

# Automatic Urban Area Monitoring Using Digital Surface Models and Shape Features

Houda Chaabouni-Chouayakh, Pablo d'Angelo, Thomas Krauss and Peter Reinartz  
 Remote Sensing Technology Institute (IMF)  
 German Aerospace Center (DLR)  
 Wessling, Germany  
 houda.chaabouni@dlr.de

**Abstract**—Accurate monitoring of urban areas using remote sensing data requires reliable change detection techniques. Nevertheless, while most of the changes are optically visible and easily detectable by an expert user, automatic processes are quite difficult to develop. That is why, the interpretation of changes has remained up-to-now visual in most operational applications in remote sensing. This paper provides an automatic approach for 3D change detection based on the joint use of the height and spatial information. In fact, 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. The subtraction of Digital Surface Models (DSMs), acquired at different epochs, should thus provide a valuable information about the 3D urban changes occurred in the studied area. However, when at least one of the DSMs presents some artifacts, a simple DSM subtraction could result also in the detection of virtual changes. To remove these virtual changes, we propose in this work to include, in addition to the height information, some shape features that could be of a great help in describing the geometry of the constructed or demolished man-made structures. After that, the Support Vector Machine (SVM) classifier is used to differentiate real from virtual changes. Evaluation of the proposed approach in terms of completeness, correctness, overall accuracy, etc has been performed proving its efficiency and relatively high accuracy.

## I. 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. New innovative tools are thus required for better monitoring 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.

The overall goal of this paper lies in the development of automatic urban growth monitoring using height and spatial or shape information. In fact, urban changes are in general either related to building construction/demolition or vegetation growth. These two issues should be well described if the height and shape information are available. In the literature, to monitor height changes, subtraction of Digital Surface Models (DSMs) has been widely used (e.g. [1]–[3]). This simple approach could provide reliable results if accurate DSMs are available. However, if at least one of the used DSMs shows some artifacts (which is quite often the case), we are in

general faced by the problem of significant height differences over some complex 3D structures. This would result in the detection of virtual changes, generally characterized by strange shapes. Therefore, this work suggests the additional use of several shape features in order to describe the geometry of the spatial extent of the different constructed/demolished buildings, generally characterized by quite regular shapes. Similar approach has been used in [4], [5] to detect respectively buildings and building changes using Lidar and Laser DSMs. Accurate building detection maps have been obtained. This is somehow due to the very good quality of the used DSMs and to the successful tuning of the different thresholds. In this article, the used DSMs are generated from two pairs of Ikonos stereo images acquired at different epochs. Since the quality of the DSMs is not as good as the Lidar/Laser ones, some post-processing steps have been included so that the proposed approach does not remain limited only to high quality DSMs. Also, still in the frame of the automatization of our change detection approach and to avoid the manual tuning of the different thresholds, after the feature extraction step, we suggest to separate the real changes from the virtual ones, using the Support Vector Machine (SVM) classifier which has shown great efficiency and robustness in various pattern recognition applications ([6]).

The organization of this paper is as follows: Sections II and III describe the data used in this work and the different steps of the proposed 3D change detection approach, respectively. Section IV assesses the accuracy of our method using different objective metrics, while section V gives some conclusions and perspectives.

## II. PRESENTATION OF THE DATA

In this work, we perform a multi-temporal and multi-season change monitoring of an Asiatic urban area using DSMs. Two pairs of Ikonos-2 stereo images (©EUSI provided under the EC/ESA GSC-DA) acquired in spring 2006 and winter 2010 have been used to generate the corresponding DSMs, using the Semi-Global Matching algorithm (SGM) implemented at DLR ([7], [8]). Since the resulting DSMs show holes when the matching between the two stereo images fails (e.g. over occluded areas), the delta surface fill technique has been applied to fill the DSMs with data from the corresponding SRTM DSM. Fig. 1 displays the two filled DSMs.

