Building Change Detection Based on Satellite Stereo Imagery and Digital Surface Models

Jiaojiao Tian, Shiyong Cui, and Peter Reinartz, *Member, IEEE*

Abstract—Building change detection is a major issue for urban area monitoring. Due to different imaging conditions and sensor parameters, 2-D information delivered by satellite images from different dates is often not sufficient when dealing with building changes. Moreover, due to the similar spectral characteristics, it is often difficult to distinguish buildings from other man-made constructions, like roads and bridges, during the change detection procedure. Therefore, stereo imagery is of importance to provide the height component which is very helpful in analyzing 3-D building changes. In this paper, we propose a change detection method based on stereo imagery and digital surface models (DSMs) generated with stereo matching methodology and provide a solution by the joint use of height changes and Kullback–Leibler divergence similarity measure between the original images. The Dempster–Shafer fusion theory is adopted to combine these two change indicators to improve the accuracy. In addition, vegetation and shadow classifications are used as no-building change indicators for refining the change detection results. In the end, an object-based building extraction method based on shape features is performed. For evaluation purpose, the proposed method is applied in two test areas, one is in an industrial area in Korea with stereo imagery from the same sensor and the other represents a dense urban area in Germany using stereo imagery from different sensors with different resolutions. Our experimental results confirm the efficiency and high accuracy of the proposed methodology even for different kinds and combinations of stereo images and consequently different DSM qualities.

Index Terms—Building change detection, Dempster–Shafer theory, digital surface model (DSM), Kullback–Leibler divergence, optical stereo data, stereo matching.

I. INTRODUCTION

AUTOMATIC building change detection based on satellite imagery is becoming increasingly important for city monitoring, disaster assessment, and map database updating. As manual image interpretation and vectorization are time consuming and expensive, automatic and semi-automatic change detection procedures are of high interest. With recently developed dense stereo matching methodologies [1]–[3], it is possible to extract improved height information from satellite stereo images, which is helpful for building change detection. The work presented in this article focuses on building change detection based on satellite stereo data and digital surface models (DSMs) generated from these data. Instead of extracting buildings for both dates, we directly focus on the surface change information and combine the change detected from multispectral images and from DSMs to generate a final building change map.

Traditional satellite or airborne image-based automatic building change detection methods are mainly based on radiometric information analysis [4], [5]. The images acquired at two dates are compared pixel by pixel based on the original spectral information. In this case, when the satellite images are acquired at different seasons, with different weather conditions or from different sensors, the radiometric information between the images can be quite different, which often leads to false alarms. Previous change detection methods developed for medium-resolution satellite images for large-scale land cover changes [4], [5], are not efficient for high-resolution images when the real land cover changes are mixed with irrelevant changes. In particular, if particular objects are of interest, as in our case buildings, it will be very difficult to extract those without height information. Many irrelevant changes will be mixed with building changes, particularly when the data are acquired from different sensors or acquired under different imaging geometries (e.g., viewing directions).

Previous DSM-assisted building change detection methods can be classified into three categories. The first is object-based comparison, which is commonly employed for map updating. In this type, DSMs are used to improve the classification accuracy. The extracted building class can then be compared with an existing building map [6]–[11]. The existing building maps can also be used as training data [10]–[12] in the classification procedure.

The second type of approach employs feature-based methods. Height information from DSMs is normally used as change or no-change features to describe building changes [13]. With high quality DSMs from laser scanning, building changes can be extracted through DSM subtraction. The change detection accuracy can be improved if other features are also employed. A region-based method was proposed in [14], roughness and size of the segments from two Lidar data sets were used to separate trees from building changes.

The third approach is to provide "change candidates" using the height information [15]. DSM subtraction is computationally less expensive to get the initial change map [15]–[18], which can be improved to a more precise building change map when additional information from the original image is employed.

Most of the literature in DSM-assisted change detection use DSMs generated from Airborne Laser Scanning (ALS) or airborne stereo data, and only very few articles use optical stereo satellite data (since they exhibit lower resolution than

airborne data). However, since the spatial resolution of satellite stereo data [now up to 50 cm ground sampling distance (GSD)] and the availability of stereo data from space (Worldview-1/-2, GeoEye-1) are increasing, these data are also now suited for DSM-assisted change detection. Due to the relatively low quality of the DSMs from satellite data, a robust methodology has to be developed to get higher accuracy change detection results. As more and more satellite data providers acquire stereo data with a worldwide availability, we think that satellite-based change detection including the height component will be of more importance in the future, particularly for urban areas and in case of disasters.

Therefore, in this paper, an optimized fusion method for building change detection in urban and suburban areas is developed, falling between the second approach (feature-based methods) and the third approach (change candidates) using DSMs produced by stereo matching together with multispectral satellite data. The novelty is the establishment of change and no-change indicators from several data modalities, which are combined in a robust fusion process using statistical measures. Change features from original satellite images are determined using the Kullback–Leibler (KL) divergence. In a double fusion approach, they are combined with the height changes from the DSMs and the no-change indicators from shadow and vegetation classes based on Dempster–Shafer (DS) theory to get a change probability map. Finally, changed single building objects are extracted using shape features. In the experiments section, it is shown for two different urban areas that the developed method leads to good results, but some shortcomings mainly due to DSM quality are presented as well.

II. METHOD

The accuracy of the DSMs generated from stereo images largely depends on the radiometric data quality, the GSD of the data, the convergence angle of the stereo data, and the amount of available stereo pairs. In most cases, the DSM exhibits 3-D information but also different kinds of inaccuracies [19], [20]. The aim of this paper is to generate a high quality change map by fusing the building change indicators and no-building change indicators from DSMs and the original images. Therefore, as shown in Fig. 1, we divide the proposed method into two parts: the fusion of pixel-based building change indicators and refinement of building change result using no-building change indicator. Finally, an object-based building extraction based on shape features is employed.

A. Building Change Indicator

The following change indicators are used to distinguish the building changes from other changes in the fusion procedure.

1) Height Change: DSM is generated using a robust stereo matching algorithm based on semi-global matching (SGM) using a combination of census and mutual information as cost functions [1]–[3]. After DSMs from two dates are generated, first a co-registration between them is necessary to remove any shift in three dimensions. Depending on data availability and accuracy, the DSMs can be of similar or quite different
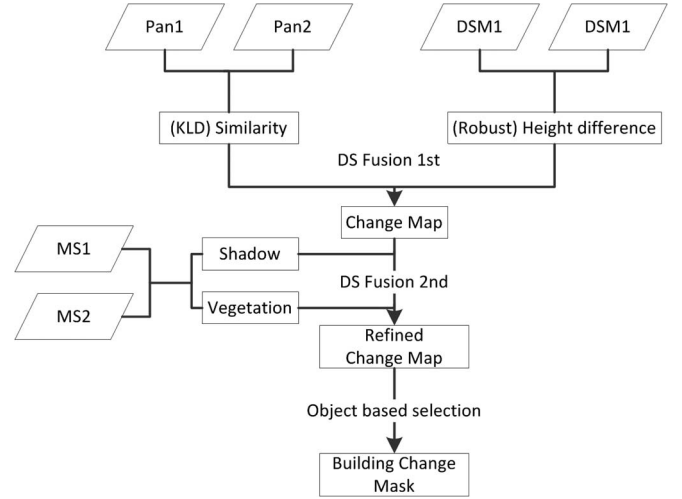

Fig. 1. Flow chart of the proposed method.

quality. If the DSMs have good quality, a pixel-wise subtraction of the two DSMs can already lead to good change results, but for spaceborne stereo data, such a subtraction is generally not applicable, particularly in the case of different sensor, illumination, and stereo angle properties. These factors lead to different matching qualities and therefore different DSM qualities. Fig. 2(b) compares the results of DSM provided by ALS (dashed line) versus SGM-based DSM from high-resolution satellite images (solid line) along the white line in Fig. 2(a). As shown, the SGM-based DSM has a much rougher surface and more blurred features. Therefore, for the comparison of DSMs from various resolutions or various sensors, first a robust height differencing method is necessary.

