https://www.tandfonline.com/doi/full/10.1080/19479832.2018.1513957

RESEARCH ARTICLE

Fusion of multispectral imagery and DSMs for building change detection using belief functions and reliabilities

ARTICLE HISTORY
Compiled November 5, 2018

ABSTRACT
The extraction of building changes from very high resolution satellite images is an important but challenging task in remote sensing. Digital surface models (DSMs) generated from stereo imagery have proved to be valuable additional data sources for this task. In order to efficiently use the change information from the DSMs and spectral images, belief functions have been introduced. In this article, two-step building change detection fusion models based on both Dempster-shafer theory (DST) and Dezert-Smarandache Theory (DSmT) frameworks are proposed. In the first step, basic belief assignments (BBAs) of the change indicators from images and DSMs are calculated by using a refined sigmoidal BBA model. Then these BBAs are employed for the new proposed building change detection decision fusion approach. In order to cover the miss-detections introduced by the wrong height values of the DSMs and incomplete information from images, disparity maps from the DSM generation procedure and shadow maps from the multispectral channels are adopted to generate reliability maps, which are further integrated to the fusion models. In the last step, building change masks are generated based on four decision-making criteria. In the experimental part of this work, we evaluate the performance of this new building change detection method on real satellite images thanks to a building change reference mask representing the ground truth. Substantial accuracy improvements are achieved when comparing the new results with those obtained from classical 3D change detection approaches.

KEYWORDS
Change detection, belief functions, DSmT, DST, DSM

1. Introduction

Efficient and accurate detection of building changes using remote sensing data is of great importance for urban monitoring and disaster monitoring. It is one of the fundamental tasks in remote sensing and is attracting more interests due to the high and accelerated rate of urban growing and more frequent natural disasters with climate changes.

In the last decades, 2D change detection methods on large scale land cover monitoring have been extensively studied and applied on satellite images (Lu *et al.* 2004, Tewkesbury *et al.* 2015, Liu *et al.* 2018). There are many excellent approaches available which can extract landcover changes from multi-temporal images (Bruzzone and Prieto 2000, Nielsen 2007). However, highlighting only building changes in urban area remains difficult due to the mixture of other background changes, for instance the changes introduced by different illumination conditions or human activities. The influence of these changes is growing as higher resolution images show more details of the landcover objects. In addition, even with very high resolution data it is sometimes

impossible to distinguish buildings and roads using simple spectral change.

Therefore, height information derived from Digital Surface Model (DSM) is posing new possibilities for building change detection. Benefiting from improved data quality and advanced computer vision techniques, the accuracy of the DSMs from satellite stereo imagery has been largely improved and enables building change detection in a larger region and with high frequency. However, the DSMs may exhibit some inaccurate height values resulting from failed matching and occlusions within the stereo and multiple views. Thus the fusing of changes from multispectral image and DSMs would be an effective solution for building change detection. The comparison of DSMs can locate the changes of high-level objects efficiently and robustly and the spectral images have rich spectral and texture feathers which can highlight more changes among the multi-temporal datasets. On the other hand, as the DSMs have been generated from the multispectral data, there is no time difference between them. The 2D and 3D information can be combined through post-refinement, region-based approaches or decision fusion (Qin *et al.* 2016a). In more recent researches, DSMs from multi-sensors and time-series data were involved (Li *et al.* 2017, Pang *et al.* 2018).

Regarding to feature fusion due to the diverse building characteristics and background information, the urban building monitoring approaches may perform variedly for different test regions. Thus recently some researches are trying to combine different change features and change classification methods, and fuse the results with a decision model. For instance, Sirmacek and Unsalan (2011) have proposed a probabilistic framework to fuse the results from four local feature vectors for building detection. Based on an adaptive network-based fuzzy inference system, Janalipour and Taleai (2017) have fused the change detection results from different feature combinations. Besides fusing the detection result, decision fusion can be also directly used for classification and change detection (Le Hégarat-Mascle *et al.* 1997, Rottensteiner *et al.* 2005, Rottensteiner 2007, Liu *et al.* 2014, Liu 2014).

Thus until now there is no decision fusion model that directly takes the change indices from images and height maps for building change detection. In our previous research works (Tian *et al.* 2014a), belief functions have performed very well for 3D building change detection. As aforementioned, the accuracy of 2D change detection of specific objects is limited due to the misdetections caused by irrelevant changes. These irrelevant changes have a larger effect on very high resolution (VHR) images than on low and moderately high resolution images, since in VHR images their higher details are more sensitive to viewing and solar angle differences. DSMs generated from satellite stereo imagery can largely help to solve this problem. Unfortunately, the fusion model proposed in Tian *et al.* (2014a) is rather basic, and it is not robust in dealing with high conflict situations. Therefore, the belief functions have been further investigated and improved in this article. Besides Dempster-Shafer Theory (DST) (Dempster 1967, Shafer 1976), an extended Dezert-Smarandache Theory (DSmT) (Smarandache and Dezert 2004–2015) will be adopted in this article to generate the building change detection models. One of the difficulties of using Dempster-Shafer theory is the definition of uncertainty and the calculation of the basic belief assignments (BBAs). Tian *et al.* (2014a) used one sigmoid function to distribute the values of one change feature to the BBAs ranging from 0 to 1. The symmetry point which indicates a certainty of 50% was automatically calculated with a thresholding method. However, the accuracy and robustness of the thresholding approach will directly influence the correctness of the obtained BBAs. Thus, as well as the fusion models, the BBAs construction approach should be updated to further improve the change detection result. These problems have been well addressed in our modified approach (Tian *et al.* 2015a). In addition,

the uncertainty of change indicators was measured in order to improve the accuracy of BBAs. Due to space limitation constraint of conference paper format, the methodology part has been only shortly described in Tian *et al.* (2015a) and only small patches have been tested in the experimental part. A better description of this methodology with more experiments of our approach is presented in this article with the improvement of the reliability discounting approach.

Focusing on building change detection by fusing spectral and height information extracted from satellite stereo imagery, this article is organised as follows. First, the belief functions of DST and DSmT are briefly reviewed. Then, the building change models are proposed for these theoretical frameworks. The belief functions are used in both BBAs preparation and change detection procedure. Two sigmoid functions are simulated for each change feature to obtain the BBAs. In order to further improve the BBAs values reliability discounting techniques are presented. We use the unfilled disparity map and shadow maps to generate the reliability map of the changes from the height and 2D images, respectively. The reliability maps are then used in the fusion process to refine the initial BBAs. We generate four sets of global BBAs. With four decision criteria the final change detection masks can be generated. In the end, these refined fusion models are tested on four sets of real satellite images, and a comprehensive comparison is included to validate the new approaches.

2. Belief functions, DST and DSmT

2.1. *Basics of belief functions*

The details of DST and DSmT have been presented by Shafer (1976), Smarandache and Dezert (2004–2015) and Dezert and Tchamova (2014). Let Θ be a frame of discernment of a problem under consideration. $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$ consists of a list of N exhaustive and mutually exclusive elements θ_i, $i = 1, 2, \ldots, N$. Each θ_i represents a possible state related to the problem we want to solve. The assumption of exhaustivity and mutual exclusivity of elements of Θ is classically referred as *Shafer's model* of the frame Θ. A BBA also called a belief mass function (or just a mass for short), is a mapping $m(.) : 2^\Theta \to [0, 1]$ from the power set[1] of Θ (denoted 2^Θ) to $[0, 1]$, that verifies Shafer (1976):

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{X \in 2^\Theta} m(X) = 1. \tag{1}$$

$m(X)$ represents the mass of belief exactly committed to X. An element $X \in 2^\Theta$ is called a focal element if and only if $m(X) > 0$. The belief and plausibility functions based on DST theory are defined respectively as:

$$Bel(A) = \sum_{B \in 2^\Theta, B \subseteq A} m(B). \tag{2}$$

$$Pl(A) = \sum_{B \in 2^\Theta, B \cap A \neq \emptyset} m(B). \tag{3}$$

[1]The power set is the set of all subsets of Θ, empty set included.

In DST, the combination (fusion) of several independent sources of evidences is done with Dempster-Shafer[2] (DS) rule, assuming that the sources are not in total conflict[3]. DS combination of two independent BBAs $m_1(.)$ and $m_2(.)$, denoted symbolically by $DS(m_1, m_2)$, is defined by $m^{DS}(\emptyset) = 0$, and for all $X \in 2^\Theta \setminus \{\emptyset\}$ by:

$$m^{DS}(X) = \frac{1}{1 - K^{DS}} \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2), \tag{4}$$

where the total degree of conflict K^{DS} is given by

$$K^{DS} \triangleq \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2). \tag{5}$$

A discussion on the validity of DS rule and its incompatibility with Bayes fusion rule for combining Bayesian BBAs can be found in the literature (Dezert and Tchamova 2014, Dezert *et al.* 2012, Tchamova and Dezert 2012). To circumvent the problems of DS rule, Smarandache and Dezert (Smarandache and Dezert (2004–2015), Vol. 2, Chap. 1), then Martin and Osswald (Smarandache and Dezert (2004–2015), Vol. 2, Chap. 2) have developed in DSmT (Smarandache and Dezert 2004–2015) two fusion rules called PCR5 and PCR6 based on the proportional conflict redistribution (PCR) principle which consists

(1) apply the conjunctive rule
(2) calculate the total or partial conflicting masses
(3) then redistribute the (total or partial) conflicting mass proportionally on non-empty sets involved in the conflict according to the integrity constraints one has for the frame Θ.

This PCR principle transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is not degraded. Because the proportional transfer can be done in different ways, this has yielded to several different fusion rules. It has been proved by Smarandache and Dezert (2013) that only PCR6 rule is compatible with frequentist probability estimation, and that is why we recommend its use in the applications. PCR5 and PCR6 rules simplify greatly and their formulas coincide for the combination of two sources. In this case, the PCR6 combination is obtained by taking $m^{PCR6}(\emptyset) = 0$, and for all $X \neq \emptyset$ in 2^Θ by

$$m^{PCR6}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) + \sum_{\substack{Y \in 2^\Theta \setminus \{X\} \\ X \cap Y = \emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)}\right], \tag{6}$$

[2]Although the rule has been proposed originally by Dempster, we call it Dempster-Shafer rule as it has been widely promoted by Shafer in DST (Shafer 1976).

