



Sub-Graph Matching-Based Building Changes Detection Using High-Resolution Remote Sensing Images

FENG-HUA HUANG¹⁻⁵, ZHENG-YUAN MAO³⁻⁵ AND WEN-ZAO SHI³⁻⁵

¹Postdoctoral Programme of Electronic Science and Technology, Fuzhou University, Fuzhou 350116, China

²Yango College, Fuzhou 350015, China

³Key Lab of Spatial Data Mining & Information Sharing of Ministry of Education, Fuzhou University, Fuzhou 350002, China

⁴National Engineering Research Centre of Geospatial Information Technology, Fuzhou University, Fuzhou 350002, China

⁵Spatial Information Engineering Research Centre of Fujian Province, Fuzhou University, Fuzhou 350002, China

Email: fenghuait@sina.com

Abstract: High-resolution remote sensing (HRRS) images of urban regions have large viewing angle variations, significant noise jamming, and obvious building shadows. Hence, deviation and distortion usually occur to the buildings in HRRS images collected at different phases (multi-temporal HRRS images). The traditional detection methodology is ineffective for accurate and efficient building changes detection in multi-temporal HRRS images. In order to address these problems, this paper proposes a sub-graph matching-based building changes detection (SGMBCD) scheme. First, this paper presents a Graphcut-based buildings extracting (GCBE) method from multi-temporal HRRS images. Next, a sub-graph matching-based registration (SGMR) method is devised to register the previously extracted buildings from multi-temporal HRRS images and to obtain matched ASIFT feature point pairs and singular points. Finally, singular points and overlay analysis-based method (SPOA) is developed to detect building changes in multi-temporal HRRS images. The types of building changes included in this paper are changes (e.g., erection, dismantling, repairing, and reconstruction) and non-changes. In order to demonstrate the effectiveness of the proposed SGMBCD scheme, it is compared with five typical algorithms (i.e., BCDBPM, SDBBCD, BCDBICO, NCUTBCD, and RSMBCD) on three sets of multi-temporal WorldView2 test images. Experimental results show that in comparison with the other methods, SGMBCD can effectively address the challenging problem of building changes detection in multi-temporal HRRS images. The average recall ratio, precision ratio and F value is 91.47%, 86.49% and 88.91% respectively, and the average time consumption is 60.3 s. This demonstrates SGMBCD can detect building changes in multi-temporal HRRS images accurately and efficiently.

Keywords: High-Resolution Remote Sensing Images, Sub-graph Matching, Building Changes Detection, Overlay Analysis.

1. Introduction

Building changes detection via remote sensing images refers to the technique of detecting building changes on the ground using multi-temporal remote sensing data. Building changes include reconstruction and extension due to land use and coverage variation, as well as collapse and damage due to natural disasters [1]. Building changes detection is of great significance in urban planning, GIS data upgrading, smart cities, and military surveillance [2,3]. Currently, the data source that underlies building changes detection is primarily high-resolution remote sensing (HRRS) images. Although HRRS images provide the benefit of building recognition, the images are subject to large angle variations, significant noise jamming, and apparent building shadows. Deviations and distortions usually occur to the same building in multi-temporal HRRS images, which makes it more difficult to detect building changes in HRRS images

than those of other objects on the ground, such as water bodies, vegetation, and roads. Other factors that added to the difficulty of building changes detection include: lack of directly relevant 3D data, possibility of the same object exhibiting different spectrums, diversity in building appearance, and complexity of the scene surrounding buildings in urban areas [1-4]. As a result, traditional methods cannot accurately and efficiently detect building changes in multi-temporal HRRS images. Fruitful studies have been done worldwide on this issue. Typical algorithms include the probabilistic model method [5], shadow analysis method [6], inter-class overlay analysis method [7], graph segmentation method [8], and image matching method [9]. These algorithms have shortcomings in terms of self-adaption, accuracy and efficiency, especially for HRRS images with complex backgrounds. In order to address these problems, this paper proposes a sub-graph matching-based building

changes detection (SGMBCD) method, which provides improved accuracy and efficiency.

2. Sub-Graph Matching-Based Registration (SGMR) Method for HRRS Images

Due to large variations in angle and scale among multi-temporal HRRS images, it has been proven that matching images via scale-invariant feature transform (SIFT) features results in few feature points and inaccuracy [10-13]. Substituting affine-SIFT (ASIFT) features for SIFT features can increase the number of feature points, but it is inefficient and prone to generate an uneven distribution of feature points. ASIFT is more suitable for matching images with large viewpoint variations. The advents of H-SIFT [11], PIF [12] and AIF [13] in recent years have provided effective approaches for accurate and efficient matching of inclined aerial images. But, it is difficult to estimate exterior orientations and determine accurate camera angles for satellite remote sensing images, and the estimation accuracy has a large influence on matching results. Hence, these algorithms are not suitable for matching multi-temporal HRRS images. In this context, this paper proposes a sub-graph matching-based registration (SGMR) method for multi-temporal HRRS images. The steps are illustrated in Figure 1.

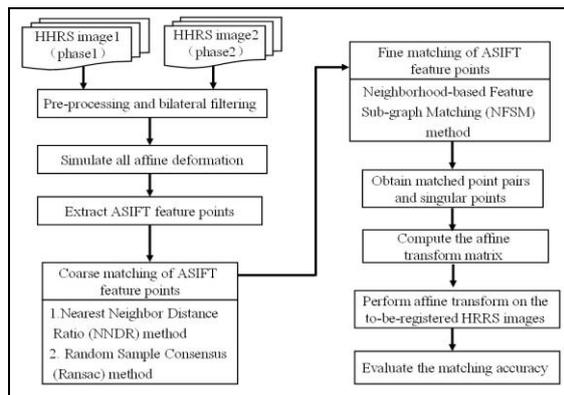


Figure 1: Process of sub-graph matching-based registration of HRRS images

Figure 1 shows that, in SGMR, firstly, the multi-temporal HRRS images are pre-processed and bilateral-filtered, and then they are transformed by simulating all possible affine deformations; Secondly, extract ASIFT feature points from the multi-temporal HRRS images after processed above, and perform Coarse matching on ASIFT feature points using the Nearest Neighbor Distance Ratio method (NNDR) [14] and the Random Sample Consensus method (Ransac) [15] in order to obtain coarsely matched ASIFT feature point pairs; Thirdly, a Neighborhood-based Feature Sub-graph Matching (NFSM) method is proposed to perform fine matching on the coarsely matched ASIFT feature point pairs. In this manner, we obtain the final ASIFT matched point pairs and singular points of the multi-temporal HRRS images. Finally, we compute the affine transform matrix using

the least squares method with the set of final matched point pairs. The to-be-registered HRRS images are reconstructed using bilinear interpolation to achieve the final registration of multi-temporal HRRS images.