A robust difference between the initial DSM x_1 and the second DSM x_2 for the pixel (i, j) can be defined as the minimum of differences computed between the pixel $x_2(i, j)$ in the second DSM and a certain neighborhood (with window size $2 \times w + 1$) of the pixel $x_1(i, j)$ in the first DSM x_1. The robust positive and negative differences $X_{Pdif}(i, j)$ and $X_{Ndif}(i, j)$ relative to the pixel (i, j) are defined as written in (1) and (2), respectively

$$X_{\text{Pdif}}(i, j) = \min_{\substack{p \in [i-w, i+w], \\ q \in [j-w, j+w]}} \{(x_2(i, j) - x_1(p, q)),$$

$$(x_2(i, j) - x_1(p, q)) > 0\} \quad (1)$$

$$X_{\text{Ndif}}(i, j) = \max_{\substack{p \in [i-w, i+w], \\ q \in [j-w, j+w]}} \{(x_2(i, j) - x_1(p, q)),$$

$$(x_2(i, j) - x_1(p, q)) < 0\}. \quad (2)$$

It means that only the minimum value (greater than zero) in case of positive change, or the maximum value in case of negative change is taken, all within the defined window size. Typically used window sizes are 3×3 pixels up to 7×7 pixels depending on the DSM qualities and the difference in resolution between the two available DSMs.

2) Similarity Measurement: If at least one of the DSMs exhibits larger errors introduced in the matching procedure or through interpolation, false alarms will still be produced in the height difference map. Therefore, in addition to the robust
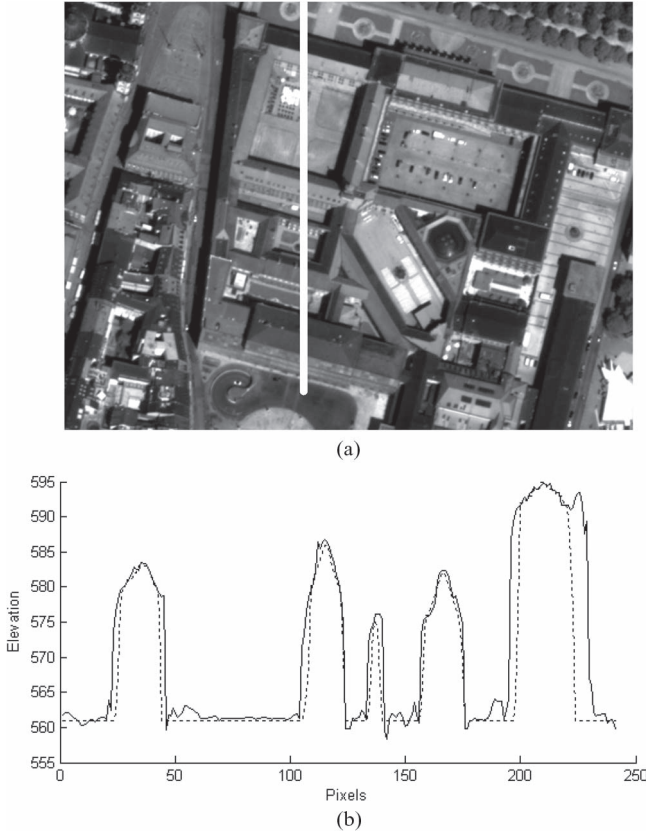
Fig. 2. Comparison of DSMs generated from ALS and stereo matching (Dashed line: ALS DSM; Solid line: SGM-based DSM) (a) shows the location of the test area and (b) is the profile comparison result along the while line in (a).

differencing, a robust change indicator using original satellite images can be very helpful to improve the accuracy. As we have discussed in Section I, radiometric information for the same land cover class can be very different due to the different acquisition circumstances. Instead of comparing gray values directly, we use an information similarity measure to highlight building changes. Information similarity measures, such as Mutual Information, have been widely used in the image processing community [21]–[23]. One prominent work [24] proposed KL divergence for multi-temporal change detection based on the evolution of the local statistics of the image between two dates. The local statistics are estimated by using 1-D Edgeworth series expansion, which approximates probability density functions (PDFs) in the neighborhood of each pixel in the image. In [25], this method was extended for object-based change detection by computing the KL divergence of the two corresponding objects derived by image segmentation.

Assuming the local neighborhood of two corresponding pixels is considered as two random variables X and Y, with f_X and f_Y being the marginal PDF, then the KL divergence of these two random variables is defined as

$$K(Y|X) = \int \log \frac{f_X(x)}{f_Y(x)} f_X(x) dx. \tag{3}$$

The remaining point is to estimate the marginal distributions, which can be achieved using several approaches. The simplest

way is to make use of histograms. However, to do that, a large amount of samples is required for proper estimation [24], which means a relatively large window size. In this paper, we use the cumulant-based KL approximation for PDF estimation as proposed in [24]. This method works under the assumption that the PDF is close to a Gaussian distribution. In this case, the marginal density function $f_X(x)$ can be expressed by the Edgeworth series expansion, which relies only on the first four cumulants [26]. The cumulants can be estimated from moment $\mu_{X;i}$, which is the centered moment of order i of variable X. The Edgeworth series expansion can be expressed as in [24], [27]

$$KL_{Edgeworth}(X,Y)$$
$$= \frac{1}{12} \frac{\kappa_{X';3}^2}{\kappa_{X';2}^3}$$
$$+ \frac{1}{2} \left(\log \frac{\kappa_{Y;2}}{\kappa_{X;2}} - 1 + \frac{1}{\kappa_{Y;2}} \left(\kappa_{X;1} - \kappa_{Y;1} + \kappa_{X;2}^{1/2} \right)^2 \right)$$
$$- \left(\kappa_{Y';3} \frac{a_1}{6} + \kappa_{Y';4} \frac{a_2}{24} + \kappa_{Y';3}^2 \frac{a_3}{72} \right)$$
$$- \frac{1}{2} \frac{\kappa_{Y';3}^2}{36} \left(c_6 - 6 \frac{c_4}{\kappa_{X;2}} + 9 \frac{c_2}{\kappa_{Y;2}^2} \right)$$
$$- 10 \frac{\kappa_{X';3}\kappa_{Y';3}(\kappa_{X;1} - \kappa_{Y;1})(\kappa_{X;2} - \kappa_{Y;2})}{\kappa_{Y;2}^6} \tag{4}$$

where

$$\kappa_{X;1} = \mu_{X;1}$$
$$\kappa_{X;2} = \mu_{X;2} - \mu_{X;1}^2$$
$$\kappa_{X;3} = \mu_{X;3} - 3\mu_{X;2}\mu_{X;1} + 2\mu_{X;1}^3$$
$$a_1 = c_3 - 3\frac{\alpha}{\kappa_{Y;2}}$$
$$a_2 = c_4 - 6\frac{c_2}{\kappa_{Y;2}} + \frac{3}{\kappa_{Y;2}^2}$$
$$a_3 = c_6 - 15\frac{c_4}{\kappa_{Y;2}} + 45\frac{c_2}{\kappa_{Y;2}^2} - \frac{15}{\kappa_{Y;2}^3}$$
$$c_2 = \alpha^2 + \beta^2; \quad c_3 = \alpha^3 + 3\alpha\beta^2$$
$$c_4 = \alpha^4 + 6\alpha^2\beta^2 + 3\beta^4$$
$$c_6 = \alpha^6 + 15\alpha^4\beta^2 + 45\alpha^2\beta^4 + 15\beta^6$$

and $\alpha = \dfrac{\kappa_{X;1} - \kappa_{Y;1}}{\kappa_{Y;2}}; \quad \beta = \dfrac{\kappa_{X;2}^{1/2}}{\kappa_{Y;2}}.$

Because of the asymmetric property of KL divergence, the symmetric KL divergence

$$KLD = KL_{Edgeworth}(X,Y) + KL_{Edgeworth}(Y,X) \tag{5}$$

is used as one change indicator in this paper.

