information derivation from large data sets.

1. INTRODUCTION

With the launch of more and more high resolution satellites like the Sentinel series the available amount of earth observation data increases so rapidly that new fully automatic methods for information derivation from these data is increasingly essential. Very high resolution satellites like the commercial WorldView, GeoEye or Pléiades satellites and their predecessors QuickBird and Ikonos expanded the ground sampling distance for the first time below one metre. Such resolutions allow the automatic generation of high resolution DSMs even from urban areas. Through the newly available dense matching methodologies very high detailed DSMs can be derived for urban areas, especially if several viewing directions are acquired by the satellite within one orbit. Consecutive monitoring of areas and automatic derivation of the digital surface models make an automatic detection of changes more efficient than just using the spectral information. So e.g. the image based automatic change detection between vegetation and snow covered regions delivers too much false positive results where volume changes like new buildings, excavations or dumps mostly will not be detected correctly. In such cases the generation of DSMs give an additional information for detection of changes.

A newly developed method for automatic bundle adjustment of VHR stereo or multistereo images and the subsequent derivation of a DSM of high accuracy is presented, using the DLR developed semi global matching (SGM) method. By developing a methodology of fusing the information from imagery with the high resolution DSM from different epochs an automatic detection of three dimensional changes together with a detailled volume estimation can be performed.

1.1. Overview

In this article we first present an operational method for generating a digital surface model (DSM) from two or more satellite images from very high resolution (VHR) satellites – i.e. satellites with a ground sampling distance of about 2.5 m or below. Using two DSMs and the associated ortho images of different epochs allow for a more sophisticated change detection than only using imagery. In the second part we describe a new approach for an also operational and fully automatic method for deriving change detection maps.

In the Experiments section we apply the presented change detection method to different VHR image sets of same or different sensors and discuss the results finally in the Results section.

1.2. Preliminary work

The presented system is based on a method for dense stereo matching called Semi global matching

Proc. 'ESA Living Planet Symposium 2013', Edinburgh, UK
9–13 September 2013 (ESA SP-722, December 2013)

described in [1] and refined in [2] which generates a dense digital surface model (DSM). Based on this resulting DSM a digital terrain model (DTM) containing only the ground height and no objects like buildings and trees is derived using approaches as described in [3] or [4]. Also ortho images can be calculated using the stereo imagery and the resulting DSM. In [5] an approach for fusing information of DSMs and multispectral images for improving 3D change detection performance is presented. In our method the ortho images are classified and objects get extracted as described in [6]. Afterwards a robust change detection method like the one described in [7] is applied to the imagery and DSM data. For building change detection also shapes and and object based change detection as presented in [8] will lead to better results. On the other hand the forest change detection based on only one-channel panchromatic Cartosat images is very challenging. As shown in [9] a special region based method has to be developed to cope with the missing multispectral information.

2. METHOD

2.1. DSM generation

Before applying the semi global matching [1] to a set of different multitemporal (multi-) stereo image pairs first a bundle block adjustment as described in [2] has to be done to assure a highly accurate (below one pixel) relative orientation of all of the images. These relatively corrected images are transformed to epipolar geometry (see fig. 1).

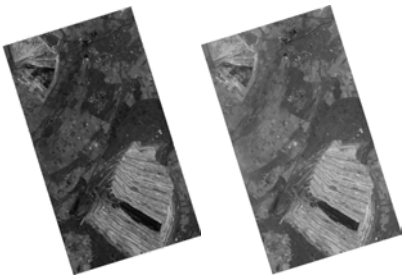

Figure 1. Pan channels of a GeoEye stereo pair transformed to epipolar geometry; area 16×8 *km, Jülich, Germany, acquired 2012-08-12; north to the upper right corner*

In this geometry height changes show only up in horizontal shifts. So the result of applying the semi global matching (see fig. 2) to each stereo pair is a so called dense disparity map $D(p)$ fitting on one of the stereo images and containing a relative pixel distance of each pixel to it's mate in the stereo partner image in epipolar (horizontal) direction.