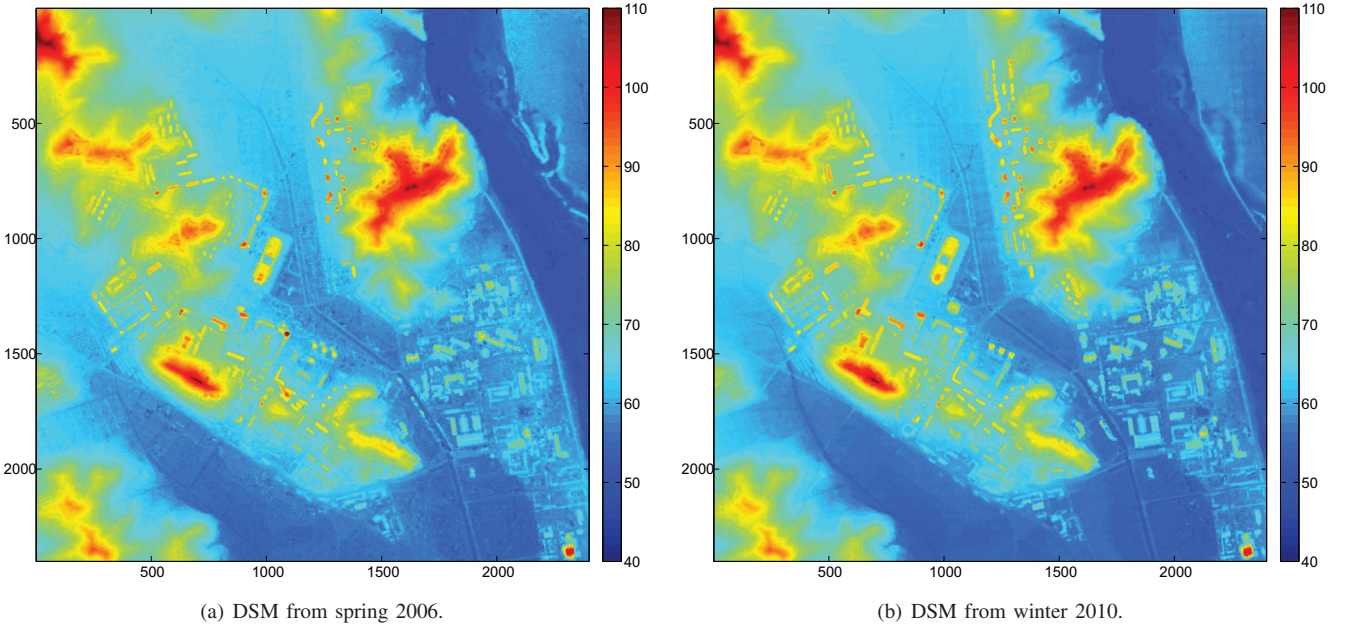


Fig. 1. The two DSMs used in this work to perform a multi-temporal and multi-season change monitoring of an Asiatic urban area.

### III. CHANGE DETECTION METHODOLOGY

A typical solution to detect positive and negative changes consists in subtracting one DSM from the other. Such an approach provides generally good results when every pixel in the image represents the real height of the corresponding point in the studied area. However, when at least one of the DSMs exhibits artifacts, this simple approach can not be reliable. In this work, after the subtraction of the two DSMs depicted in Fig. 1, we propose to introduce some post-processing steps in order to generate accurate change detection results. Actually, after examining deeply the DSM difference image depicted in Fig. 2, we have found out that:

- most of the virtual changes come from the DSMs artifacts caused either by the SRTM-based filling over shadowed areas or by some precision errors in the height computation. To overcome the first problem, we propose to eliminate shadows over the wrongly filled areas. Whereas, a histogram-based thresholding and shape feature extraction were included in our change detection procedure to remove automatically the virtual changes caused by height computation errors.
- the real changes are mainly linked to the construction of several new buildings or to varying levels of vegetation growth since the data have been acquired in different seasons. Therefore, in our change detection procedure, we will focus on building detection using shape features and vegetation detection using the Normalized Differenced Vegetation Index (NDVI) computed from the multispectral images.

#### A. Shadow elimination

One of the common artifacts in DSMs over urban areas comes from the SRTM-based filling. Exemplarily, neighboring

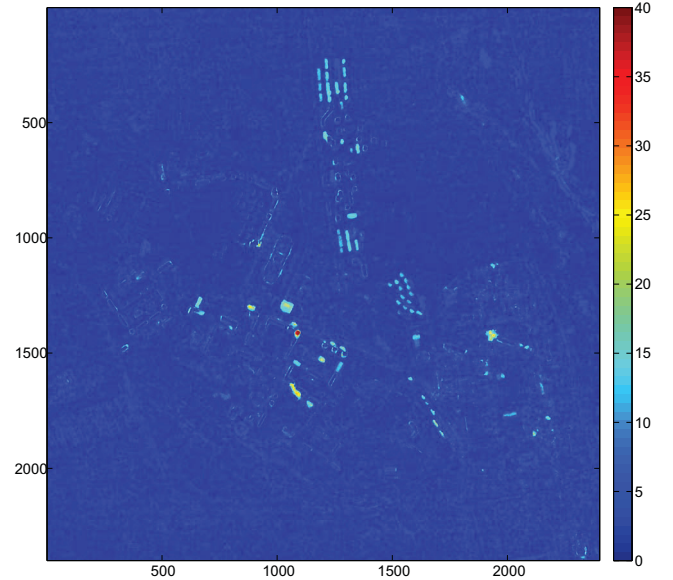


Fig. 2. Absolute difference between the DSMs of Fig. 1.