[3]otherwise DS rule is mathematically not defined because of 0/0 indeterminacy.

where all denominators in Eq. (6) are different from zero. If a denominator is zero, that fraction is discarded.

If a denominator, e.g., $m_1(X) + m_2(Y)$ tends towards 0, then also the conflicting mass $m_1(X)m_2(Y)$ that is transferable tends to zero because $m_1(X)$ and $m_2(Y)$ tend to zero (since they are positive); therefore, the redistribution of masses also tends to zero. That reflects the continuity of PCR6.

2.2. *Reliability discounting*

The reliability discounting has been described and discussed in the references (Mercier *et al.* 2005, Smarandache *et al.* 2010). Briefly, if an additional knowledge about the reliability (α) of certain source of evidence is available, it can be adopted to refine the initial BBAs. For instance the height change and image change indicators may not perform well under some situations. This situation can be measured, and used as reliability factors. Each factor α would be a value ranging from 0 to 1. And $\alpha = 1$ means fully reliable, while $\alpha = 0$ means the indicator is totally unreliable. And all the remaining discounted mass are transferred to the full ignorance Θ. Based on Shafer's discounting model (Shafer 1976), the reliability discounting factor α is introduced to discount any BBA $m(.)$ defined on the power set 2^Θ as follows $\forall X \in 2^\Theta$:

$$\begin{cases} m_\alpha(X) = \alpha \cdot m(X), \text{for } X \neq \Theta \\ m_\alpha(\Theta) = \alpha \cdot m(\Theta) + (1 - \alpha). \end{cases} \tag{7}$$

3. Building change detection fusion model

3.1. *Choice of the frame of discernment*

Focusing on change detection, as a data preparation step, DSMs are calculated from satellite stereo imagery based on semi-global matching approach (d'Angelo and Reinartz 2012, Tian *et al.* 2013). It follows two main steps. First, the epipolar image pair is generated through a pyramidal local least squares matching. Then the matching is cast into dynamic programming to minimise the cost function. We use census feature to measure the similarity between two pixels (d'Angelo and Reinartz 2012). The challenges and opportunities of the DSMs assisted building change detection have been well described in Tian *et al.* (2014a). The geo-information is employed to co-register these data, which enables a sub-pixel accuracy. Focusing on building change detection, two change indicators, one from images and one from DSMs are extracted. Changes from spectral images are highlighted by using the Iteratively Reweighted Multivariate Alteration Detection (IRMAD) (Nielsen 2007). Consequently, height changes from DSMs are shown after robust height difference (Tian *et al.* 2010, 2014a). We suppose that new, demolished or rebuilt buildings may exhibit both height and spectral changes. But the spectral changes can also be introduced by seasonal changes and other irrelevant changes. After excluding building changes, changed pixels exclude building regions are named here as *OtherChange*. Therefore, three classes are considered to define the frame of discernment satisfying Shafer's model (i.e. the elements of the frame

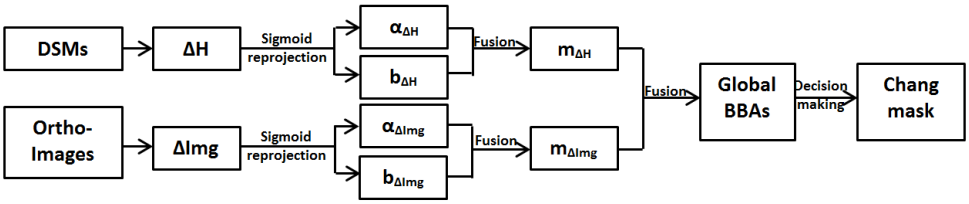

Figure 1. Workflow of the proposed method.

of discernment are disjoint):

$$\begin{aligned}\Theta = \{&\theta_1 \triangleq \text{Pixel} \in \text{BuildingChange},\\ &\theta_2 \triangleq \text{Pixel} \in \text{OtherChange},\\ &\theta_3 \triangleq \text{Pixel} \in \text{NoChange}\},\end{aligned} \tag{8}$$

and

$$\theta_1 \cap \theta_2 \cap \theta_3 = \emptyset. \tag{9}$$

In image domain, each pixel represents a single sample, thus in Eq. (8), we have directly used the word 'Pixel'. Based on the three exclusive classes, the set of potential focal elements FE that enter in our application is:

$$FE = \{\theta_1, \theta_2, \theta_3, \theta_1 \cup \theta_2, \theta_1 \cup \theta_3, \theta_2 \cup \theta_3, \theta_1 \cup \theta_2 \cup \theta_3\}. \tag{10}$$

It is worth noting that even if we work with Shafer's model of the frame of discernment for this application (which is the basis of DST), we can also use PCR6 rule developed in DSmT because PCR6 works also with Shafer's model as shown in Smarandache and Dezert (2004–2015).

The whole procedure of the proposed building change detection model is shown in Fig. 1. After the changes from DSMs and images are extracted, they will be reprojected using the sigmoid function to calculate the concordance index a and discordance index b. Then the decision fusion rules will be performed to generate the BBAs for height change and image change, respectively. After that, global BBAs can be calculated by using both DST and DSmT fusion rules. Finally, change mask can be obtained with various decision-making criteria.

3.2. *BBAs construction for building change detection*

In Tian *et al.* (2015b) a sigmoidal model for both concordance and discordance indexes has been briefly presented. The details and advantages of this approach are described in Dezert and Tacnet (2012). The concordance index measures the concordance of change indicator and BBA in the assertion, while the discordance measures the opposition of change indicator to the BBAs in the assertion. In our previous works (Tian *et al.* 2014a), the BBAs were built based on sigmoid curves related with the concordance index only. As explained in Tian *et al.* (2014a), the original sigmoid curve is defined as

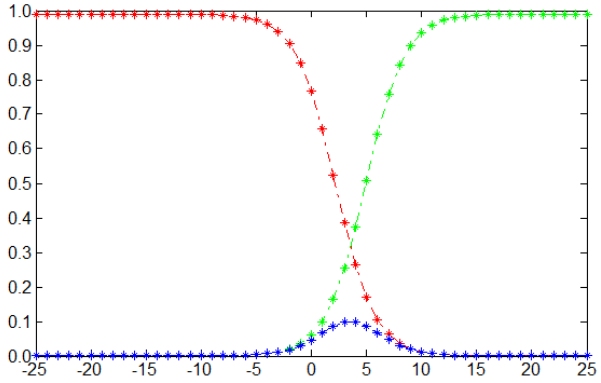

Figure 2. Concordance and discordance index.

$$f_{(\tau,T)}(x) = 0.99/(1 + e^{-\frac{x-T}{\tau}}), \tag{11}$$

where x is the original value of each indicator (ΔH, ΔImg),where ΔH means the change in the height and ΔImg means the change between two spectral images at a given pixel location. Two parameters T and τ are used to control the symmetry point and the slope of the sigmoid function. The symmetry point indicates a certainty of 50%. In this article, we improve our model to construct the BBAs thanks to sigmoidal models for both concordance and discordance indexes following the idea proposed by Dezert and Tacnet (2012). The concordance index is similar as the indicator of our previous research. The green line in Fig. 2 shows an example of the concordance index from height changes. A higher height change indicator leads to a higher probability to be building change. The discordance index is defined as an indication for the opposite argument. The discordance index in Fig. 2 is shown in red color, which means that a higher height change reflects a lower probability to be *not building change*. The blue curve shows the conflict between the concordance and the discordance index. Both concordance index and discordance index are projected to the sigmoid curve distribution characterised by parameters T and τ.

In Dezert and Tacnet (2012) these two parameters T and τ were manually selected. Here, as an improvement the multi-level Otsu's thresholding method (Otsu 1975, Liao *et al.* 2001) is used for automatically getting the symmetry points for both concordance index and discordance index. Otsu's algorithm assumes that an image is composed of objects and background. A discriminant analysis is performed by minimising the intra-class variance. When three classes are of interest, two thresholds T_1 and T_2 are expected, and Otsu's method can be extended to

$$\sigma^2_\omega(T_1, T_2) = \omega_1\sigma^2_1(T_1, T_2) + \omega_2\sigma^2_2(T_1, T_2) + \omega_3\sigma^2_3(T_1, T_2). \tag{12}$$

The weights ω_i are the probabilities obtained from the image histogram that are separated by the thresholds T_1 and T_2. σ_i is the standard deviation of the i-th class, for $i = 1, 2, 3$. T_1 and T_2 can be used as the symmetry points of discordance and concordance index, respectively. Thus, using the height change index as in the example, the BBAs for concordance and discordance height change index are functions of values

Table 1. BBA construction for height change indicator ΔH. [Conflict: $K_{\Delta H} = a_{\Delta H} b_{\Delta H}$]

Focal Elem.	$m_1(.)$	$m'_1(.)$	$m_1^{DS}(.)$	$m_1^{PCR6}(.)$
θ_1	$a_{\Delta H}$	0	$\frac{a_{\Delta H}(1-b_{\Delta H})}{1-K_{\Delta H}}$	$a_{\Delta H}(1-b_{\Delta H}) + \frac{a_{\Delta H} K_{\Delta H}}{a_{\Delta H}+b_{\Delta H}}$
θ_2	0	0	0	0
θ_3	0	0	0	0
$\theta_1 \cup \theta_2$	0	0	0	0
$\theta_2 \cup \theta_3$	0	$b_{\Delta H}$	$\frac{(1-a_{\Delta H})b_{\Delta H}}{1-K_{\Delta H}}$	$(1-a_{\Delta H})b_{\Delta H} + \frac{b_{\Delta H} K_{\Delta H}}{a_{\Delta H}+b_{\Delta H}}$
$\theta_1 \cup \theta_2 \cup \theta_3$	$1-a_{\Delta H}$	$1-b_{\Delta H}$	$\frac{(1-a_{\Delta H})(1-b_{\Delta H})}{1-K_{\Delta H}}$	$(1-a_{\Delta H})(1-b_{\Delta H})$

Table 2. BBA construction for image change indicator ΔImg. [Conflict: $K_{\Delta Img} = a_{\Delta Img} b_{\Delta Img}$]