B. No-Building Change Indicator

In binary change detection, a very important final step is to separate the relevant changes associated with buildings from other changes (noise and other areas which might contain irrelevant change, here e.g., no-buildings). Vegetation growth can also create a height change, and if these areas are located around buildings, it will largely influence the building change detection result. To alleviate this effect, we use normalized difference vegetation index to indicate vegetation cover. Another important effect is shadowing which can significantly influence the quality of the DSMs. It has been shown in our previous paper [17] that shadow areas usually result in relatively bad matching results. Matching failures in shadow areas, which often represent ground level, are displayed in the original generated DSMs partly as holes, and through interpolation methods, they often get higher height values than the ground level. If this kind of error exists only in data of one epoch, building change alarms are produced. Here, the shadow mask is extracted with the method introduced by Marchant and Onyango [28] in which the relationship of red/blue channel and green/blue channel can be used to extract the shadow class.

In urban areas with simple building structures, vegetation and shadow areas can be successfully detected in the multi-spectral data in order to substantially reduce their influence to the results. For urban areas with more complicated building structures, the differentiation can be more difficult as, e.g., vegetation can also be found on the roof of buildings. Therefore, these features are only used to give a proper probability of no-building indication, while not an absolute decision.

C. Pixel-Based DS Fusion

Information similarity measures can be employed to extract changes between two images, but the task here is to extract only the building changes. In order to combine two kinds of uncertain information to one building change map, we adopt the DS fusion theory [29], which has been proven to give convincible results in combining various uncertain indicators in classification and change detection [5], [10], [30], [31].

DS theory is a belief function-based combination method. Each indicator gives different certainty to decision class $A(A \in 2^\Theta)$; 2^Θ represents the object classes of interest (like building change class) and all possible combination of them. The certainties of the decision class A is called probability masses $(m(A))$ and have values between 0 and 1

$$\begin{cases} m(\phi) = 0 \\ \sum_{A \in 2^\Theta} m(A) = 1 \end{cases}. \tag{6}$$

When p indicators are considered, each indicator will give a probability to the classes in $B(B \in 2^\Theta)$, who have a fuzzy relationship with A. The fused certainty of A can be represented with DS fusion theory [31] based on

$$m(A) = \frac{\sum_{B_1 \cap \ldots B_p = A} \prod_{i=1}^{p} m_i(B_i)}{1 - K} \tag{7}$$

i indicator, and $1 \leq i \leq p$,
m_i probability mass
p amount of indicators

$$K = \sum_{B_1 \cap \ldots B_p = \varnothing} \prod_{i=1}^{p} m_i(B_i).$$

K is a measure of conflict among different indicators, meaning that a decision cannot be, e.g., change and no-change at the same time. $m_i(B_i)$ represent mass functions of indicator i to class B. The probability masses shown in (6) are the original belief certainties. For instance, if one pixel has less than 1 m height change, then it is given a small probability to indicate low significance of real building changes, since this is in the range of DSM noise. In contrast, if it has more than 10 m of height change, it will have a high probability indicating real building change. In paper [10], a cubic parabola was used to model the mass function for the probability between these extremes, which worked well in DSM-assisted classification. In that model, several experimental-based thresholds are needed for each indicator. Therefore, for this investigation, we prefer the sigmoid curve [32] and use it to get a similar "S" shape with less parameters

$$P_i(x) = \frac{0.99}{1 + e^{-\frac{x-T}{\tau}}}. \tag{8}$$

In (8), x is the original value of each indicator. For our purpose, we add two parameters T and τ to control the symmetry point and shape of sigmoid function. The symmetry point indicates a certainty of 50%. In this paper, we employed minimum error thresholding method proposed by Kittler [33] to select T. This thresholding method assumes both objects and the background in the image to follow a Gaussian distribution. The threshold T is chosen to optimally separate these two distributions, which means the pixels at this threshold point have a probability of 50% of being an object. τ is used to control the slope of this distribution, which can be estimated using one given sample. According to the rule of the DS fusion theory, no indicator can give 100% certainty; therefore, a maximum value of 0.99 is designed for this model.

Table I shows the pixel-based decision model used in this paper. The purpose of this procedure is to distinguish the building changes (B) from other land surface changes (S) and from no-change (N). Here, we assume that height changes (ΔH) indicate building change, and the KLD similarity (ΔSM) indicates building change and surface change. We suppose that for new, demolished, or changed buildings together with the height change, also some spectral change is likely.

The second step is the fusion of the no-building change indicators to the building change probability map. Here, we employ the result from step one as one change indicator (i) and fuse it with the other two no-building change indicators separately based on (7). The no-building change indicators are considered and accepted to the fusion procedure when they have a mass values $m(A)$ greater than threshold T. According to the decision rule mentioned in [30], $T = 50\%$ is used in

TABLE I
DECISION MODEL FOR THE CLASSES (B: BUILDING CHANGE;
S: LAND SURFACE CHANGE; N: NO-CHANGE;
ΔH: HEIGHT CHANGE; ΔSM: SIMILARITY)

Type	ΔH	ΔSM	Combined Mass
B	$P_{\Delta H}$	0	$P_{\Delta H} \cdot P_{\Delta SM}/1\text{-}C$
S	0	0	$(1\text{-}P_{\Delta H}) \cdot P_{\Delta SM}/1\text{-}C$
N	0	$1\text{-}P_{\Delta SM}$	$(1\text{-}P_{\Delta H}) \cdot (1\text{-}P_{\Delta SM})/1\text{-}C$
B $\cup$ S	0	$P_{\Delta SM}$	
S $\cup$ N	$1\text{-}P_{\Delta H}$	0	
B $\cup$ N	0	0	

Where: $C = P_{\Delta H} \cdot (1\text{-}P_{\Delta SM})$

this procedure. Three fusion cases are considered in the fusion procedure:

1) $m(i) > T$, while $m(j) < T \Rightarrow$ Building change
2) $m(i) < T$, while $m(j) > T \Rightarrow$ No building change
3) $m(i) > T$, and $m(j) > T \Rightarrow$ DS fusion

In the case 1) of the above fusion rules, if corresponding pixels exhibit a high certainty of building change $m(i)$ and a low certainty to be vegetation or shadow $m(j)$, the decision is building change. If both kinds of indicators are higher than the assumed threshold (T), as for case 3), the DS fusion is performed based on (7) [31].

A building change probability map is achieved after these two fusion steps. This intermediate result might be already important for change indication in a semi-automatic procedure (not considered here) because it can save time and manual work if an interpreter uses this pre-information. Here, we are aiming to get the final changed building object in an automatic procedure. Therefore, a threshold on the building change probability map is needed to get an initial building change mask before the object-based building extraction. Since a further refinement is considered after we get the pixel-based building change map, in order to leave more candidates for the subsequent object filtering, we set a low threshold of 0.45 for the building change probability in the building change mask.

D. Region-Based Refinement

After generating the building change mask, it is still required to separate "changed building" from false change alarms. We therefore apply an edge-based building extraction method and improve the output by extracting the undesired objects based on their shape properties [17]. The three most important features to differentiate building areas from other objects are height, area (size), and convexity. Herein, an object level-based refinement is proposed to combine the three change indicators.

Height: For our purpose, height means the average height for each object, in order to get only one vertical change value for each constructed/destructed building defined by a single mask. We average the pixel values in the fusion result of the "difference image" belonging to the same changed object, and define this value as the vertical change of each building. As follows, we exclude all pixels which have "0" value (no height in the changed area), as well as very low values or very high values which can be attributed to potential blunders in one or

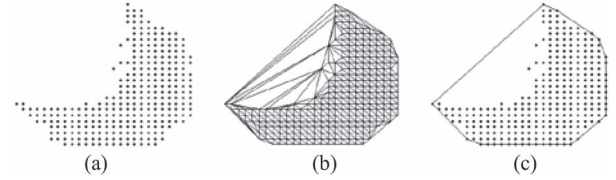Fig. 3.   Vertical change value evaluation.