buildings separated by a narrow road appear usually as one connected structure in the filled version of the DSM as can be seen in Fig. 3. In order to recover this problem, we propose to apply the following shadow-hole mask on the DSM difference image:

$$Mask = (Mask_{Shadow}^L \cup Mask_{Shadow}^R) \cap Mask_{Hole}, \quad (1)$$

where  $Mask_{Shadow}^L$  and  $Mask_{Shadow}^R$  are the shadow masks computed, according to the method of [9], from the left and right stereo images, respectively, and  $Mask_{Hole}$  represents the hole mask calculated from the unfilled DSM.

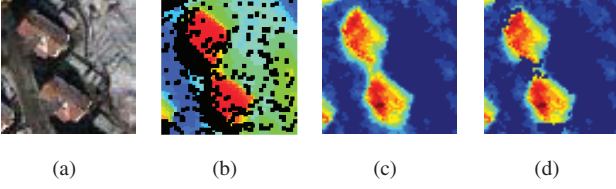


Fig. 3. Illustration of the usefulness of the shadow detection in improving the DSM quality inside urban areas. (a) Multi-spectral image. (b) Unfilled DSM. (c) DSMs subtraction. (d) Improved DSMs subtraction.

### B. Histogram-based thresholding

As done in [10], we perform a histogram-based thresholding on the DSM difference image after applying the shadow-hole mask to remove the virtual changes coming from height computation errors. Observing the histogram of the difference image (Fig. 4), possible changes are located far away from the average value. To determine the threshold  $\text{Thres}_{\text{pos}}$  relative to the possible positive changes, a histogram-based thresholding is applied as follows:

$$\text{Thres}_{\text{pos}} = \min_{k \in [0, h_{\text{pos}}^{\text{max}}]} \left( \frac{\sum_{i=0}^k h_{\text{pos}}(i)}{\sum_{i=0}^{h_{\text{pos}}^{\text{max}}} h_{\text{pos}}(i)} \right) > 0.99, \quad (2)$$

where  $h_{\text{pos}}$  is the histogram relative to the positive changes and  $h_{\text{pos}}^{\text{max}}$  is the maximal height difference. The threshold relative to the possible negative changes is similarly computed.

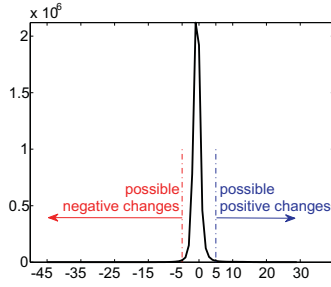


Fig. 4. Histogram of the DSM difference image whose absolute value is depicted in Fig. 2 after applying the shadow-hole mask.

### C. Feature extraction and SVM-based real change detection

After thresholding, the changes are treated as segments and different features have been computed for each segment in order to describe at best the real positive and negative changes. The mainly observed 3D changes are linked either to vegetation changes since the two stereo images have been acquired in two different seasons (spring and winter), or to building construction since a quite long period (4 years) separates the acquisition years of the two stereo image pairs. Therefore, for feature extraction, the following steps have been followed:

- First, NDVI masks are computed from the multi-spectral images and their overlaps with the change segments are examined in order to detect changes relative to vegetation.
- After that, assuming that the rest of the change segments are linked to building construction, we suggest to compute the following shape features for each segment: area,

elongation (ratio of the major axis length and the minor axis one), eccentricity (ratio of the distance between the foci of the ellipse that has the same second-moments as the segment, and its major axis length), solidity (proportion of the pixels in the convex hull that are also in the segment), extent (ratio of pixels in the segment to pixels in the total bounding box) and compactness (ratio of the square root of the area to the perimeter of the segment). In addition to these features, we compute the mean and standard deviation of the height over each segment.