Focal Elem.	$m_2(.)$	$m'_2(.)$	$m_2^{DS}(.)$	$m_2^{PCR6}(.)$
θ_1	0	0	0	0
θ_2	0	0	0	0
θ_3	0	$b_{\Delta Img}$	$\frac{(1-a_{\Delta Img})b_{\Delta Img}}{1-K_{\Delta Img}}$	$(1-a_{\Delta Img})b_{\Delta Img} + \frac{b_{\Delta Img} K_{\Delta Img}}{a_{\Delta Img}+b_{\Delta Img}}$
$\theta_1 \cup \theta_2$	$a_{\Delta Img}$	0	$\frac{a_{\Delta Img}(1-b_{\Delta Img})}{1-K_{\Delta Img}}$	$a_{\Delta Img}(1-b_{\Delta Img}) + \frac{a_{\Delta Img} K_{\Delta Img}}{a_{\Delta Img}+b_{\Delta Img}}$
$\theta_2 \cup \theta_3$	0	0	0	0
$\theta_1 \cup \theta_2 \cup \theta_3$	$1-a_{\Delta Img}$	$1-b_{\Delta Img}$	$\frac{(1-a_{\Delta Img})(1-b_{\Delta Img})}{1-K_{\Delta Img}}$	$(1-a_{\Delta Img})(1-b_{\Delta Img})$

$a_{\Delta H}$ and $b_{\Delta H}$ defined by

$$a_{\Delta H} = f_{\tau,T_1}(\Delta H), \quad \text{and} \quad b_{\Delta H} = f_{-\tau,T_2}(\Delta H). \tag{13}$$

The discordance index can be considered as a reflection of the concordance index along the mirror line. Therefore, they are sharing the same τ. Here, the factor τ is calculated with a sample value ($\Delta H = 1$, $a_{\Delta H} = 0.1$), which means 1 m height change indicates 10% probability to be building changes. The BBAs for discordance and concordance image change index are built similarly. Differences appearing in 2D images give a concordance indication for all changes, which include the building changes or other changes ($\theta_1 \cup \theta_2$). In this article, the changes from images are named ΔImg.

In the Tables 1 and 2, we present the two ways of construction of the BBAs from the sources of evidence based either on DS or on PCR6 rules of combination for the height change indicator (i.e. the first source of evidence) and the image change indicator (i.e. the second source of evidence). It has to be noted that $\theta_1 \cup \theta_3$ is not mentioned in the fusion model, as they do not share similar characters within the used feature space. In Table 1, $m_1(.)$ and $m'_1(.)$ represent the concordance and discordance BBAs from ΔH, whereas in Table 2 $m_2(.)$ and $m'_2(.)$ represent the concordance and discordance BBAs from images. $K_{\Delta H}$ is the total conflicting mass value between $m_1(.)$ and $m'_1(.)$, and $K_{\Delta Img}$ in Table 2 is the total conflicting mass value between $m_2(.)$ and $m'_2(.)$,

3.3. *Reliability discounting*

In the DSM assisted building change detection, false alarms arise if wrong heights are presenting in the DSM for large regions (Tian *et al.* 2014a). And these wrong

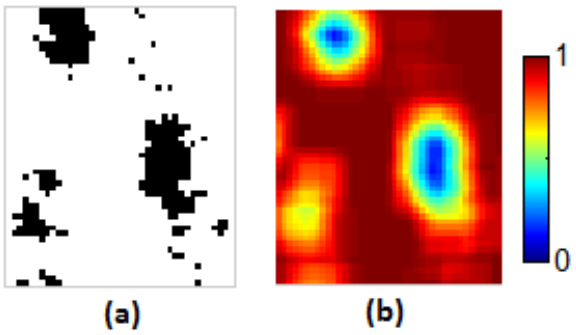

Figure 3. Reliability map (b) generated from the gaps mask (a).

heights are mostly introduced not in the stereoscopic images matching procedure, but in the gaps filling step. In the last step of the DSM generation procedure, the height of un-matched pixels is interpolated using the height values of neighbourhood pixels. Normally a reliable height value can be achieved for small gaps. But when large gaps appear in the disparity map, for example, for a whole building roof, the height of that building can not be correctly interpolated. Thus, the percentage of available successfully matched pixels inside a predefined neighbourhood region can be used to generate the height reliability. Fig. 3 shows an example of the generated reliability map. Fig. 3a is the gap mask. The gaps region of the disparity map is represented with black colour. Pixels with proper elevation values are displayed with white colour. It can be observed, based on our approach that pixels in the centre of a gap get lower reliability factor values than pixels next to the gap boundary (see Fig. 3b).

In the building change detection procedure, the reliability map of two DSMs (α_{DSM1} and α_{DSM2}) are calculated, respectively. They are then fused together to generate a final reliability map $\alpha_{\Delta H}$ for the height change mass.

$$\alpha_{\Delta H} = \alpha_{DSM1} \cdot \alpha_{DSM2}. \tag{14}$$

Shadow has played an important role when analysing very high resolution images in urban region. Both of the changes of shadow and coverage of shadow will bring false alarms for change detection. Therefore, the 2D changes that are detected in shadow regions are less reliable than in non-shadow regions. Benefit to this character, we can adopt the shadow map as the reliability map of BBA from the image change indicator. For this purpose, first a shadow map is generated by calculating the average brightness of the multi-spectral image, as normally a dark colour indicates the existence of shadows. We take an easy and fast shadow detection approach as shown in Eq. (15) to highlight the shadow class. It is a pixel-based approach, therefore, B_k in Eq. (15) represents the intensity values at one pixel location in different multi-spectral band images. And n is the number of the multispectral bands. The detected shadow map from brightness is enough for our purpose. In this shadow map, a smaller value indicates higher probability to be shadows; thus, the 2D changes detected in these regions are less reliable.

$$\text{Brightness} = \frac{1}{n}\sum_{k=1}^{n} B_k. \tag{15}$$

A further process is proposed to obtain a valid reliability map from the shadow map. First, it has been projected to a sigmoid curve. The lower threshold value from the two-level Ostu threshold is used as the symmetry point of the sigmoid curve. The obtained probability map is denoted as $ShadowMap$. In order to control the influence of the $ShadowMap$, we have only kept the values less than 0.5.

$$\alpha_{img} = \begin{cases} 0.5 + I_{ShadowMap}, \text{if } I_{ShadowMap} < 0.5, \\ 1, \text{otherwise.} \end{cases} \tag{16}$$

where $I_{ShadowMap}$ is the pixel intensity of the shadow map in [0, 1]. The reliability map generated from the shadow map is then recorded as $\alpha_{\Delta Img}$, and it is the combination of the shadow maps of two dates.

$$\alpha_{\Delta Img} = \alpha_{img1} \cdot \alpha_{img2}. \tag{17}$$

3.4. *Global BBAs*

The BBAs related with the concordance and discordance indexes are combined to get the global BBA regarding each source of evidence. These global BBAs will then be used as input for solving the change detection problem thanks to their combination. From the previous step of BBAs modelling, each pixel will get two sets of BBAs to combine results from Table 1 and 2. More precisely, we will have to combine either $\{m_1^{DS}(.), m_2^{DS}(.)\}$ if DS rule is preferred for the BBA modeling, or $\{m_1^{PCR6}(.), m_2^{PCR6}(.)\}$ if the PCR6 rule is adopted. These BBAs from Table 1 and 2 are represented by a_1, a_2, a_3 and b_1, b_2, b_3. In this article, the mass values a_1, a_2 and a_3 are further discounted by the generated reliability map $\alpha_{\Delta H}$ and denoted respectively as A_1, A_2 and A_3. The mass values from the image change indicator b_1, b_2 and b_3 are discounted by the vegetation and shadow indicators $\alpha_{\Delta Img}$ obtained in formula Eq. (17) to B_1, B_2 and B_3.

More precisely, one computes

$$\begin{cases} A_1 = \alpha_{\Delta H} \cdot a_1 \\ A_2 = \alpha_{\Delta H} \cdot a_2 \\ A_3 = \alpha_{\Delta H} \cdot a_3 + (1 - \alpha_{\Delta H}). \end{cases} \tag{18}$$

$$\begin{cases} B_1 = \alpha_{\Delta Img} \cdot b_1 \\ B_2 = \alpha_{\Delta Img} \cdot b_2 \\ B_3 = \alpha_{\Delta Img} \cdot b_3 + (1 - \alpha_{\Delta Img}). \end{cases} \tag{19}$$

Table 3 and Table 4 describe the final building change detection models based either on DS or on PCR6 rules. Here, the discounted height change indicator is denoted as $m_{1\alpha_{\Delta H}}(.)$, and the discounted image change indicator is denoted as $m_{2\alpha_{\Delta Img}}(.)$.

$m_{1\alpha_{\Delta H}}(.)$ can be obtained from the discounting of the fusion results presented in Table 1. Thus they have been denoted respectively as $m^{DS}_{1\alpha_{\Delta H}}(.)$ and $m^{PCR6}_{1\alpha_{\Delta H}}(.)$. These discounted height change indicators are fused in the second step with the image change

Table 3. DS fusion model for building change detection.

Focal Elem.	$m_{1\alpha_{\Delta H}}(.)$	$m_{2\alpha_{\Delta Img}}(.)$	$m_{12}^{DS}(.)$
θ_1	A_1	0	$\frac{A_1(B_1+B_3)}{1-A_1B_2}$
θ_2	0	0	$\frac{A_2B_1}{1-A_1B_2}$
θ_3	0	B_2	$\frac{(A_2+A_3)B_2}{1-A_1B_2}$
$\theta_1 \cup \theta_2$	0	B_1	$\frac{A_3B_1}{1-A_1B_2}$
$\theta_2 \cup \theta_3$	A_2	0	$\frac{A_2B_3}{1-A_1B_2}$
Θ	A_3	B_3	$\frac{A_3B_3}{1-A_1B_2}$

Table 4. PCR6 fusion model for building change detection.

Focal Elem.	$m_{1\alpha_{\Delta H}}(.)$	$m_{2\alpha_{\Delta Img}}(.)$	$m_{12}^{PCR6}(.)$
θ_1	A_1	0	$A_1(B_1+B_3)+\frac{A_1A_1B_2}{A_1+B_2}$
θ_2	0	0	A_2B_1
θ_3	0	B_2	$(A_2+A_3)B_2+\frac{B_2A_1B_2}{A_1+B_2}$
$\theta_1 \cup \theta_2$	0	B_1	A_3B_1
$\theta_2 \cup \theta_3$	A_2	0	A_2B_3
Θ	A_3	B_3	A_3B_3

indicator $m_{2\alpha_{\Delta Img}}(.)$ to generate the final global BBAs. From the Tables 3 and 4, four sets of global BBAs can be computed based on different BBAs and fusion models. The flow diagram in Fig. 4 summarises the different fusion schemes tested in our application.

