

## A BUILDING CHANGE DETECTION METHOD BASED ON A SINGLE ALS POINT CLOUD AND A HRS IMAGE

Z. Yang<sup>a,b</sup>, J. X. Chai<sup>a</sup>, Y.S. Zhang<sup>\*b</sup>, S. H. Mei<sup>a</sup>

<sup>a</sup>Hunan Provincial Key Laboratory of Hydropower Development Key Technology, PowerChina Zhongnan Engineering Corporation Limited, Changsha City, China.

<sup>b</sup>School of Geosciences and Info-Physics, Central South University  
#932 South Road Lushan, Changsha, Hunan, China. Email: [zhangys@csu.edu.cn](mailto:zhangys@csu.edu.cn)

### Commission II, WG II/3

**KEY WORDS:** Building change detection, Shadow extraction, GrabCut, Point cloud data, Satellite image

### ABSTRACT:

For common remote sensing image change detection based on different time phases, it is difficult to solve the problem of surface tilt caused by the different shooting angle and time, which makes it difficult to complete the accurate registration. Some scholars utilized three-dimensional data for change detection to avoid registration problem, however, it costs very high for using 3D data to detect changes. Aim at this problem, this paper proposes a method of combining a single phase of ALS (Airborne Laser Scanning) data and HRS (High Resolution Satellite) image. It is composed of the following four steps: (1) Extracting the shadow area of the new-phase optical image, and superimposing it with the classified old-time point cloud data, and then the disappeared building area and the approximate building area are obtained according to the associated relationship between the shadow area and buildings. (2) Determining the unchanged building areas as positive samples, and taking the remaining area and vegetation area after removing the approximate buildings from the image as negative samples, and then GrabCut algorithm is used to segment new-phase HRS images to obtain building areas. (3) Comparing between building areas obtained in the previous step and the registered old-time ALS data to obtain a new building area with noise. (4) Denoising the results of the previous step to obtain the final new building area. Two datasets are used to verify the method. The detection accuracy of the disappeared buildings is over 85%, and the detection accuracy of the newly added buildings is over 70%.

### 1. INTRODUCTION

As a result of urbanization, many buildings are experiencing constructing or demolishing each year. To automatically monitor building change situation is very important for the government. If a field survey is carried out for each building, it will inevitably lead to a undesirable cost on manpower and material resources. Therefore, it is important to automatically complete urban building change detection at a lower cost. Li (2013) referred automation and real-time tracking of change detection was needed. Therefore, the realization of automation for the change detection of buildings has become a hot research topic.

During the past decades, various building change detection methods have been developed. Most scholars directly use the spectrum and texture information of the building itself on remote sensing images for change detection. For example, Tao et al. (2017) proposed a building change detection method based on the Morphological Building Index (MBI); Zhang et al. (2018) proposed a method of combining pixel-level and object-level building change detection for high-resolution remote sensing images; Meng et al. (2008) proposed a building change detection method based on similarity calibration; Yu et al. (2018) proposed a multi-feature building change detection method combining MBI, texture features and edge features. When

facing high spatial resolution images, a building roof shifts in different directions in different images, and the change detection method directly based on image contrast is susceptible to the accuracy of image registration.

Some scholars have combined other data sources for building change detection. José et al. (2013) applied the support vector machine (SVM) classification algorithm to joint satellite and laser datasets for building change detection. Tian et al (2019) fused multi-spectral images and Digital Surface Models (DSM) to reliability conduct building change detection by using a reliability function.

However, in the existing technology, the accuracy of change detection using only the spectrum and texture of the building itself on remote sensing images is generally not high; and when combined with other data, there are problems such as large data volume and complex models.