Transforming these disparities $D(p)$ to heights using the sensor model and reprojecting the resulting disparity map back on one of the stereo images result in a so called height-map $H(p)$. This height-map fits exactly on one of the stereo images and contains for each pixel the derived absolute ellipsoid height or no-value if the stereo processing fails for this pixel. The latter occurs usually in so called "occlusions". These are areas which are only visible in one of the stereo images like walls of houses or narrow streets. See for clarification of the terms used also fig. 3.

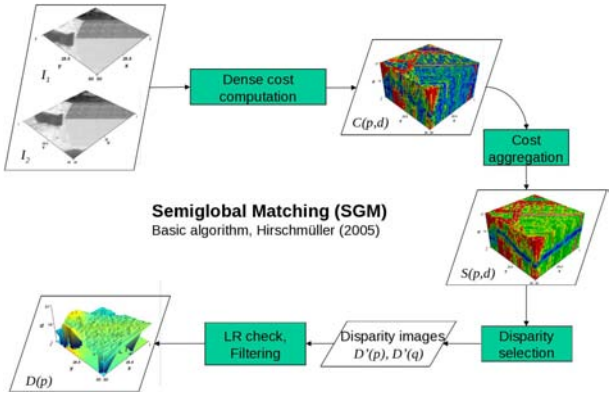

Figure 2. Semi global matching method: for a disparity range $[-d, d]$ *a cost cube* $C(x, y, d)$ *(e.g.* $|I_1(x, y) - I_2(x + d, y)|$*) of size* $w \times h \times (2d + 1)$ *(*w*,* h *are the width and height of the images* I_i*) is calculated, afterwards these costs are aggregated from different directions (semi-globally) and the disparity* d *containing the lowest aggregated cost for each pixel* $p = (x, y)$ *is selected; applying this method exchanging* I_1 *and* I_2 *and doing a left-right-check filters out most of the outliers (mostly occluded areas)*

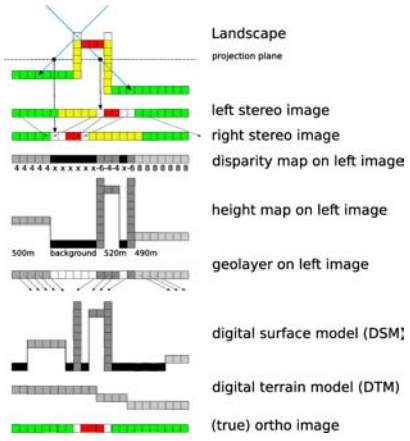

Figure 3. Terms and definitions of all components used throughout this paper illustrated on one image line

Using the original imagery together with a filled height map allows the generation of ortho images. Applying this method to stereo pairs of different epochs results in one filled and one unfilled DSM together with one ortho image for each of the epochs (see figs. 4 and 5).

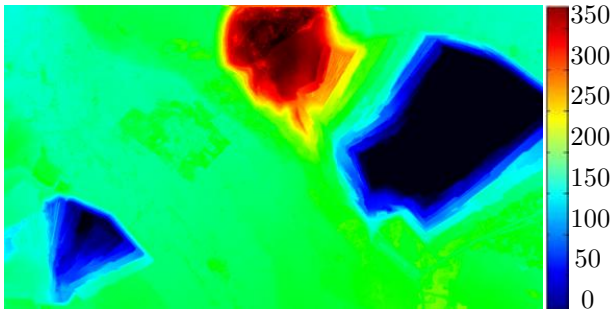

Figure 4. DSM of the GeoEye scene from 2012, 16×9 *km, 0 to 350 m height*

Figure 5. Ortho image of the GeoEye scene from 2012, 16×9 *km*

2.2. Data preparation

For applying robust change detection methods from the DSMs and ortho imagery derived in the DSM generation step some more information has to be extracted in advance. This contains the digital terrain model (DTM) which is in turn together with the DSM used for detection of elevated objects. Also the multispectral imagery has to be converted to top of atmosphere (TOA) reflectances for classification of ground objects.