NFSM regards feature points in coarsely matched point pairs as a node in the graph. For each coarsely matched point pair and the corresponding nodes in the graph, NFSM determines correspondence accuracy by checking the similarity between topological relationships of the nodes with other nodes in their respective graphs. Consider two sets of coarsely matched point pairs that consist of the same number of matched points from the reference image and the to-be-matched image, respectively. We construct two graphs, Graph X and Graph Y , based on the neighborhood relationships of feature vectors corresponding to feature points in the two sets. As shown in Figure 2, $i = 1..n$ and $j = 1..n$, where n denotes the number of coarsely matched point pairs. Let matrices $D_X(i,j)$ and $D_Y(i,j)$ denote the length of directed edges from node i to node j in X and Y . Their values are the Mahalanobis distance between the feature vectors of ASIFT feature points i and j . Let matrices $\beta_X(i,j)$ and $\beta_Y(i,j)$ denote the counterclockwise direction angle of directed edges from node i to node j in X and Y . Their values are the difference between the principal direction angles for the two ASIFT feature points i and j . Given the symmetry of the four matrices, it is sufficient to compute half the elements, i.e., $D_X(j,i) = D_X(i,j)$, $D_Y(j,i) = D_Y(i,j)$, $\beta_X(j,i) = 180 + \beta_X(i,j)$, and $\beta_Y(j,i) = 180 + \beta_Y(i,j)$.

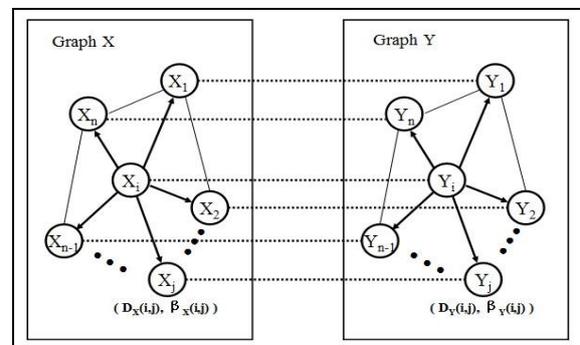


Figure 2: NFSM-based fine matching of ASIFT feature points

The neighborhood features for each node in X and Y can be effectively described using the length vectors (Mahalanobis distance) and direction angle vectors of all edges that start with the node. This description strategy is evidently too strict. Despite it can greatly improve the matching accuracy of feature point pairs, the number of obtained matched point pairs is substantially reduced, which results in a non-uniform distribution of matched point pairs. Consequently, the to-be-matched images suffer significant distortions or deformations in the case of the affine transform. Furthermore, a large value of n means that the matching process is inefficient and computationally intensive. In order to further improve matching

efficiency without compromising matching accuracy, this paper constructs a feature sub-graph, SX or SY , by using the top m ($m < n, i = 1..n$) closest nodes (with the shortest Mahalanobis distance from node i) in X (or Y) and the m directional edges that start with i . The neighborhood of each node in X and Y can be effectively described using the length vectors and direction angle vectors of each directional edge in SX and SY . For any two nodes in X and Y , if their respective sub-graphs SX and SY are matched, then the two nodes are also matched, and vice versa. The steps for selecting feature sub-graph SX from X are given in Figure 3, which is also true for the steps required to generate SY . The value of m has a large influence on fine matching accuracy and efficiency of ASFIT feature points. The optimal value of m is determined through experimentation.

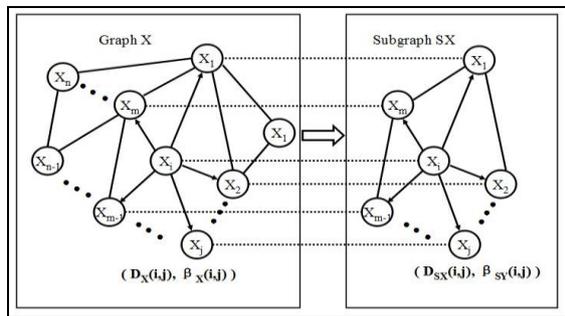


Figure 3: Process of selecting SX from X

Let matrices $D_{SX}(i,j)$ and $D_{SY}(i,j)$ denote the length of directed edges from node i to node j in SX and SY , where $i = 1..m$ and $j = 1..m$. Their values are the Mahalanobis distance between feature vectors of ASIFT feature points i and j . Let matrices $\beta_{SX}(i,j)$ and $\beta_{SY}(i,j)$ denote the counterclockwise direction angles of directed edges from node i to node j in SX and SY . Their values are the difference between the principal direction angles for the two ASIFT feature points i and j . The following procedure shows how to register X and Y using sub-graphs.

First, we compute the two descriptive feature vectors of node i in X , i.e., the length vectors and direction angle vectors of m directional edges in sub-graph SX , $D_{SX}(i)$ and $\beta_{SX}(i)$. Next, we sort the elements in $D_{SX}(i)$ and $\beta_{SX}(i)$ in ascending order to obtain the feature vectors that describe node i , $D'_{SX}(i)$ and $\beta'_{SX}(i)$. Similarly, we can compute feature vectors $D'_{SY}(i)$ and $\beta'_{SY}(i)$ that describe node i in Y . In order to determine whether the i^{th} coarsely matched nodes in X and Y are matched, we compute the difference between the distance feature vectors $\Delta D(i)$ and the difference between the direction feature vectors $\Delta \beta(i)$. $\Delta D(i)$ and $\Delta \beta(i)$ can be computed as shown in Equations (2-1) and (2-2), where $Dot()$ refers to the dot product of vectors.

$$\Delta D(i) = Dot(D'_{SX}(i), D'_{SY}(i)) \quad (2-1)$$

$$\Delta \beta(i) = Dot(\beta'_{SX}(i), \beta'_{SY}(i)) \quad (2-2)$$

In special cases, if the i^{th} nodes in X and Y and all other nodes agree on the directional difference of feature vectors, i.e., $Sum(\beta'_{SX}(i)) = 0$ and $Sum(\beta'_{SY}(i)) = 0$, where $Sum()$ is the 1D matrix summation function, then the i^{th} feature points in X and Y are matched. Ordinarily, false matches that remain in coarsely matched point pairs can be eliminated by thresholds. That is, if $\Delta D(i) > T_D$ and $\Delta \beta(i) > T_\beta$ (T_D and T_β are thresholds), then the i^{th} feature points in X and Y are matched.