To solve the above problems, this paper proposes a method of combining the classified old-time point cloud data and the new-time optical image for building change detection. The workflow of the proposed method is shown in Figure 1. Firstly, the shadow area is extracted from the remote sensing image, and then the image is superimposed with the classified and registered point cloud data to obtain the disappeared buildings

\* Corresponding author: Yunsheng Zhang  
([zhangys@csu.edu.cn](mailto:zhangys@csu.edu.cn))

and the unchanged buildings; Secondly, the unchanged buildings are treated as positive samples, and the combination of the remaining area (after removing the approximate buildings) and the vegetation areas are treated as negative samples. After

that, the GrabCut algorithm is used to classify the new-phase remote sensing image. Finally, the unchanged buildings in the classification results are removed to obtain the newly added buildings.

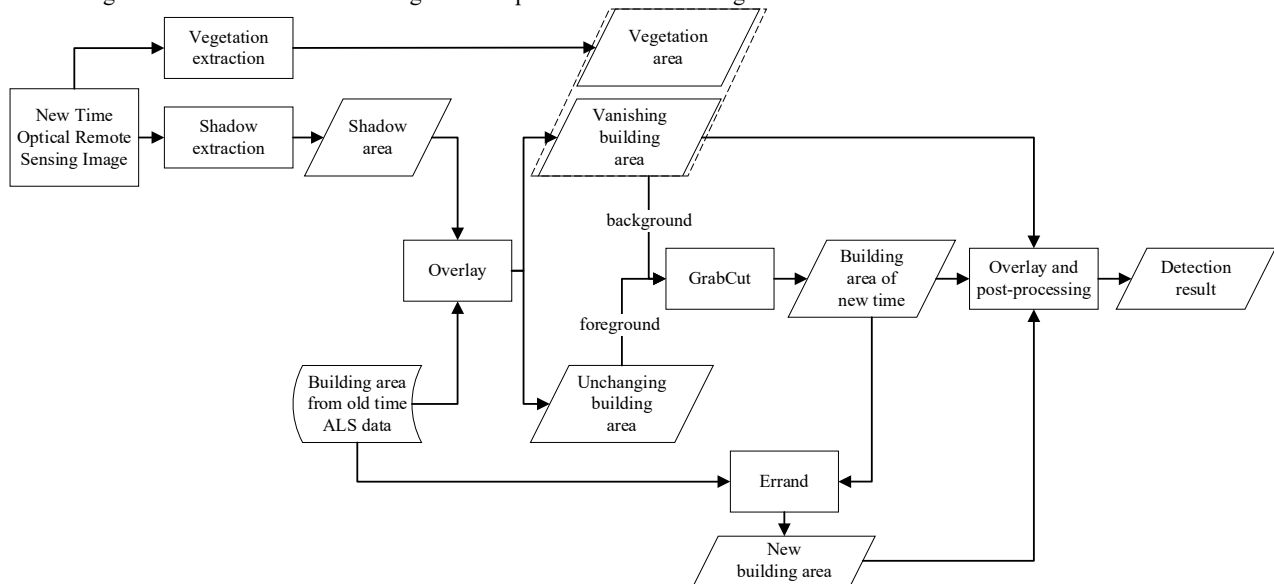


Figure. 1 The workflow of the proposed method

## 2. METHOD

### 2.1 Shadow extraction

Due to the existence of obstructions, the radiation energy of the radiation source (the sun) cannot reach certain areas on the ground. These areas are the shadow areas on remote sensing images, and they usually have an associated relationship with obstructions, for example buildings. Therefore, there will be shadow areas near the building areas, and the positional relationship between the building area and the corresponding shadow area is related to the azimuth of the sun. In view of this observation, this paper uses the positional relationship between buildings and shadow areas to detect the disappeared buildings, at the same time to obtain the approximate building areas. The shadow extraction method includes the following four steps:

