A NOVEL BUILDING DETECTION METHOD USING ZY-3 MULTI-ANGLE IMAGERY OVER URBAN AREAS

Huijun Chen[1], Xin Huang[1], Chun Liu[2], Jiayi Li[1], Jianya Gong[1]*

1 School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, P. R. China
2 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, P. R. China

ABSTRACT

This paper presents a new building indicator based on the multi-angle images, the angular difference feature (ADF), which characterizes angular properties from high-resolution ZY-3 multi-view images. The method for detecting buildings based on ADF consists of two main steps: ADF feature extraction and a post-processing step to refine the results by simultaneously incorporating the spectral and geometrical information. Experiments are conducted with three ZY-3 images acquired over Chinese cities. The proposed ADF achieves promising building detection performance over both highly dense urban areas and suburban areas, with an overall accuracy of better than 89% for all the three data sets.

Index Terms—multi-angle, building detection, high spatial resolution, ZY-3

1. INTRODUCTION

The precise detection of buildings is a prerequisite for application of building modeling, urban management and planning, and the estimation of human population. Nowadays, the advanced high-resolution satellite images with multispectral bands enable a more detailed observation of the earth in fine scales, which provide new opportunities for object detection, land-cover/use mapping, and human activity analysis. However, due to the heterogeneous and complex urban scenes, building detection remains a challenging task. It is difficult to separate buildings from other man-made objects (e.g., roads and squares) because of the spectral similarities between different classes. Although structural and textural features have been investigated to facilitate building detection, the complexities of urban scenes (especially in the vertical dimension), shadows and occlusions still seriously decrease the detection accuracy.

To address aforementioned obstacles, some studies have incorporated 3D information for extracting buildings. Light-detection-and-ranging (LiDAR) and multi-angle images are commonly utilized to acquire 3D information. However, LiDAR data are costly and not always accessible in urban areas. Spaceborne multi-angle images with a high resolution are promising sources to collect large-area 3D data with relatively low cost [1].

The ZY-3 01 satellite is China's first civilian high-resolution satellite specifically designed for along-track stereo imagery collection, with a spatial resolution of 2.1 m for the nadir angle and 3.5 m for the forward and backward angles. More recently, ZY-3 02, the second satellite in the ZiYuan-3 series, was launched successfully on May 2016. Compared with the ZY3-01 satellite, the spatial resolution of the forward and backward images of ZY-3 02 was improved from 3.6 m to 2.5 m. The ZY-3 three-line array images are acquired nearly simultaneously. This unique merit makes it particularly suitable for vertical feature extraction of the Earth's surface. Most of the existing studies generated a digital surface model (DSM) from the multi-angle images to retrieve the height information of objects, and hence extract the buildings. However, in such a way, the multi-angle information is not effectively exploited, which is mainly due to the errors and difficulties of the multi-view image matching and the inaccuracy of the generated DSM over complex and dense urban scenes, especially in Chinese cities where buildings with different heights are mixed together [2]. Therefore, in this study, a new building index is designed to detect buildings in complex image scenes, by describing the angular properties based on multi-angle images. Moreover, we investigate the potential for the building detection in complex urban scenes (especially in the vertical dimension) with ZY-3 multi-view images by incorporating the angular difference information.

2. PROPOSED METHOD

A flowchart that describes the overall building detection methodology is presented in Fig. 1. Firstly, the angular difference feature (ADF) is extracted from the ZY-3 panchromatic multi-angle images. Then, initial building detection result is obtained by applying a threshold on the ADF. In the next step of the algorithm, along with the water and vegetation indexes extracted from the nadir pan-sharpened image, the initial building detection results are refined using spectral constraints and shape constraints.
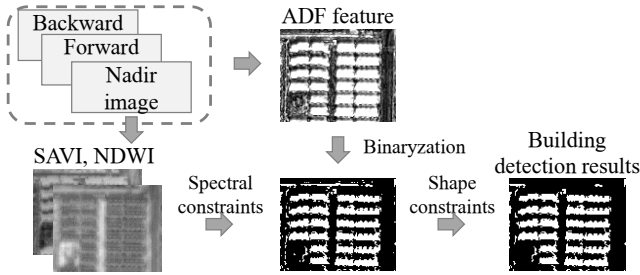
Fig. 1.Methodology Overview.