Fig. 4.   Convex Hull calculation (a) Pixels in original mask; (b) Generated Delaunay triangulation; and (c) Convex hull.

both of the DSMs, so that these pixels will not be involved in the mean value calculation procedure. As displayed in Fig. 3, only the pixels of the middle part (gray line section) of the height difference values are analyzed.

Area: The area of the generated region is calculated by counting the number of pixels inside a single object included in the generated mask

$$\text{Area} = \text{Number of pixels} \times (\text{Resolution})^2.$$

Convexity: The convexity of a region is defined in this paper as the area ratio of this region and the smallest convex polygon that can contain the region. The computation of the smallest convex polygon [34] is shown in Fig. 4. Fig. 4(a) is the original mask from the difference image. After connecting all pixels in the original mask using Delaunay Triangulation [Fig. 4(b)], the object is represented by the generated edges, which constitute the smallest convex hull of the original region [shown as the black line in Fig. 4(c)].

III. DATA SETS FOR EXPERIMENTS

The experimental data sets for this research work consist of two pairs of stereo imagery captured on different dates and the corresponding DSMs. Two test areas are chosen for this procedure. In the first test area, the stereo images from two dates are both acquired by the IKONOS satellite on 12 February, 2006 and 2 May, 2011 respectively. The area represents an industrial region in Korea with an area size of 0.6×0.6 km^2 containing a factory center with mainly well separated, regularly shaped smaller size buildings. The other test data sets shows the city center of Munich, Germany (1.3×1.3 km^2). In this data sets, the first pair of the stereo imagery was acquired by IKONOS-2 on 15 July 2005, while the other pair features WorldView-2 data acquired on 12 July, 2010. This test area contains more complicated building shapes as well as buildings which cannot be separated from each other easily. Figs. 5(a)–(e) and 6(a)–(e) show the orthorectified images, building change reference data, and DSMs of the two test areas.
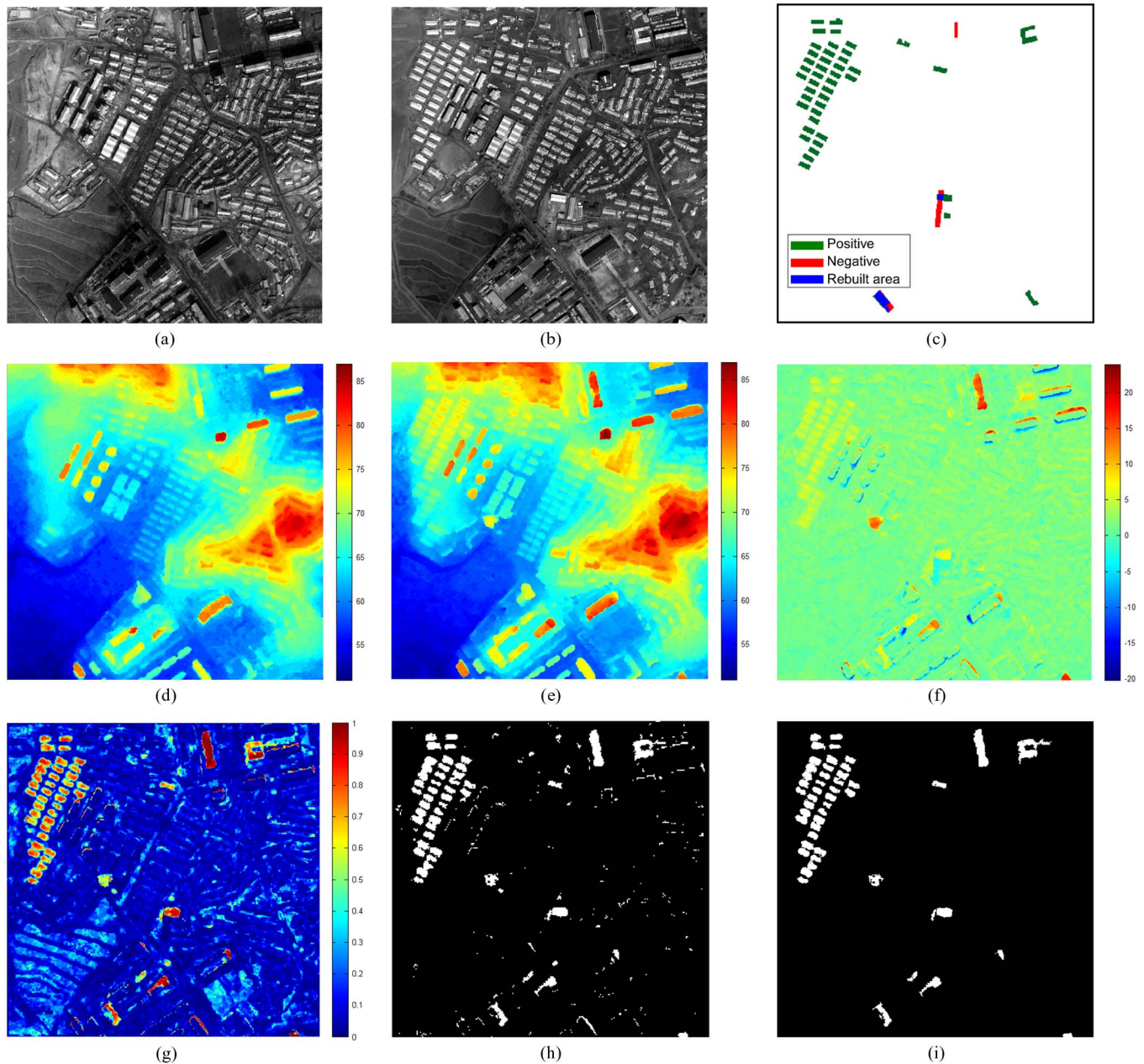
Fig. 5. Test area 1 (a) Panchromatic image from year 2006; (b) Panchromatic image from year 2011; (c) Reference change map; (d) DSM from year 2006; (e) DSM from year 2011; (f) Height change map; (g) Fusion result after second step; (h) Thresholding result; (i) Object-based filter result.

Stereo Imagery

The stereo imagery in the first test area consist of two IKONOS multispectral and panchromatic images. Both data sets are along-track stereo pairs. The multispectral IKONOS imagery has a GSD of 4 m containing four bands (red, green, blue, and near infrared). The panchromatic images have a GSD of 1 m [Fig. 5(a) and (b)]. The orthorectifed panchromatic and multispectral images are generated using the DSMs, meaning that the pixels in the images and DSMs are co-registered precisely. The stereo imagery in the second test area (cf. Fig. 6) consist of one IKONOS image pair (1 m GSD pan and 4 m multispectral) and one WorldView-2 image pair with a GSD of 0.5 m pan and 2 m multispectral (8 bands), respectively.