The DTM is generated using the morphological method described in [4] based on the un-interpolated DSM. Afterwards a shadow mask is calculated following the algorithm of [10] using the blue, green and red channel of the ortho photo. Also the land cover is derived from the pan-sharpened multispectral TOA images using the spectral fuzzy classification method described in [6].

Combining the DSM and DTM allows the extraction of high objects. Adding the shadow and classes gives as result the class map shown in fig. 6. Using these classes the changes can be classified as lined out in tab. 1.

In mining areas the classification becomes more complex since the DTM does not follow the DSM correctly, so the mining area can show up any of the

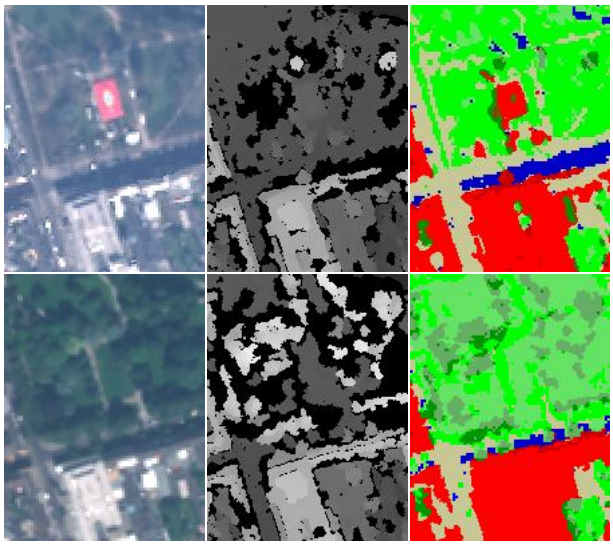

Figure 6. Ortho image, unfilled DSM and class map of the WorldView-2 (2011, top) and the GeoEye (2012, bottom) scenes (section 150×200 *m); Classes: red/dark red: high non-vegetation objects (mainly buildings)/with shadows, tan/blue: low non-vegetation objects (mainly bare soil)/with shadows, dark green/olive: high vegetation objects (e.g. trees)/with shadows, light green/olive: low vegetation objects (grassland)/with shadows*

Table 1. Possible changes depending on classified areas, empty fields mean "no change", v *and* $\bar{v}$ *denote vegetation or non-vegetation, bld. stands for building*

	low → low	low → high	high → low	high → high
$v \rightarrow v$		trees growed	trees logged	
$v \rightarrow \bar{v}$	harvested	bld. erected	trees logged	trees → bld.
$\bar{v} \rightarrow v$	planted	trees growed	bld. demolished	bld. → trees
$\bar{v} \rightarrow \bar{v}$		bld. erected	bld. demolished	

height changes. In this case only the real height change is of interest. Also the phenology of the trees have a great influence on the classification. In the WorldView-2 image taken in March most of the forest areas do not show up as high vegetation since the DSM generation process delivers mainly the more dominant ground.

2.3. Change detection

For applying the 3D change detection first the DSMs containing no-data values (mismatches and occlusions) have to be filled. This is done by segmenting the no-data areas. These holes are filled using the inverse distance weighted interpolation from the lowest neighbouring heights.

Afterwards the so called Robust difference [7] is used

for detecting height changes. The robust difference for pixel $p = (x, y)$ between the two DSMs D_1 and D_2 is defined as

$$d_p(x,y) = \min_{x',y'} ((D_1(x,y) - D_2(x',y')) \geq 0) \quad (1)$$

$$d_n(x,y) = \max_{x',y'} ((D_1(x,y) - D_2(x',y')) \leq 0) \quad (2)$$

with $x' \in [x-w, x+w]$ and $y' \in [x-w, x+w]$ in a small window around p and $d = d_p$ if $|d_p| < |d_n|$ or $d = d_n$ otherwise.