In visible remote sensing images, most of the radiant energy is composed of sunlight, and the chromaticity of shadow areas should be the same as it is directly illuminated. Therefore, neither the shadow area nor the high brightness area will be affected by the normalized colour space (Xu et al., 2005). Therefore, we characterize the shadow feature  $f1$  by obtaining the difference between the original colour space and the normalized colour space on the image;

$$r = R/(R + G + B) \quad (1)$$

$$g = G/(R + G + B) \quad (2)$$

$$f1 = \text{mean}(|r - R| + |g - G|) \quad (3)$$

Due to the presence of obstructions, shadow areas show lower brightness on the optical remote sensing image. The brightness can be used as the feature  $f2$ . There are many ways to calculate brightness, such as mean value method, maximum value method, calculation in a HSV space, etc. But these methods do not take the sensitivity of the human eye to the three types of RGB light into account. Therefore, according to the human eye's sensitivity to the three types of RGB light. The following brightness calculation methods are designed for different sensitivities (Liu, 2011).

$$f2 = 0.46 * R + 0.5 * G + 0.04 * B \quad (4)$$

Due to the existence of gaps between leaves in the vegetation area, spot-like shadow areas will be formed in the vegetation area. Therefore, it is necessary to construct the characteristics of the vegetation area to remove these spots to obtain a purer shadow area. In the visible light range, vegetation is generally green, so it can be characterized by feature  $f3$  as calculated by equation (3) (Shen, 2014).

$$f3 = G - \min(R, B) \quad (5)$$

$$f3(f3 < 0) = 0 \quad (6)$$

In the comprehensive decision-making part, after the above analysis, for feature  $f1$ , the shaded area and the highlight area present larger values. For feature  $f2$ , the shaded area and vegetation area present smaller values, and the highlighted area presents larger values. For feature  $f3$ , the vegetation area shows a larger value, and the other parts show a smaller value. Therefore, the following formula is constructed as the final decision item (the three features have been normalized to [0, 1] before calculation), and then the Otsu threshold segmentation is used to obtain a better shadow extraction result.

$$f = \alpha f1 - \beta f2 - \lambda f3 \quad (7)$$

$$\text{Shadow} = \begin{cases} 1 & \text{if } (f > T) \\ 0 & \text{if } (f \leq T) \end{cases} \quad (8)$$

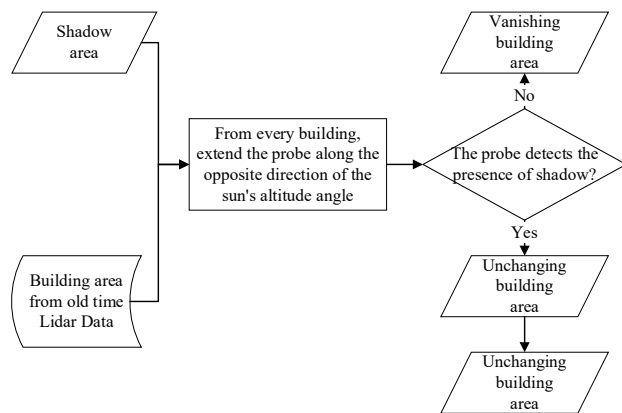
Among them,  $\alpha$ ,  $\beta$ , and  $\lambda$  are the corresponding weights of the three features, which is adjusted by a try-and-error method.

### 2.2 Disappear building extraction and positive sample selection for GrabCut

The disappeared buildings are those that exist on the old-time point cloud data, but do not exist on the new-time remote sensing images. This paper uses the method shown in Figure 2 to extract the disappeared buildings and select positive samples for the subsequent GrabCut algorithm (Rother et al., 2004).

Firstly, the registered point cloud data and the shadow extraction results of the remote sensing image are superimposed, and then an adjustable "probe" (the length of the "probe" in this

paper is 10 pixels) is extended along the direction opposite to the azimuth of the sun for each building. If the probe touches the shadow, the building is considered unchanged and regarded as the positive sample of the GrabCut algorithm. If the probe does not touch the shadow, the building is considered to disappear.



