Article

Towards Automatic 3D Change Detection inside Urban Areas by Combining Height and Shape Information

Houda Chaabouni-Chouayakh & Peter Reinartz, Oberpfaffenhofen

Summary: Monitoring of urban areas using remote sensing data requires reliable change detection techniques. While most of the changes are optically visible and easily detectable by an expert user, automatic processes that remain valid even when different kinds of input data are considered, are quite difficult to develop. This paper provides new solutions for semi-automatic 3D change detection of buildings based on the joint use of height and spatial information. It is an attempt to build a reliable scheme for change detection able to process high as well as lower quality Digital Surface Models (DSMs). The subtraction of DSM, computed from stereo pairs acquired at different epochs, provide valuable information about 3D urban change. However, when at least one of the DSMs presents some artifacts, a simple DSM subtraction may result also in the detection of virtual changes. Several post-processing steps are proposed in this paper and adapted to different DSM qualities in order to quantify real changes. Shape features are introduced to describe the geometry of the detected changes and a Support Vector Machine (SVM) classifier is used to differentiate real from virtual changes. Evaluation of the proposed approach on object and pixel level in terms of completeness, correctness, overall accuracy, etc is performed, proving its efficiency and relatively high accuracy for different kind of stereo images and consequently different DSM qualities.

Zusammenfassung: *Automatische 3D-Veränderungsanalyse in städtischen Gebieten durch die Kombination von Höhen und Form Information.* Das Monitoring städtischer Gebiete mit Fernerkundungsdaten erfordert zuverlässige Verfahren der Veränderungsanalyse. Obwohl die meisten Änderungen visuell durch einen erfahrenen Bildinterpreten leicht erkennbar sind, erweist sich die Entwicklung automatischer Verfahren, die auch dann verlässliche Ergebnisse liefern, wenn unterschiedliche Qualitäten von Input-Daten vorliegen, als sehr problematisch. Dieser Artikel zeigt neue Ansätze zur semi-automatischen Erkennung von 3D Veränderungen von Gebäuden, die auf der gemeinsamen Nutzung der räumlichen sowie der Höhen-Information basieren. Ziel ist die Erstellung eines Verfahrens zur Erkennung von Veränderungen in Digitaler Oberflächenmodellen (DSM) welches auch mit Daten geringerer Qualität zuverlässig arbeitet. Die Subtraktion von DSM erstellt aus Stereo-Bildpaaren aus unterschiedlichen Zeiten enthält wertvolle Information über neue oder abgerissene Gebäude. Wenn allerdings eines oder beide DSM Fehler aufweisen, wird eine einfache Subtraktion von DSMs auch virtuelle Veränderungen enthalten. Aus diesem Grund werden mehrere Verarbeitungsschritte entwickelt um virtuelle Änderungen auch in DSM schlechterer Qualität möglichst weitgehend zu eliminieren. Es wird die Verwendung von vorgegebenen Form-Eigenschaften vorgeschlagen, welche die Geometrie der erfassten 3D Objekte beschreiben und ein Klassifikator aus dem Bereich Support Vector Machines (SVM) wird zur Unterscheidung von virtuellen und echten Änderungen verwendet. Die anschließende Bewertung auf Objekt- und Pixel-Level im Hinblick auf Vollständigkeit, Richtigkeit und Genauigkeit zeigt die Effizienz des vorgeschlagenen Ansatzes für verschiedene Arten von Stereo-Bilddaten mit unterschiedlichen DSM Qualitäten.

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. New innovative tools are thus required for better monitoring of such areas. Remotely sensed imagery in some cases may be a very 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 a reliable automatic urban growth monitoring scheme valid even when input remote sensing data with different qualities and/or from different sources are considered. The proposed processing scheme is based on the joint use of height and spatial or shape information. In fact, urban changes are in general either related to building and road construction/demolition or vegetation growth. These issues should be well described if height and shape information are available. Many works have been dedicated to detect 2D changes using remote sensing data as reviewed in (Lu et al. 2004, Radke et al. 2005), but the issue of 3D change detection has been seldom tackled. In the literature, to monitor height changes, subtraction of Digital Surface Models (DSMs) has been widely used (Gong et al. 2000, Heller et al. 2001, Hollands et al. 2007). 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 may 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. A similar approach has been used in the works of (Ekhtari et al. 2008, Matikainen et al. 2010) to detect buildings and building changes using Lidar and Laser DSMs. Accurate building detection maps

have been obtained, but this is somehow due to the very good quality of the used DSMs and to the successful tuning of the different thresholds. This height and spatial information joint exploitation has also been the focus of some of our previous publications (Tian et al. 2011, Chaabouni-Chouayakh et al. 2011) and has shown quite promising change detection results for DSMs generated from satellite images. This article aims at generalizing this approach so that it remains valid for a wider range of DSM qualities and sources. In here, two different DSM datasets have been involved in order to evaluate the ability of the proposed method to detect changes from various input remote sensing data. The first DSM dataset has been produced from airborne images acquired at two different epochs from the DLR 3K camera system (Kurz et al. 2007). Whereas, the second DSM dataset is generated from two pairs of Ikonos satellite stereo images acquired at different epochs. Since the quality of the considered DSM datasets is not as good as the Lidar/Laser ones, some adaptive post-processing steps have been included so that the proposed approach does not remain limited only to very high quality DSMs. Also, still in the frame of the automation of the proposed change detection approach and to avoid the manual tuning of the different thresholds, after the feature extraction step, this work suggests to separate the real changes from the virtual ones, using the Support Vector Machine (SVM) classifier which has shown high efficiency and robustness in various pattern recognition applications (Burges 1998).

