3D CHANGE DETECTION INSIDE URBAN AREAS USING DIFFERENT DIGITAL SURFACE MODELS

Houda Chaabouni-Chouayakh, Thomas Krauss, Pablo d'Angelo and Peter Reinartz

Remote Sensing Technology Institute
German Aerospace Center (DLR)
Oberpfaffenhofen, 82234 Weling, Germany
houda.chaabouni@dlr.de

Commission III/3

KEY WORDS: DSM from different sensors, 3D change detection, urban area, DSM subtraction, DSM class-for-class differencing.

ABSTRACT:

The development of reliable change detection techniques from remote sensing data is one of the main challenges in urban growth and change monitoring research. One of the main open issues consists in automatizing the detection of changes whose interpretation has remained up-to-now visual in most operational applications in remote sensing. When dealing with urban areas, one possibility to cope with the automatic growth monitoring is the exploitation of the height information relative to the different man-made objects that exist in the scene. In fact, the comparison of Digital Surface Models (DSMs), acquired at different epochs, should provide a valuable information about the 3D urban changes occurred in the studied area. Nevertheless, while most of the changes are optically visible and easily detectable by an expert user, automatic processes are difficult to achieve due to the problems of co-registration and the significant height difference that may occur between DSMs acquired from different sensors and/or generated by different algorithms. This results in the detection of virtual or irrelevant changes. This article proposes two semi-automatic methods for 3D change detection using DSMs obtained from different sources. While the first method is based on the simple subtraction of DSMs from two epochs, the second one consists in comparing the classification maps of these two DSMs, through class-for-class differencing. In both cases, adaptative post-processing steps have been introduced in order to distinguish real from virtual changes. Evaluations of the proposed approaches have been carried out to detect the 3D changes that have occurred in the city center of Munich in Germany from 2003 to 2005.

1 INTRODUCTION

In the last few decades, the constantly intensive global urbanization has made the urban and suburban areas among the most dynamic sites on Earth. Therefore, new innovative tools are required for better monitoring and finer description of such areas. Remotely sensed imagery in some cases may be the only reliable source for better understanding of urban areas. In fact, satellite imagery can significantly improve the monitoring of cities in a wide range of applications, e.g. urban growth monitoring, disaster damage assessment, urban change detection, etc.

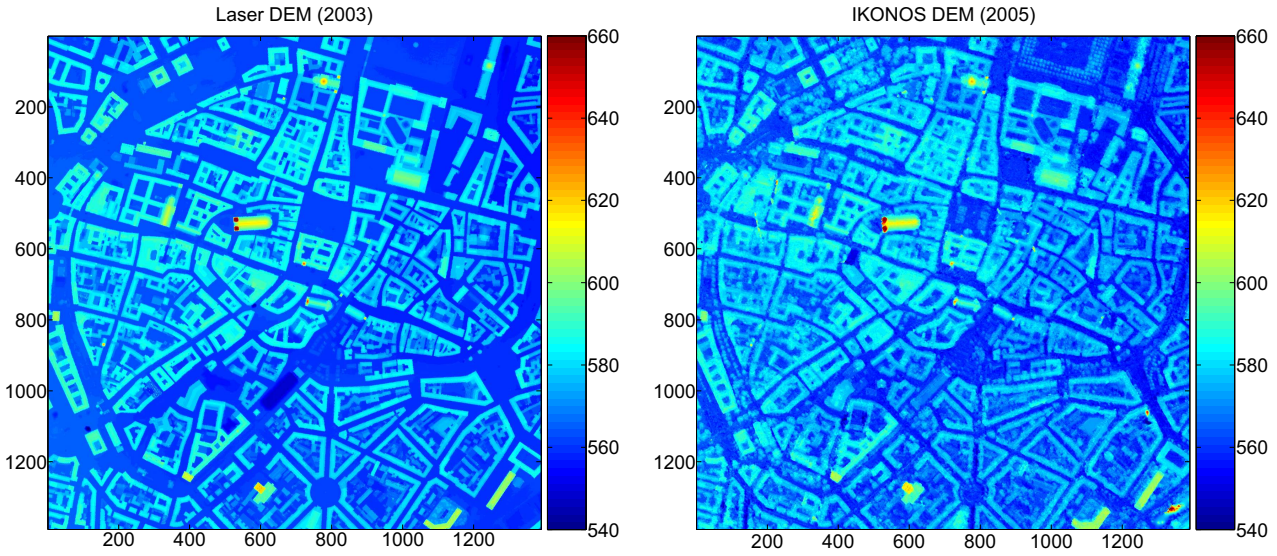
Our overall objective in this paper is the development of reliable techniques to detect automatically 3D changes inside urban areas. Since the height information is very important in characterizing man-made structures, we suggest to use Digital Surface Models (DSMs) from two different epochs in order to identify 3D changes. Indeed, very high resolution stereo data is a very promising source to extract 3D surface models. Many works have been dedicated to detect 2D changes using remote sensing data (e.g. (Lu et al., 2004)), but the issue of 3D change detection using DSMs has been seldom tackled in the literature. DSM-based change detection researches (e.g. (Gong et al., 2000, Heller et al., 2001, Hollands et al., 2007)) mainly propose to compute simple difference between DSMs from different epochs. Such an approach could provide reliable results if the same sensor is used. However, with different sensors (which is quite often the case) or various DSM generators, we are in general faced by the problem of co-registration and significant differences at building outlines. Post-processing steps are thus highly required to distinguish real from virtual changes. In other publications (e.g. (van der Sande et al., 2008, Champion et al., 2009)), the authors proposed to improve 2D change detection, obtained when comparing images acquired at different epochs, by including the 3D information extracted from DSMs. However, careful post-processings should also be adopted to avoid the extension of the 2D errors in the 2D change maps into the 3D ones.