3. Graphcut-Based Buildings Extraction (GCBE) from HRRS Images

A Graphcut-based buildings extraction method from HRRS images is proposed in this section. First, we regard each HRRS image pixel as a node in the undirected graph G , and segment HRRS images using Graphcut (max flow/min cut algorithm) to generate a set of objects. Next, we filter the objects by comparing three metrics (e.g., rectangle degree, aspect ratio, and area) to extract buildings effectively. The process of GCBE is given in Figure 4.

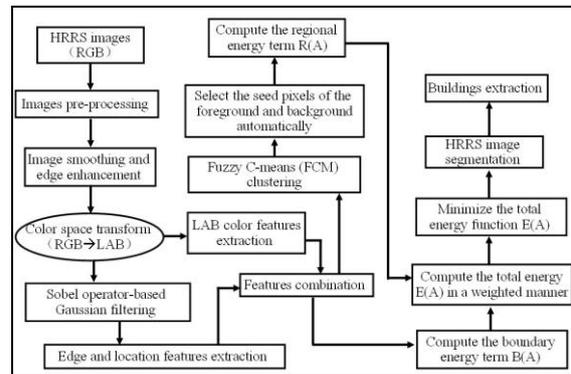


Figure 4: Process of Graphcut-based buildings extraction from HRRS images

First, in order to eliminate calibration error, registration error, and much noise from initial HRRS images, this paper pre-processes initial HRRS images through geometrical and radiation corrections. Next, a bilateral filter[16] that is capable of removing noise while maintaining edges is used to eliminate noise (e.g., burs and small holes) from initial HRRS images and to enhance edges of artificial objects on the ground in urban areas. Afterwards, the RGB color space of HRRS images is converted to LAB color space, which has a wider gamut and more colors. The LAB color model consists of a lightness component, I_L , and two color components (I_A and I_B). A convolution Gaussian filter based on the Sobel operator is used to extract edges from LAB images and to obtain features (e.g., edges and locations) of LAB images. Building roofs in HRRS images have the characteristics high brightness, special colors, regular shapes and clear edges. Therefore, this paper constructs a feature vector $T = [I_L, I_A, I_B, DS, GI]$, where I_L denotes the brightness component in LAB, I_A and I_B denote the two color components in LAB, DS denotes the Euclidean distance between neighboring

pixels, and GI denotes the edge intensity. At first, this feature vector is used to extract feature samples. Later, based on the set of extracted feature samples, we compute a regional energy term $R(F)$ and a boundary energy term $B(F)$ in the overall segmentation cost function of image F , $E(F)$. The two terms are combined in a weighted manner to obtain $E(F)$. Finally, HRRS images are segmented using Graphcut (maximum flow/minimum cut algorithm) to generate the set of objects, and segmented objects are filtered based on the three metrics (rectangle degree, aspect ratio, and area), in order to effectively extract the buildings.

For the Graphcut-based HRRS image segmentation in this paper, the overall segmentation cost function, $E(F)$, is defined as the equation (3-1):

$$E(F) = \lambda R(F) + B(F) \tag{3-1}$$

Where $R(F)$ and $B(F)$ denote the regional energy term and boundary energy term, and the weighting parameter λ denotes the relative importance of the regional energy term and boundary energy term. A large λ means that the regional term accounts for a large proportion in the overall cost (energy) and a small λ means that the boundary term accounts for a large proportion in the overall cost (energy). $R(F)$ and $B(F)$ are computed as follows.

3.1 Calculation of regional energy term R(F)

Unsupervised learning is performed on the extracted feature samples using the effective clustering algorithm fuzzy C-means (FCM) in order to automatically obtain the seed pixels of the foreground and background. By computing the Euclidean distance between other non-label pixels and the two seed pixels (which can be divided into two types: foreground is building and background is other objects on the ground), we can compute the regional energy term of the overall cost function, $R(F)$, is defined as the equation (3-2):

$$R(F) = \sum_{i \in V} R_i(F_i) \tag{3-2}$$

Let $R_i(obj)$ and $R_i(bkg)$ be the possibility that a non-label pixel belongs to the foreground and background, respectively. Then, $R_i(F_i)$ can be computed as follows:

If F_i belongs to the foreground, then $R_i(obj) = 0$ and $R_i(bkg) = \infty$;

If F_i belongs to the background, then $R_i(obj) = \infty$ and $R_i(bkg) = 0$;

If F_i does not belong to the foreground or background, then $R_i(obj)$ and $R_i(bkg)$ are computed as equation (3-3) and equation (3-4):

$$R_i(obj) = D_{obj} / (D_{obj} + D_{bkg}) \tag{3-3}$$

$$R_i(bkg) = D_{bkg} / (D_{obj} + D_{bkg}) \tag{3-4}$$

Where D_{obj} and D_{bkg} denote the Euclidean distance between the center v_i of F_i and the clustering centers of

the foreground and background, v_{obj} and v_{bkg} . That is, $D_{obj} = |v_i - v_{obj}|$ and $D_{bkg} = |v_i - v_{bkg}|$.