DSMs From Satellite Stereo Imagery

The DSMs of all stereo pairs are generated using the SGM algorithm implemented at DLR [1], [2] for dense image matching. In the first test area, the convergence angle is $17.39°$ for the data from 2006 and $22.28°$ for the data of year 2011, which means that the quality of both DSMs should be similar [as can be seen in Fig. 5(d) and (e)]. The DSMs generated by SGM exhibit 1 m GSD; the resulting orthoimages also have a GSD of 1 m. The WorldView-2 data set used in the second test area exhibit a GSD of 0.5 m and is comprised of two stereo pairs acquired on the same day and orbit in July 2010 with an overlap in the test area, leading to four different viewing directions with stereo angles between $12°$ and $24°$. The IKONOS data set was acquired in 2005 and has a convergence angle of approximately
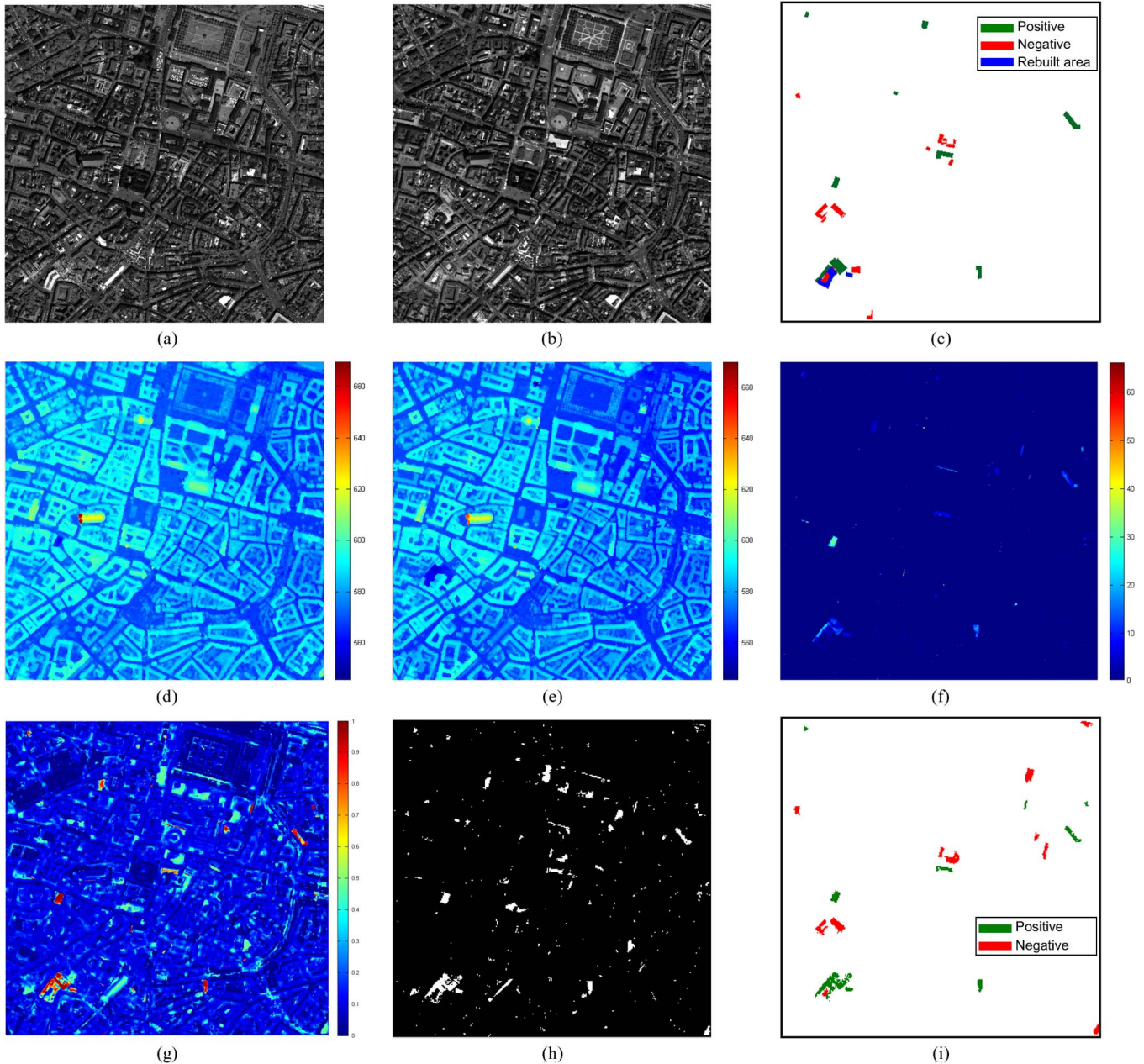
Fig. 6.   Test area 2 (a) Panchromatic image from year 2005; (b) Panchromatic image from year 2010; (c) Reference change map; (d) DSM from year 2005; (e) DSM from year 2010; (f) Height change map; (g) Fusion result after second step; (h) Thresholding result; (i) Object-based filter result (with both positive and negative building change results).

$10°$. Since the WorldView-2 panchromatic data exhibit 0.5 m resolution, the quality of the generated DSM is better than the DSM from IKONOS images [as can be seen in Fig. 6(d) and (e)]. For easier comparison, change detection is evaluated with 1 m GSD for both DSMs and orthorectified satellite images also for the second test area.

Reference

The reference is generated based on manual/visual interpretation using the orthorectified panchromatic and multispectral data sets [shown in Figs. 5(c) and 6(c)]. The reference data includes high buildings, normal houses, and tents. A tent with an area size less than 100 m^2 and low height cannot be detected

accurately enough with 1 m resolution satellite images and DSMs from stereo matching technique. Thus, in this research, only the buildings larger than 100 m^2 are considered in the reference data, procedures, and experiments.

IV. EXPERIMENTS AND RESULTS EVALUATION

A. Overall Description

In this section, we apply the proposed DSM-assisted building change detection workflow introduced in Section II. For the first data sets acquired by the same IKONOS-2 sensor, we subtract the DSM of year 2006 from DSM of year 2011 to get a first height difference map. As the second data sets are from different sensors, the DSM quality is also different.

That is why in this case, we use the robust difference with a window size of 5×5 pixels. The similarity map based on the original images is generated according to Section II-A. In the first test area, we choose a 9×9 window size in the KL divergence calculation procedure. In the second area, since the two panchromatic images are from different sensors with different radiometric characteristics and spatial resolution, a window size of 17×17 pixels is selected, and a Gaussian filter is employed in the preprocessing procedure to alleviate the multi-resolution problems. In the DS feature fusion procedure, for the sigmoid function, all of the threshold values are computed automatically. One sample datum is chosen for each change indicator to control the slope of the sigmoid distribution. For instance, $(P(x) = 0.1|x = 0)$ is used for the KL divergence value in Munich area to calculate τ. As stated in Section II-A, DSMs derived from satellite data are often erroneous and can hardly be used to detect small objects until now. Therefore, for the object-based building extraction procedure from Section II-D, we choose masks with a mean height change of more than 5 m, an area size of more than 100 m^2 for industrial area, and more than 200 m^2 for urban area, and a convexity value of more than 0.5 and 0.55, respectively, for industrial area and urban area. A sensitivity analysis to justify these values is shown in Section IV-C (Fig. 8).

B. Evaluation Method

In order to evaluate the effectiveness of the proposed method for the detection of position and size of the changed buildings and the overall change situation, we compare our results with reference data. The evaluation is performed for each step of the process: only height difference, only KL divergence, fusion of the two measures, combination of the two with the additional no-change indicators, and including the region-based refinement. The assessment of the change detection results is carried out both at pixel level and object level.

1) Pixel-Based Evaluation: For the pixel level evaluation, the results are displayed in terms of Receiver Operating Characteristics (ROC) curve analysis [35]. We use the area under the ROC curve to evaluate the quality of each change index and the generated result [35]. For each threshold between the maximum and minimum value in the result, we calculate the percentage of true positives (the building change pixels that are correctly detected as building change) and the percentage of the false positives (no-building change pixels that are falsely detected as building change pixels). The ROC curve shows the relationship of the true positive against the false positive. The area under the ROC curve is used to measure the ability of single or combined features and difference map to detect the real building changes [35].

2) Object-Based Evaluation: As a higher level of the analysis, the changed buildings are treated as single objects without consideration of their size. Only the effectiveness of the detection of distinct changed areas (buildings) in the change map is considered in the assessment. Therefore, four parameters are measured to evaluate the object-based change detection result:

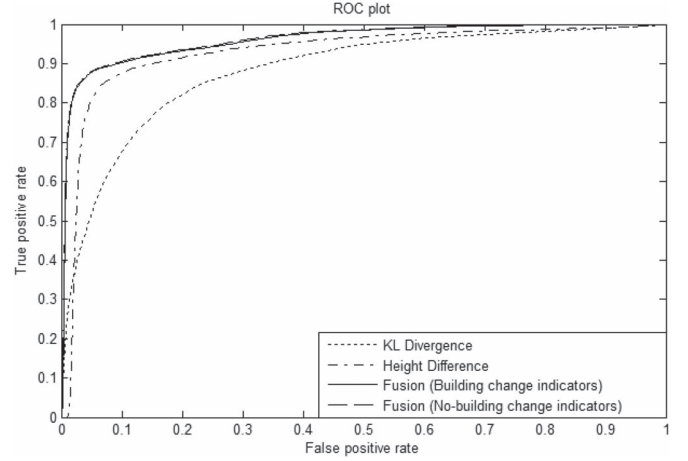1) True detected number (TDN): The number of changed objects that are correctly detected as changed.