As one sees, if both the BBA modelling procedure and global BBAs are constructed based on DS fusion rule, the generated global BBA is recorded as G_1. If the global BBAs are constructed based on PCR6 fusion rule, they are recorded as G_2. The basic BBAs can also be calculated with PCR6 fusion rule, as shown in Table 2. Based on these BBAs, the global BBAs can be also constructed using DS theory G_3 and PCR6 rule G_4. It has to be mentioned that these four fusion schemes have different computational cost and G_1 is the simplest one and G_4 is the most expensive one in terms of computational burden.

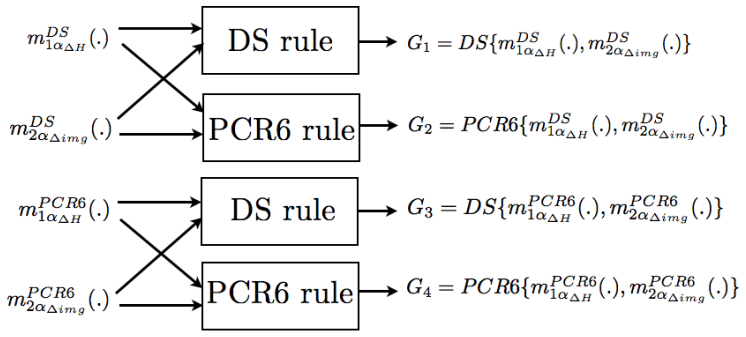

Figure 4. Four fusion schemes based DS and PCR6 rules.

3.5. *Change mask generation*

The final building change mask is our decision-making procedure. After the second step of fusion, each pixel in the images will get a certain degree of belief for all focal elements. The value of global BBAs in θ_1 gives a direct building change probability map. A decision criterion is required in generating the final building change detection masks. A change mask can be generated after giving a threshold value (Tian *et al.* 2014a). However, BBAs on the partial ignorance and full ignorance set should also be considered in the decision-making procedure. The building change probability map is only a part of the global BBAs. DST and DSmT propose different approaches to make the final decision. Several decision criteria are available. In this article, four decision criteria are tested. They are: 1) maximum of belief (Max_Bel), 2) maximum of plausibility (Max_Pl), 3) maximum of betting probabilities (Max_BetP) and 4) the maximum of DSmP (Max_DSmP) (Shafer 1976, Smarandache and Dezert 2004–2015)(Vol. 3, Chap. 3).

3.5.1. Maximum of Belief (Max_Bel)

Valid for different strategies of BBA modelings and fusions according to Fig. 4. More precisely, for a strategy G generating a combined mass $\in G_1, G_2, G_3, G_4$, the label (decision) is obtained by comparing the final global mass values obtained from Table. 3 and 4.

$$Label = \text{argmax}\{G(\theta_1), G(\theta_2), G(\theta_3)\}. \tag{20}$$

3.5.2. Maximum of plausibility (Max_Pl)

Plausibility is defined in Eq. (3). Max_Pl compares the plausibility of each class.

$$Label = \text{argmax}\{Pl(\theta_1), Pl(\theta_2), Pl(\theta_3)\}. \tag{21}$$

3.5.3. Maximum of betting probabilities (Max_BetP)

The pignistic probabilities, denoted as $BetP$, is making decisions on the pignistic level. In the betting probabilities, global masses of joint focal elements are averagely redistributed to each class.

$$BetP(A) = \sum_{B \in \Theta} \frac{|A \cap B|}{|B|} m(B), A \in \Theta. \tag{22}$$

$$Label = \text{argmax}\{BetP(\theta_1), BetP(\theta_2), BetP(\theta_3)\}. \tag{23}$$

3.5.4. Maximum of DSmP (Max_DSmP)

DSmP probabilistic transformation is an important alternative to the pignistic transformation (Dezert and Smarandache 2008). The basic idea of DSmP is to redistribute

the mass of (partial and total) ignorances proportionally to the masses of singletons involved in the ignorances.

$$DSmP_{\varepsilon}(A) = \sum_{B \in \Theta} \frac{\sum\limits_{\substack{Z \subseteq A \cap B \\ |Z|=1}} m(Z) + \varepsilon |A \cap B|}{\sum\limits_{\substack{Z \subseteq B \\ |Z|=1}} m(Z) + \varepsilon |B|} m(B). \tag{24}$$

where $\varepsilon \geq 0$ is a small positive number (typically 0.001) that avoids numerical indeterminacies in very degenerated cases occurring if the mass in the denominator of Eq. (24) is zero. More detailed information about DSmP is described in Dezert *et al.* (2011) and Dezert and Smarandache (2008).

$$Label = \text{argmax}\{DSmP(\theta_1), DSmP(\theta_2), DSmP(\theta_3)\}. \tag{25}$$

Among the four decision-making rules, max of belief or max of plausibility have the advantage to be very simple to calculate but they represent respectively two extreme pessimistic or optimistic decisional attitudes. The choice of one of these extreme attitudes depends on the consequence of decision error we are ready to take which is conditioned by the type of application under concern. Moreover, it has been shown by Dezert and Smarandache (2008) that the more sophisticate transformation DSmP outperform BetP transformation at a price of much higher computational complexity, which can be a bottleneck in some real-time image processing applications.

4. Experiments

4.1. *Datasets*

The belief function-based building change detection models have been tested on four pairs of satellite images. Each of the first three experimental datasets consist of two pairs of IKONOS stereo imagery captured in February 2006 and May 2011 over an industrial region in Dong-an, North Korea. These three sub-test regions are shown in Fig. 5 and 6 and 7, respectively. The original IKONOS stereo imagery has 1 m pixel size in the panchromatic band and 4 m pixel size in the multispectral bands. The fourth experimental dataset (shown in Fig. 8) was captured over the centre of Munich, Germany, which is a typical European urban region. The two pairs of stereo data of this dataset were captured by IKONOS on July 15, 2005 and WorldView-2 on July 12, 2010, respectively. In Fig. 5 to Fig. 8, the first two images are the panchromatic images of before- and after-change. (c) and (d) are the generated DSMs. They have been generated based on the method explained by d'Angelo and Reinartz (2012). The elevation values from low to high are represented with the colours from dark blue to dark red as described in the colour bar. These images are co-registered through camera model parameter corrections before the DSM generation procedure with block adjustment among all datasets (d'Angelo and Reinartz 2012, Qin *et al.* 2016b). A sub-pixel accuracy in planimetry and 1 to 2 m in height can be achieved. The Gram-Schmidt pan-sharpening method which has been widely used and implemented in ENVI software is applied to the multispectral channels of all three test regions (Laben

and Brower 2000). In the first three subsets the generated DSMs have been re-sampled to 1 m resolution. As the IKONOS and WorldView-2 data for the Munich test region have different resolutions, the IKONOS images are up sampled to 0.5 m resolution, to be equal to WorldView-2 data. Instead of down-scale (Klaric *et al.* 2013), an up-scale re-sampling is selected here to keep the sharp boundaries in the WorldView-2 data. The resulting DSMs also have a resolution of 0.5 m.

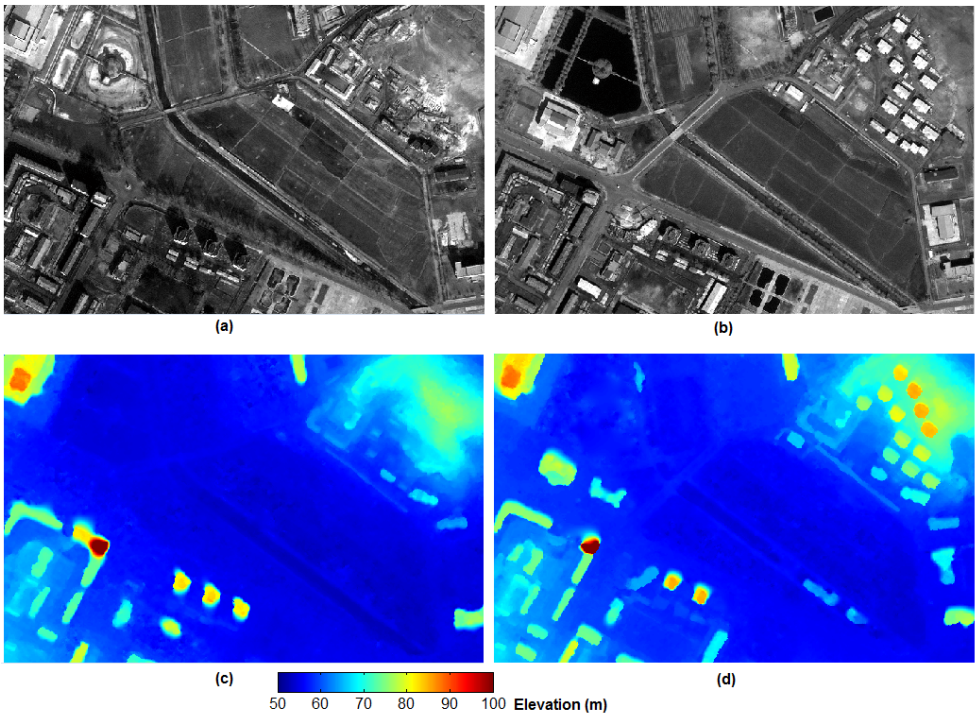

Figure 5. Datasets of the test region 1: a) panchromatic image from date1; b) panchromatic image from date2; c) DSM from date1; (d) DSM from date2.

Fig. 5 and Fig. 6 show normal building change examples with DSMs in high accuracy. The size of these two test regions are 450×700 m^2, and 1000×400 m^2, respectively. In Fig. 5 some seasonal changes are visible. The generated DSMs are displayed in Fig. 5c and 5d. The second test region (Fig. 6) shows much larger sized buildings, and these buildings are well separated from each other. The third test region consists of two images with the size of 160×340 pixels. This region is characterised by small sized buildings (Fig. 7). It has to be mentioned, the largest building with a dark colour roof does not have the correct height in the first DSM, as is shown in Fig. 7c. This test region is especially selected to prove the robustness of our fusion models. The image size of the fourth test region is 1600×1600 pixels, which is 640,000 m^2. It has mainly large size buildings with complex roof shapes. From 2005 to 2010, besides newly constructed buildings, there are also rebuilt/demolished buildings. Especially, many roofs have been renovated with another material. Without height information, it is very difficult to separate the newly constructed buildings from other kinds of changes.