3.2 Calculation of boundary energy term B(F)

The boundary energy term $B(F)$ is computed as equation (3-5):

$$B(F) = \sum_{(i,j) \in E} B(i,j) |F_i - F_j| \tag{3-5}$$

This equation implies that $B(F)$ mainly depends on the weight of the edge connecting pixels i and j , $B(i,j)$. In this paper, $B(i,j)$ is computed as the weighted combination of the brightness weight $B_l(i,j)$, color weight $B_{ab}(i,j)$, pixel-spacing weight $B_d(i,j)$, and edge intensity weight $B_e(i,j)$. Thus, we let $B(i,j)$ be computed as equation (3-6):

$$B(i,j) = \sqrt[3]{B_l(i,j) \times B_{ab}(i,j) \times B_d(i,j) + \alpha B_e(i,j) + \beta B_d(i,j)} \tag{3-6}$$

Based on Equations (3-5) and (3-6), the boundary energy term $B(F)$ can be computed as equation (3-7):

$$B(F) = \sum_{(i,j) \in E} |F_i - F_j| (\sqrt[3]{B_l(i,j) \times B_{ab}(i,j) \times B_d(i,j) + \alpha B_e(i,j) + \beta B_d(i,j)}) \tag{3-7}$$

Where α and β denote the weighted coefficients of edge intensity and pixel spacing. The brightness weight $B_l(i,j)$, color weight $B_{ab}(i,j)$, pixel-spacing weight $B_d(i,j)$, and edge intensity weight $B_e(i,j)$ are defined as follows:

(1) Brightness weight, $B_l(i,j)$:

Brightness weight is defined as equation (3-8):

$$B_l(i,j) = \exp\left(-\frac{(L_i - L_j)^2}{\sigma_l^2}\right) \tag{3-8}$$

Where L_i and L_j denote the LAB brightness components of pixels i and j , respectively, and σ_l denotes the variance of the global brightness. If the brightness values of the two pixels are close to one another and B_l is large, then there is a high probability that the two pixels belong to the same type.

(2) Color weight, $B_{ab}(i,j)$:

Color weight is defined as equation (3-9):

$$B_{ab}(i,j) = \exp\left(-\frac{\sqrt{((FA_i - FA_j)^2 + (FB_i - FB_j)^2)}}{\sigma_{ab}^2}\right) \tag{3-9}$$

where FA_i and FA_j denote components A of pixels i and j in the LAB color space, respectively; FB_i and FB_j denote components B of pixels i and j in the LAB color space, respectively; and σ_{ab} denotes the average global variance of components A and B in the LAB color space. If the color space of the two pixels are close to one another and B_{ab} is large, then there is a high probability that the two pixels belong to the same type.

(3) Pixel-spacing weight, $B_d(i,j)$:

Pixel-spacing weight is defined as equation (3-10):

$$B_d(i, j) = \exp\left(-\frac{\|C_i - C_j\|^2}{\sigma_d^2}\right) \quad (3-10)$$

Where C_i and C_j denote the locations of pixels i and j , and σ_d denotes the variance of the global distance between pixels. If the distance between the two pixels is short and B_d is large, then there is a high probability that the two pixels belong to the same type.

(4) Edge intensity weight, $B_e(i, j)$

Edge intensity weight is defined as equation (3-11):

$$B_e(i, j) = \exp\left(-\max_{x \in \text{line}(i, j)} \left(\frac{\|SBL(x)\|^2}{\sigma_e^2}\right)\right) \quad (3-11)$$

Where $\text{line}(i, j)$ denotes the line connecting pixels i and j , $SBL(x)$ denotes the edge intensity of pixel x in $\text{line}(i, j)$ which intersects with the edge, and σ_e denotes the variance of the global edge intensity. If the line connecting pixels i and j does not intersect with the edge, then a large value of $B_e(i, j)$ means there is a high probability that the two pixels belong to the same type.

4. Sub-Graph Matching-Based Building changes detection (SGMBCD) Method

Due to the variation of camera angles in multi-temporal HRRS images, rotations, deviations and distortions may occur to the same buildings, which adds to the difficulty of building changes detection. This paper proposes a sub-graph matching-based building changes detection method using the process given in Figure 5.

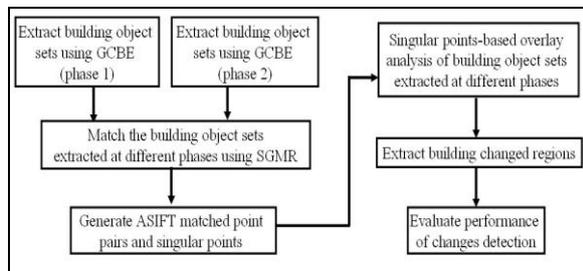


Figure 5: Process of sub-graph matching-based building changes detection

The above figure shows that GCBE is used first to extract buildings from multi-temporal HRRS images and to obtain the set of building objects. Next, SGMR is used to register the set of building objects and to generate ASIFT matched point pairs and singular points. Finally, overlay analysis is performed on the matched buildings, ASIFT matched point pairs and singular points in order to obtain matched regions and singular regions. SGMBCD only registers sets of buildings in multi-temporal HRRS images, which results in improved matching efficiency and lower interference from other objects on the ground. In this paper, the types of building changes include changes (e.g., erection, dismantling, repair, and reconstruction)

and non-changes. Singular regions of buildings belong to changed areas, and these changes happen due to erection or dismantling of buildings. Matched regions of buildings are likely to be repaired, reconstructed, or unchanged regions. This paper determines if the union of two matched regions is changed or not by checking whether $\Delta EG = \Delta E + \Delta \bar{G}$ is less than the threshold T_{EG} , where $\Delta \bar{G} = |\bar{G}_2 - \bar{G}_1|$ is the absolute value of the average gradient difference of building objects in matched regions, \bar{G}_1 and \bar{G}_2 denote the average gradient of matched building objects in two phases, $\Delta E = |E_2 - E_1|$ is the absolute value of the entropy difference, and E_1 and E_2 denote the entropy of matched building objects at different phases. If $\Delta EG \geq T_{EG}$, the union of two matched regions is a changed region; otherwise, the union is an unchanged region. For matched building objects in this paper, the average gradient \bar{G} and entropy E are defined as equation (4-1) and equation (4-2):

$$\bar{G} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n \sqrt{\Delta x f(i, j)^2 + \Delta y f(i, j)^2} \quad (4-1)$$

$$E = -\sum_{i=1}^N P(c_i) \ln P(c_i) \quad (4-2)$$

Where $\Delta x f(i, j)$ and $\Delta y f(i, j)$ denote the first-order difference of pixel (i, j) along the x and y directions, m and n denote the number of rows and columns of relevant images, N is the number of pixel gray levels in relevant images, and $P(c_i)$ is the probability of the i^{th} level pixel gray value in relevant images.