In this article, we propose two methods to detect changes inside the city center of Munich in Germany by using DSMs acquired from two different sensors: Laser and Ikonos stereo data (section 2). While in the first method (section 3), we detect changes at pixel-level, in the second one (section 4), our focus is rather targeted at class-level. It is worth to note that in both cases, we apply adaptative post-processing steps and we include contextual knowledge in order to distinguish real from virtual changes. In fact, when urban scenes have to be interpreted, many works have been combining contextual features to the geometric and textural ones to determine plausibility and enforce constraints in label assignments (Hoiem et al., 2008, Chaabouni-Chouayakh and Datcu, 2010). Since contextual knowledge has been useful in the identification of different urban structures using aerial or satellite images, it may also be the clue in our case to generate accurate 3D change detection maps inside urban areas when using DSMs.

2 DESCRIPTION OF THE TEST DATA

In this paper, we study the potential changes that have occurred in the city center of Munich in Germany from 2003 to 2005 by using DSMs. These DSMs have been co-registered after estimating and then eliminating the 3D shift between them. Figure 1 illustrates the two co-registered DSMs used in this work. They are:

- a 1m resolution Laser DSM from February 2003 ; and
- a 1m resolution DSM generated from Ikonos stereo images acquired in July 2005, using the Semi-Global Matching algorithm (SGM) implemented at DLR ((Hirschmüller, 2008,



(a) Laser DSM from 2003.

(b) Ikonos DSM from 2005.

Figure 1: The two DSMs used in this work to monitor the potential changes that have occurred in Munich from 2003 to 2005.

d'Angelo et al., 2008)). The resulting DSM has been filled using inverse distance weighted interpolation in order to remove the occluded areas.

From Figure 1, we can see how a difference regarding the DSM sources (here Laser and Ikonos stereo data) influences the generation and consequently the appearance of the different man-made structures in the DSMs. In fact, already from a visual DSMs comparison, many interesting observations can be made:

- The building outlines appear sharp in the Laser DSM, while they are quite smooth in the Ikonos DSM. Two reasons could explain this smoothness effect: 1) First, the presence of shadows on buildings outlines in the stereo Ikonos images results in the rise of occluded areas in the generation of the DSM by the SGM algorithm. These occluded areas are then replaced by smooth slopes during the interpolation step. 2) The second reason for these smoothness artifacts consists in the fact that only one pair of stereo data is used in the generation of the Ikonos DSM. This produces more occluded areas and therefore smoother slopes on buildings outlines. Actually, the problem of occlusion is in general less pronounced when more pairs of stereo images are involved in the DSM generation, as stated in (Hirschmüller, 2008). This different Laser vs. Ikonos DSM behaviors disturbs our 3D change detection when this latter is based on computing a simple pixel or object based DSM subtraction, since this will unfortunately result in the detection of virtual changes. In order to solve this first problem, adaptative post-processing steps have been applied to each of our 3D change detection algorithms, so that these virtual changes are removed and only the real ones are kept. More details about our different post-processing steps will be given in sections 3 and 4.