These height changes are filtered afterwards using morphological closing and opening operation to eliminate noise. Also for a more accurate change detection of buildings the changes can be masked using the shadow mask and or other land cover masks. But this masking of the height changes is not applied in the experiments described subsequently.

Combining the classification results (vegetation/non-vegetation and high/low masks) with the detected changes allow for a more detailled classification of the changes.

3. EXPERIMENTS

3.1. WorldView-2 – GeoEye

The first example is a multi-season multi-sensor 3D change detection of an area near Jülich, Germany containing urban areas, open land and the largest open-cast mining area for brown coal in Germany. For the change detection analysis two scenes were available: A WorldView-2 stereo pair acquired on 2011-03-24 at 10:58 and a GeoEye stereo pair acquired on 2012-08-12 at 10:39. Based on these stereo images first the DSMs were derived using the SGM algorithm as described above in fig. 2 and afterwards the DTM, the filled DSM and the TOA ortho image together with the classification images were genereated. Fig. 5 shows the ortho image and fig. 4 the DSM of this area.

Fig. 7 shows the changes resulting from the robust difference calculation and appropriate morphological filtering. Fig. 8 shows a small detail of the changes. Here a newly erected building can be seen in the left top edge (in green) while a cut down forest is visible in the right bottom (in red). The types of objects (building, forest) and type of change (erection and logging) is hereby derived from the class-map and following tab. 1.

Taking only the open-cast mining area in the left bottom edge (Braunkohletagebau Inden) which is defined as a positive or negative height change from

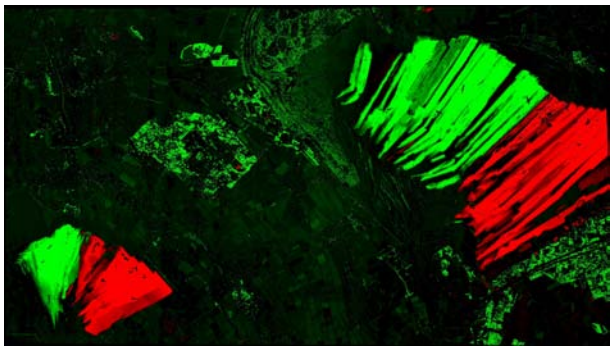

Figure 7. Volume changes between the WorldView-2 and GeoEye scene, 16 × 9 km, red: negative changes, green: positive changes (0–60 m)

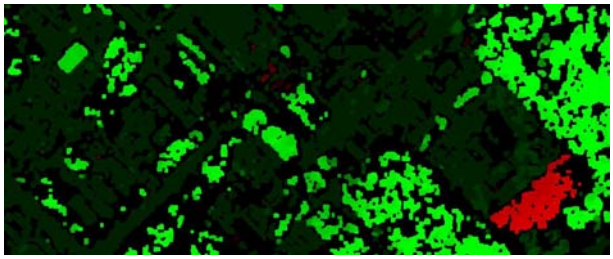

Figure 8. Volume changes between the WorldView-2 and GeoEye scene, small section 1100 × 450 m, top left (green): newly erected building, bottom right (red): cut down forest, rest (green): trees grown up or only detected in the summer DSM

any kind of vegetation to non-vegetation allows us to calculate the mined volumes. For this the positive (green) and negative (red) volumes of fig. 7 are calculated as 79.87 and 122.76 millions of cubic metres. The negative changes are the whole digged out volume containing the coal and also the mine wastes. The positive changes are however only the dumped mine wastes. The mine wastes consist only of earth and stone with a density of $\rho_{soil} = 1500$ kg/m^3 while the brown coal has a density of $\rho_{coal} = 1000$ kg/m^3. This allows us to convert the volume changes to mass changes as shown in tab. 2.