This paper evaluates the performance of building changes detection method in terms of precision (P), recall (R), F-score (F) and time consumption (T). Let $N1$ denote the number of changed buildings correctly detected in the experimental zone, $N2$ denote the total number of changed buildings detected in the experimental zone, and $N3$ denote the actual number of changed buildings existing in the experimental zone. Then, P , R , and F are defined as equation (4-3), equation (4-4) and equation (4-5):

$$P = N1 / N2 \times 100\% \quad (4-3)$$

$$R = N1 / N3 \times 100\% \quad (4-4)$$

$$F = \frac{2 \times P \times R}{P + R} \times 100\% \quad (4-5)$$

F is computed as the harmonic mean of P and R , so it combines two metrics to represent overall algorithm performance. F can alleviate the possible conflict between P and R . The higher the value of F , the better the algorithm [17]. In addition to detection accuracy, T is an important metric that describes the efficiency of building changes detection.

5. Experimental Results and Analysis

Experimental data in this paper comes from RGB color images (WorldView2) of Shenzhen collected at a resolution of 0.5 m in two phases (November 2012 as phase 1 and August 2013 as phase 2). Due to large camera angle variations, objects in the two images on the ground (especially the buildings) have significant deviation and distortion. In order to perform detection of building changes, three sets of representative sub-images (i.e., A1-B1, A2-B2 and A3-B3) of the experimental zone (totaling 6 subimages) are selected from the images at phases 1 and 2. Their sizes are 1796×1721, 1870×1646, and 1796×1646, respectively, as shown in Figure 6.

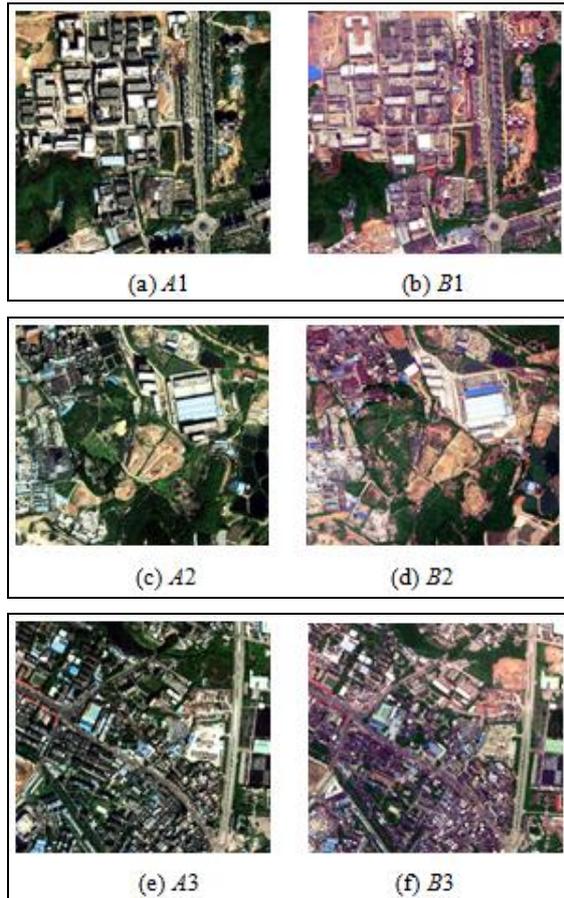


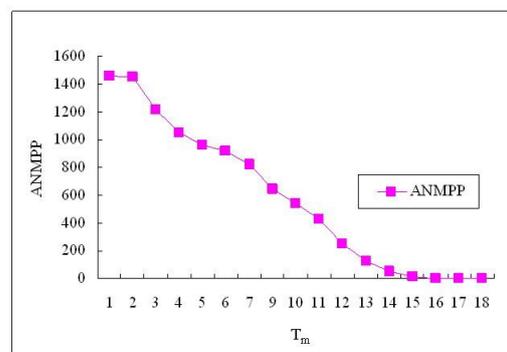
Figure 6: Set of test images

Figure 6 shows that A1-B1 consists mainly of buildings, nudation, and vegetation. The building shapes are regular (approximately rectangular) and the spacing between buildings is small. But some buildings are aligned in a row, their shapes are irregular, and the rooftop color and brightness are distinct. Furthermore, shadows and side walls of buildings, cement ground and nudation are important sources of noise interference. In addition to the features in A1-B1, A2-B2 has large water bodies, building sizes are more distinct, roads are more complex, and nudation is larger, making it more difficult to detect building changes. Unlike A1-B1 and A2-B2, A3-B3 consists mainly of buildings, nudation, vegetation and few water bodies. Building heights in

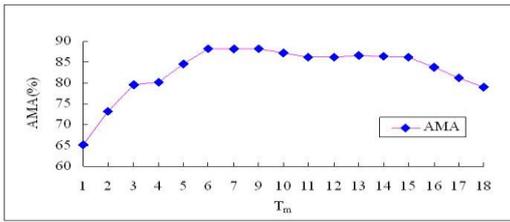
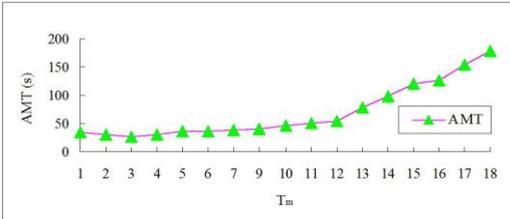
A3-B3 are low, shadows and side walls of buildings are not obvious, buildings are densely distributed with similar dimensions, but color and brightness at the top of buildings are distinct. The influence of roads and nudation in A3-B3 on detection is not negligible. Detection of building changes in A3-B3 is challenging. Due to the difference in camera angle, interference from shadows in A1 and A2 is obvious, and some buildings are mixed with shadows. But images B1 and B2 contain much information on the side walls of buildings. These factors add to detection difficulty.