- The class of vegetation is also seen differently from the Laser and Ikonos stereo images. As a matter of fact, the Laser data of Figure 1 corresponds to the last pulse range and therefore the vegetation pixels are mostly not available. To cope with this second problem, a vegetation mask has been applied on both DSMs. This mask has been generated by computing the Normalized Differenced Vegetation Index (NDVI) from the red and infra-red bands of the Ikonos image. Actually, eliminating the vegetation from our 3D change detection scheme is not so critical since our focus in this paper is only 3D changes relative to building construction and destruction.

3 PIXEL-BASED 3D CHANGE DETECTION

In this first method, we detect changes that occurred in Munich from 2003 to 2005 at pixel level, by using only the height information. We start by computing the simple difference between the two DSMs of Figure 1. The result is given in Figure 2.

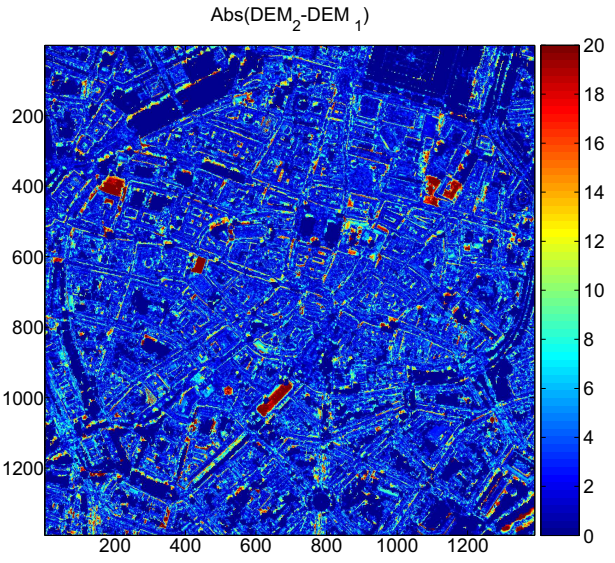


Figure 2: Pixel-based change detection: Absolute difference between the DSMs of Figure 1. The simple DSMs difference is not enough robust to distinguish real from virtual changes.

From the absolute difference of Figure 2, we notice that a simple difference between DSMs at different epochs is not enough robust to detect the real changes. As a matter of fact, it highlights also virtual changes coming from the different nature of the DSM sensors (here Laser and Ikonos stereo data). Indeed, the fact that the buildings outlines appear sharp in the Laser DSM and rather smooth in the Ikonos one, results in the rise of polyline-like changes on the borders of the buildings in the difference map. Moreover, some virtual changes arise from errors in the DSM computation resulting in small differences in the height. Actually, they are not very critical since they could be eliminated by simple thresholding.
We propose to apply several post-processing steps to keep only the real changes and remove the virtual ones (caused by the different nature of the DSM sources and also from the computation errors). The complete flowchart of this pixel-based 3D change detection algorithm is depicted in Figure 3.

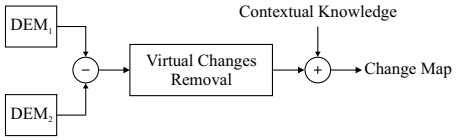


Figure 3: Flowchart of the pixel-based 3D change detection approach.

3.1 Virtual changes removal

Our virtual changes removal is done in two steps. First, we apply a thresholding so that height differences less than a threshold T are considered as computation errors and not detected as 3D changes. Then, we binarize the thresholded map into positive (construction) and negative (destruction) changes. For each pixel (i, j), we compute the new binary change map BC as follows:

$$\text{if } \text{DSM}_2(i,j) - \text{DSM}_1(i,j) \geq T \quad \text{BC}(i,j) = + \quad (1)$$
$$\text{if } |\text{DSM}_2(i,j) - \text{DSM}_1(i,j)| < T \quad \text{BC}(i,j) = 0 \quad (2)$$
$$\text{if } \text{DSM}_2(i,j) - \text{DSM}_1(i,j) \leq -T \quad \text{BC}(i,j) = - \quad (3)$$

The binary change map BC obtained when applying a threshold $T = 5$m is given in Figure 4.