2.1. Angular difference feature

The multi-view images (e.g., ZY-3 images) provide new opportunities to acquire three-dimensional information, which is particularly useful for building detection in urban scenes with complex vertical structures. It has been demonstrated that buildings show characteristic angular differences in the multi-angle images that can efficiently be used to delineate buildings in urban scenes [3]. Due to the solar observational cross-section, an effect responsible for changes in the reflectance of the objects with non-flat surfaces [4], buildings exhibit strong differences under different acquisition angles, while the low-lying objects such as roads and soil, present relatively consistent structures in different viewing angles. To describe the angular differences, angular difference feature (ADF) is defined as the maximum value from the difference between the stereo images:

$$ADF = \max(X_b - X_n, X_n - X_f, X_b - X_f) \tag{1}$$

where X_f, X_n and X_b are forward, nadir and backward panchromatic images acquired over the same area, respectively. Prior to ADF extraction, the backward and forward imagery are resampled at the same spatial resolution as the nadir imagery. The multi-angle panchromatic images are normalized by histogram matching, taking the nadir images as the reference. ADF is able to highlight the pixels associated with significant angular differences, and thus can be used to identify and delineate the off-ground objects (e.g., buildings) and distinguish them from the low-lying objects such as roads and soil. Then, initial building detection result is obtained by applying a threshold on the ADF. The threshold needs to be determined for each image based on the specific characteristics of the sensor and the scene.

2.2. Post-processing

After obtaining the initial building extraction results, irrelevant objects (e.g., bridges) and small false alarms caused by noises can be retained. Moreover, ADF can be sensitive to other objects with distinct vertical structures such as trees. To address above limitations and classify building pixels more accurately, we use the spectral constraints and shape constraints to refine the initial building detection result.

2.2.1. Spectral constraints

Vegetation and water indexes are derived from the pan-sharpened nadir image to refine the building detection result by incorporating spectral information. Since vegetation with vertical structures (e.g., trees and bushes) can be misidentified as buildings by ADF, therefore, the soil-adjusted vegetation index (SAVI) is used to reduce the confusion between the buildings and vegetation. Moreover, we mask out the water areas using normalized difference water index (NDWI), because some water waves may also cause commission errors in ADF. The rule of spectral constraints can be expressed as:

$$\textbf{IF } \text{SAVI}(x) \geq V_{th} \textbf{ OR } \text{NDWI}(x) \geq W_{th,}$$
$$\textbf{THEN } \text{ADF}(x)=0 \tag{2}$$

where V_{th} and W_{th} denote the vegetation and water index thresholds, respectively. x is a pixel.

2.2.2. Shape constraints

In order to alleviate commission errors, the shape constraint is utilized to remove small and narrow objects (e.g., roads and bridges) by considering the area and length-with ratio of the detected objects. Connected component analysis is first applied on the building extraction results to divide the detected area into separate objects. Subsequently, all objects smaller than A_{th} or have a larger length-with ratio than R_{th} are removed from the result.
The rule of shape constraint can be expressed as:

$$\textbf{IF } \text{Area}(obj) < A_{th} \textbf{ OR } \text{LWR}(obj) > R_{th,}$$
$$\textbf{THEN } obj \text{ is identified as nonbuilding structure} \tag{3}$$

where A_{th} is the threshold of the Area of object obj, and R_{th} is the threshold of its length-with ratio (LWR).