**Figure. 2** Flow chart for the extraction of disappeared buildings and the positive sample selection of GrabCut algorithm

### 2.3 Negative sample selection

In Section 2.1, it is analyzed that buildings and shadows have an accompanying relationship. Therefore, the certain range of the azimuth of the sun in the shadow area is the approximate area of the building. The removal of these approximate building areas from the remote sensing image is a pure negative sample. However, due to the presence of vegetation shadows, these approximate building areas will be mixed with many vegetation areas, which would result in the selected negative samples with little or no vegetation. Therefore, it is necessary to extract part of the vegetation separately as additional negative samples. This paper uses a visible light remote sensing image vegetation index  $f$  calculated by equation (9) to extract vegetation (Meyer and Neto, 2008). After that,  $f$  is compared to a given threshold  $T$  as shown in equation (11) (in order to obtain a pure negative sample,  $T$  is set to a large value) to obtain a purer vegetation area.

$$f = EXG - EXR \quad (9)$$

Where, the  $EXG$  is an extended green index, and  $EXR$  is an extended red index calculated by equation (10).

$$\begin{aligned} EXG &= 2 * G - R - B \\ EXR &= 1.4 * R - G \end{aligned} \quad (10)$$

$$Vegetation = \begin{cases} 1 & \text{if } (f > T) \\ 0 & \text{if } (f \leq T) \end{cases} \quad (11)$$

### 2.4 Extract buildings by GrabCut

After selecting some positive and negative samples for buildings, the GrabCut algorithm is employed to segment the images to classify and obtain buildings. The GrabCut algorithm is an effective image segmentation algorithm for extracting the target foreground from the complex background. It has three improvements on the basis of the GraphCut algorithm (Ding and Zhang, 2012): (1) Gaussian mixture model (hereinafter referred to as GMM) is used to replace the histogram figure; (2) In the process of GMM parameter learning and estimation, it replaces a minimum estimation with an evolvable iterative algorithm to complete the energy minimization and improve the

segmentation accuracy; (3) Reduce interaction through incomplete labelling operate. However, when the algorithm processes larger images, iterative processing will slow down the image segmentation speed. Therefore, super-pixels are used to replace the previous single pixels to build an s-t network graph model. It can not only maintain high segmentation accuracy, but also speed up the segmentation speed.

The GrabCut algorithm based on graph cut theory transforms computer vision problems into pixel labelling problems. The image segmentation problem is a typical binary label combination optimization problem that marks pixels as foreground/background. The basic idea is to determine the label of the pixel by constructing an energy function about the label and obtaining the minimized energy function (Qiu and Wang, 2012). The energy function is shown in equation (12), which is composed of data constraint items and smoothing constraint items.

$$E(g) = E_{data}(g) + E_{smooth}(g) \quad (12)$$

If all pixels form a vertex set  $V$ , the foreground and background are regarded as endpoints  $s$  and  $t$ , and each vertex  $v$  is connected to  $s$  and  $t$  to form an edge, and the two edges are given different weights and summed to form the energy function of the vertex. The sum of the energy functions of all vertices constitutes the data constraint item  $E_{data}(g)$  in the above formula, which indicates the degree of similarity between each pixel and the two endpoints, and is used to determine the possible mark of the pixel; In addition, adjacent vertices also form edges. The sum of the energy functions of these edges is the smoothing constraint item  $E_{smooth}(g)$  in the above formula, which indicates the degree of similarity between adjacent pixels to ensure the smoothness of the result. By minimizing the energy function  $E(g)$ , the vertex set  $V$  is divided into two vertex sets  $S, T (s \in S, t \in T)$  connected to the source point  $s$  and the sink point  $t$  (Boykov and Jolly, 2001).

In this paper, the invariant buildings are used as the foreground, and the remaining part of the image after removing the buildings and the vegetation area are used as the background, and the building change detection is converted into a binary classification problem. We use GrabCut algorithm to extract the building area from the new time optical image, then we remove the unchanged building area, after that we can get the new building area. In this paper, the invariant buildings are used as the foreground samples of the GrabCut algorithm; the vegetation area and the shadow area are used as the background samples of the GrabCut algorithm, and the GrabCut algorithm is implemented using the python language combined with the OpenCV package programming. The algorithm performs five iterations to extract the building area. What's more, to improve efficiency, the optical remote sensing image is segmented into super-pixels firstly by SLIC (Achanta et al., 2012), and then these super-pixels are used as nodes to construct a graph model for segmentation by GrabCut. After segmentation, the building area was obtained.