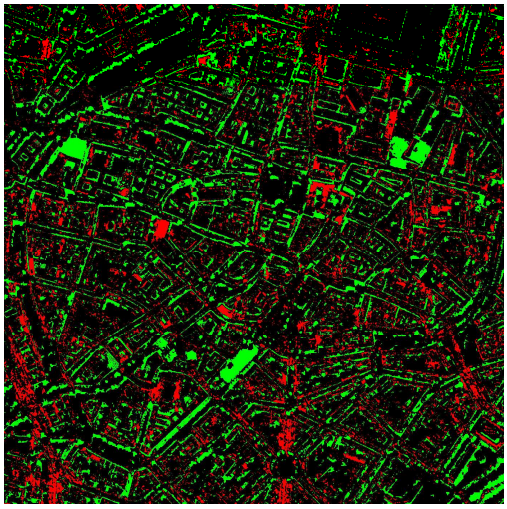


Figure 4: Pixel-based change detection: Binary change detection map. The positive and negative changes are highlighted in green and red, respectively. The black pixels correspond to no change.

In comparison to the absolute difference map, the detection of the different changes from the binary map of Figure 4, could be more easily and straightly performed. In fact, the thresholding step succeeds in removing most of the virtual changes caused by the DSM computation errors. Moreover, we notice that the different constructed and destructed buildings correspond to the large green and red segments, respectively. Nevertheless, polyline-like virtual changes do still remain and need to be removed. Therefore, more adaptative post-processing steps are required, so that only real changes are kept. For this purpose, we propose in this work to apply some opening and closing morphological operations in order to make the real changes more compact and the virtual ones thinner. In our case, we have used an opening filter with a 7×7 kernel followed by a closing one with a 5×5 kernel. This results in a much cleaner change map as could be seen in Figure 5.

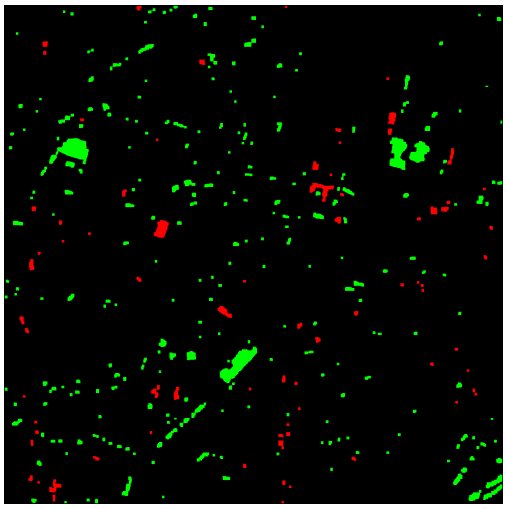


Figure 5: Pixel-based change detection: Binary change detection map after applying morphological operations: the positive and negative changes are highlighted in green and red, respectively. The black pixels correspond to no change. While the real changes become more compact and the virtual ones are thinner.

3.2 Contextual knowledge introduction

When the goal is to detect urban changes (constructed and destructed buildings), it is better to treat each change as single object (or vector). That is why, we propose at this final step of our approach to vectorize the binary change map of Figure 5. After that, we suggest to include some contextual knowledge aiming at describing the geometry of the changed objects. Indeed, when dealing with man-made structures such as building, geometry is one of the most important features that should be included for a better characterization. We mainly used two shape features: the size and the compactness as indicatives of largeness and homogeneity, respectively. The compactness is defined as :

$$\text{compactness} = 2\frac{\sqrt{\pi A}}{P}, \quad (4)$$

where A and P denote respectively the area and the perimeter of the corresponding vector. Based on these two shape features, we perform a query on the thresholded binary vectorized map selecting the large compact polygons since buildings are in general characterized by such property. In our case, polygons whose (size $>$ 1000 and compactness $>$ 0.5) or (size $>$ 500 and compactness $>$ 0.9) have been selected. The resulting change map is illustrated in Figure 6.