3. EXPERIMENTAL RESULTS

The proposed ADF is compared with three methods: 1) morphological building index (MBI) [5]; 2) enhanced building index (EBI) [6]; and 3) normalized digital surface model (nDSM) [7]. MBI is an accurate indicator that con-siders the characteristics of buildings (brightness, size, contrast, directionality, and shape) by integrating multiscale and multidirectional morphological operators. EBI is an improved version of MBI, with an effective post-processing framework which simultaneously considering the spectral, geometrical, and contextual information. The parameters of the MBI and EBI are set according to [5] and [6]. DSM is calculated from the multi-angle images using the semi-global matching algorithm. Subsequently, nDSM is derived from DSM to represents the actual object height above the earth's surface. The threshold value for nDSM is set to 3 m, which is the height of one building floor in the datasets. To reduce the false alarms caused by trees, nDSM is also refined by the post-processing mentioned above. The parameters of the post-processing steps are set according to the suggestions of

[6]. The evaluation of the proposed method is performed on three ZY-3 image acquired over the urban area of three Chinese cities (see Table 1). The reference data are manually delineated from the corresponding nadir images through field investigation and visual interpretation of the ZY-3 images and Google Earth. We evaluate algorithm performance by overall accuracy (OA), omission error (OE), commission error (CE) and kappa coefficient (Kappa).

Table 1 Data characteristics for all datasets.

Sensor	Date	Coverage (km^2)	Location	Major coverage
ZY-3 01	12/08/2013	3×3	Wuhan	highly dense urban
	15/10/2014	3×3	Shanghai	suburban
ZY-3 02	29/05/2017	3×3	Harbin	Various buildings

3.1. Wuhan data set

The area is highly urbanized, and most of the elements in the scenes are roads and highly compact buildings, except for some lakes and parks. By comparing the MBI, EBI, nDSM and ADF visually [see Fig. 1(a)], ADF shows better performance in detecting the buildings, leading to the highest overall accuracy of 89.8% (Table 2). Due to low local contrast, the MBI and EBI feature can ignore dark roofs since they are more sensitive to detecting buildings with bright roofs and large local contrast. However, the dark roofs are successfully identified by the ADF, since the ADF can capture the angular variation characteristics with the aid of multi-view images. The performance of nDSM building detection strongly depends on the quality of nDSM, related to the inaccurate building shape, position and orientation, or even missing buildings due to the errors of multi-angle matching. Moreover, densely distributed buildings have blob-like shapes in the nDSM, and do not show clear boundaries for each building. The blurry boundaries of buildings derived from ZY-3 DSM are not conducive to the building detection.

3.2. Shanghai data set

The area is located in the outskirts of the main urban area of Shanghai, and is characterized by sparse settlements, mainly small buildings less than 3 floors. The building detection results show that these small buildings can be well extracted using ADF [see Table 2 and Fig. 2(b)]. It can be observed from Fig. 2(b) that there are many false alarms caused by agriculture land in the MBI. Compared to the MBI, EBI shows much less false alarms due to the effectiveness of the post-processing, and the commission error is reduced from 20.3% to 6.9%. However, both MBI and EBI omit quantity of scattered buildings located at the upper side of the image due to the low contrast between detached buildings and the

surrounding natural features. The low height of the buildings (less than 3 floors) in this data set also creates a challenge in detecting buildings using nDSM, resulting in a large omission error of 38.9%. The results demonstrate that ADF is superior in capturing the scattered and small buildings in suburban areas with regards to the compared methods.

3.3. Harbin data set

The Harbin data set was acquired by the recently launched ZY-3 02 satellite. This dataset represents typical urban scenes with a mixture of various man-made structures with heterogeneous heights, sizes and shapes, including roads and different types of buildings. All the four features show low omission errors, however, the commission errors (i.e., false alarms) caused by the low-lying objects (e.g., squares and roads) are obviously smaller in ADF compared with MBI, EBI and nDSM. Most of omission in MBI and EBI refer to heterogeneous and dark roofs, while the commission errors are related to bright soil, open areas and roads since they are also brighter than their surroundings. Due to the improved resolution of backward and forward images of the ZY-3 02, nDSM shows better details and sharper building boundaries. However, compared with ADF, MBI and EBI, it is still difficult for nDSM to retain the building boundaries (see lower-middle part of the Fig. 1c). With the ability to extract buildings with different heights and colors, ADF achieves the highest accuracy.