The organization of this paper is as follows: Section 2 gives a short overview about the DSM processor adopted to generate the DSMs. Section 3 describes the data used in this work while Section 4 provides the different steps of the proposed 3D change detection approach. Section 5 assesses the accuracy of the method using several objective metrics. Finally, Section 6 gives some conclusions and proposes some perspectives.
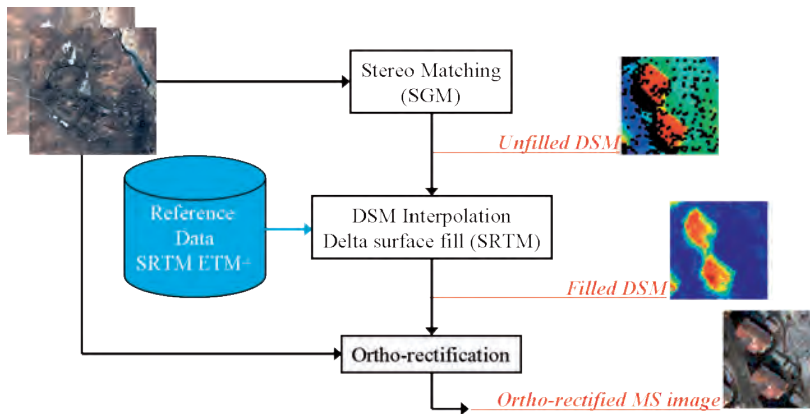
Fig. 1: DSM processor overview.

2 DSM Processor Overview

To generate the different DSMs investigated in this work, the Semi-Global-Matching (SGM) algorithm implemented at DLR (HIRSCHMÜLLER 2008, D'ANGELO et al. 2008) has been used. The high performance of SGM in comparison to other DSM processors has been highlighted in (HEIPKE et al. 2007) and its overall processing scheme is displayed in Fig. 1.

In the following, the different steps of the DSM processor of Fig. 1 are explained in more detail. For each pair of stereo images:

1. A stereo matching is first applied using SGM. This results in a first DSM version, usually called *"unfilled DSM"* since it shows some holes when the matching between the two stereo images fails (e. g., in occluded areas).

2. Then, the Delta surface fill technique is adopted to fill the *"unfilled DSM"* with data from the corresponding DSM in the SRTM ETM+ reference data. This has the notion that a very coarse DSM is used to interpolate a high resolution DSM but this technique is adapting the SRTM DSM to the local height values of the SGM-DSM and therefore leads to a meaningful filling. This step provides the final DSM version, called *"filled DSM"*. Although this interpolation step is highly required to get higher DSM quality, it might be omitted when the unfilled DSM does not present too many oc-cluded areas. This is generally the case when more than one pair of stereo images[1] is used as input to the DSM processor. Actually, it is worth to note that one of the common artifacts in DSMs over urban areas comes from the interpolation over some problematic areas such as shadowed areas from buildings. This will be explained in more detail in Section 4.1.

3. Additionally, ortho-rectified Multi-Spectral (MS) images (according to the DSM) might also be provided, when necessary.

3 Presentation of the Data

The evaluation of the proposed change detection approach has been carried out using two DSM datasets with different quality:

- **Munich dataset: High quality unfilled DSMs (Fig. 2):**
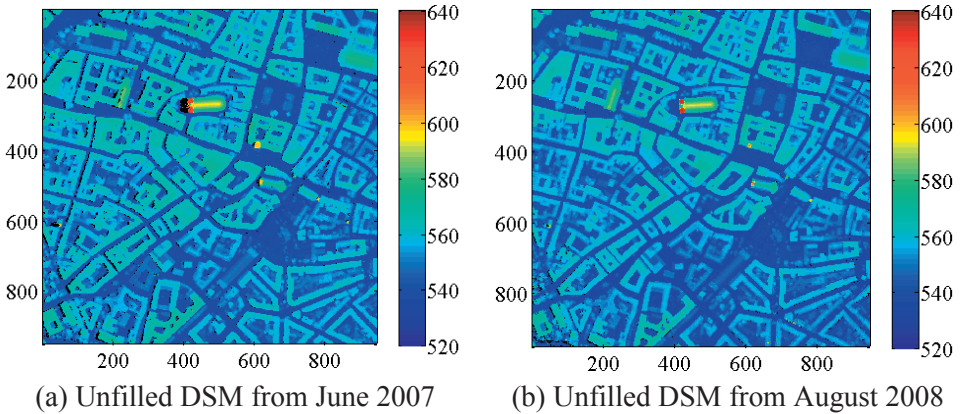 These DSMs have been generated after applying SGM on airborne images from June 2007 and August 2008 over the city of Munich in Germany. These images have been provided by the DLR 3K camera system (Kurz et al. 2007). The 2007 data are composed of 7 nadir and 7 side-looking images