5.1 Parameter analysis and settings

Based on the above discussions, the proposed SGMBCD has four thresholding parameters, i.e., a threshold for the number of neighbors in NSFM, T_m , a threshold for the difference $\Delta D(i)$ between the distance feature vectors, T_D , a threshold for the difference $\Delta\beta(i)$ between the direction feature vectors, T_β , and a threshold for changes of union matched regions of buildings, T_{EG} . The settings for T_D , T_β , and T_{EG} influence mainly on the accuracy of SGMR. Hence, the optimal values of these three parameters can be determined through OTSU (Maximization of interclass variance) algorithm; the optimal values obtained using OTSU are $T_D=0.3$, $T_\beta=0.4$ and $T_{EG}=5.5$. The setting of T_m influences the matching accuracy of SGMR, and also has a large impact on the matching efficiency of SGMR. Thus, the value of T_m should be determined experimentally. In what follows, experiments are performed using SGMR on images A1-B1, A2-B2 and A3-B3. Figure 7 shows the curves that describe the influence of T_m on the average number of matched point pairs (ANMPP), average matched accuracy (AMA), and average matched time (AMT). Figure 7 indicates that, with an increase in T_m , ANMPP gradually decreases and AMT gradually increases. But, AMA first increases, peaks at $T_m = 6$, and then begins to slowly fall. This is because when the number of neighboring nodes T_m increases, the requirements for matching ASIFT feature points are stricter, and more feature points are regarded as substandard and eliminated.



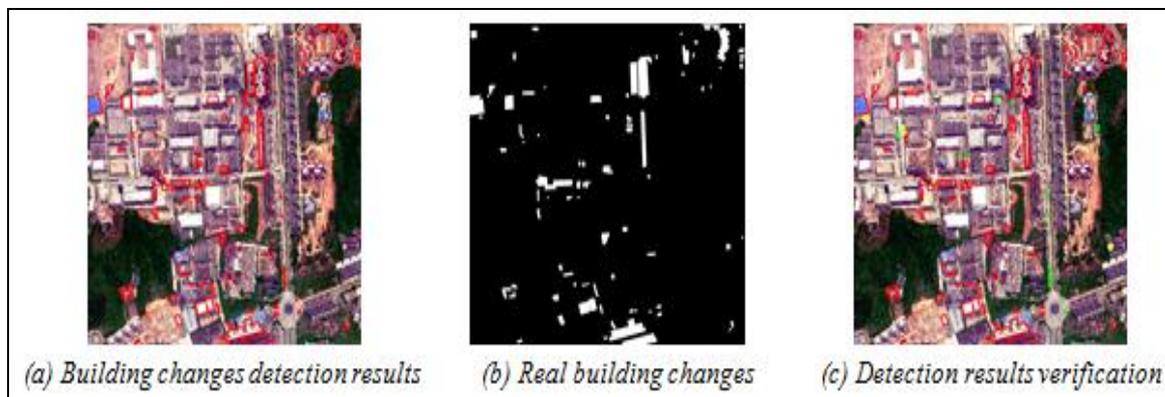
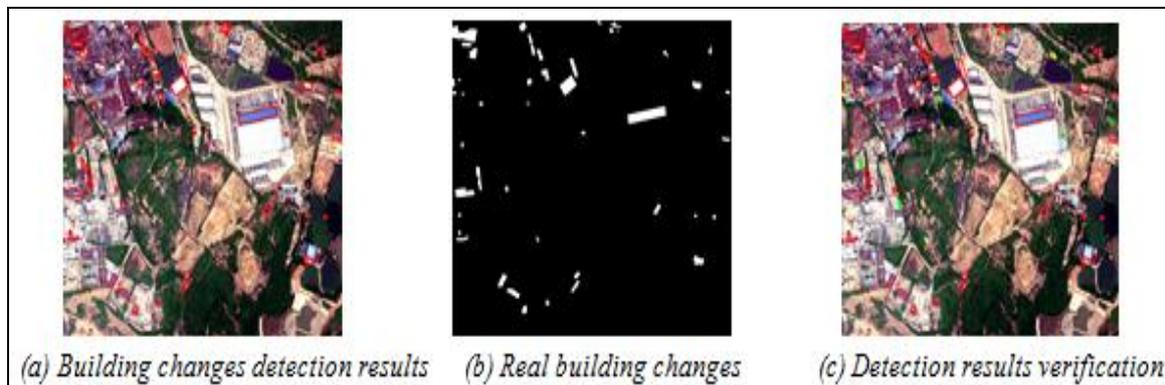
(a) Influence of T_m on ANMPP

(b) Influence of T_m on AMA(c) Influence of T_m on AMT**Figure 7:** Influence of T_m on SGMR performance

Thus, the absolute number of matched point pairs falls on average, but the accuracy of matched point pairs increases substantially. Furthermore, in the case of $T_m = 6$, SGMR achieves the highest matching accuracy, the average number of matched point pairs is large, and the average matching time is short. Hence, the optimal value of T_m is 6 in this paper.

5.2 Analysis of results from sub-graph matching-based building changes detection

Building changes in this paper include changes (e.g., erection, dismantling, repair, and reconstruction) and non-changes. SGMBCD is used to detect building changes in three sets of multi-temporal images: A1-B1, A2-B2, and A3-B3. Detection results are shown in Figures 8-10. These figures indicate that, although the three multi-temporal test images are complex to a varying degree and there are some misdetections and omissions, most building changes are detected accurately by SGMBCD. The misdetected objects on the ground in the three sets of images are small in size and similar to the buildings in terms of shape, brightness, and color. In addition, in the HRRS images collected at different phases with varying camera angles, the locations or shapes of the misdetected objects (e.g., large vehicles, sides of buildings, nudation with regular shapes, and small temporary buildings) are prone to change, making it possible to mistake them as building changes. Moreover, the same buildings in multi-temporal HRRS images are significantly and differently interfered by surrounding objects on the ground, and this may also cause misdetection of building changes. For omitted buildings, some fail to be detected due to the buildings themselves, but the major reason is that the rooftop changes slightly after reconstruction, and that there are similar surface objects near the newly erected or dismantled buildings.