4.2. *Results*

The proposed DS fusion model and PCR6 fusion model have been applied to all datasets. In the first step, the four sets of global BBAs for all three focal elements and

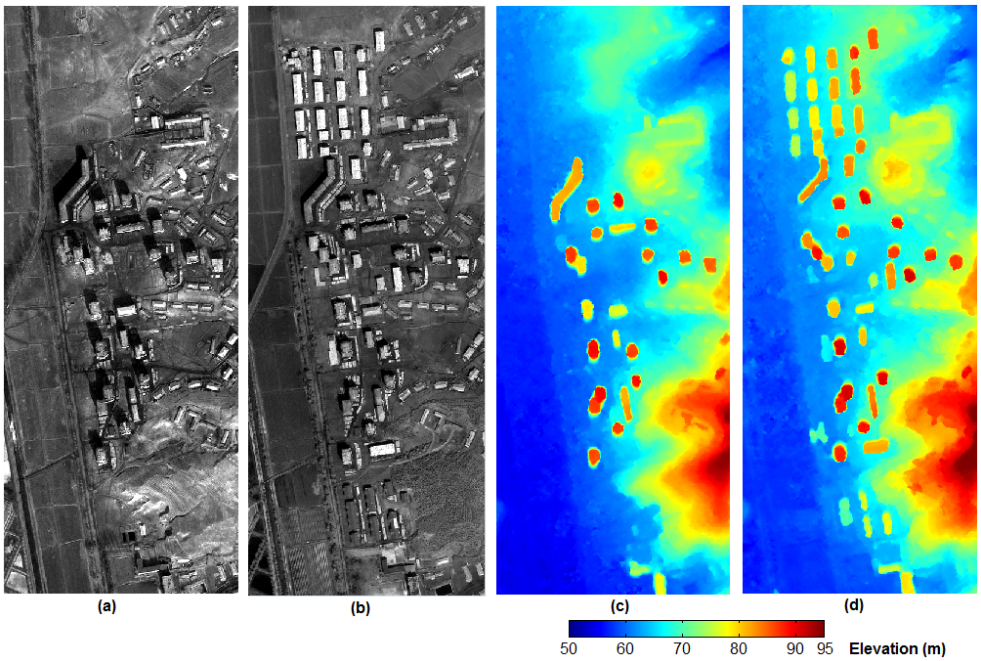

Figure 6. Datasets of the test region 2: a) panchromatic image from date1; b) panchromatic image from date2; c) DSM from date1; (d) DSM from date2.

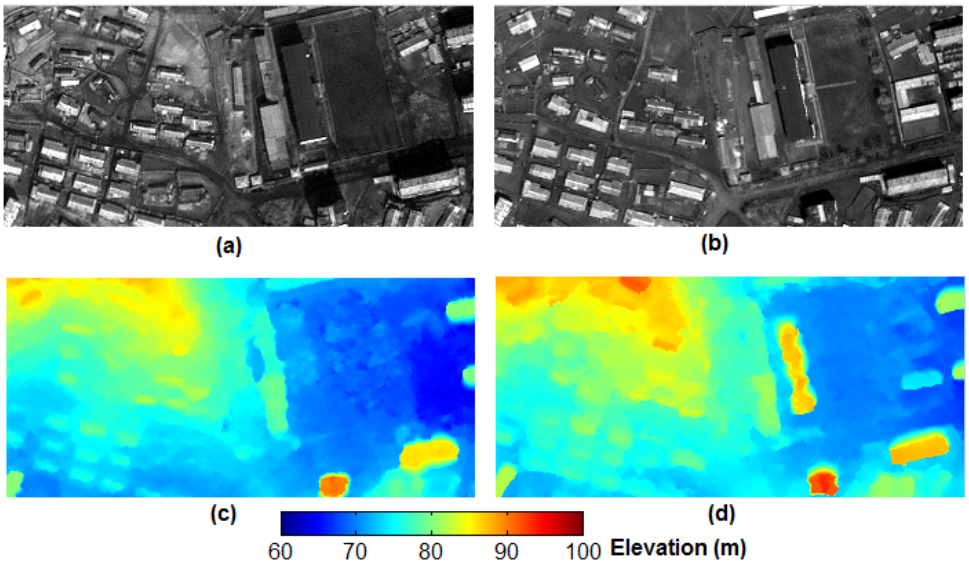

Figure 7. Datasets of the test region 3: a) panchromatic image from date1; b) panchromatic image from date2; c) DSM from date1; (d) DSM from date2.

joint elements are generated based on various fusion rules and fusion rule combinations. In the second step, building change masks are generated by using four decision criteria. All three classes including *BuildingChange*, *OtherChange* and *NoChange* are generated. But this article focuses on the newly constructed buildings, thus only the *BuildingChange* results are analysed and evaluated. The proposed models have two novel properties. The first one is the improved fusion model, and the second one is the reliability discounting. In the experimental part, the minimal value of the reliability

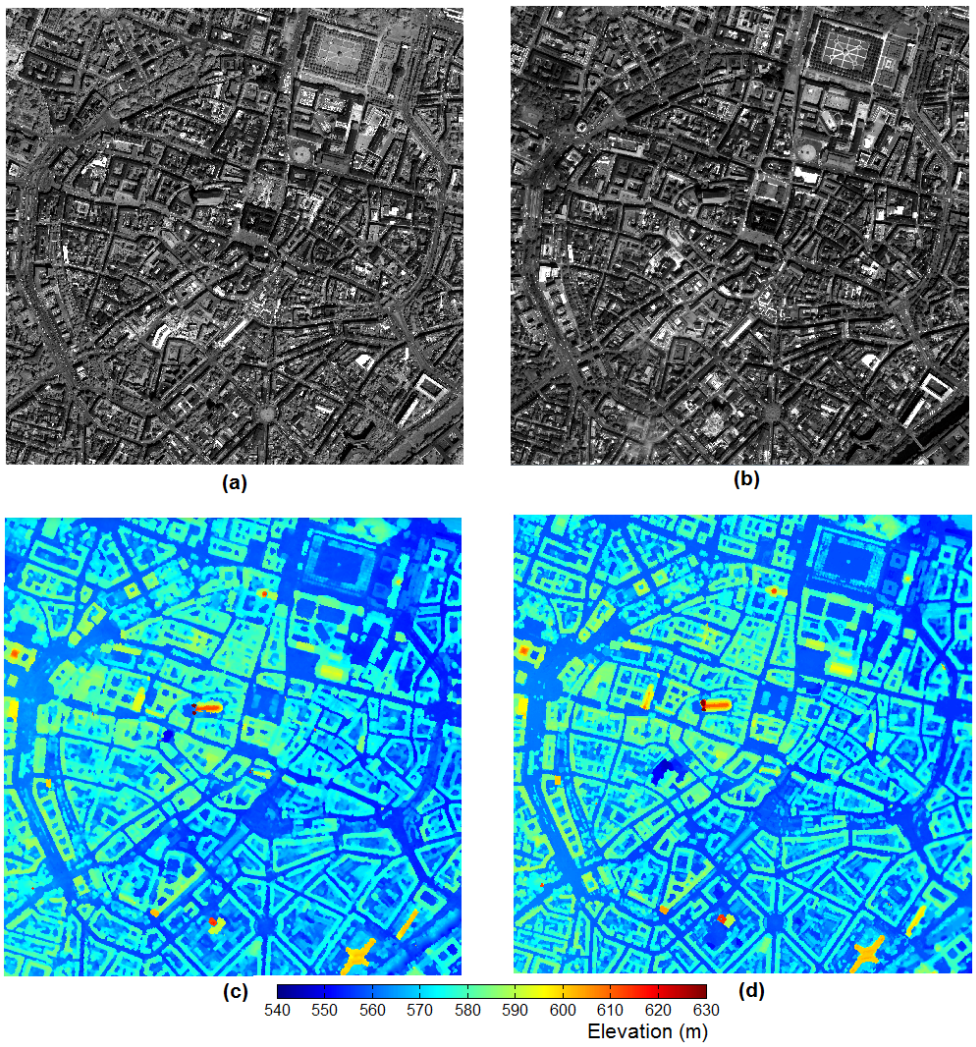

Figure 8. Datasets of the test region 4: a) panchromatic image from date1; b) panchromatic image from date2; c) DSM from date1; (d) DSM from date2.

map generated from DSM gaps is manually modified to 0.1 to remove too small values. In the height change reliability map generation procedure, a window size of 9×9 is selected.

To prove the advantages of the proposed method, firstly the best building change detection results are displayed together with the original height change map. The results of all four test regions are displayed in Fig. 9, Fig. 10, Fig. 11 and Fig. 12, respectively. In each figure, different colours represent different height changes in Figs. 9-12(a). Figs. 9-12(b) are the generated building change masks. To show the quality of these building change masks, these masks have been overlaid with the change reference data, which have been manually extracted for all four test regions. In Figs. 9-12(b) the green colour represents the correctly detected building changes. The false alarms which indicate pixels that are wrongly detected as building changes are presented with red colours. The blue colour objects are the misdetected changed buildings, which are named as *false negatives* in this article.

Generally speaking, the proposed models are able to extract the newly constructed buildings in high accuracy. Noise effects from the height change map have been largely

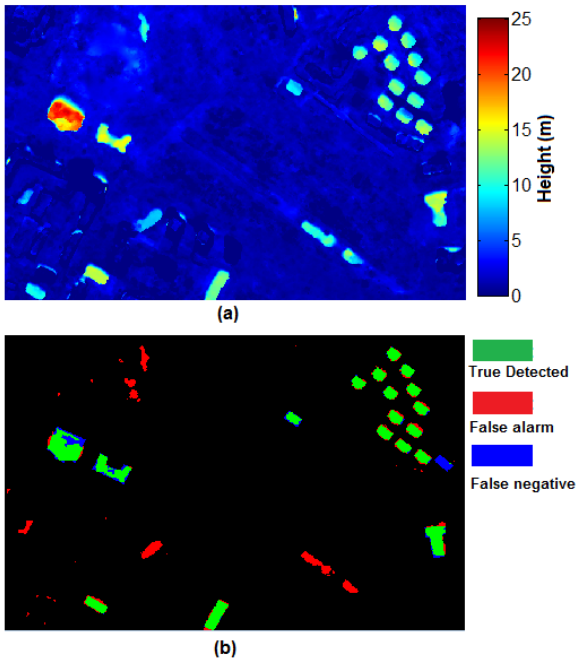

Figure 9. Change detection results of test region 1 (a) original height change map (b) building change result $Max\_Pl(G_4)$ overlaid with change reference data.