---

[1] When more than one pair of stereo images is available, the SGM stereo matching is first performed on each pair of these pairs. Then, the resulting *"unfilled DSMs"* are combined using a median value to generate one final *"unfilled DSM"*.
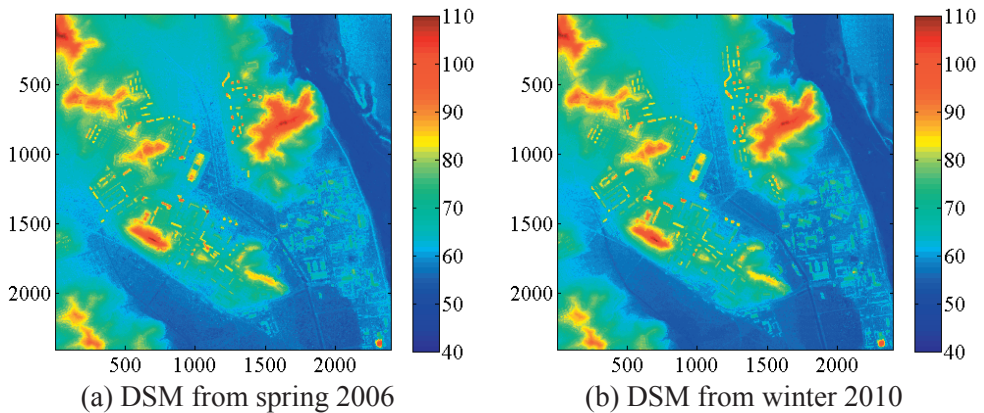
with a side overlap of 66%. Neighboring na-dir and side-looking pairs were processed as described in Fig. 1 to generate the final unfilled DSM. However, the 2008 data are made of 18 nadir images with an along-track overlap of 80%. The corresponding unfilled DSMs have been generated in this case after processing several neighboring nadir pairs. The two results are displayed in Fig. 2. It can be noticed that the unfilled DSM of 2008 shows much less occluded areas (black pixels) than the one of 2007. This can be explained by the fact that a mix-ture of nadir and side-looking images has been used to generate the 2007 data while only nadir images (which are often prefer-red to the side-looking images) have been involved in the generation of the 2008 data. Another possible explanation consists in the larger number of image pairs used in the case of the 2008 data. Despite its occluded areas, the 2007 unfilled DSM is still consi-dered to have a tolerable quality and the Delta surface fill-based interpolation of the DSM processor could consequently be omitted for both 2007 and 2008 data.

It should be noted that the Munich dataset has been used in this work to check the ro-bustness of the proposed approach in per-forming a multi-temporal urban area change monitoring using high quality DSM without integrating the potential DSM computation

(a) Unfilled DSM from June 2007   (b) Unfilled DSM from August 2008

Fig. 2: Munich dataset: The two unfilled DSMs used in this work to perform a multi-temporal change monitoring of the city of Munich in Germany. The black pixels in the DSM images corre-spond to no-data values (occluded areas).

(a) DSM from spring 2006   (b) DSM from winter 2010

Fig. 3: Yeonwon dataset: The two filled DSMs used in this work to perform a multi-temporal and multi-season change monitoring of the city of Yeonwon in North Korea.

Tab. 1: DSM Quality evaluation.

	Munich Dataset		Yeonwon Dataset	
	DSM 2007	**DSM 2008**	**DSM 2006**	**DSM 2010**
Percentage of no-data pixels	2%	0.4%	29.6%	29.7%
STD [m]	0.45	0.48	0.64	0.62

errors coming from the interpolation step as described in Section 2.

- **Yeonwon dataset: Lower quality filled DSMs (Fig. 3):**
  These DSMs have been generated from two pairs of Ikonos-2 stereo images (©EUSI provided under the EC/ESA GSC-DA) acquired in spring 2006 and winter 2010 over the city of Yeonwon in North Korea. The generated filled DSMs (Fig. 3) exhibit a sampling distance of 1 m and a quite good quality. Indeed, most of the man-made structures are well reconstructed.

  Note that the Yeonwon dataset has been used to check the robustness of the proposed approach in performing a multi-temporal and multi-season urban area change monitoring using lower DSM qualities generated from satellite data. Since also multispectral information is available, it will be later included in the proposed change detection scheme in order to improve its performance.

The quality of the two DSM datasets has been measured in terms of 1) percentage of no-data pixels (over the occluded areas and due to other matching problems) and 2) Standard deviation (STD) for completely flat areas. Tab. 1 provides the quantitative evaluation of the quality of the DSMs used to detect changes in the cities of Munich and Yeonwon.

4   Change Detection Procedure

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.