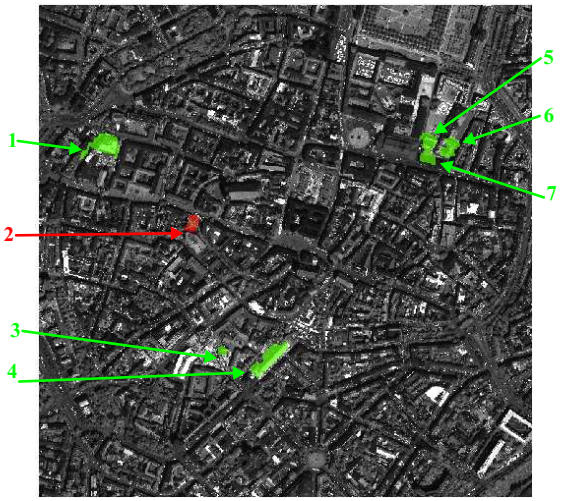


Figure 6: Pixel-based change detection: Final change detection map. Real positive and negative changes are correctly detected and highlighted in green and red, respectively.

3.3 Evaluation of the final change map

Figure 6 shows a cleaner change map. The introduction of the contextual knowledge has thus succeeded in removing the polyline-like virtual changes coming from the different sensor nature. Only large compact segments, generally corresponding to buildings, are kept. However, the detected segments do not fit exactly with the borders of the buildings. This is due to the smoothness artifacts that we have explained in section 2, and also to the use of morphological operations which remove unfortunately parts from real changes. The choice of the opening and closing kernels size should be very carefully done so that the maximum of real changes is kept and the maximum of virtual ones is removed. In order to evaluate the accuracy of our 3D change detection algorithm, its correctness and completeness have been verified. For the correctness, an *in-situ* identification has been carried out:

- **Change 1** corresponds to a real positive change. It is the shopping center named "Oberpollinger". According to (Oberpollinger, 2010, Wikipedia, 2009a), an extension of the building has been done between 2003 and 2006.
- **Change 2** corresponds to a real negative change. According to (Schörghuber-Unternehmensgruppe, 2005), this building was destroyed in November 2004 and replaced later with another one whose big opening was in April 2006.
- **Change 3** corresponds to a real positive change. It is the Ohel Jakob synagogue which was built between 2004 and 2006 as the new main synagogue of the Munich Jewish community. According to (Wikipedia, 2009b), the synagogue was inaugurated in November 2006.
- **Change 4** corresponds to a real positive change. It is the "Schrannenhalle" in Munich. The building was still in construction in 2003. A big opening has been organized on September 5, 2005 to celebrate the end of the constructions. More details could be found at (Schrannenhalle, 2010).
- **Change 5**[1] and **6** seem to be real positive changes because from the DSMs depicted in Figure 1, we could already detect a height elevation over a quite large area from 2003 to 2005. Unfortunately, no reference has been found to affirm the accuracy of our detection approach in the case of these two changes. However, our *in-situ* verification confirms the modernity of the corresponding buildings style.
- **Change 7**[1] corresponds to a real positive change. It is the "Maximilianhöfe" store which has been in construction between 2002 and 2005 according to (Maximilianhöfe, 2007).

[1]Changes 5 and 7 have been combined in a same polygon in our approach but they correspond to two different buildings separated by a narrow street in reality. That is why they have been linked in the Ikonos DSM and therefore clustered in a same polygon in our approach.

To evaluate the completeness of the proposed approach, a visual comparison of the contents of the Laser DSM and the Ikonos stereo images has been performed. This comparison has showed that no large changes have been missed during our 3D change detection procedure.

4 CLASS-BASED 3D CHANGE DETECTION

In this second method, the changes are detected at class-level. The complete flowchart is depicted in Figure 7.