**Figure 8:** Analysis of building changes detection results (A1-B1)**Figure 9:** Analysis of building changes detection results (A2-B2)

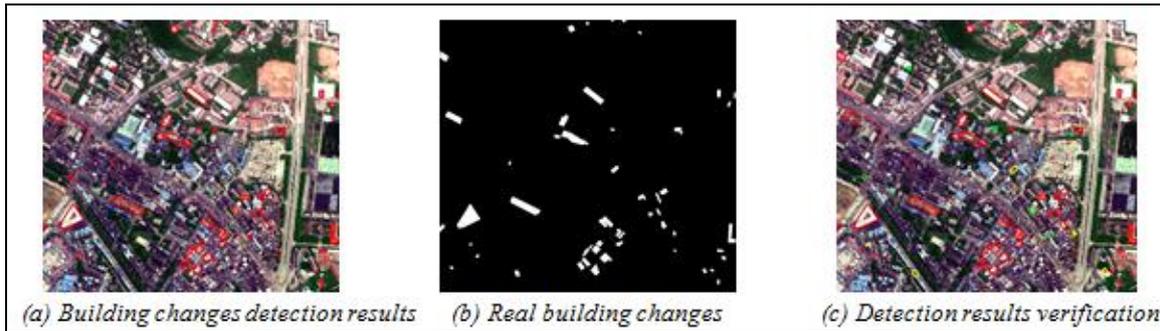


Figure 10: Analysis of building changes detection results (A3-B3)

These factors add difficulty to the detection of building changes. When SGMBCD is used to detect building changes in A1-B1, A2-B2, and A3-B3, the

obtained performance parameters are given in Table 1.

Table 1: Analysis of experimental results using SGMBCD

Test images	N1	N2	N3	Omissions	Misdetections	Recall R (%)	Precision P (%)	F (%)	Time T(s)
A1-B1	75	85	80	10	5	93.75	88.24	90.91	69.5
A2-B2	38	45	42	7	4	90.48	84.44	87.36	51.2
A3-B3	46	53	51	7	5	90.20	86.79	88.46	60.2
Average	53	61	58	8	5	91.47	86.49	88.91	60.3

Table 1 shows that, when building changes detection in A1-B1, A2-B2 and A3-B3 using SGMBCD, the average recall and precision ratios are 91.47% and 86.49%, the average value of F is 88.91%, and the average time consumption is 60.3 s. This demonstrates the ability of SGMBCD to detect building changes in multi-temporal HRRS images accurately and efficiently. Recall and precision ratios for A1-B1 are higher than the other two image sets. This is mainly due to more interference around the buildings in A2-B2 and A3-B3. Compared with A1-B1, A2-B2 consists of more distinct building dimensions, more complicated roads and larger nudation. In A3-B3, although buildings are relatively low and shadow interference is unobvious, the buildings are densely concentrated, the color and brightness at the top of buildings change greatly, and roads and nudation have a large influence on detection. Hence, it is challenging to detect building changes in A2-B2 and A3-B3.

5.3 Comparative analysis of experimental results

In order to evaluate the performance of the proposed SGMBCD, the method is compared with the following typical algorithms of the same type: probabilistic model-based method (BCDBPM) by W.Liu, et al. in literature[5], shadow detection-based method (SDBBCD) by S.P.Ji and X.X.Yuan in literature[6], inter-class overlay analysis-based method (BCDBICO) by Z.S.Wang in literature[7], graph segmentation-based method (NCUTBCD) by W.Z.Shi and Z.Y. Mao in literature[8], and image matching-based method (RSMBCD) by X.L.Song and W.M.Li in literature[9]. The comparison is made on the three sets of test images A1-B1, A2-B2 and A3-B3. Results are given in Table 2. Overall performance of the algorithms is evaluated in terms of average precision ratio (P), average recall ratio (R), average value of F , and average time consumption (T).

Table 2: Results of experimental comparison

Detection algorithms	Average recall ratio R (%)	Average precision ratio P (%)	Average value of F (%)	Average time consumption T(s)
SGMBCD	91.47	86.49	88.91	60.3
BCDBPM	85.65	83.32	84.47	53.5
SDBBCD	86.25	82.35	84.25	45.21
BCDBICO	85.75	83.65	84.69	55.45
NCUTBCD	89.85	84.68	87.19	57.6
RSMBCD	87.93	83.61	85.72	59.3

Table 2 shows that, although SGMBCD is slightly more time consuming than the other methods, it is superior in terms of average recall, average precision and average value of F . The accuracy of BCDBPM, SDBBCD, BCDBICO and RSMBCD is less than

SGMBCD. In the case of large parallax and distortion, BCDBPM causes significant errors when extracting building contours using the probabilistic model. The accuracy of SDBBCD is limited when used for buildings with large shadows or that have overlap

with shadows. BCDBICO detects building changes by performing inter-class overlay analysis, but the strategy for extracting buildings is inaccurate and prone to error accumulation. Hence, the accuracy of BCDBICO is limited for building changes detection. Like the proposed SGMBCD, RSMBCD relies on matching algorithms to register multi-temporal HRRS images and then detect building changes. But for multi-temporal HRRS images that have large camera angle differences, building deviations and distortions, the HRRS image matching method that relies on Maximally Stable External Regions (MSER) affine-invariance is inferior to SGMR of SGMBCD (i.e., matching images based on ASIFT feature points) in terms of accuracy. Furthermore, the MSRE region is not definitely within buildings, and RSMBCD is less accurate than SGMBCD. But, RSMBCD does not need to extract buildings, and there are fewer MSRE regions than ASIFT feature points for image matching. Hence, RSMBCD is usually more efficient than the proposed SGMBCD. The accuracy of NCUTBCD is high and next to SGMBCD. This is mainly because building objects extracted from multi-temporal HRRS images are not matched effectively. Significant error exists during overlay analysis of multi-temporal buildings whose locations and contours change greatly, which results in limited accuracy of building changes detection. But, compared with SGMBCD, NCUTBCD does not need to perform automatic matching on building objects, and thus is less time-consuming. Hence, the proposed SGMBCD outperforms BCDBPM, SDBBCD, BCDBICO, NCUTBCD, and RSMBCD overall.