### 2.5 New building detection based on GrabCut classification results

The results obtained by GrabCut segmentation include newly added buildings and unchanged buildings. As long as the unchanged buildings in the classification result are subtracted, the newly added buildings are obtained, and then a median filtering and small blocks removing process are performed on

the difference result. Finally, a denoising processing is performed to get accurate new buildings.

### 3. EXPERIMENTS AND ANALYSIS

In order to verify the effectiveness of the proposed method, this paper selects old-time ALS data and new-time high-resolution optical remote sensing images of two regions in a certain area, detect building changes, and uses the confusion matrix as the accuracy evaluation criterion. The experimental results show that the detection accuracy of the two experimental regions reaches 89% and 84%, respectively.

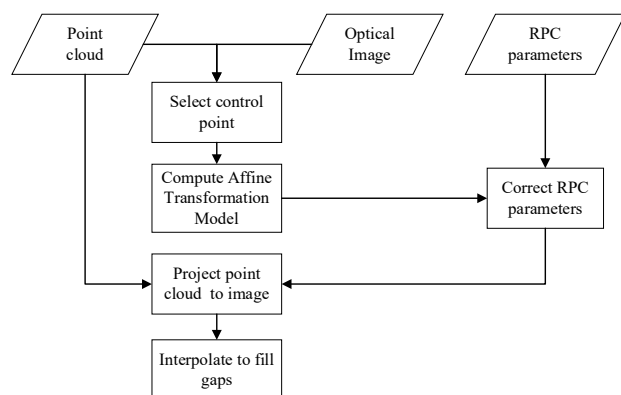
#### 3.1 Experimental data

Two datasets shown in Figure 4 are used for experiments. The details are as follows:

(1) ALS point cloud data: the aerial laser point cloud data was obtained in 2014, the point cloud density is about 10 points/m<sup>2</sup>. As shown in Figure 4 (c) and (d), the building areas are extracted from the point clouds by the method proposed in (Du et al., 2018), the extraction results are shown in Figure 4 (e) and (f). From Figure 4 (e) and (f), we can see that the building extraction results are accurate and can represent the building area of the experimental area in 2014.

(2) Remote sensing image: it was obtained in 2018. It is generated by fusing WorldView3 panchromatic and multi-spectral images. The resolution is about 0.31 meters.

Since the point cloud data obtained in 2014 and the remote sensing images in 2018 have different coordinate systems, it is necessary to perform the registration of these two datasets. The registration process is shown in Figure 3. Firstly, select the control points in the remote sensing image and point cloud data, calculate the affine transformation model between the two datasets. Secondly, correct the RPC parameters of the remote sensing image, and then use the corrected RPC parameters to project the point clouds. Finally, interpolation is performed to fill in some holes to complete the registration of point cloud data and remote sensing images.