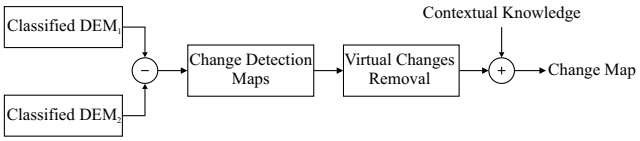
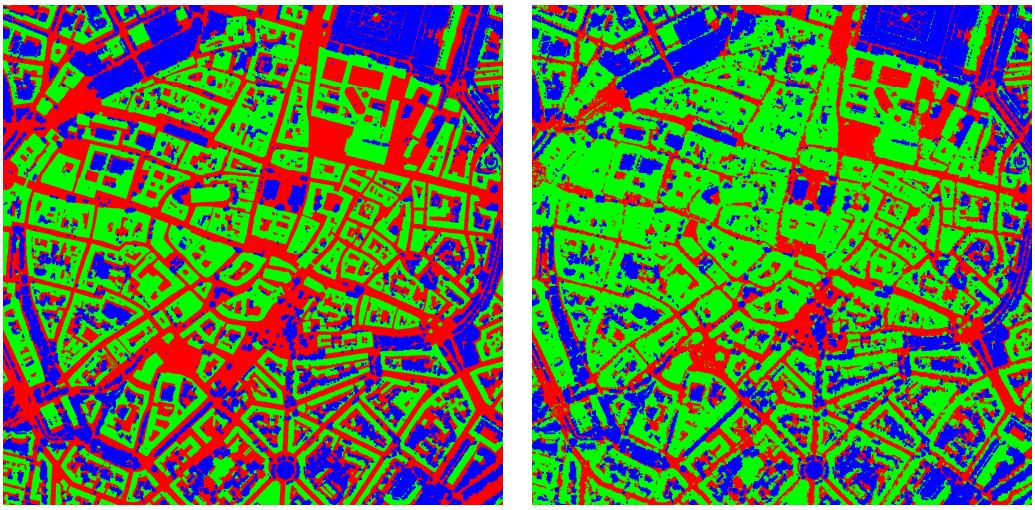


Figure 7: Flowchart of the class-based 3D change detection approach.

The first step of this second 3D change detection method consists in performing an IsoData unsupervised classification on the two DSMs of Figure 1. The IsoData algorithm is a variation of the K-means clustering algorithm which uses splitting and merging clusters methods to do the clustering (Tou and Gonzalez, 1974). To get a robust class separation, we have tuned many IsoData parameters regarding the iterations (e.g. maximum number of iterations, minimum change between iterations), the clusters (e.g. maximum and minimum number of clusters) and the splitting and merging parameters (e.g. minimum number of pixels per region). The resulting classification maps are illustrated in Figure 8. Three clusters have been generated, that match mainly with the classes of buildings, ground and vegetation. Here also, we could see that the different nature of the sensors (Laser and Ikonos) has also an impact on the classification maps: the building outlines do not appear in the same position in the classification maps of 2003 and 2005, although the two DSMs are well co-registered in the same UTM projection. Adaptative post-processing steps are thus required to detect only the real changes. This effect is also illustrated in the change detection maps presented in Figures 9 and 10. In this analysis, we do not apply a simple differencing of the two classification maps. We rather perform a class-for-class difference to identify, for each initial state class (in the classification map of 2003), the classes into which those pixels changed in the final state image (in the classification map of 2005). Change detection statistics are also reported in Table 1 in terms of percentages.

		Initial State		
		Vegetation	Ground	Building
Final State	Vegetation	100 %		
	Ground		65.67%	1.9%
	Building		34.33%	98.1%

Table 1: Change detection matrix.



(a) Laser DSM from 2003.

(b) Ikonos DSM from 2005.

Figure 8: Unsupervised classification of the DSMs. The red, green and blue clusters match with the ground, buildings and vegetation classes, respectively.

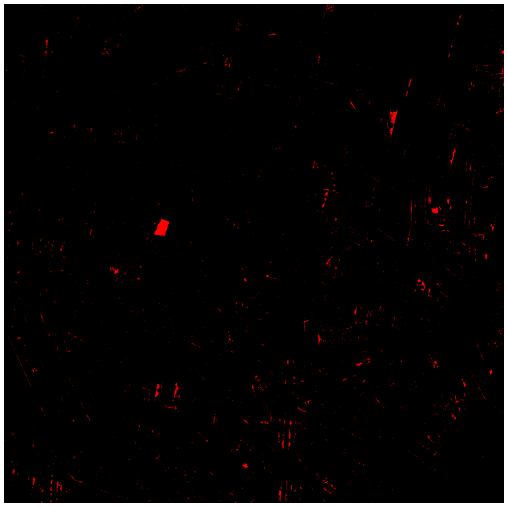


Figure 9: Building to ground change map.

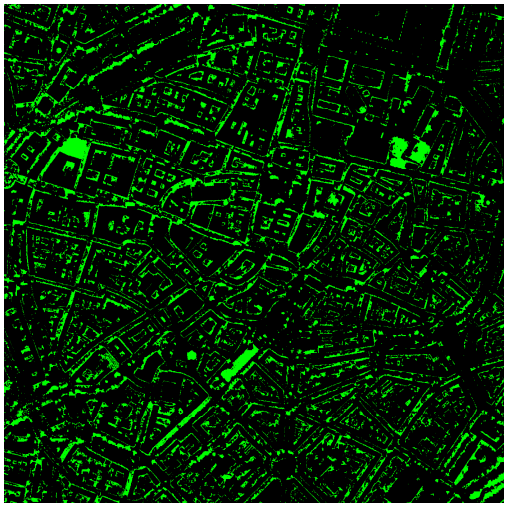


Figure 10: Ground to building change map.