After feature extraction, we propose to use SVM to classify the segments into real and virtual changes. We run SVM 10 times with different training and testing data to avoid any dependency between the choice of the training data and the classification results. In Fig. 5, we provide the mean of the 10 repetitions. The degrees of redness and blueness represent the probability of each segment to be a real change: the higher the mean value of each segment is, the more probable the segment to a real change corresponds.

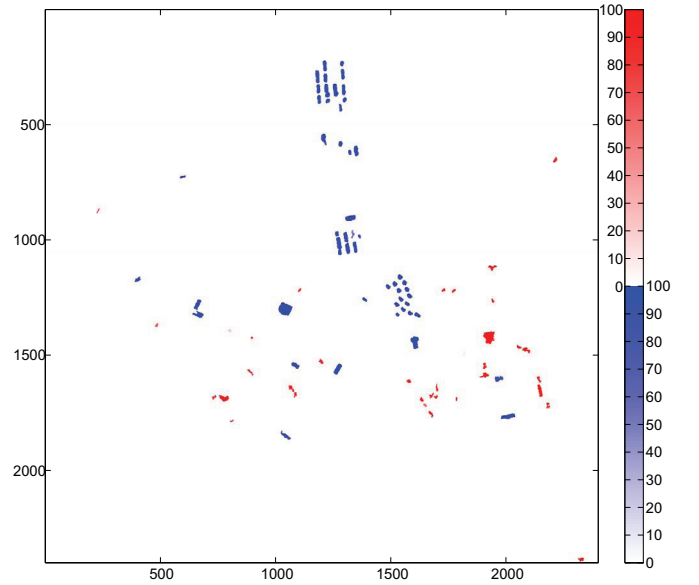


Fig. 5. Positive (in blue) and negative (in red) change map. The degrees of blueness and redness represent the probability of each segment to be a real change.

## IV. ACCURACY ASSESSMENT

To assess the accuracy of our change detection results, we have compared a sub-image of the change map (Fig. 6(a)) to a Ground Truth (GT) map (Fig. 6(b)) that has been manually derived from the stereo images. We could notice that only two changes (whose centroids are located approximately at (1360, 1250) and (1400, 1280)) out of 32 have not been detected. They correspond actually to 3m height differences which have been removed during the thresholding step. Also the changes whose centroids are located at (1275, 1010) and (1278, 1055) have been detected as one connected component since their boundaries are separated only by 4 pixels.



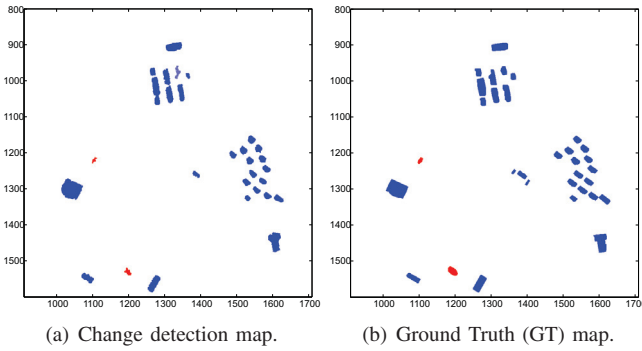


Fig. 6. Change detection results versus GT.

A pixel-to-pixel evaluation of our change detection approach in terms of confusion matrix, is summarized in Table I where TP (True Positive) and TN (True Negative) are the numbers of pixels classified as "Change" and "Non-change" by both maps, respectively, and FP (False Positive) and FN (False Negative) are the numbers of pixels classified as "Change" only in our change detection map or only in the GT one, respectively.

TABLE I

PIXEL-TO-PIXEL EVALUATION OF OUR CHANGE DETECTION ALGORITHM.

Our results	Ground truth	
	Change	Non-change
Change	13530 (TP)	1362 (FP)
Non-change	3083 (FN)	622025 (TN)

Based on the quantities computed in Table I, the following objective metrics ([4], [11]) were employed to provide a quantitative assessment of our change detection algorithm:

$$\text{Branching Factor} = \text{FP/TP}, \quad (3)$$

$$\text{Miss Factor} = \text{FN/TP}, \quad (4)$$

$$\text{Completeness(\%)} = 100 \times \text{TP}/(\text{TP} + \text{FN}), \quad (5)$$

$$\text{Correctness(\%)} = 100 \times \text{TP}/(\text{TP} + \text{FP}), \quad (6)$$

$$\text{Quality Percentage(\%)} = 100 \times \text{TP}/(\text{TP} + \text{FN} + \text{FP}), \quad (7)$$

$$\text{Overall Accuracy(\%)} = 100 \times (\text{TP} + \text{TN})/\#\text{pixels}. \quad (8)$$