After the DSM subtraction, this work proposes to introduce some adaptive post-processing steps to generate better change detection results. Note that the subtraction of unfilled DSMs consists in differentiating the data only over the areas where the stereo matching did not fail, as described in the following statement:
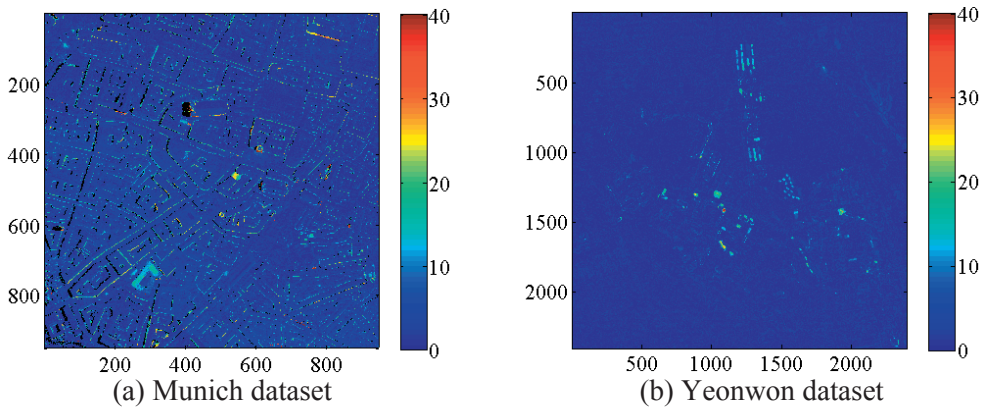
For each pixel (i, j):

$$\text{if} \quad (\text{uDSM}_1(i, j) \neq \text{NoData}) \text{ and}$$
$$(\text{uDSM}_2(i, j) \neq \text{NoData}), \qquad (1)$$
$$\text{then } \text{Diff}_{\text{uDSM}}(i, j)$$
$$= \text{uDSM}_2(i, j) - \text{uDSM}_1(i, j),$$
$$\text{else } \text{Diff}_{\text{uDSM}}(i, j) = 0$$

where $\textbf{uDSM}_k$, $(\textbf{k}=1,2)$ are the two unfilled DSMs displayed in Fig. 3 and $\textbf{Diff}_{\textbf{uDSM}}$ denotes the difference image while **NoData** corresponds to the value adopted to identify the not matched areas.
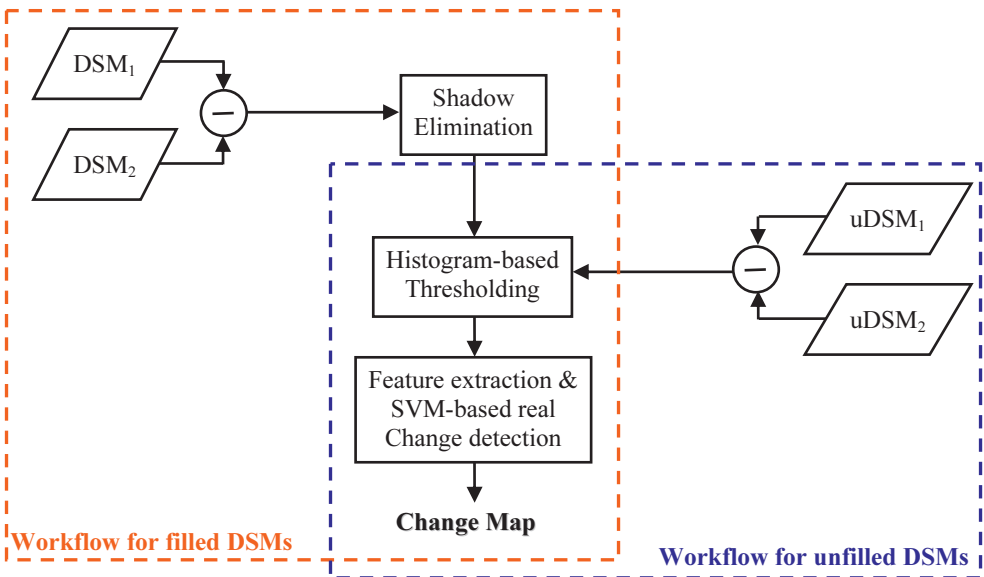
Fig. 4 depicts the two absolute DSM differences, when the statement (1) is applied to the Munich dataset unfilled DSMs of Fig. 3 and a simple subtraction of the Yeonwon dataset filled DSMs of Fig. 4 is performed.

After examining deeply the DSM difference images depicted in Fig. 4, the following observations can be stated:

- Most of the virtual changes come from the DSM artifacts caused either by the SRTM-based filling over certain occluded areas, specially the ones corresponding to shadowed regions of buildings (in the case of the Yeonwon dataset), or by some precision errors in the height computation of the DSM. In order to overcome the first problem, elimination of the shadowed regions over the wrongly filled areas is suggested in this work (Section 4.1). Whereas, a histogram-based thresholding and shape feature extraction are included in the overall proposed change detection procedure to remove automatically the virtual changes caused by

(a) Munich dataset

(b) Yeonwon dataset

Fig. 4: DSMs absolute difference. The black pixels in (a) correspond to no-data values in at least one of the unfilled DSMs of Fig. 2.
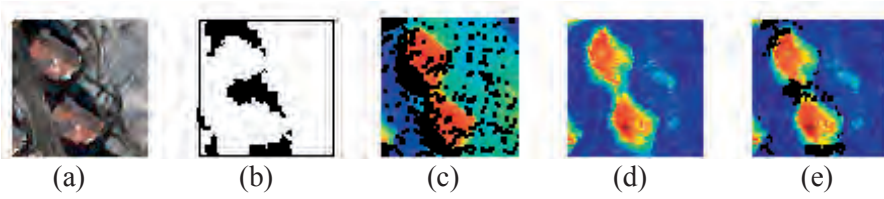

Fig. 5: Overall workflow of the proposed change detection procedure.

height computation errors (Sections 4.2 and 4.3).