4.1 Virtual change removal and contextual knowledge introduction

According to Figures 9 and 10 and Table 1, the problem of virtual changes detection is more pronounced in the ground to building change map than in the building to ground one. Indeed, a quite high percentage of pixels (34.33%) has been converted from ground to buildings from 2003 to 2005. They are more likely to be a mixture of real and virtual changes. Thus, a contextual knowledge introduction as done in section 3 should lead to more accurate change results. To remove virtual changes from the ground to building change map, we have applied a 7×7 opening kernel morphological filter followed by a 5×5 closing one. A cleaner map, more suitable to the step of the contextual knowledge introduction, has thus been provided, as seen in Figure 11.

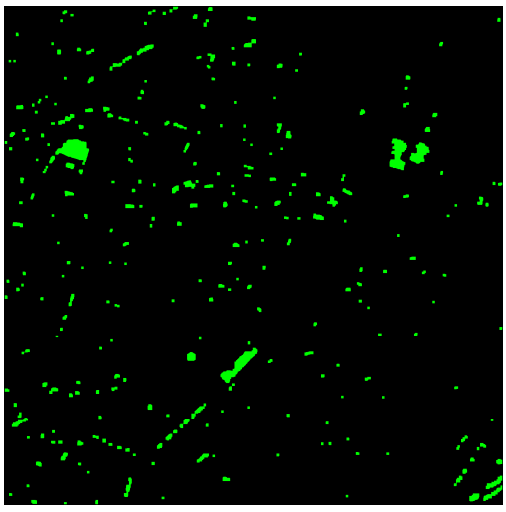


Figure 11: Ground to building change map after applying morphological operations. A cleaner map is obtained where real and virtual changes are more compact and thinner, respectively.

To move to the object-level and finalize our 3D change detection scheme, we have introduced, as in the first method, the contextual knowledge derived by the two shape parameters (compactness

and size). Polygons satisfying these properties: (size > 1000 and compactness > 0.5) or (size > 500 and compactness > 0.9), have been selected. The final result is given in Figure 12.

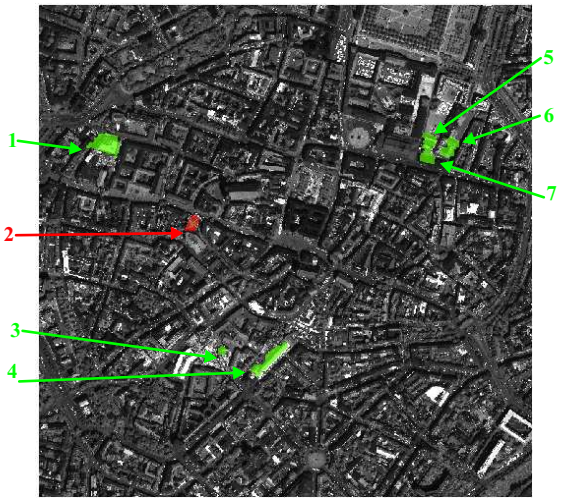


Figure 12: Class-based change detection: Final change detection map. Real positive and negative changes are highlighted in green and red, respectively. Similar change map has been obtained with the first approach. However, in this second approach, more parameters have to be tuned to get accurate results.

4.2 Evaluation of the final change map

Figure 12 shows similar accurate change map as the one obtained by the first method (Figure 6). All the real positive and negative changes described in subsection 3.3, have been identified. The completeness of this second approach has been evaluated through a visual comparison of the contents of the Laser DSM and the Ikonos stereo images; and no large changes have been missed. However, here also the change segments do not fit exactly with the building borders. This could still be explained by the smoothness effect pronounced in the Ikonos DSM and also by the fact that morphological operations do not remove only virtual changes but also parts from some buildings outlines. Although similar change maps have been obtained, a possible comparison between the two approaches could be done in terms of parametrization. Indeed, the classification in the second approach requires a fine tuning of a larger number of parameters to get accurate results.