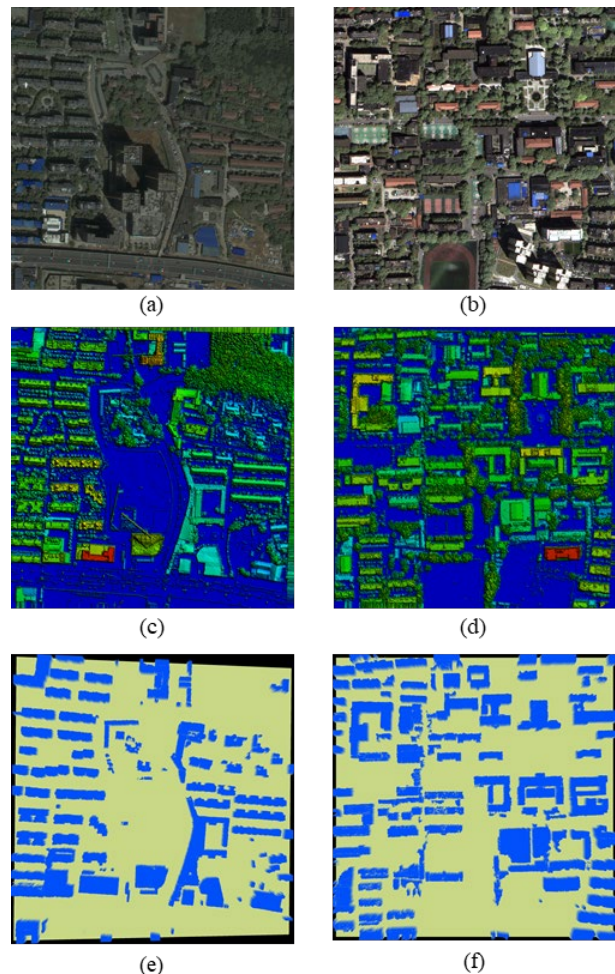


**Figure 3** Registration process of old-time point cloud data and new-time optical remote sensing image

#### 3.2 Evaluation criteria

In this paper, the confusion matrix is used to evaluate the accuracy. Confusion matrix, also known as error matrix, is a standard format for expressing accuracy evaluation, which is represented by a matrix with  $n$  rows and  $n$  columns. Each column of the confusion matrix represents the predicted category, and the total number of each column represents the number of data predicted to belongs to each category; each row

represents the true attribution category of the data, and the total amount of data in each row represents the number of data instances of each category. The specific evaluation indicators include overall accuracy, mapping accuracy and user accuracy. These accuracy indicators reflect the accuracy of image classification from different aspects. Through the confusion matrix, the accuracy of the classification results can be seen very intuitively, so we use it to detect the accuracy of the proposed method.



**Figure 4** Experimental data: (a) and (b) is the remote sensing image in 2018; (c) and (d) is the aerial laser point cloud data in 2014; (e) and (f) is the extracted building areas

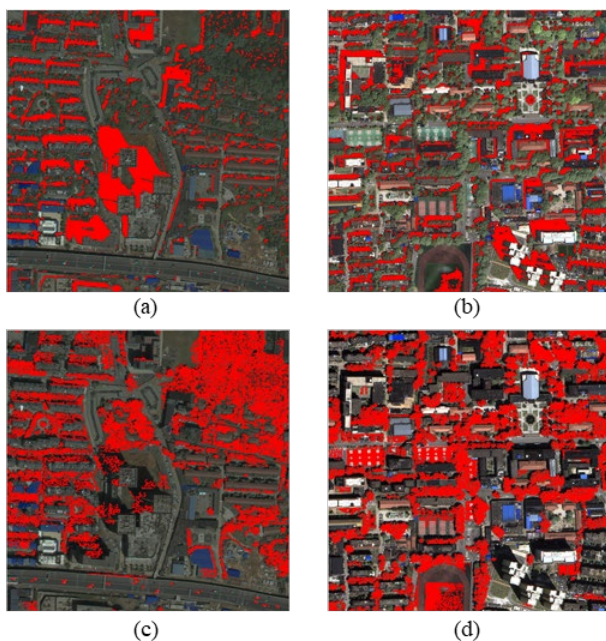
#### 3.3 Experimental results

Firstly, the shadow extraction method was performed on the experimental remote sensing images, the results were shown in Figure 5 (a) and (b), and then superimposed on the building obtained from the 3D ALS data. Then the unchanging building areas and the disappearing building areas were obtained. Secondly, the vegetation area was extracted on remote sensing images of the two experimental. The vegetation extraction results are shown in Figure 5 (c) and (d). After that, the vegetation areas and the disappeared building areas were treated as negative samples, and the unchanged building areas were treated as positive samples, and the GrabCut algorithm is used to segment the whole image, and finally get the newly added building areas.