- The real changes correspond to either construction/destruction of buildings or to varying levels of vegetation growth. For the Munich dataset, only changes linked to building construction/destruction are considered using several shape features. However, for the Yeonwon dataset, vegetation changes are additionally monitored since the data have been acquired in completely dif-

ferent seasons (spring and winter). Also, unlike the Munich dataset, in this case the availability of the multi-spectral data should allow accurate vegetation detection through the simple computation of the Normalized Differenced Vegetation Index (NDVI). The extracted features in both cases are then used to feed a SVM classifier to distinguish between real and virtual changes.

- The overall workflow of the proposed change detection procedure is displayed in Fig. 5.

(a)   (b)   (c)   (d)   (e)

Fig. 6: Illustration of the usefulness of the shadow detection in improving the DSM quality inside urban areas. (a) Multi-spectral image. (b) Shadow mask. (c) Unfilled DSM. (d) DSMs subtraction. (e) Improved DSMs subtraction (the masked pixels are highlighted in black).

4.1 Shadow Elimination

One of the common artifacts in DSMs over urban areas comes from the filling with a coarser DSM in locally high varying height regions. Exemplarily, neighboring buildings separated by a narrow road appear usually as one connected structure in the filled version of the DSM with a quite strange shape, as can be seen in Fig. 6(a-c). Such an artifact bothers the characterization of the real changes using shape features. To recover this problem, the following shadow-hole mask has been applied to the DSM difference image:

$$\text{Mask} = (\text{Mask}^{\text{L}}_{\text{Shadow}} \cup \text{Mask}^{\text{R}}_{\text{Shadow}}) \cap \text{Mask}_{\text{Hole}}, \quad (2)$$

where $\textbf{Mask}^{\textbf{L}}_{\textbf{Shadow}}$ and $\textbf{Mask}^{\textbf{R}}_{\textbf{Shadow}}$ are the shadow masks computed, according to the method of (MARCHANT & ONYANGO 2000), from the left and right stereo images, respectively, and $\textbf{Mask}_{\textbf{Hole}}$ represents the hole mask calculated from the unfilled DSM.

Fig. 6(d) illustrates well the usefulness of the shadow detection step in improving the DSM quality inside urban areas. The characterization of changes relative to single building construction/destruction becomes therefore easier through shape feature extraction.
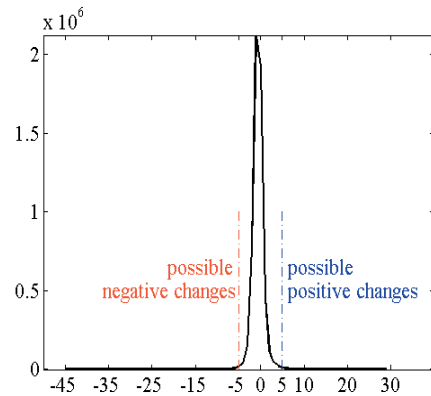
4.2 Histogram-based Thresholding

As done in the work of (TUONG THUY et al. 2004), in order to remove the virtual changes coming from height computation errors, a histogram-based thresholding is performed on:

- the unfilled DSM difference in the case of the Munich dataset (Fig. 4(a)); and
- the DSM difference image (Fig. 4(b)) in the case of the Yeonwon dataset after applying the shadow-hole mask described by Eq. (2).

In fact, observing exemplarily the histogram of the Yeonwon dataset difference image after shadow elimination depicted in Fig. 7, it could be noticed that possible real changes are located in the tail of the histogram (far from the average value, which is close to zero).

To determine the threshold $\textbf{Thres}_{\textbf{pos}}$ relative to the possible positive changes, a histogram-based thresholding is applied as follows:

$$\text{Thres}_{\text{pos}} = \min_{k \in [0, h^{\max}_{\text{pos}}]} \left(\frac{\sum_{i=0}^{k} h_{\text{pos}}(i)}{\sum_{i=0}^{h^{\max}_{\text{pos}}} h_{\text{pos}}(i)} \right) > \alpha_{\text{pos}}, \quad (3)$$

Fig. 7: Histogram of the DSM difference image whose absolute value is depicted in Fig. 4(b) after applying the shadow-hole mask. Possible real changes are located in the tail of the histogram.

where h_{pos} is the histogram relative to the positive changes, h^{max}_{pos} is the maximal height difference and α_{pos} denotes the bound over which a height difference is identified as real change. The threshold **Thres**$_{neg}$ relative to the possible negative changes is similarly computed. In the case of the Munich and Yeonwon dataset, α-bounds of 0.7 and 0.99 have been found to be a good compromise in detecting real changes, maintaining 9 and 1.2% of the pixels for the next steps, respectively. Although these bounds have been here customized to each of the considered dataset, more automation in this thresholding step is expected in future work.

4.3 *Feature Extraction and SVM-based Real Change Detection*

After the thresholding step, the changes are no more treated as single pixels, but as objects which are represented through connected pixel regions after applying Eq. (3). Different features are then computed for each object in order to describe at best the real positive and negative changes.