Each metric mentioned above provides its own quantitative measure for evaluating the overall performance of the algorithm. The branching and miss factors describe the two types of potential mistakes (FP and FN) that may occur in the automatic process. The completeness represents the percentage of "Change" pixels which are correctly detected while the correctness shows the percentage of detected "Change" pixels which belong indeed to the "Change" class. The quality percentage describes how likely a "Change" pixel produced by the automatic approach is true, and is the most stringent measure of the overall results of the six statistics. The overall accuracy shows the percentage of correctly classified pixels. The suggested 3D change detection algorithm shows a branching factor of 0.1 and a miss factor of slightly poor performance (0.22). This indicates that the number of over-classified "Change" pixels is less than the number of missed "Change"

pixels. A rate of 81.4% in completeness and a higher rate of 90.8% in correctness have been obtained. This is also due to the tendency of our algorithm to produce less FP pixels than FN ones. Finally, the proposed change detection technique shows a quality percentage of 75.2% and an overall accuracy of 99.3%, proving its efficiency and relatively high accuracy.

## V. CONCLUSIONS AND PERSPECTIVES

This paper suggested an automatic urban area monitoring technique by extracting height and spatial information from DSMs generated from two pairs of stereo data acquired at different epochs. Height changes are computed through DSM subtraction. Whereas, spatial information is extracted by computing several shape features for each change. Finally, the separation between real and virtual changes is performed through SVM-based classification. To provide a quantitative assessment of the developed change detection algorithm, different common objective metrics such as completeness, correctness, etc have been computed. 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 for instance a fusion between the multi-spectral data and the DSMs is considered in the overall change detection process.

## ACKNOWLEDGMENT

The present material was produced in the context of the GMOSAIC project, co-funded by the European Commission within the 7<sup>th</sup> Framework Programme. It is intended only for evaluation purposes and not for redistribution or resale.

## REFERENCES

- [1] T. Hollands, G. Boström, J. G.M. Goncalves, K. Gutjahr, I. Niemeyer and V. Sequeira, "3D Scene Change Detection from Satellite Imagery", Proc. 29th Symp. on Safeguards and Nuclear Material Management, Aix-en-Provence, France, 2007, pp. 1-6.
- [2] P. Gong, G.S. Biging and R. Standiford, "Use of Digital Surface Model for Hardwood Rangeland Monitoring", J. of Range Management, 2000, Vol. 53, No. 6, pp. 622-626.
- [3] A.J. Heller, Y.G. Leclerc and Q.T. Luong, "A Framework for Robust 3-D Change Detection", Proc. of SPIE, Toulouse, France, 2001.
- [4] N. Ekhtari, M.R. Sahebi, M.J. Valadan Zoej and A. Mohammadzadeh, "Automatic Building Detection From Lidar Point Data", The International Arch. of the Photogrammetry, Remote Sens. and Spatial Information Sciences, Vol. XXXVII, Part B4, Beijing 2008.
- [5] L. Matikainen, J. Hyypä, E. Ahokas, L. Markelin and H. Kaartinen, "Automatic Detection of Buildings and Changes in Buildings for Updating of Maps", Remote Sens. 2010, Vol. 2, pp. 1217-1248.
- [6] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, 1998.
- [7] H. Hirschmüller, "Stereo Processing by Semiglobal Matching and Mutual Information", IEEE TPAMI, Vol. 30, No. 2, 2008, pp. 328-341.
- [8] P. d'Angelo, M. Lehner, T. Krauss, D. Hoja and P. Reinartz, "Towards Automated DEM Generation from High Resolution Stereo Satellite Images", In: Proc. of ISPRS 2008, Peking, China, 2008, pp. 1137-1342.
- [9] J. A. Marchant and C. M. Onyango, "Shadow-Invariant Classification for Scenes Illuminated by Daylight", J. Opt. Soc. Am. A, Vol. 17, No. 11, 2000.
- [10] Vu Tuong Thuy, M. Matsuoka and F. Yamazaki, "LIDAR-based Change Detection of Buildings in Dense Urban Areas", IGARSS 2004, Vol. 5, pp. 3413-3416.
- [11] G. Sohn and I. Dowman, "Data Fusion of High-resolution Satellite Imagery and LiDAR Data for Automatic Building Extraction", ISPRS J. of Photogrammetry and Remote Sens. 62 (2007), pp. 43-63.