5 CONCLUSIONS AND FUTURE WORK

In this paper, the problem of 3D change detection using DSMs generated from different sensors has been tackled. Two methods have been proposed and evaluated to detect changes in the city of Munich in Germany using a Laser DSM from 2003 and an Ikonos DSM from 2005. However, while the first one operates at pixel-level by performing a simple DSMs subtraction, the second one detects rather changes at class-level through class-for-class difference. In both cases, post-processing steps have been necessary to remove virtual changes coming from DSM computation errors and from the different nature of DSM sources. Similar change detection results have been successfully obtained by the two methods. Their accuracy has been evaluated in terms of correctness and completeness by means of an *in-situ* verification of the different changes and a visual comparison of the contents of the Laser DSM and the Ikonos stereo images, respectively. Despite the similarity of the resulting change maps, a comparison between the two techniques was done in terms of parametrization. In fact, in the class-based change detection method, more parameters need to be tuned when classifying the two DSMs. The results are globally satisfying and promising, although some of them could still be improved and completed. They can be considered as preliminary results for some higher level urban area monitoring where more 3D change scenarios such as forestration/deforestration are involved, or DSM quality analysis.

REFERENCES

Chaabouni-Chouayakh, H. and Datcu, M., 2010. Coarse-to-Fine Approach for Urban Area Interpretation Using TerraSAR-X Data. IEEE GRSL 7(1), pp. 78–82.

Champion, N., Rottensteiner, F., Matikainen, L., Liang, X., Hyyppä, J. and Olsend, B. P., 2009. A Test of Automatic Building Change Detection Approaches. IAPRS, Paris, France XXXVIII(3/W4), pp. 145–150.

d'Angelo, P., Lehner, M., Krauss, T., Hoja, D. and Reinartz, P., 2008. Towards Automated DEM Generation from High Resolution Stereo Satellite Images. ISPRS 2008 , Peking, China XXXVII(B4), pp. 1137–1342.

Gong, P., Biging, G. and Standiford, R., 2000. Use of Digital Surface Model for Hardwood Rangeland Monitoring. Journal of Range Management 53(6), pp. 622–626.

Heller, A., Leclerc, Y. and Luong, Q., 2001. A Framework for Robust 3-D Change Detection. Proc. of SPIE, Toulouse, France.

Hirschmüller, H., 2008. Stereo processing by semiglobal matching and mutual information. IEEE TPAMI 30(2), pp. 328–341.

Hoiem, D., Efros, A. A. and Hebert, M., 2008. Closing the Loop on Scene Interpretation . CVPR.

Hollands, T., Bostrm, G., Goncalves, J. G., Gutjahr, K., Niemeyer, I. and V.Sequeira, 2007. 3D Scene Change Detection from Satellite Imagery. Proc. 29th Symp. on Safeguards and Nuclear Material Management, Aix-en-Provence, France pp. 1–6.

Lu, D., Mausel, P., Brondzio, E. and Moran, E., 2004. Change Detection Techniques. International Journal of Remote Sensing 25(12), pp. 2365–2407.

Maximilianhöfe, 2007. Maximilianhöfe. `http://www.rijo.homepage.t-online.de/pdf/DE_MU_TO_maximilian.pdf`.

Oberpollinger, 2010. Oberpollinger. `http://www.oberpollinger.de/en/store-overview/history/`.

Schörghuber-Unternehmensgruppe, 2005. Kaufingerstrasse 15. `http://www.xn--schrghuber-unternehmensgruppe-i5c.de/html/presse/meldung_alles.php?flash=&js=&press_id=4&actual=&category_id=5&year=2005`.

Schrannenhalle, 2010. Schrannenhalle. `http://historie.schrannenhalle.de/de/baustelle/baukalender.php`.

Tou, J. and Gonzalez, R., 1974. Pattern Recognition Principles. Addison-Wesley Publishing Co., Reading, Massachusetts.

van der Sande, C., Zanoni, M. and Gorte, B., 2008. Improving 2D Change Detection by Using Available 3D Data. ISPRS Commission VII, Beijing, China.

Wikipedia, 2009a. Oberpollinger. `http://de.wikipedia.org/wiki/Oberpollinger`.

Wikipedia, 2009b. Ohel Jakob Synagogue. `http://en.wikipedia.org/wiki/Ohel_Jakob_synagogue_(Munich,_Germany)`.