In the case of the Munich dataset, only changes that correspond to building construction/destruction have been studied. That is why, the following shape features have been computed for each object in this work:

- area
- elongation (ratio of the major axis length and the minor axis one)

$$\text{Elongation} = \frac{\text{Major Axis Length}}{\text{Minor Axis Length}} \qquad (4)$$

- 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)

$$\text{Eccentricity} = \frac{\text{Distance between the foci of the ellipse that has the same second-moments}}{\text{Major Axis Length}}. \qquad (5)$$

- solidity (proportion of the pixels in the convex hull that are also in the segment)

$$\text{Solidity} = \frac{\text{Area}}{\text{Area of convex hull}}. \qquad (6)$$

- extent (ratio of pixels in the segment to pixels in the total bounding box)

$$\text{Extent} = \frac{\text{Area}}{\text{Major Axis Length} \times \text{Minor Axis Length}} \qquad (7)$$

- compactness (ratio of the square root of the area to the perimeter of the segment)

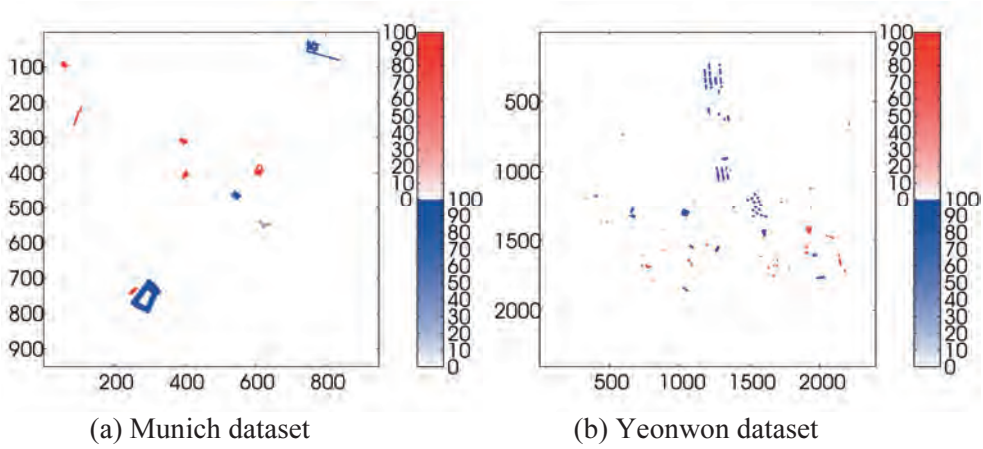$$\text{Compactness} = \frac{2\sqrt{\pi \times \text{Area}}}{\text{Perimeter}}. \qquad (8)$$

In addition to the features described above, the mean and standard deviation of the height over each object have been computed.

In the case of the Yeonwon dataset, the mainly observed 3D changes are linked either to varying levels of vegetation appearance 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, in this case, it is first proposed to compute NDVI masks from the multi-spectral images and examine their overlaps with the change segments in order to detect changes relative to vegetation. After that, assuming that the rest of the change segments are linked to building construction, the previously described features are computed for each non-vegetation segment.

After the feature extraction step, SVM has been used to classify the segments into real and virtual changes. SVM has been run 10 times with different training and testing data to avoid any dependency between the choice of the training data and the classification results.

Fig. 8 provides the final classification maps obtained for the two studied datasets in terms

(a) Munich dataset   (b) Yeonwon dataset

Fig. 8: Change maps: Positive (in blue) and negative (in red) changes. The degrees of blueness and redness represent the probability of each segment to be a real change.
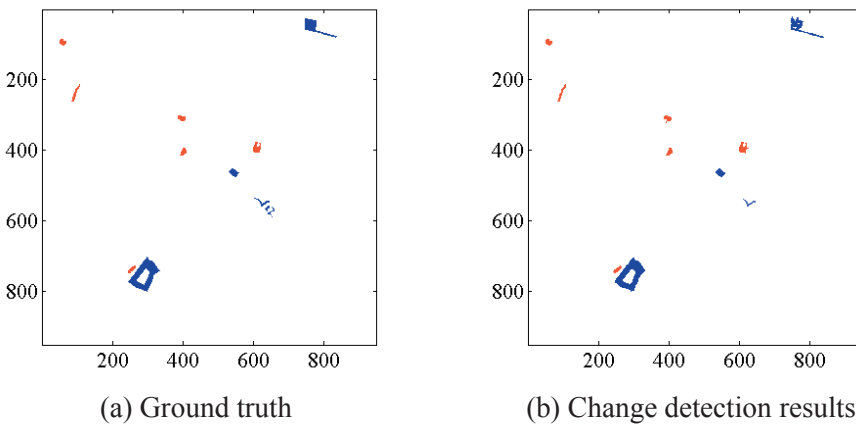
of mean of the 10 repetitions. The degrees of redness and blueness represent the probability of each segment to be a real change: the higher the mean value of each segment is, the more probable the segment corresponds to a real change.
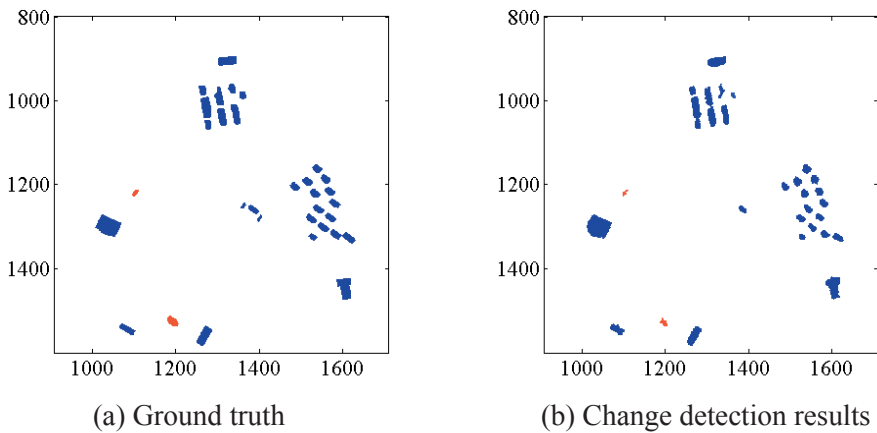
5 Accuracy Assessment