6. Analysis and Discussion

Due to various uncertainties (e.g., camera angle, sensor type, weather, illumination and uncertainty of remote sensing data), buildings are prone to be distorted, inclined and deviated across multi-temporal HRRS images. Due to the interferences from building shadows, side walls, roads, and nudation, building changes detection in HRRS images is complicated and is not yet to be addressed completely. Compared with other methods of the same type, the proposed SGMBCD is more accurate and efficient. But, it also has limitations. First, SGMBCD has four major parameters (thresholds) and these parameters exhibit obvious patterns in influencing detection performance. But this paper determines the optimal values of these parameters manually in the experiment, and has not yet to completely automate the optimization; Second, SGMBCD is inaccurate for buildings that have irregular shapes or whose top surface is broken. Also, it is unable to effectively eliminate interfering surface objects (e.g., vehicles and containers) whose appearance is similar to buildings and whose locations change frequently. Hence, SGMBCD suffers from some misdetections and omissions. This paper has not yet to find a

complete solution, and these two problems above are the focus of future work.

7. Conclusions

In order to detect building changes in multi-temporal HRRS images more accurately and efficiently, this paper proposed a sub-graph matching-based building changes detection (SGMBCD) scheme which is capable of effectively removing the influence of many factors (e.g., camera angle, noise, deviation, distortion, and shadows) on detection. SGMBCD is compared with five other algorithms (i.e., BCDBPM, SDBBCD, BCDBICO, NCUTBCD, and RSMBCD) on three sets of multi-temporal WorldView2 test images. Experimental results show that SGMBCD has higher accuracy and efficiency overall, with average recall and precision ratios of 91.47% and 86.49%, average value of F of 88.91%, and average time consumption of 60.3 s. SGMBCD is also well suited for building changes detection in HRRS images with different resolutions. However, the proposed SGMBCD also has some limitations. For example, SGMBCD has not yet to completely automate the optimization of parameters (thresholds). Also, SGMBCD is inaccurate for buildings whose shapes are irregular, whose top surface is broken, whose appearance is similar to buildings and whose locations change frequently. In future work, we will take these problems into account to find a more accurate, efficient and automatic detection method.

8. Acknowledgements

This work was funded by Program for Outstanding Youth Scientific Research Talents Cultivation in Fujian Province University (2015) and China Postdoctoral Science Foundation Project (2015M571963). The authors would like to thank R.Xu and Q.Weng from Spatial Information Engineering Research Centre of Fujian Province in Fuzhou University for useful assistance, suggestions, and discussions.

References

- [1] C.Tao, *Research on Urban and Building Detection from High Resolution Remotely Sensed Imagery*, Doctoral thesis, Huazhong University of Science and Technology, Wuhan, 2012.
- [2] B.Sirmacek and C.Unsalan, "Urban Area Detection Using Local Feature Points and Spatial Voting", *IEEE Geoscience and Remote Sensing Letters*, Vol.7, No.1, PP.146-150, 2010.
- [3] W.Su, J.Li, Y.H.Chen, Z.G. Liu, J.S.Zhang, T.M.Low, I.Suppiah and S.A.M.Hashim, "Textual and Local Spatial Statistics for the Object-oriented Classification of Urban Areas Using High Resolution Imagery", *International Journal of Remote Sensing*, Vol.29, No.11, PP.3105-3117, 2008.
- [4] B.Sirmacek and C.Unsalan, "Urban-Area and Building Detection Using SIFT Keypoints and

- Graph Theory”, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.47, No.4, PP.1156-1167, 2009.
- [5] W.Liu and V. Prinet, “Building Detection from High-Resolution satellite Image Using Probability Modeling”, *International Geosciences and Remote Sensing Symposium (IGARSS2005)*, PP.3888-3891, Seoul, Korea, 2005.
- [6] S.P.Ji and X.X.Yuan, “A Method for Shadow Detection and Change Detection of Man-made Objects”, *Journal of Remote Sensing*, Vol.11, No.3, PP.323-329, 2007.
- [7] C.L.Liu, *The Technology Study of Buildings Change Detection Based on Multi-temporal Remote Sensing Images*, Master thesis, Harbin Institute of Technology, Harbin, 2014.
- [8] W.Z.Shi and Z.Y. Mao, “The Research on Building Change Detection from High Resolution Remotely Sensed Imagery Based on Graph-cut Segmentation”, *Journal of Geo-Information Science*, Vol.18, No.3, PP.423-432, 2016.
- [9] X.L.Song and W.M.Li, “Local Image Matching towards Building Change Detection in Remote Sensing Images”, *Microcomputer Information*, Vol.24, No.19, PP.270-272, 2008.
- [10] J.M.Morel and G.S.Yu, “ASIFT: A new framework for fully affine invariant image comparison”, *Siam Journal on Imaging Sciences*, Vol.2, No.2, PP.438-469, 2009.
- [11] X.Zhao, Q.Zhu, X.W.Xiao, D.R.Li, B.X.Guo, P.Zhang, H.Hu and Y.L.Ding, “Automatic matching method for aviation oblique images based on homography transformation”, *Journal of Computer Application*, Vol.35, No.6, PP.1720-1725, 2015.
- [12] X.W.Xiao, D.R.Li, B.X.Guo, Y.Jiang, Y.F.Yu and J.C.Liu, “A Robust and Rapid Viewpoint-Invariant Matching Method for Oblique Images”, *Geomatics and Information Science of Wuhan University*, Vol.40, No.6, PP.1-9, 2015.
- [13] X.W.Xiao, B.X.Guo, D.R.Li, X.Zhao, W.S.Jiang, H.Hu and C.S.Zhang, “A Quick and Affine Invariance Matching Method for Oblique Images”, *Acta Geodaetica Et Cartographica Sinica*, Vol.44, No.4, PP. 414-421, 2015.
- [14] J.Cao, W.J.Wang, F.Han and Y.S.Liu, “Study on Local Feature-based Image Target Recognition”, *Computer Engineering*, Vol.36, No.10, PP.203-205, 2010.
- [15] T.W.Qu, B.An and G.L.Chen, “Application of Improved RANSAC Algorithm to Image Registration”, *Journal of Computer Applications*, Vol.30, No.7, PP. 1849-1851, 2010.
- [16] Z.Q.Zhang and W.Y.Wang, “A Modified Bilateral Filtering Algorithm”, *Journal of Image & Graphics*, Vol.14, No.3, PP.443-447, 2009.
- [17] A.O.Ok, “Automated Detection of Buildings from Single VHR Multispectral Images Using Shadow Information and Graph Cuts”, *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol.86, No.12, PP.21-40, 2013.