Table 2 Accuracies of the building detection results for the data sets.

Test set	Method	OE (%)	CE (%)	OA (%)	Kappa
Wuhan	MBI	19.6	12.2	82.3	0.724
	EBI	20.1	7.6	84.6	0.756
	nDSM	16.5	6.6	87.1	0.788
	ADF	11.7	6.3	89.8	0.826
Shanghai	MBI	26.4	20.3	77.0	0.660
	EBI	26.9	6.9	83.5	0.742
	nDSM	38.9	4.2	78.7	0.684
	ADF	11.8	8.5	89.7	0.826
Harbin	MBI	15.9	17.8	86.8	0.777
	EBI	16.6	8.1	90.7	0.831
	nDSM	9.4	27.1	83.4	0.738
	ADF	6.6	4.2	95.8	0.919

4. CONCLUSIONS

In this paper, we introduce a novel angular feature for detecting buildings from the ZY-3 01 and ZY-3 02 multi-view images. ADF feature is first extracted from the multi-view images for characterizing angular properties. Then, post-processing is applied to further refine the building detection results by simultaneously considering the spectral and geometrical information. It is evident from our experiments that the proposed ADF shows good performance with small false alarms in dense urban area.

Moreover, ADF is also suitable for detecting the small and scattered buildings in suburban areas. Preliminary results are very promising and the potential of the proposed ADF to be employed at larger scale will be further assessed in our future research.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1]   G. A. Licciardi, A. Villa, M. D. Mura, L. Bruzzone, J. Chanussot, and J. A. Benediktsson, "Retrieval of the Height of Buildings From WorldView-2 Multi-Angular Imagery Using Attribute Filters and Geometric Invariant Moments," *Sel. Top. Appl. Earth Obs. Remote Sensing, IEEE J.*, vol. 5, no. 1, pp. 71–79, 2012.

[2]   X. Huang, D. Wen, J. Li, and R. Qin, "Multi-level monitoring of subtle urban changes for the megacities of China using high-resolution multi-view satellite imagery," *Remote Sens. Environ.*, vol. 196, pp. 56–75, 2017.

[3]   X. Huang, H. Chen, and J. Gong, "Angular difference feature extraction for urban scene classification using ZY-3 multi-angle high-resolution satellite imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 135, pp. 127–141, 2018.

[4]   N. Longbotham, C. Chaapel, L. Bleiler, C. Padwick, W. J. Emery, and F. Pacifici, "Very high resolution multiangle urban classification analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1155–1170, 2012.

[5]   X. Huang and L. Zhang, "A Multidirectional and Multiscale Morphological Index for Automatic Building Extraction from Multispectral GeoEye-1 Imagery," *Photogramm. Eng. Remote Sensing*, vol. 77, no. 7, pp. 721–732, 2011.

[6]   X. Huang, W. Yuan, J. Li, and L. Zhang, "A New Building Extraction Postprocessing Framework for High-Spatial-Resolution Remote-Sensing Imagery," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 10, no. 2, pp. 654–668, 2017.