Table 2. Validation of 3D change detection results for the open-cast mining Inden (left bottom mining area); reference data from http://de.wikipedia.org/wiki/Tagebau_Inden

Value	*Mine wastes*	*Digged brown coal*
Reference amount per year	80–85 Mio t/a	20–25 Mio t/a
Amount between images (506 days)	111–118 Mio t	28–35 Mio t
Volumes (m/ρ)	74–79 Mio m^3	28–35 Mio m^3
Measured volume changes	80 Mio m^3	43 Mio m^3

In tab. 2 first the reference amounts per year have

to be scaled to the time distance of the two scenes which are 506 days apart. Afterwards the amount is converted to volumes and compared to the measured volumes from the 3D change detection between the two scenes. Finally we see that the measured volume changes from the described automatic change detection method comply very well with the amounts listed on Wikipedia for this mine.

3.2. Ikonos building detection in summer and winter scenes

For an example of a building change-detection in different seasons two sections of Ikonos scenes from Dong-An (North-Korea) acquired 2006 and 2010 (snow covered) are shown in fig. 9 and the extracted DSMs in fig. 10.

Figure 9. Ortho images of Ikonos scenes of 2006 (left) 2010 (right, covered with snow), both sections 1000×1000 *m*

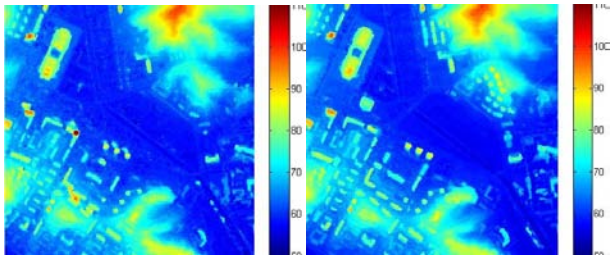

Figure 10. DSMs from the scenes of 2006 (left) 2010 (right), both sections 1000×1000 *m, color coded heights in m*

The described method is applied to the DSMs masked with the results from the ortho images and deliver the change probility map shown in fig. 12, left. Using a manually created reference change mask and thresholding the changes allows for a evaluation of the results (see fig. 11 for explanation of the colors). The overall accuracy for this example is 0.9920 while the kappa index of agreement is 0.7502. These measures are calculated pixel-based between the results and the manually created change mask. Here the true positives (TP, green) are all pixels found by the algorithm and in the reference mask, true negatives (TN, black) are the number of pixels detected as unchanged and also marked as unchanged in the reference. False positives (FP, red) denote unchanged pixels found erroneously as changes by the algorithm while false negatives (FN, blue) are changed pixels wrongly classified as unchanged. From this the overall accuracy (OA) and the kappa index of agreement κ are calculated as [11]:

$$OA = \frac{TP + TN}{N} \quad \text{and} \quad \kappa = \frac{OA - PA}{1 - PA} \tag{3}$$

with $N = TP + TN + FP + FN$ as the total number of pixels and PA as the hypothetical probability of agreement:

$$PA = \frac{(TP + FP) \cdot (TP + FN)}{N^2} + \frac{(TN + FP) \cdot (TN + FN)}{N^2} \tag{4}$$

Reference data / Change Mask	Change	No-Change
Change	True detected	False alarm
No-Change	Missed alarm	

Figure 11. Color coding scheme of changes shown in change masks

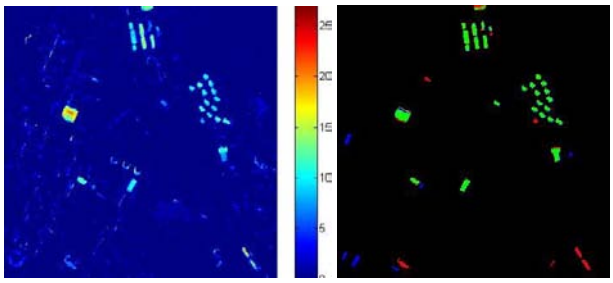

Figure 12. Detected changes in the Ikonos stereo DEMs of Dong-An, left: change probability (in %), right: change mask (green: true detected changes, blue: missed changes, red: erroneously detected changes)

3.3. Cartosat – Forest

Using the image segmentations of the pre and post images and merging these together allows also for forest change detection in pan images as delivered from the indian Cartosat-1 (IRS-P5) satellite. These images are acquired with two stereo cameras (forward: 26°, backward: −5°) with a resolution of 2.5 m.

