A New Building Extraction Postprocessing Framework for High-Spatial-Resolution Remote-Sensing Imagery

Xin Huang, *Senior Member, IEEE*, Wenliang Yuan, Jiayi Li, *Member, IEEE*, and Liangpei Zhang, *Senior Member, IEEE*

***Abstract*—In conjunction with the recently developed morphological building index (MBI), the proposed postprocessing framework describes the characteristics of buildings by simultaneously considering the spectral, geometrical, and contextual information, and can be successfully applied to large high–spatial-resolution images. In this way, the proposed framework can alleviate the amount of false alarms to a remarkable extent, which mainly come from the bright soil and vegetation in rural and mountainous areas. Validated on a series of large test images obtained by the widely used commercial satellite sensors, the experiments confirm the promising performance of the proposed framework over various areas, including urban, mountainous, rural, and agricultural areas. Furthermore, the proposed framework increases the quality index by 11% and 9% on average compared to the performance of the original MBI and DMP-SVM, respectively. In addition, the parameter sensitivity is analyzed in detail and appropriate ranges of the parameters are suggested. The proposed building detection framework is designed to be of practical use for building detection from high-resolution imagery.**

***Index Terms*—Building detection, building index, feature extraction, high resolution, mathematical morphology.**

I. Introduction

MONOCULAR building extraction has been an active research topic in photogrammetry and computer vision for many years [1]. The precise detection of buildings is of great importance to urbanization evaluation, urban management and planning, and the estimation of human population [2]. In China, with the rapid development of the economy, extremely dynamic urbanization has resulted in new social and environmental issues that need to be addressed [3]. Nowadays, commercial high-resolution satellite images with multispectral channels can provide more detailed ground information, and, hence, are better suited for object detection, land-cover/use monitoring, and human activity analysis.

Manuscript received February 02, 2016; revised June 13, 2016; accepted June 28, 2016. This work was supported by the National Natural Science Foundation of China 91338111, the China National Science Fund for Excellent Young Scholars 41522110, and the Foundation for the Author of National Excellent Doctoral Dissertation of PR China 201348. *(Corresponding author: Xin Huang.)*

X. Huang and J. Li are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China (e-mail: xhuang@whu.edu.cn; zjjerica@163.com).

W. Yuan and L. Zhang are with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2016.2587324

Although the high-resolution imagery provides more abundant detailed information in the spatial domain, detecting buildings or built-up areas are still a challenging task. The major obstacle is the high intraclass variance and low interclass variance for the spectral statistics of ground features [4], and it is widely agreed that the high spatial resolution does not necessarily facilitate high interpretation accuracy [5]. Previous studies focusing on building detection can be categorized from the following aspects. Supervised machine-learning approaches [6]–[8], which combine spectral information with spatial features, are among the most widely used approaches, e.g., differential morphological profiles (DMPs) [9], wavelet-based textures [10], the pixel shape index [11], and gray-level cooccurrence matrix textures [12]. It is also worth mentioning that such approaches are subject to the trivial selection of training samples and the time-consuming process of machine learning. Meanwhile, researchers have also focused on automatic building detection from high-resolution images, using spectral/spatial priors [13]–[16], marked point process models [17], and semantic-based shadow/building symbiosis [18]–[20]. A number of building feature indices have also been proposed, which can be used to indicate the possible presence of buildings. These indices are effective for supporting the automatic processing of massive high-resolution satellite data. For instance, Pesaresi *et al*. [12] proposed a procedure for the calculation of a texture-derived built-up presence index, namely PanTex, which is based on the assumption that buildings and their surrounding shadow can lead to high local contrast. Several improved versions of PanTex [21], e.g., approaches considering the vegetation components or highlighting the morphological characteristics of built-up structures, have also been proposed for building extraction from high-resolution images [12], [21] and synthetic aperture radar images [22].

More recently, Huang and Zhang [23] proposed the morphological building index (MBI) for automatic building extraction, the fundamental principle of which is to apply the morphological operators (e.g., top-hat by reconstruction, granulometry, and directionality) to describe the implicit characteristics of buildings (e.g., brightness, size, and contrast). Furthermore, an improved version of the original MBI, with the aid of object-based analysis and shadow/building symbiosis (i.e., the morphological shadow index) [24], has been proposed to alleviate the commission and omission errors in urban areas. The effectiveness of the MBI, as well as its variants, in high building density urban areas has been verified by a lot of research, including studies

of urban land-use classification [25] and building change detection [26], [27]. Nevertheless, it should be mentioned that the aforementioned methods based on the MBI have been focused on urban areas and they have not considered the difficulty of building detection in suburban, mountainous, agricultural, and rural areas. Although there are several automatic shadow-based building detection approaches which can be applied to complex scenes [18]–[20], they are limited by the shadow detection performance.