To assess the accuracy of the obtained change detection results, a comparison has been carried out between the change maps of Fig. 8 and the manually derived Ground truth maps

(Figs. 9 and 10). In the case of the Yeonwon dataset, only an 800×800 pixels sub-image of the change map displayed in Fig. 8(b) has been investigated in this step since the whole change map is quite large (2400×2400 pixels).

Fig. 9 shows that two of twelve changes, which can be found at the (x,y)-coordinates (640,560) and (670,570), have been missed in Munich study site. They actually correspond to some scaffoldings used during the restoration of the carillon in the tower of the old town hall in the centre of Munich. They have unusual shapes (in comparison to the rest of the changes), which makes their retrieval as real

(a) Ground truth   (b) Change detection results

Fig. 9: Munich dataset accuracy assessment: Change detection results versus GT.

(a) Ground truth

(b) Change detection results

Fig. 10: Yeonwon dataset accuracy assessment (for part of the data): Change detection results versus GT.

changes with SVM based on shape features quite difficult to perform.

In the case of the Yeonwon dataset (Fig. 10), 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.

A pixel-to-pixel evaluation of the proposed change detection approach in terms of confusion matrix, is summarized in Tab. 2 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 Ground truth one, respectively.

Based on the quantities summarized in Tab. 2, the following objective metrics widely used in the literature (SOHN & DOWMAN 2007, EKHTARI et al. 2008), were computed to provide a quantitative assessment of our change detection algorithm:

$$\text{Branching Factor} = \text{FP/TP} \qquad (9)$$
$$\text{Miss Factor} = \text{FN/TP},$$
$$\text{Completeness(\%)} = 100 \times \text{TP/(TP + FN)},$$
$$\text{Correctness(\%)} = 100 \times \text{TP/(TP + FP)},$$
$$\text{Quality Percentage(\%)}$$
$$= 100 \times \text{TP/(TP+FN+FP)},$$
$$\text{Overall Accuracy(\%)}$$
$$= 100 \times \text{(TP + TN)/(TP + TN + FN + FP)}.$$

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

Tab. 2: Accuracy assessment: Pixel-to-pixel evaluation of the change detection results versus Ground truth.

Munich dataset	Ground truth	
Our results	Change	Non-Change
Change	7085 (TP)	276 (FP)
Non-Change	737 (FN)	894402 (TN)

Yeonwon dataset	Ground truth	
Our results	Change	Non-Change
Change	13530 (TP)	1164 (FP)
Non-Change	3083 (FN)	622223 (TN)

"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 is another metric which evaluates the accuracy of any classification process. It shows the percentage of correctly classified pixels.

Tab. 3 provides an evaluation of the proposed change detection algorithm in the case of the Munich and Yeonwon datasets by means of the objective metrics described in Eq. (4).

According to Tab. 3, the proposed change detection algorithm performs better in the case of the Munich dataset, especially in terms of miss factor, completeness and quality percentage. This is already expected since the corresponding DSMs have a higher quality.

From the first four objective metrics of Tab. 3, we can notice the tendency of the proposed algorithm to produce less FP pixels than FN ones. Indeed, branching factors of 0.04 and 0.08 and miss factors of slightly poor performance (0.1 and 0.22) have been obtained. Also, the two correctness rates are higher than the corresponding completeness ones. This indicates that the number of over-classified "Change" pixels is less than the number of missed "Change" pixels. Finally, the proposed change detection technique could reach a quality percentage of more than 87% when high quality DSMs are involved and a high overall accuracy of 99.89%, proving its efficiency and relatively high accuracy.

Tab. 3: Accuracy assessment: Objective metrics based evaluation of the obtained change detection results versus ground truth.

Objective Metrics	Munich dataset	Yeonwon dataset
Branching Factor	0.04	0.08
Miss Factor	0.1	0.22
Completeness (%)	90.58	81.44
Correctness (%)	96.25	92.08
Quality percentage (%)	87.49	76.11
Overall Accuracy (%)	99.89	99.34

The comparison between the object based and the pixel based change detection shows that, if one excludes the scaffolding change in Munich (which can also be seen as a virtual one), the absolute number of detected changed objects is quite high (Completeness and Correctness are 100% for the Munich case and about 94% for the Yeonwon case). So on object level this method is highly reliable even for the lower quality data set, but if one goes into the details of single pixel comparison then probably some parts of buildings are missing, the forms are not exactly given in the DSMs and therefore quality values are lower.