Two ortho images of a region near Oberammergau (Germany) are shown in fig. 13. The DSMs created from the stereo imagery can be seen in fig. 14.

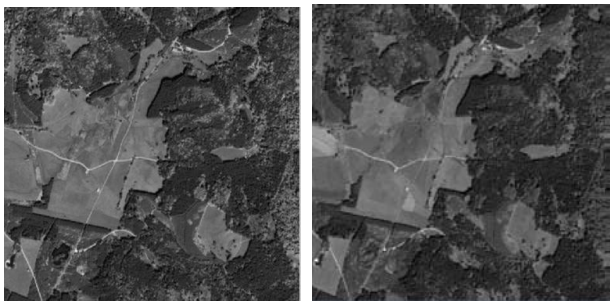

Figure 13. Cartosat ortho images of a forest area near Oberammergau 2008 (left) and 2009 (right)

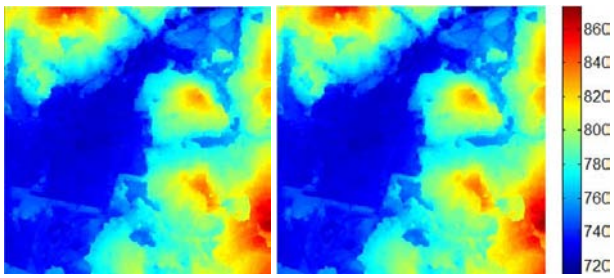

Figure 14. DSMs created from the Cartosat stereo images of a forest area near Oberammergau 2008 (left) and 2009 (right)

Applying a mean-shift over-segmentation to the ortho imagery and a region merging on the combination results in a land cover segmentation. To analyze the changes the height changes from the DSMs and robust multi-level change features from the images are used. Putting this all together results in a change probability map shown in fig. 15, left.

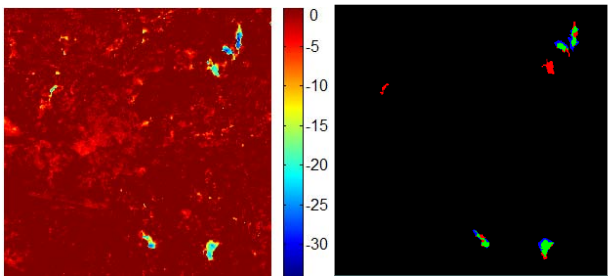

Figure 15. Resulting changes from the Cartosat forest images/DSMs, left: change probability map, right: change mask (green: true detected changes, blue: missed changes, red: erroneously detected changes)

For the change detection between the two Cartosat images the overall accuracy is 0.9945 and the kappa index 0.5879.

4. RESULTS

The 3D change detection results from the described automatic chain are promising but yet not perfect. This is mostly due to the errors in the creation of the DSMs. Better DSMs can be generated using three or more stereo images. First experiments already conducted with multi-stereo WorldView-2 and Pléiades imagery show that a third image improves the quality of the DSM drastically while a fourth or fifth image lead only to minor further improvements. Also scenes acquired in the same season lead to a more consistent classification for the change detection.

The volume calculation of the mining area shows the usability of the method for such applications. But the 3D change detection delivers many layers of information so that the results are also somehow dependend on the searched changes. If the main goal is the detection of changed buildings the usage of shadow masks allow for a more precise modeling of the buildings and improvement of the DSM. While when searching for mining activities the generated DSMs should be changed (improved) as few as possible to keep the volume information valid. Also the DSMs generated in forest areas or even worse on water areas are containing much noise and outliers. So for such applications sophisticated segmentation, classification and DSM filtering has to be conducted.