Fig. 7. ROC curve comparison for the positive change in test area 1.

2) True detected rate (TD): True detected objects number in percentage TD = $\text{TDN}/\text{N}_\text{R} \times 100$.
3) False detected number (FDN): The number of unchanged objects that are incorrectly detected as changed.
4) False detected rate (FD): False detected objects number is percentage FD = $\text{FDN}/\text{N}_\text{D} \times 100$,

where N_R and N_D are the total number of changed objects of the reference data and the detected map, respectively.

C. Results

The building change detection results in each procedure are shown in Figs. 5 and 6. The height difference image is shown in Figs. 5(f) and 6(f) where red pixels indicate large height change values (more than 20 m change), and blue color highlights no-change or small height change values. Fig. 6(f) shows the robust positive change detection result, most of the noise coming from different sensor and different resolution is already eliminated. Figs. 5(g) and 6(g) show the DS fusion result including all the change indices and no-change indices, the values in the images represent the probability of each pixel to belong to the building change. The changed buildings are mostly highlighted in red color. Figs. 5(h) and 6(h) are the results after thresholding. The white masks represent the possible changed buildings according to the threshold in Section II-C. A combination of area, height, and convexity is applied to extract the real changed buildings as shown in Figs. 5(i) and 6(i). In test area 2, we extracted both, the positive change buildings and negative change buildings, which are shown in Fig. 6(i) with green and red color, respectively, while only positive change is evaluated in test area 1. As shown in the reference data, only three negative change buildings are located in the test area, one is a reconstructed building; the newly built building, at the same place, is much higher than the original one. Therefore, this building is correctly detected as positive change. Due to statistical reasons, a change for two buildings, within the category "negative change" is not included in the result evaluation.

In order to evaluate the change indicators and efficiency of procedures in each step, the ROC curve is applied in this procedure (Fig. 7). For better comparison, the area under the

TABLE II
AREAS UNDER ROC CURVES WITH DIFFERENT METHODS

No.	METHOD	TEST AREA1		TEST AREA2	
		POSITIVE CHANGE	NEGATIVE CHANGE	POSITIVE CHANGE	NEGATIVE CHANGE
1	HEIGHT DIFFERENCE	0.9299	0.5460	0.8745	0.6781
2	KL DIVERGENCE	0.8823		0.4955	
3	FUSION(CHANGE INDICATORS)	0.9558	0.7480	0.8642	0.7977
4	FUSION(NO-CHANGE INDICATORS)	0.9621	0.7549	0.9013	0.8525

TABLE III
OBJECT-BASED RESULT EVALUATION

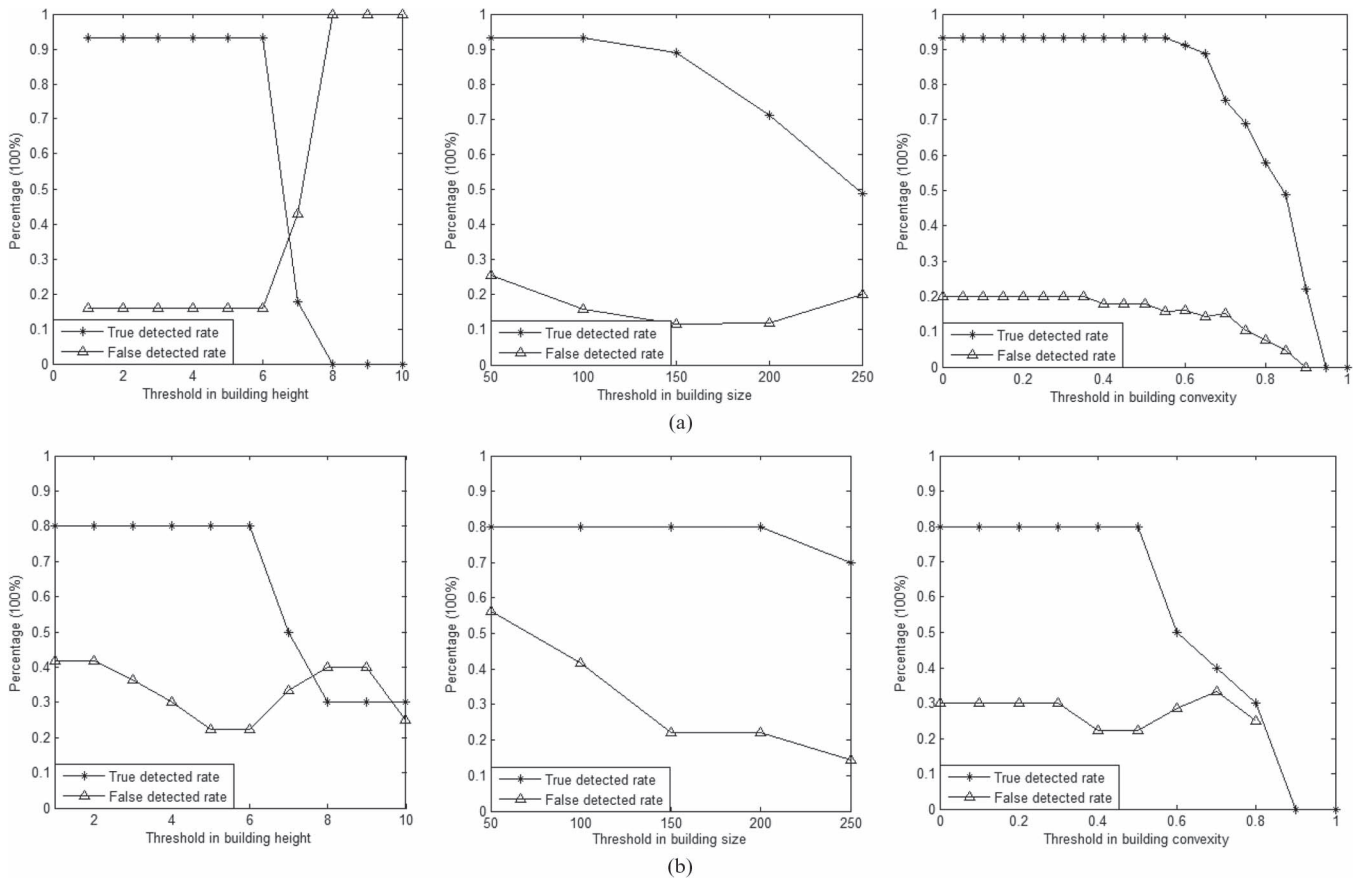| Type		True Detected(Object)		False Detected (Object)	
		Number	Rate	Number	Rate
Test area 1	Pos	42	93.33%	6	15.79%
	Neg	-----------			
Test area 2	Pos	8	80.00%	2	22.22%
	Neg	5	55.00%	5	45.00%

Fig. 8.   Sensitivity of parameters to threshold values in (a) Test area 1 and (b) Test area 2.