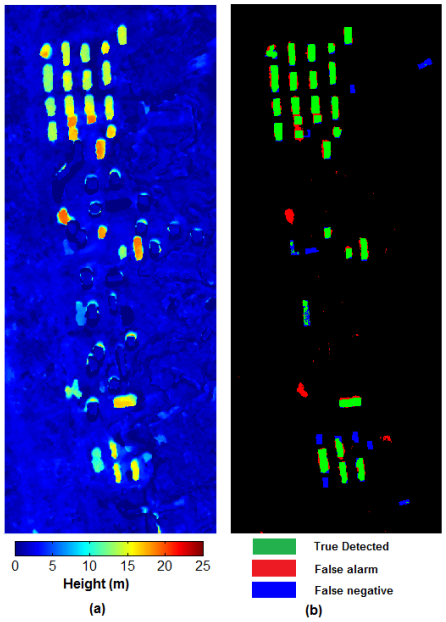

Figure 10. Change detection results of test region 2 (a) original height change map (b) building change result $Max\_Pl(G_4)$ overlaid with change reference data.

reduced in the final change results. The four selected test regions present four different situations. In the first test region most of the buildings are relatively low in the height and well separated from each other. The second test region has much higher and larger buildings, which produce large regions of shadow. The third test region is a special case. As we observe, in the first DSM of test region 3 the height of one big building is not correctly extracted. Actually the same building has been detected as false alarm

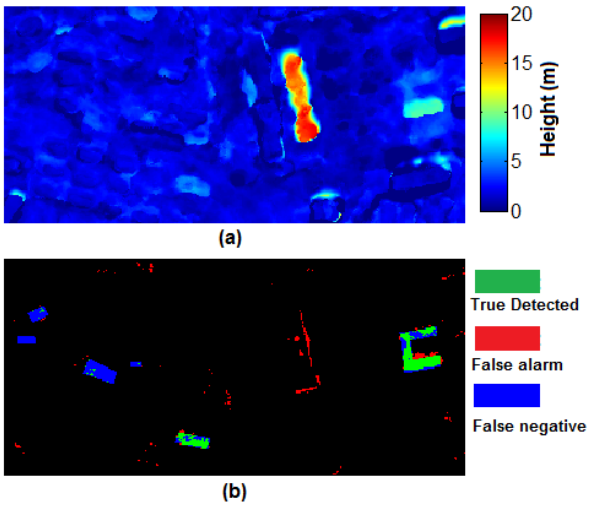

Figure 11. Change detection results of test region 3 (a) original height change map (b) building change result $Max\_DSmT(G_2)$ overlaid with change reference data.

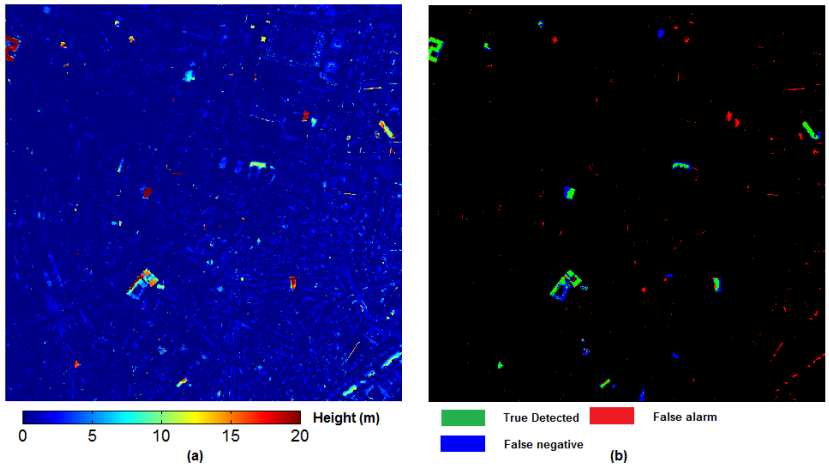

Figure 12. Change detection results of test region 4 (a) original height change map (b) building change result $Max\_DSmT(G_2)$ overlaid with change reference data.

and been discussed in the reference (Tian *et al.* 2014a). It has been explained in Tian *et al.* (2014a), due to the large region size and height change values, the false alarm can not be avoided. The fourth test region is much more complicated than the others, exhibiting very high building density, complex roof shapes and various building change types.

Benefiting from the improved fusion models and the reliability discounting procedure, some false alarms can be successfully avoided. Especially for the building in the left-bottom corner in the test region 1 and that big building in test region 3. In both situations, the first DSM is not able to get the correct height values. Based on the traditional feature fusion approach or our initial fusion model (Tian *et al.* 2014a), this kind of buildings will very possibly be detected as *BuildingChange*. However, as we observe in the presented change detection results, these buildings are correctly detected as *NoChange*. It has to be noted that vegetation change is not considered in this model. Thus, in the centre of the first region, these two large regions of false

alarms, which are newly planted trees from visual interpretation, are not able to be avoided. Another difficult to detect region is one building in construction. Half of the building has been finished in the after-change data; thus, this region has both height and spectral changes. As it can not be called a finished building yet, we did not include it as *BuildingChange* in our reference data.

Many false negatives (blue regions/pixels) in Fig. 10(b) are visible. Most of these false negatives can be explained by the quality of the DSMs. A subset of the gaps mask of test region 2 in date2 is displayed in Fig. 13. As it shows all of the four missed buildings (shown in blue colour) are actually gaps in the unfilled DSM. After gaps filling, they are not interpreted with correct height values, as shown in Fig. 6 (d). Thus, these four buildings only feature spectral changes, therefore are falsely identified as *OtherChange*.

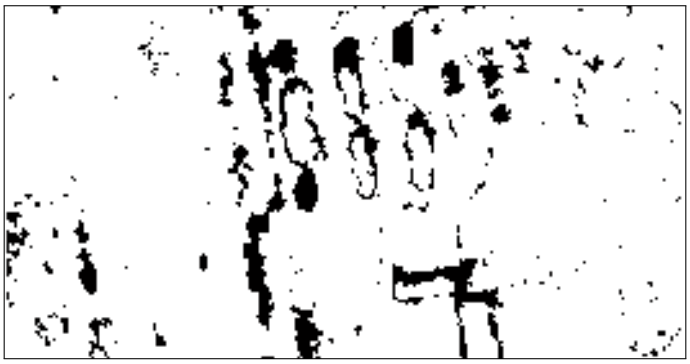

Figure 13. DSMs gaps of part of test region 2 (black holes).

4.3. *Results evaluation*

To further understand the quality of these results and the advantages of the proposed method, more evaluation and analysis are proposed. First, the building change masks extracted from these four global BBA sets are compared and evaluated. Each global BBA set results four building change masks based on the four decision criteria. The building change masks are compared with the masks from Tian *et al.* (2015b). The accuracy of these results have been evaluated by comparing them with ground truth images, which have been manually prepared by visually comparing the pre- and post-event images and referring additional Google Earth history data (GoogleEarth 2018). The similarity between the obtained result and the ground truth is measured in terms of Kappa Accuracy (KA) (Congalton 1991). The evaluation results of test region 1, 2, 3 and 4 are shown in Table. 5, 6, 7 and 8, respectively. Limited to the available reference data, only the *BuildingChange* class is evaluated.

Table 5. Change masks evaluation from four global BBAs of test region 1 (KA).

	G_1		G_2		G_3		G_4	
	Original	Refined	Original	Refined	Original	Refined	Original	Refined
Max_Bel	0.7392	0.7150	0.7369	0.7138	0.7419	0.7144	0.7391	0.7130
Max_Pl	0.7619	**0.7648**	0.7607	**0.7642**	0.7623	**0.7652**	0.7609	**0.7641**
Max_BetP	0.7533	0.7442	0.7515	0.7423	0.7541	0.7428	0.7522	0.7412
Max_DSmP	0.7468	0.7200	0.7450	0.7189	0.7490	0.7190	0.7465	0.7181

Table 6. Change masks evaluation from four global BBAs of test region 2 (KA).

	G_1		G_2		G_3		G_4	
	Original	Refined	Original	Refined	Original	Refined	Original	Refined
Max_Bel	0.7401	0.7821	0.7399	0.7816	0.7401	0.7826	0.7401	0.7821
Max_Pl	0.7380	0.7800	0.7391	0.7818	0.7380	0.7812	0.7393	0.7831
Max_BetP	0.7413	**0.7853**	0.7409	**0.7853**	0.7412	**0.7868**	0.7409	**0.7867**
Max_DSmP	0.7402	0.7842	0.7403	0.7841	0.7405	0.7857	0.7403	0.7855

Table 7. Change masks evaluation from four global BBAs of test region 3 (KA).

	G_1		G_2		G_3		G_4	
	Original	Refined	Original	Refined	Original	Refined	Original	Refined
Max_Bel	0.3356	**0.5432**	0.3356	**0.5418**	0.3351	**0.5415**	0.3345	**0.5419**
Max_Pl	0.2396	0.3689	0.2416	0.3703	0.2391	0.3694	0.2409	0.3713
Max_BetP	0.2860	0.4726	0.2885	0.4756	0.2869	0.4761	0.2882	0.4786
Max_DSmP	0.3043	0.5082	0.3057	0.5094	0.3008	0.5072	0.3030	0.5066

Table 8. Change masks evaluation from four global BBAs of test region 4 (KA).

	G_1		G_2		G_3		G_4	
	Original	Refined	Original	Refined	Original	Refined	Original	Refined
Max_Bel	0.5158	0.5217	0.5159	0.5219	0.5154	0.5193	0.5158	0.5195
Max_Pl	0.5122	0.5229	0.5125	0.5232	0.5120	0.5224	0.5128	0.5232
Max_BetP	0.5137	0.5268	0.5140	0.5267	0.5135	0.5258	0.5137	0.5258
Max_DSmP	0.5161	**0.5285**	0.5163	**0.5284**	0.5157	**0.5274**	0.5162	**0.5275**

4.3.1. Comparison of the fusion and decision rules

Table. 5 to 8 mainly aim to describe and compare the performance of the DS fusion and DSmT fusion rules and the four decision criteria. Unfortunately, the differences among these four global BBA sets of all four test regions are indistinguishable. Our quantitative evaluations results allow comparing the different fusion and decision-making strategies for building change mask construction in different types of region under analysis. As we have observed, there is no unique best fusion and decision strategy working for all types of regions which is an interesting result to be aware of and the different fusion methods (with a chosen decision strategy) perform always better with our refined approach than the previous (original) works which is the main contribution of this work for all type of regions tested.