[7]   R. Qin, "Rpc Stereo Processor (Rsp) – a Software Package for Digital Surface Model and Orthophoto Generation From Satellite Stereo Imagery," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. III–1, no. July, pp. 77–82, 2016.
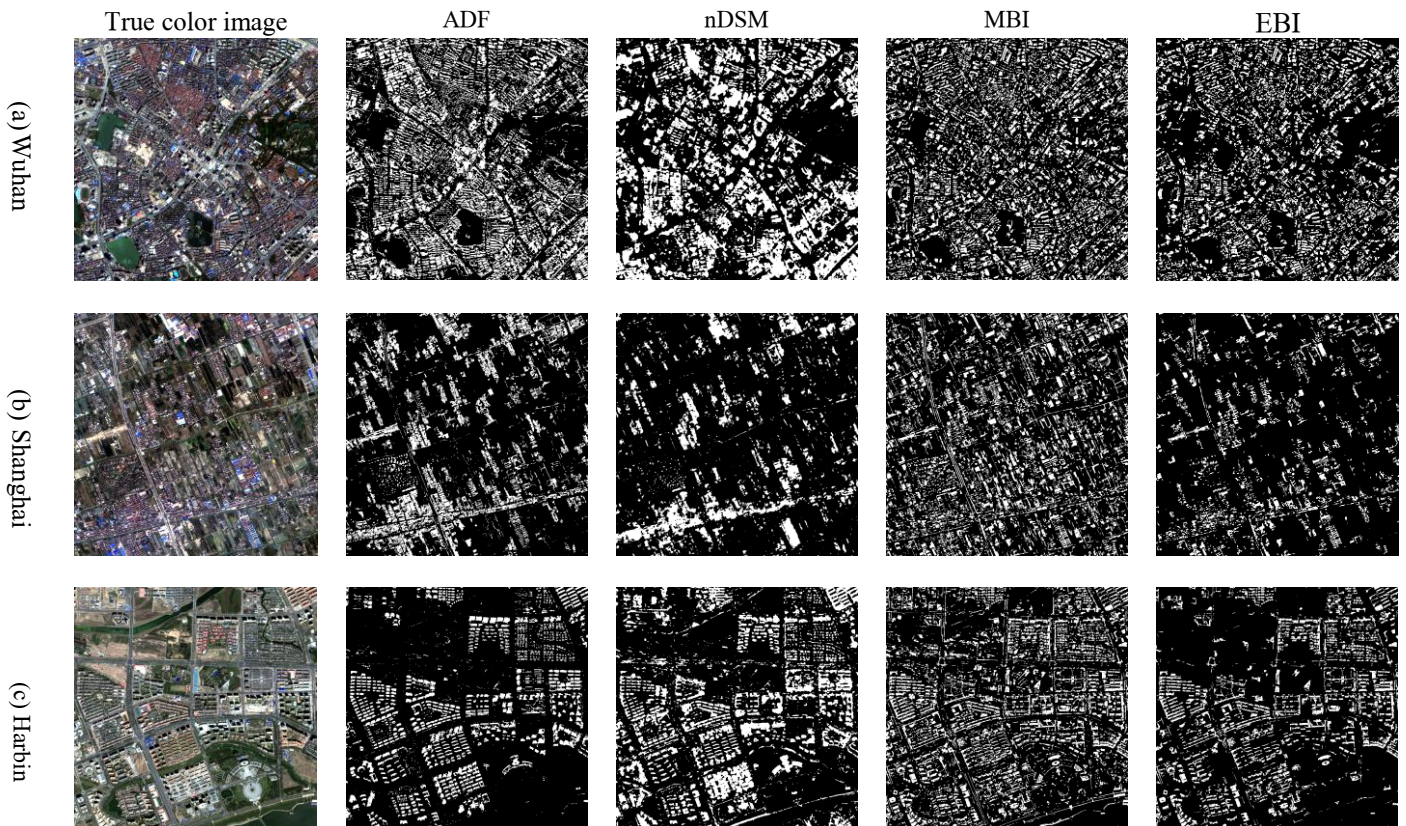
True color image	ADF	nDSM	MBI	EBI
(a) Wuhan				
(b) Shanghai				
(c) Harbin				

Fig. 2. The building detection results for the three datasets.