ROC curve as explained in Section IV-B is used and illustrated in Table II. It shows that with each step in the fusion process the reliability of the change detection is improved. Although the KL divergence measure does not lead to a high difference for the positive change in test area 2, it improves the result for true negative alarms substantially. Table III shows the object-based evaluation results for both test areas. In the first test area, 42 out

of 45 changed buildings are correctly detected. In the Munich test site, 8 out of 10 changed buildings are detected correctly. A sensitivity analysis of the three parameters in the object-based filter step is presented in Fig. 8. When the threshold lies within a specific range, it does not significantly influence the change detection result, and therefore the specific value is not very sensitive to the result.
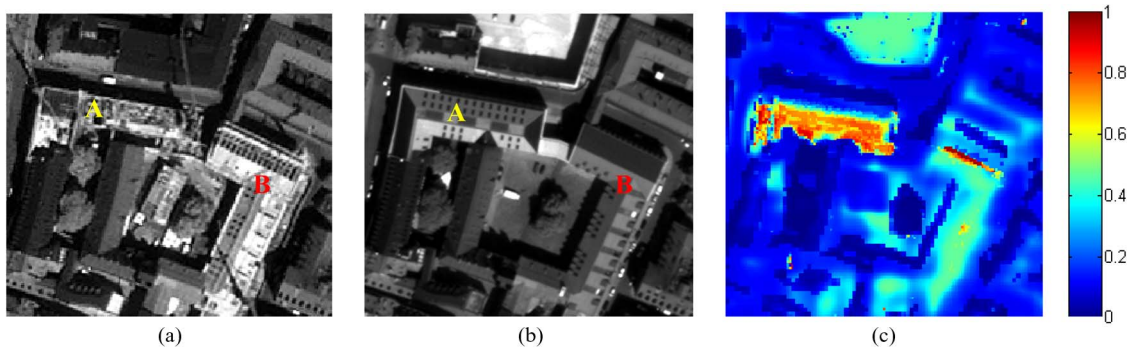
Fig. 9.   Correct detected change buildings example. (a) Panchromatic-date 1; (b) Panchromatic-date 2; (c) DS Fusion result after step-2.
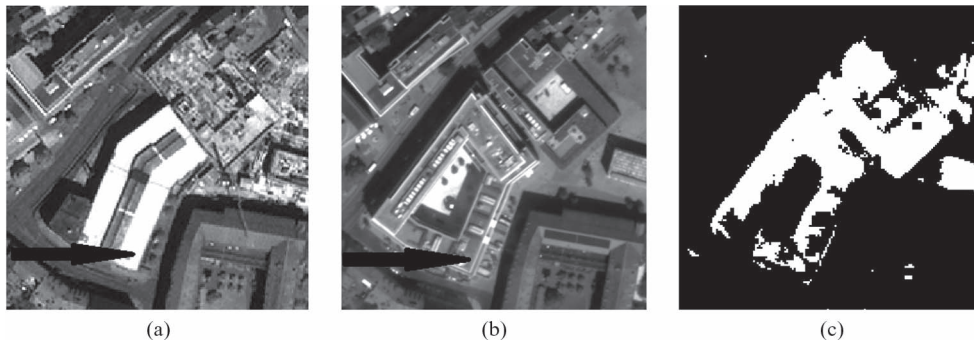

Fig. 10.   Example of reconstructed building. (a) Panchromatic-date1; (b) Panchromatic-date2; (c) Pixel-based change detection result.

D. Discussion of Results

The detection results in Figs. 5(f)–(i) and 6(f)–(i) and the evaluation outcome in Tables II and III show a good performance for the industrial area (test area 1) and medium performance for the urban area (test area 2). In the first test area, when comparing Fig. 5(f) with Fig. 5(g), it can be seen that the buildings are better separated with much sharper edges which is the result of the improvement through the KL divergence measure. Moreover, several false alarms of large height differences in Fig. 5(f) are eliminated after DS statistics. Although some season/vegetation changes are still obvious in Fig. 5(g) (at the lower left part of the image), they exhibit only 20% to 30% probability to be changed buildings. Therefore, they can be easily removed from the positive changes after thresholding. On the top left of Fig. 5(f), the four changed buildings do not appear very clear; by combining the KL divergence measure, they are all highlighted and well separated. From the pixel-based evaluation result, the first step fusion result is not improved by using the KL divergence in the second test area, but it still improves the final result. The reason is that in this test area, with 5 years of time difference, some buildings have been reconstructed with a similar height at nearly the same place as the former buildings. In this case, height change is not really helpful, which in turn influences the feature fusion result. However, after the second steps of the fusion, the area under ROC curve results exceeds 0.9, and the final object-based refinement missed two small buildings and found several falsely changed buildings. The reasons are discussed in the following.

One problem for an automatic procedure arises if buildings are still in construction. In Fig. 9, building A is only half

constructed at date 1 [shown in Fig. 9(a)], while finished at date 2 [shown in Fig. 9(b)]; building B is a building without change in height. Both buildings show different gray values in the panchromatic images. With our detection method, building A has a much higher change probability after the pixel-based DS fusion [shown in Fig. 9(c), color bar shows the building change probabilities], and can be extracted successfully.

Further, it should be pointed out that a false alarm will be detected if one of the DSMs exhibits large area matching errors. This is shown in top middle of Fig. 5(a) and (b). From the images, we can clearly see the buildings are not changed. Due to matching error of stereo images from date 1, the large building in the center of the images is missing in the first DSM [Fig. 5(d)]. While in the DSM of the later date [Fig. 5(e)], the building is represented correctly. The color indicates the extracted height information in Fig. 5(d) and (e). Those kinds of mistakes cannot be handled by our method. Also, for the second test area (Munich) the DSM from 2010 shows several matching errors in dense building areas. This leads to several false detected objects within the negative change detections.

The other missed alarms are caused by reconstruction, for which the new buildings are built at the same or similar area of the old buildings. Within 5 years of time difference, several places show this activity. One example can be seen in the center of Fig. 10. The shown building was reconstructed with a new building shape. By combining the change information from images, we get some small changed area as in the lower part of Fig. 10(c) (shown with the black arrow), but since they are relatively small and not connected with the other part of the buildings, after the object-based filtering they are not kept in the change mask.

V. CONCLUSION

Height information is essential for urban area monitoring, and it is particularly helpful for building change detection. Many previous studies using laser scanning data for this purpose show the potential but laser data are acquired less frequent and hardly can be used for large and remote areas. In particular, in disaster situations satellite stereo data are usually much easier, faster, and cheaper to acquire. Using new dense stereo matching algorithms and DSM generation technology, it is shown in this paper that DSMs generated with spaceborne stereo data can be a reliable source for efficient building change detection. To fully use all of the change information contained in original panchromatic images, multispectral images, and the height information, we use the DS fusion theory for a fusion process to extract real building changes. The sigmoid function is employed in the automatic initial mass function generation procedure. Furthermore, the whole procedure works unsupervised, and the probability calculation is much faster than model-based clustering methods. The generated probability maps can also be used effectively in a semi-automatic procedure. In the end, an object-based filtering is used to extract the building footprints. The results in each step are compared with a reference map, and the evaluation shows the efficiency of the combined methodology. The extracted change maps show the surface changes correctly for most parts of the test areas. Shortcomings appear if one of the DSMs does not meet the required quality and cannot show the real situation through which the result is negatively influenced.

As future work, an investigation of the DSM quality in various situations will be performed and might be added to our fusion model. Moreover, in this research, only pixel-based features are employed; more texture features will also be tested in future follow up work if they can give indication for building change or no building change. For the object-based building change map refinement procedure, more and also 3-D shape features could be used to improve the overall change detection result.

ACKNOWLEDGMENT

REFERENCES

[1] H. Hirschmüller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 1–14, Feb. 2008.

[2] P. d'Angelo, M. Lehner, and T. Krauss, "Towards automated DEM generation from high resolution stereo satellite images," in *Proc. ISPRS Congr.*, Beijing, China, 2008, vol. 37, pp. 1137–1342, Part B4.

[3] P. d'Angelo and P. Reinartz, "Semiglobal matching results on the ISPRS stereo matching benchmark," in *Proc. ISPRS Workshop*, Hannover, Germany, Jun. 2011, vol. 38-4/W19, pp. 79–84.

[4] A. Singh, "Digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, 1989.

[5] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2407, Jun. 2004.