4.3.2. Validation of the reliability discounting

The global BBAs obtained with and without reliability discounting are listed under the name of *Refined* and *Original* in Table 5 to 8. *Original* refers to the approach presented by Tian *et al.* (2015b), in which the reliability discounting is not involved.

In the first test region, the advantage of the reliability discounting is not obvious. By using the Max_Pl and Max_BetP decision rules, the refined models perform better than the original models. However, the original models get higher KA values when using the Max_Pl and Max_BetP decision rules. This can be partly explained by the shadow detection results, as one dark colour building roof (middle left in the test region) get higher probability to be shadow; thus, a lower probability to be *BuildingChange*.

The second test region is characterised mainly by large and high buildings; thus, the influences of shadows are stronger than in the first test region. The refined models with reliability discounting get generally better accuracy than the original fusion models. Here, we will compare the Max_DSmP of G_4 of this test region, as it shows the highest difference among these four decision criteria in Table 6. Fig. 14 shows building change

masks of the top left part of the test region 2. Fig. 14 (a) and (b) display the change masks obtained from the original model and the refined model overlaid with the change reference mask respectively. The same as Fig. 10 (b), the green colour represents the true detected, the red colour shows the false alarms, while the blue colour pixels are the false negatives. As it shows, based on the refined model, building boundary regions of the change mask obtain less false alarms than the results from the original fusion model.

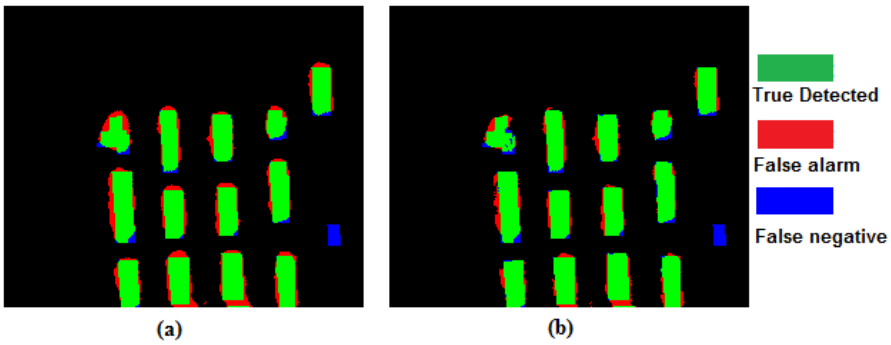

Figure 14. Building change masks from the original model (a) and refined model (b) of a subset of test region 2.

The advantage of the improved decision fusion models has been well proved by Table 7. The first DSM of this test region contains a large region of pixels with incorrect height introduced by stereo image matching failures. The improved models can solve this problem by adopting the reliability map of height change. Therefore, the increase of KA value of this region is much higher than for the other two test regions. More precisely, under all fusion rules the KAs have improved from around 0.30 to 0.50. For better understanding of this improvement, the global BBAs of *BuildingChange* without and with reliability discounting are displayed in Fig. 15 (a) and (b), respectively. We display here only the $Prob(\theta_1)$ of G_1. Both probability maps are less noisy than the original height change map, which are displayed in Fig. 11. By observing the original panchromatic images in Fig. 7, it is not difficult to find out that this building exists in the panchromatic images of both dates. This is the same building that has been mentioned in Tian *et al.* (2014a), for which only the DSM of pre-change contains the correct height values. In Fig. 7 (c), this building can not be recognised as a high-level object. A higher value in $m_1(.)$ leads to a larger global BBA in the class of *BuildingChange*. Thus, this building would be incorrectly detected as *BuildingChange* if no reliability discounting is applied (Fig. 15(a)). Fig. 16 shows the generated height change reliability map. As can be seen, that building region get very low reliability values, that means the height changes of this region cannot be trusted. Therefore, the proposed model is able to remove this kind of errors and correctly recognise this region as *NoChange* (Fig. 11).

The Munich test region has a much larger size and includes several kinds of building changes. The proposed method is able to fuse the spectral and height information efficiently; thus, to identify the newly constructed buildings. The main false negatives are produced in the rebuilt buildings and construction sites. As shown in Fig. 17, the labelled four buildings represent four types of changes. Building A is labelled as a newly constructed building in our reference data. However, half of that building has similar shape and height as the original one, which brings false negatives to our result. Building B, C and D are buildings in different construction phases. By referring

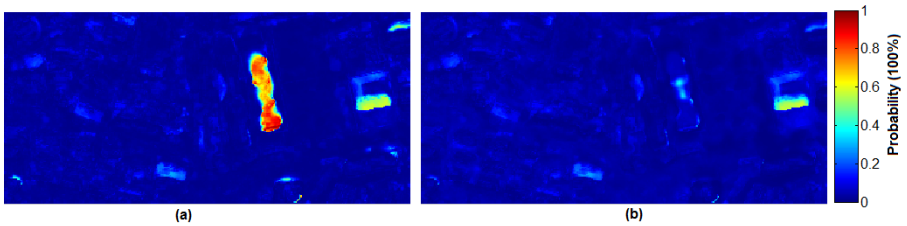

Figure 15. Comparison of building change global BBAs $Prob_{\theta_1}$ of G_1 based on the fusion models without reliability discounting (a) and with reliability discounting (b).

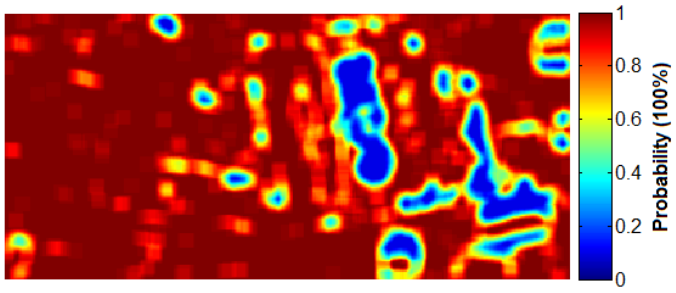

Figure 16. Reliability discounting map of the height changes of test region 3.

GoogleEarth (2018), in the reference data only D is identified as *OtherChange* as it is almost completed in Fig. 17(a). In the result we are able to correctly identify B as a newly constructed building and D as *OtherChange*. But building C is falsely labeled as *OtherChange* due to low height change values.

Figure 17. Example of the various building change types in test region 4.

A windowsize of 9×9 has been used to generate the $\alpha_{\Delta H}$. In order to test the sensitivity of our fusion model to the window width used, we have changed the width parameter from 3 to 13 by steps of 2. For each size, we generate the global BBA G_1. Thus four final building change masks based on the four decision criteria can be regenerated. We provide the KA for each mask as show in Fig. 18. As a comparison, we have also provided the KAs without using $\alpha_{\Delta H}$. This test shows that the final results benefit largely from the reliability discounting procedure, but the KA rate did not change significantly with various window sizes for all four test regions.

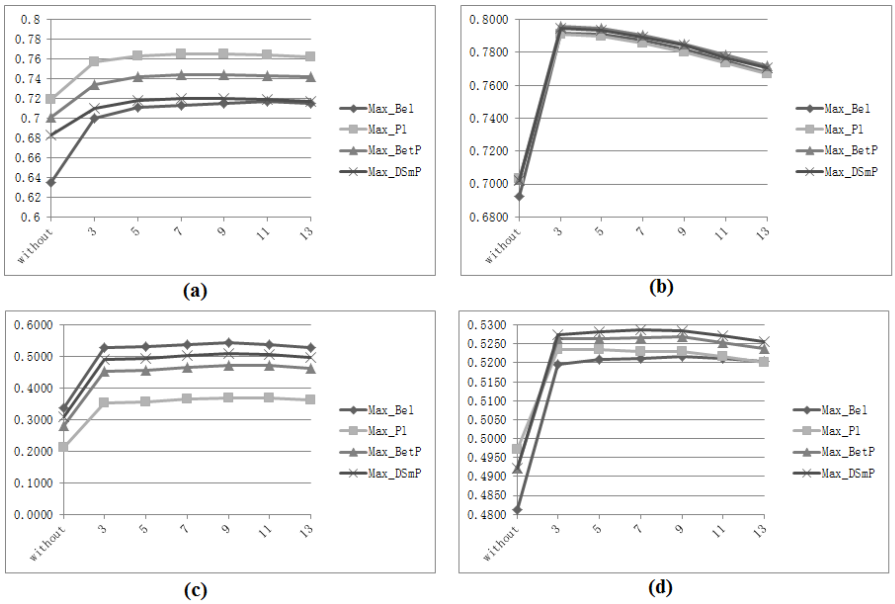

Figure 18. Effect of windowsize on KA for test region 1 (a), test region 2 (b), test region 3 (c) and test region 4 (d).

4.4. *Comparison with existing methods*

In this section, the improved belief fusion models are compared with the directly feature fusion method (Tian *et al.* 2014b) and the initial fusion model that described in Tian *et al.* (2014a).

As a typical feature fusion approach, Tian *et al.* (2014b) adopted the kernel Minimum Noise Fraction (kMNF) approach to fuse change features from the DSMs and panchromatic images. Based on the resulting kMNF components, a change mask was extracted with iterated canonical discriminant analysis (ICDA). Tian *et al.* (2015b) randomly selected the training data from the ground truth, as the experiments were devoted to algorithm comparison. However, it was not a practical procedure, because in real situations the ground truth is unknown. Therefore, in this article as well as using the set of random pixels from the ground truth, another set of training data for each test region is prepared by manually selecting changed regions. All pixels in these regions are then used as training samples.

The results generated based on these two sets of training data are described as $kMNF_{random}$ and $kMNF_{manual}$, respectively, in Table 9 in the term of KA and Overall Accuracy (OA). All training data in the first three test regions contain around 200 pixels/samples. In the fourth test region, 500 pixels are used to fit with the large image size. If the training data are selected from the ground truth, the newly proposed approach can deliver a slightly better result than the approach in Tian *et al.* (2014b). When using the manually selected training data, the advantages of the newly developed approach are obvious. As in real applications the ground truth is normally unknown, we conclude that the proposed fusion method is more robust for larger test regions with diverse characterised objects.