6 Conclusions and Perspectives

This paper suggested solutions for automatic 3D change detection inside urban areas using stereo remote sensing data with different quality. Our solutions are based on the joint use of height and spatial information. Height changes are computed through DSM subtraction followed by histogram-based thresholding. 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.

The approach is tested and evaluated using data taken from different sensors. While the first data set consists of high quality DSMs, each of which is computed from a large number of airborne image pairs, the second data set consists of lower quality DSMs, each of which is generated from only one pair of high resolution satellite stereo images. A number of common objective metrics (branching factor, miss factor, completeness, correctness, quality percentage, overall accuracy) are computed in the frame of the quantitative assessment of the developed change detection algorithm. They show that although the algorithm performs better in the case of the high quality data set, it presents quite accurate change maps also for the lower quality data set.

The results show a real progress towards automatic 3D change detection inside urban areas, although some of them might still be improved and completed. For instance, more automation of the procedure could still be

done at the histogram-based thresholding and classification levels to avoid any parameter tuning. The results can also be considered as preliminary results for some higher level applications, such as refining urban area monitoring through a more intensive exploitation of the multi-spectral information so that a wide range of man-made structures is involved. Moreover, DSMs could be further processed to identify several levels of building or vegetation growth. Future work will include such approaches.

Acknowledgements

References

Burges, C., 1998: A Tutorial on Support Vector Machines for Pattern Recognition. – Kluwer Academic Publishers, Boston, Data Mining and Knowledge Discovery **2**: 121–167.

Chaabouni-Chouayakh, H., d'Angelo, P., Kauss, T. & Reinartz, P., 2011: Automatic Urban Area Monitoring Using Digital Surface Models and Shape Features. – Joint Urban Remote Sensing Event, Munich, Germany, on CD.

d'Angelo, P., Lehner, M., Krauss, T., Hoja, D. & Reinartz, P., 2008: Towards Automated DEM Generation from High Resolution Stereo Satellite Images, – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B4): 1137–1342.

Ekhtari, N., Sahebi, M.R., Valadan Zoej, M.J. & Mohammadzadeh, A., 2008: Automatic Building Detection From Lidar Point Data. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B4): 473−478.

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

Heipke, C., Oberst, J., Albertz, J., Attwenger, M., Dorninger, P., Dorrer, E., Ewe, M., Gehrke, S., Gwinner, K., Hirschmüller, H., Kim, J.R., Kirk, R., Mayer, H., Muller, P., Rengarajan, M., Rentsch, M., Schmidt, R., Scholten, F., Shan, J., Spiegel, M., Wählisch, M., Neukum, G. & HRSC Co-Investigator Team, 2007: Evaluating planetary digital terrain models – The HRSC DTM test. – Planetary and Space Science **55** (14): 2173–2191.

Heller, A.J., Leclerc, Y.G. & Luong, Q.T., 2001: A Framework for Robust 3-D Change Detection. – SPIE, Toulouse, France, on CD.

Hirschmüller, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (2): 328−341.

Hollands, T., Boström, G., Goncalves, J.G.M., Gutjahr, K., Niemeyer, I. & Sequeira, V., 2007: 3D Scene Change Detection from Satellite Imagery. – 29th Symposium on Safeguards and Nuclear Material Management, on CD.

Kurz, F., Müller, R., Stephani, M., Reinartz, P. & Schroeder, M., 2007: Calibration of a Wide-Angel Digital Camera System for Near Real Time Scenarios. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **36** (I/W51): on CD.

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

Marchant, J.A. & Onyango, C.M., 2000: Shadow-Invariant Classification for Scenes Illuminated by Daylight. Journal of the Optical Society of America A: Optics, Image Science, and Vision **17** (11): 1952−1961.

Matikainen, L., Hyyppä, J., Ahokas, E., Markelin, L. & Kaartinen, H., 2010: Automatic Detection of Buildings and Changes in Buildings for Updating of Maps. – Remote Sensing **2**: 1217−1248.

Radke, R. J., Andra, S., Al-Kofahi, O. & Roysam, B., 2005: Image Change Detection Algorithms: A Systematic Survey. – IEEE Transactions on Image Processing **14** (3): 294−307.

Sohn, G. & Dowman, I., 2007: Data Fusion of High-resolution Satellite Imagery and LiDAR Data for Automatic Building Extraction. – ISPRS Journal of Photogrammetry and Remote Sensing **62** (1): 43−63.

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, on CD.

Tuong Thuy, Vu, Matsuoka, M. & Yamazaki, F., 2004: LIDAR-based Change Detection of Buildings in Dense Urban Areas. – IEEE International Geoscience and Remote Sensing Symposium **5:** 3413−3416.

Address of the Authors:

Dr. Houda Chaabouni-Chouayakh, Prof. Dr.-Ing. Peter Reinartz, German Aerospace Center (DLR), Remote Sensing Technology Institute, 82234 Wessling, Tel.: +49-8153-28-2756, -2757, Fax: -1444, e-mail: Houda.Chaabouni@dlr.de, Peter.Reinartz@dlr.de