In order to verify the effect of the proposed method, the real change area is marked by means of human-computer interaction,



as shown in Figure 6 (a) and (c), and the detection result is shown in Figure 6 (b) and (d). It can be seen from the results that the unchanging building areas of the two experimental areas can be accurately detected; for the experimental area I, the detection of the disappearing buildings is more accurate, but there are large new areas near the boundary of the experimental area which was not detected by the method in this paper. For the experimental area II, the detection of the newly added buildings is more accurate, but the detected boundary of each newly added building area is incomplete, and there are vanishing buildings near the boundary were false detected.



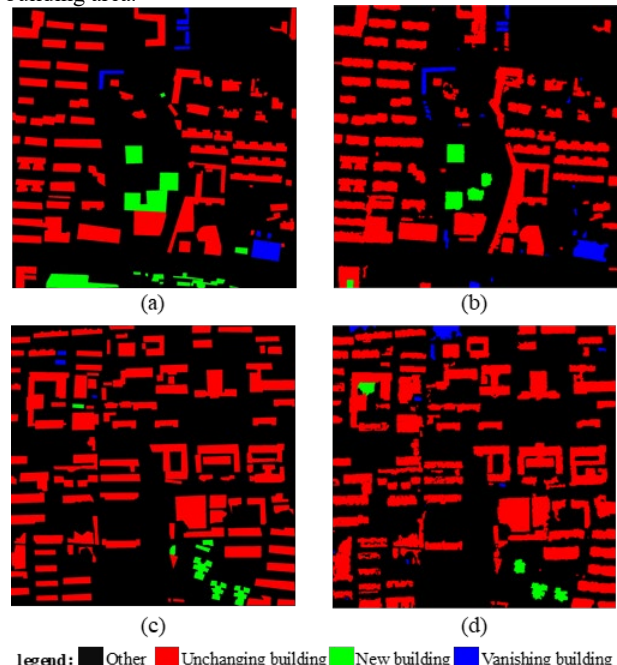
**Figure. 5** Shadow extraction results and vegetation extraction results of the two experimental areas, (a) and (b) are the shadow extraction results of the two experimental areas; (c) and (d) are the vegetation extraction results of the two experimental areas

In order to further verify the effect of the method in this paper, the confusion matrix of the change detection results is

#### 4. CONCLUSIONS

This paper proposes a novel building change detection method that combines a single-phase ALS point cloud and satellite image. It does not require manual selection of training samples, and uses the associated relationship between shadows and buildings to achieve high-precision extraction of disappeared buildings. The building area and non-building area on the new-phase image are acquired, and the building extraction is modelled as a segmentation problem of foreground and background. The GrabCut algorithm is employed to achieve the extraction of buildings in uncertain areas. Experimental results

calculated. The results of the experimental areas (a) and (b) are shown in Table. 1 and Table. 2 respectively. It can be seen from the confusion matrix that the overall accuracy of the change detection in the two experimental areas can reach more than 84%; for the experimental area (a), the detection result of the newly added building area is poor, and there are some missed detection cases, which is consistent with the intuitive observation results from Figure 6; for the experimental area (b), the detection result of the newly added building is also slightly worse due to the incomplete boundary of the detected new building area.



**Figure. 6** Comparison of building change detection results and manual annotations in this paper: (a) and (c) is the manual annotation result; (b) and (d) is the extraction result of the proposed method

show that the proposed method achieves satisfying performance, especially for the extraction of disappeared buildings. The accuracy is up to 88%.

However, there are also some shortcomings. For example, some low-rise buildings have very small shadow areas (only a few pixels), and some buildings are surrounded by vegetation. No shadows can be found on it, and they will be misjudged as disappearing buildings. Therefore, in the future, LBP texture features, MBI index, depth features would be added to improve the classification results.