In addition, the approach proposed by Tian *et al.* (2014a) is tested on the same test data, and the results are shown in the third and fourth columns. In that approach, after the fusion approach a shape-based refinement was proposed to reach the final building change mask. Thus, the resulting masks before and after the refinement procedure

are both calculated and evaluated. In the North Korea test region, we have used $T_{height} = 3m$, $T_{area} = 50m^2$ and $T_{convexity} = 0.55$ as thresholds. And in the Munich test region, as the buildings have a larger size and complicated roof shapes than North Korea, we manually modified these threshold values to $T_{area} = 100m^2$ and $T_{convexity} = 0.50$ to improve the results. The accuracies are recorded in the columns Tian *et al.* (2014a)$_{before}$ and Tian *et al.* (2014a)$_{after}$ in Table 9. The refinement is not included in this article to avoid unnecessary threshold parameters; thus, to achieve an automatic and robust work-flow. By comparing the KAs with Tables 5, 6 and 7, one can see that the shape-based refinement can further improve the result accuracy. But the fusion model in Tian *et al.* (2014a) performs rather weakly. All obtained KAs are lower than values from the proposed refined decision fusion approaches, especially for test regions 2 and 3.

Table 9. Comparison with existing methods.

	$kMNF_{random}$		$kMNF_{manual}$		(Tian *et al.* 2014a)$_{before}$		(Tian *et al.* 2014a)$_{after}$	
	KA	OA	KA	OA	KA	OA	KA	OA
Region1	**0.7178**	0.9799	0.5477	**0.9803**	0.5929	0.9628	0.6312	0.9683
Region2	**0.6791**	**0.9822**	0.2458	0.9688	0.6433	0.9681	0.6718	0.9718
Region3	0.2195	0.9794	0.2272	**0.9799**	0.3060	0.9375	**0.3287**	0.9447
Region4	0.2057	0.9878	0.1937	0.9876	0.4909	0.9912	**0.5641**	**0.9941**

It has to be mentioned that vegetation change is not noted as false alarm in the improved decision fusion model. As the vegetation change and building change can be easily separated by using a vegetation index. Tian *et al.* (2014a) has adopted vegetation index as *no-building change* indicators to highlight building changes. This step is not considered in this article as not many forest changes are available in the test regions. Moreover, if forest changes are of interest, we can easily modify this model using vegetation index to separate forest changes from building changes.

5. Conclusions and perspectives

Building change detection is a difficult topic, to solve uncertain change information from images and DSMs, decision fusion methods have been introduced as a new concept and proved to be efficient and appropriate. The innovative contribution of this article is the improvement of the decision fusion models. DS as well as DSmT decision fusion models are further developed to solve the building change detection problem in this article. Another contribution lies in the BBA calculation procedure, and the sigmoid distribution is further improved by taking both concordance and discordance situations. As a third contribution, the reliability of each indicator is introduced according to the change objects of interest.

The proposed building change detection models enable an improved result by comparing to the original fusion model and other change detection methods. A comparative analysis of the results shows that there is not a so big difference of performances between DS and DSmT fusion methods based on the best decisional strategy and so we can in practise use the simplest fusion method to reduce to computational burden without degrading too much the performance. Of course the most critical question is to select beforehand the decisional strategy based on type of region under analysis, for this we need to define efficient indicators for characterising each type of region which then will help us to automatically select the best criterion to use. Our future research works will address, and hopefully help, to solve this important question.

References

Bruzzone, L. and Prieto, D.F., 2000. Automatic analysis of the difference image for unsupervised change detection. *IEEE Transactions on Geoscience and Remote sensing*, 38 (3), 1171–1182.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment*, 37 (1), 35–46.

d'Angelo, P. and Reinartz, P., 2012. DSM based orientation of large stereo satellite image blocks. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*, 39 (B1), 209–214.

Dempster, A., 1967. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38 (2), 325–339.

Dezert, J., Liu, Z.g., and Mercier, G., 2011. Edge detection in color images based on DSmT. *In*: *Proc. of FUSION 2011 conference*. 1–8.

Dezert, J. and Smarandache, F., 2008. A new probabilistic transformation of belief mass assignment. *In*: *2008 11th International Conference on Information Fusion*. 1–8.

Dezert, J. and Tacnet, J.M., 2012. Sigmoidal model for belief function-based ELECTRE TRI Method. *In*: *Belief functions: Theory and applications*. 401–408.

Dezert, J. and Tchamova, A., 2014. On the validity of Dempster's fusion rule and its interpretation as a generalization of Bayesian fusion rule. *Int. J. Intell. Syst.*, 29 (3), 223–252.

Dezert, J., Wang, P., and Tchamova, A., 2012. On the validity of Dempster-Shafer theory. *In*: *Proc. of FUSION 2012*. 655–660. Available from: https://www.onera.fr/staff/jean-dezert/references.

GoogleEarth, 2018. Available from: https://support.google.com/earth/answer/148094?hl=en.

Janalipour, M. and Taleai, M., 2017. Building change detection after earthquake using multi-criteria decision analysis based on extracted information from high spatial resolution satellite images. *International Journal of Remote Sensing*, 38 (1), 82–99.

Klaric, M.N., *et al.*, 2013. Geocdx: An automated change detection and exploitation system for high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 51 (4), 2067–2086.

Laben, C.A. and Brower, B.V., 2000. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening.

Le Hégarat-Mascle, S., Bloch, I., and Vidal-Madjar, D., 1997. Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing. *IEEE Trans. Geosci. Remote Sens.*, 35 (4), 1018–1031.

Li, W., *et al.*, 2017. A new approach to performing bundle adjustment for time series uav images 3d building change detection. *Remote Sensing*, 9 (6), 625.

Liao, P.S., Chen, T.S., and Chung, P.C., 2001. A fast algorithm for multilevel thresholding. *J. Inf. Sci. Eng.*, 17 (5), 713–727.

Liu, Z.G., 2014. *Credal classification of uncertain data based on belief function theory*. Thesis (PhD). Télécom Bretagne; Université de Bretagne Occidentale.

Liu, Z.g., *et al.*, 2014. Change detection in heterogeneous remote sensing images based on multidimensional evidential reasoning. *IEEE Geoscience and Remote Sensing Letters*, 11 (1), 168–172.

Liu, Z., *et al.*, 2018. Change detection in heterogenous remote sensing images via homogeneous pixel transformation. *IEEE Transactions on Image Processing*, 27 (4), 1822–1834.

Lu, D., *et al.*, 2004. Change detection techniques. *Int. J. Remote Sens.*, 25 (12), 2365–2407.

Mercier, D., Quost, B., and Denœux, T., 2005. Contextual discounting of belief functions. *In*: *Symbolic and quantitative approaches to reasoning with uncertainty*. Springer, 552–562.

Nielsen, A.A., 2007. The regularized iteratively reweighted MAD method for change detection in multi-and hyperspectral data. *IEEE Trans. Image Process.*, 16 (2), 463–478.

Otsu, N., 1975. A threshold selection method from gray-level histograms. *IEEE Trans. Syst., Man, Cybern.*, 9 (1), 62–66.

Pang, S., *et al.*, 2018. Building change detection from bi-temporal dense-matching point clouds and aerial images. *Sensors*, 18 (4), 966.

Qin, R., Tian, J., and Reinartz, P., 2016a. 3d change detection–approaches and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 41–56.

Qin, R., Tian, J., and Reinartz, P., 2016b. Spatiotemporal inferences for use in building detection using series of very-high-resolution space-borne stereo images. *International Journal of Remote Sensing*, 37 (15), 3455–3476.

Rottensteiner, F., 2007. Building change detection from digital surface models and multi-spectral images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (IAPRS)*, 36 (3), 145–150.

Rottensteiner, F., *et al.*, 2005. Using the Dempster–Shafer method for the fusion of LIDAR data and multi-spectral images for building detection. *Information Fusion*, 6 (4), 283–300.

Shafer, G., 1976. *A mathematical theory of evidence*. Princeton university press Princeton.

Sirmacek, B. and Unsalan, C., 2011. A probabilistic framework to detect buildings in aerial and satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, 49 (1), 211–221.

Smarandache, F. and Dezert, J., 2004–2015. *Advances and applications of dsmt for information fusion*. vol. 1–4. American Research Press, Rehoboth, NM, U.S.A.

Smarandache, F. and Dezert, J., 2013. On the consistency of PCR6 with the averaging rule and its application to probability estimation. *In*: *Proc. of FUSION 2013 conference*. 1119–1126.

Smarandache, F., Dezert, J., and Tacnet, J.M., 2010. Fusion of sources of evidence with different importances and reliabilities. *In*: *Proc. of FUSION 2010 conference*. 1–8.

Tchamova, A. and Dezert, J., 2012. On the behavior of Dempster's rule of combination and the foundations of Dempster-Shafer theory. *In*: *Proc. of IS 2012*. 108–113.

Tewkesbury, A.P., *et al.*, 2015. A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment*, 160, 1–14.

Tian, J., *et al.*, 2010. Automatic 3D change detection based on optical satellite stereo imagery. *In*: *ISPRS TC VII Symposium 100 Years ISPRS*, Vienna, Austria. vol. XXXVIII - Part 7B, 586–591.

Tian, J., *et al.*, 2013. Region-based automatic building and forest change detection on Cartosat-1 stereo imagery. *ISPRS J. Photogramm. and Remote Sens.*, 79, 226–239.

Tian, J., Cui, S., and Reinartz, P., 2014a. Building change detection based on satellite stereo imagery and digital surface models. *IEEE Trans. Geosci. Remote Sens.*, 52 (1), 406–417.

Tian, J., Dezert, J., and Reinartz, P., 2015a. Refined building change detection in satellite stereo imagery based on belief functions and reliabilities. *In*: *2015 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 160–165.

Tian, J., Nielsen, A.A., and Reinartz, P., 2014b. Improving change detection in forest areas based on stereo panchromatic imagery using kernel MNF. *IEEE Trans. Geosci. Remote Sens.*, 52 (11), 7130–7139.

Tian, J., Reinartz, P., and Dezert, J., 2015b. Building change detection in satellite stereo imagery based on belief functions. *In*: *2015 Joint Urban Remote Sensing Event (JURSE)*. IEEE, 1–4.