[6] D. Koc San and M. Turker, "Automatic building detection and delineation from high resolution space images using model based approach," in *Proc. ISPRS Working Groups I/5-I/6I Workshop (Topograph. Mapping Space)*, Ankara, Turkey, Feb. 2006, vol. 36-1-W41, pp. 1–6.

[7] Z. Liu, P. Gong, P. Shi, H. Chen, L. Zhu, and T. Sasagawa, "Automated building change detection using UltraCamD images and existing CAD data," *Int. J. Remote Sens.*, vol. 31, no. 6, pp. 1505–1517, Mar. 2010.

[8] T. Knudsen and B. P. Olsen, "Automatic change detection for updates of digital map database," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 11, pp. 1289–1296, Nov. 2003.

[9] B. P. Olsen, "Automatic change detection for validation of digital map databases," in *Proc. ISPRS Congr.*, Istanbul, Turkey, 2004, vol. 34, pp. 569–574, Part B2.

[10] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Using the Dempster-Shafer method for the fusion of LIDAR data and multispectral images for 'building detection," *Inf. Fusion*, vol. 6, no. 4, pp. 283–300, 2005.

[11] N. Champion, F. Rottensteiner, L. Matikaninen, X. Ling, J. Jyyppä, and B. P. Olsen, "A test of automatic building change detection approaches," in *Proc. CMRT09/IAPRS*, 2009, vol. 38, pp. 145–150, Part 3/W4.

[12] L. Matikainen, J. Hyyppä, E. Ahokas, L. Markelin, and H. Kaartinen, "Automatic detection of buildings and changes in buildings for updating of maps," *Remote Sens.*, vol. 2, no. 5, pp. 1217–1248, 2010.

[13] H. Murakami, "Change detection of buildings using an airborne laser scanner," *ISPRS J. Photogramm. Remote Sens.*, vol. 54, pp. 148–152, 1999.

[14] K. Choi, I. Lee, and S. Kim, "A feature based approach to automatic change detection from Lidar data in urban areas," in *Proc. ISPRS Workshop Laserscanning*, Paris, France, 2009, pp. 259–264.

[15] F. Jung, "Detecting building changes from multitemporal aerial stereopairs," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 3/4, pp. 187–201, Jan. 2004.

[16] L. Zhu, H. Shimamura, K. Tachibana, Y. Li, and P. Gong, "Building change detection based on object extraction in dense building areas," in *Proc. ISPRS Congr.*, Beijing, China, 2008, vol. 37, pp. 905–908, Part B7.

[17] J. Tian, H. Chaabouni-Chouayakh, and P. Reinartz, "3D building change detection from high resolution spaceborne stereo imagery," in *Proc. Int. Workshop M2RSM*, Xiamen, China, 2011, pp. 1–7.

[18] H. Chaabouni-Chouayakh and P. Reinartz, "Towards automatic 3D change detection inside urban areas by combining height and shape information," *Photogramm. Fernerkundung Geoinf.*, no. 4, pp. 205–217, Jan. 2011.

[19] L. Zhang, "Automatic digital surface model (DSM) generation from linear array images," Ph.D. dissertation, Inst. Geod. Photogramm., ETH Zurich, Switzerland, 2005.

[20] H. Arefi, P. d'Angelo, H. Mayer, and P. Reinartz, "Iterative approach for efficient digital terrain model production from CARTOSAT-1 stereo images," *J. Appl. Remote Sens.*, vol. 5, no. 1, p. 053527, Jan. 2011.

[21] P. Reinartz, R. Müller, P. Schwind, S. Suri, and R. Bamler, "Orthorectification of VHR optical satellite data exploiting the geometric accuracy of TerraSAR-X data," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 1, pp. 124–132, Jan. 2011.

[22] S. Suri and P. Reinartz, "Mutual-information-based registration of TERRASAR-X and IKONOS imagery in urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 939–949, Feb. 2010.

[23] L. Gueguen, P. Soile, and M. Pesaresi, "Change detection based on information measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4503–4515, Nov. 2011.

[24] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432–1445, May 2007.

[25] F. Bovolo and L. Bruzzone, "An adaptive technique based on similarity measures for change detection in very high resolution SAR images," in *Proc. IEEE IGARSS*, Boston, MA, USA, Jul. 2008, pp. III-158–III-161.

[26] S. Cui, M. Datcu, and L. Gueguen, "Information theoretical similarity measure for change detection," in *Proc. JURSE*, U. Stilla, P. Gamba, C. Juergen, and D. Maktav, Eds., Munich, Germany, 2011, pp. 69–72.

[27] J.-J. Liu, N. Saito, and R. A. Livine, "Edgeworth expansions of the Kullback-Leibler information," Div. Stat., Univ. California, Davis, CA, USA, 1999, submitted to J. Amer. Statist. Assoc..

[28] J. A. Marchant and C. M. Onyango, "Shadow-invariant classification for scenes illuminated by daylight," *J. Opt. Soc. Amer. A, Opt., Image Sci., Vis.*, vol. 17, no. 11, pp. 1952–1961, Nov. 2000.

[29] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ, USA: Princeton Univ. Press, 1976.

[30] S. Le Hégarat-Mascle, I. Bloch, and D. Vidal-Madjar, "Application of Dempster-Shafer evidence theory to unsupervised classification in multi-source remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 1018–1031, Jul. 1997.

[31] S. Le Hégarat-Mascle and R. Seltz, "Automatic change detection by evidential fusion of change indices," *Remote Sens. Environ.*, vol. 91, no. 3/4, pp. 390–404, Jun. 2004.

[32] Sigmoid Function. http://en.wikipedia.org/wiki/Sigmoid_function

[33] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, no. 1, pp. 41–47, 1986.

[34] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational Geometry: Algorithm and Application*, 3rd ed. Berlin, Germany: Springer-Verlag, 2008.

[35] D. J. Hand and R. J. Till, "A simple generalization of the area under the ROC curve to multiple class classification problems," *Mach. Learn.*, vol. 45, no. 2, pp. 171–186, Nov. 2001.

Jiaojiao Tian received the B.S. degree in geo-information systems at the China University of Geoscience, Beijing, China, in 2006 and the M.Eng. degree in cartography and geo-information at the Chinese Academy of Surveying and Mapping, Beijing, China, in 2009. Currently, she is Ph.D. student at the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany.

Since September 2009, she has been working at the "3-D group" in the Photogrammetry and Image Analysis department of IMF. In autumn 2011, she was a Guest Scientist at the Institute of Photogrammetry and Remote Sensing, ETH Zurich, Switzerland. Her research interests include 3-D change detection for urban and forest areas, digital surface model (DSM) quality assessment, and DSM-assisted building extraction and classification.

Shiyong Cui received the M.S. degree in photogrammetry and remote sensing from the Chinese Academy of Surveying and Mapping, Beijing, China, in 2009. Currently, he is working in the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, toward the Ph.D. degree in electronic engineering.

His main research interests include multi-temporal synthetic aperture radar image analysis, signal processing, image processing, pattern recognition, and machine learning.

Peter Reinartz (M'09) received the Diploma (Dipl.-Phys.) degree in theoretical physics from the University of Munich, Munich, Germany, in 1983 and the Ph.D. (Dr.-Ing.) degree in civil engineering from the University of Hannover, Hannover, Germany, in 1989. His dissertation is on optimization of classification methods for multispectral image data.

He is the Department Head of the department "Photogrammetry and Image Analysis," at the German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF) and holds a professorship for geoinformatics at the University of Osnabrück, Osnabrück, Germany. He has more than 25 years of experience in image processing and remote sensing and over 200 publications in these fields. His main interests are in direct georeferencing, stereo-photogrammetry and data fusion of spaceborne and airborne data, generation of digital elevation models, and interpretation of very high resolution data from sensors like WorldView, GeoEye, Pleidaes a.o. He is also engaged in using remote sensing data for disaster management and using high-frequency time series of airborne image data for real-time image processing and their operational use in case of disasters as well as for traffic monitoring.