The first version of the automatic change detection method described in this paper was tested in three different applications: The monitoring of mining activities shows the high accuracy and good usefulness of the method in this case. The detection of building changes and the detection of forest changes also give already very good results. Especially the change detection example of Dong-An shows the potential of 3D change detection methods between different seasons where classical intensity based change detection methods will mostly fail. But for a more correct detection of building changes better DSMs or better DSM improvement methods have still to be developed.

5. CONCLUSION AND OUTLOOK

In this paper we presented an potentially fully automatic change detection method based on the creation of dense digital surface models (DSMs), derivation of digital terrain models, a fused classification approach and finally a robust 3D change detection of objects of different kinds. The method allows a more sophisticated change detection than only image based methods. One main advantage is that the method also works in different seasons like summer and winter. The main disadvantage of the method is the need of pre- and post-elevation information.

Especially in the application of this method to disaster mapping like after the Haiti earthquake shows the problem: In most cases no pre-disaster stereo imagery or DSMs of the required resolution and region exists. But especially for such disaster mapping applications the 3D change detection could deliver much more valuable information since in earthquakes buildings often collapse vertically. Such changes can only be seen in volumetric change detection methods.

Other useful applications are the monitoring of mining activities (as shown in the Jülich example), illegal timber logging or the monitoring of the evolution of large cities. The method still has to be improved and tuned to different resulting change detection maps for different questions and applications.

Finally we propose an additional sentinel mission covering the whole earth permanently like the indian Cartosat-1 (IRS-P5) mission with a two- or even better a three-line-stereo-scanner (like ALOS-Prism), a ground sampling distance of the stereo scanners of at least 2.5 m and also a nadir looking multispectral VNIR scanner with also about 2.5 m GSD or better.

REFERENCES

1. H. Hirschmüller. Accurate and efficient stereo processing by semi-global matching and mutual information. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.

2. Pablo dAngelo, Manfred Lehner, Thomas Krauß, Danielle Hoja, and Peter Reinartz. Towards automated dem generation from high resolution stereo satellite images. In *IAPRS*, volume 37, pages 1137–1142, 7 2008.

3. Hossein Arefi, Pablo dAngelo, Helmut Mayer, and Peter Reinartz. Automatic generation of digital terrain models from cartosat-1 stereo images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 83, 6 2009.

4. Thomas Krauß, Hossein Arefi, and Peter Reinartz. Evaluation of selected methods for extracting digital terrain models from satellite born digital surface models in urban areas. In *International Conference on Sensors and Models in Photogrammetry and Remote Sensing (SMPR 2011)*, 5 2011.

5. H. Chaabouni-Chouayakha, I. Rodes-Arnau, and P. Reinartz. Towards automatic 3-D change detection through multi-spectral and digital elevation model information fusion. *International Journal of Image and Data Fusion*, 4(1):89–101, 2013.

6. Thomas Krauss, Beril Sirmacek, Hossein Arefi, and Peter Reinartz. Fusing stereo and multispectral data from WorldView-2 for urban modeling. In *Proc. SPIE 8390*, volume 83901X, page 2012, 4 2012.

7. Jiaojiao Tian and Peter Reinartz. Multitemporal 3D Change Detection in Urban Areas Using Stereo Information from Different Sensors. In *International Symposium on Image and Data Fusion*, volume 2011, pages 1–4, 8 2011.

8. J. Tian, S. Cui, and P. Reinartz. Building Change Detection Based on Satellite Stereo Imagery and Digital Surface Models. *IEEE Transactions on Geoscience and Remote Sensing*, 51:1–12, 2013.

9. J. Tian, P. Reinartz, dAngelo P., and M. Ehlers. Region-based automatic building and forest change detection on Cartosat-1 stereo imagery. *ISPRS Journal on Photogrammetry and Remote Sensing*, 79:226–239, 2013.

10. A. Makarau, R. Richter, R. Müller, and P. Reinartz. Adaptive shadow detection using a blackbody radiator model. *IEEE Trans.Geosci. Remote Sens.*, 49(6):2049–2059, 6 2011.

11. Russell G. Congalton. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote sensing of Environment*, 37(1):35–46, 1991.