**Table. 1** Confusion matrix of experimental area (a)

		Real category				User accuracy
		Unchanging building	Vanishing building	New building	Other	
Experiment category	Unchanging building	1892275	149774	3641	18175	92%
	Vanishing building	53512	417842	0	483	89%
	New building	54314	1774	28256	0	34%
	Other	9201	988	0	20148	66%
	Drawing accuracy	94%	73%	89%	52%	
Overall accuracy: 89%			Kappa coefficient: 0.703			

**Table. 2** Confusion matrix of experimental area (b)

		Real category				User accuracy
		Unchanging building	Vanishing building	New building	Other	
Experiment category	Unchanging building	1701610	230412	11439	11955	87%
	Vanishing building	135111	524569	328	4999	79%
	New building	13342	380	13141	0	49%
	Other	1820	1062	0	215	7%
	Drawing accuracy	92%	69%	53%	1%	
Overall accuracy: 84%			Kappa coefficient: 0.624			

### ACKNOWLEDGEMENTS

This work was supported in part by the by the Hunan Province Key Laboratory of Key Technologies for Water Power Resource Development under Grant PKLHD201805, the National Natural Science Foundation of China under Grant 42171440 and Scientific research projects supported by the Department of education of Hunan Province under Grant 19K099.

### REFERENCES

Achanta, R. et al., 2012. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11): 2274-2282.

Boykov, Y.Y. and Jolly, M., 2001. Interactive graph cuts for optimal boundary & region segmentation of objects in ND images, *Proceedings eighth IEEE international conference on computer vision. ICCV 2001. IEEE*, pp. 105-112.

Ding, H. and Zhang, X., 2012. Object abstraction algorithm with fast GrabCut. *Computer Engineering and Design*, 33(4): 1477-1481.

Li, D., 2003. *Change Detection from Remote Sensing Images. Geomatics and Information Science of Wuhan University*, 28.

Liu, J., 2011. *Study on automatic recognition of high resolution optic remotely sensed images based on visual features, Shanghai Jiao Tong University*.

Malpica, J.A., Alonso, M.C., Papi, F., Arozarena, A. and Martínez De Agirre, A., 2013. Change detection of buildings from satellite imagery and lidar data. *International Journal of Remote Sensing*, 34(5): 1652-1675.

Meng, Y., Zhao, Z., Du, X. and Peng, S., 2008. Building change detection based on similarity calibration, 2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery. *IEEE*, pp. 527-531.

Meyer, G.E. and Neto, J.C., 2008. Verification of color vegetation indices for automated crop imaging applications. *Computers and electronics in agriculture*, 63(2): 282-293.

Qiu, Z. and Wang, X., 2012. A ViBe Moving Object Detection Algorithm Combined with LBP. *Journal of China Jiliang University*(3): 250-256.

Rother, C., Kolmogorov, V. and Blake, A., 2004. "GrabCut" interactive foreground extraction using iterated graph cuts. *ACM transactions on graphics (TOG)*, 23(3): 309-314.

Shen, X., 2014. *Object-oriented Building Extraction from High-resolution Remote Sensing Images based on Visual Attention Mechanism, Wuhan University*.

Tao, F. and Shao, Z., 2017. Building change detection based on the enhanced morphological building index. *Science of Surveying and Mapping*, 42(5): 29-34.

Tian, J. and Dezert, J., 2019. Fusion of multispectral imagery and DSMs for building change detection using belief functions and reliabilities. *International Journal of Image and Data Fusion*, 10(1): 1-27.

Xu, L., Landabaso, J.L. and Pardàs, M., 2005. Shadow removal with blob-based morphological reconstruction for error correction, *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. IEEE*, pp. ii/729-ii/732 Vol. 2.

Yu, X., 2018. *A Method for Multi-feature Building Change Detection Based on MBI. Geomatics & Spatial Information Technology*, 11.

Zhang, Z., Zhang, X., Xing, Q. and Yang, X., 2018. Combining the Pixel-based and Object-based Methods for Building Change Detection Using High-resolution Remote Sensing Images. *Acta Geodaetica et Cartographica Sinica*.