In this study, a new building detection postprocessing framework aligned with the MBI is designed to detect buildings in complex image scenes, not only urban areas. In view of this, the drawbacks of the MBI are investigated in the case of dealing with images of mountainous, agricultural, and rural areas, e.g., the MBI is subject to commission errors when bright soil, open areas, and roads present a brighter intensity than their neighborhoods and show similar spectral characteristics to buildings. This can be attributed to the fact that the traditional MBI does not make full use of the spectral and contextual information to prune the confusing land-cover classes with similar spectral and morphological properties. To deal with such problems, by integrating the morphological, spectral, contextual, and geometrical information in a sequential procedure, the proposed postprocessing framework can efficiently remove these false alarms and extract buildings in more complex and challenging environments. The proposed framework aims to improve the MBI by introducing a set of effective constraints. Specifically, the spectral constraint is first used to eliminate the apparent false alarms caused by bright soil and agricultural fields. Shadow verification functions as a contextual filter aiming to prune the isolated and bright components without surrounding shadow. It is also worth mentioning that the role of the proposed shadow-based contextual constraint is focused on building candidate screening, which is independent of the coarse detection step [18]–[20]. The shape constraint, e.g., the area and length–width ratio, is taken as a screening to remove small errors, e.g., roads and other narrow objects. The proposed framework is efficient and easy to implement, and, thus, has potential for rapid building detection from large-area high-resolution remote-sensing imagery.

The rest of this paper is organized as follows. Section II reviews the original MBI algorithm. The proposed framework is introduced in Section III. Section IV presents the datasets, the experimental results, and a discussion, which is followed by the conclusion in Section V.

II. Morphological Building Index

Using a set of morphological operators, the MBI aims to describe the spatial characteristics of buildings. Specifically, the following key features of buildings are modeled.

1) *Local Contrast*: Buildings often cast shadows, leading to high local contrast. Therefore, the morphological white top-hat (WTH) transform, which can highlight the bright structures with high local contrast, is used to represent the characteristics of buildings.
2) *Size and Directionality*: Unlike roads, which are elongated structures, the sizes of buildings are always within a limited range, and they are relatively isotropic. Therefore, the MBI is constructed based on a set of linear structures, which are used to model the size and directionality of buildings.

The principle of the MBI is shown in flowchart format in Fig. 1, and its calculation is briefly introduced as follows. First, the maximum value of each pixel in the visible bands is recorded as the brightness. The brightness image is focused on the consideration that buildings are brighter than their neighboring pixels [12]. Second, multiscale and multidirectional WTH operators are used to construct the DMPs, which can be viewed as the shape spectra of the objects [6]. The WTH operators are able to describe the local contrast of buildings, and the multiscale and multidirectional structural elements (SE) are used to represent their spatial sizes and isotropy. In this study, the multiscale and multidirectional WTH operators are input into the DMPs to highlight the spatial patterns of buildings in different scales and directions. Finally, the multiscale DMP-WTH profiles are integrated by an averaging operator to emphasize the presence of buildings. The MBI is, therefore, defined as

$$\mathrm{MBI} = \frac{\sum_{s\in S}\sum_{d\in D}\left(\mathrm{DMP}-\mathrm{WTH}(s,d)\right)}{N_s \times N_D} \tag{1}$$

where DMP-WTH is the WTH-based DMPs, d and s represent the direction and scale of the WTH operator, and N_S and N_D are the total number of scales and directions, respectively. Since buildings show a high local contrast in different directions, they always have large MBI feature values.

Although the MBI can show satisfactory results in urban areas, it is subject to a large number of false alarms in nonurban regions, e.g., suburban, agricultural, rural, and mountainous areas, which has prevented its application in large-area building mapping. To address this problem, in this study, by considering additional constraints (e.g., spectral, shadow, and shape) on the original MBI, we propose a new building detection framework for automatic building detection in complicated image scenes and environments.

III. Proposed Framework

The proposed framework sequentially conducts spectral constraints (abbreviated as SPE), shadow verification (abbreviated as SHD), and shape constraints (abbreviated as SHA) on the initial MBI. First, the normalized difference vegetation index (NDVI), the normalized difference water index (NDWI) [28], and the hue component (H) of the images are taken as the spectral prior to refine the original MBI feature. The resulting building map can then be aggregated into a series of connected components, so as to further describe the shape attributes and contextual relationship. Second, shadow verification is used to remove the structures without shadows existing in the illumination direction. Third, shape constraints, including the area and the length–width ratio, can further alleviate small errors, roads, and other narrow objects, resulting in the final detection map of the proposed framework.

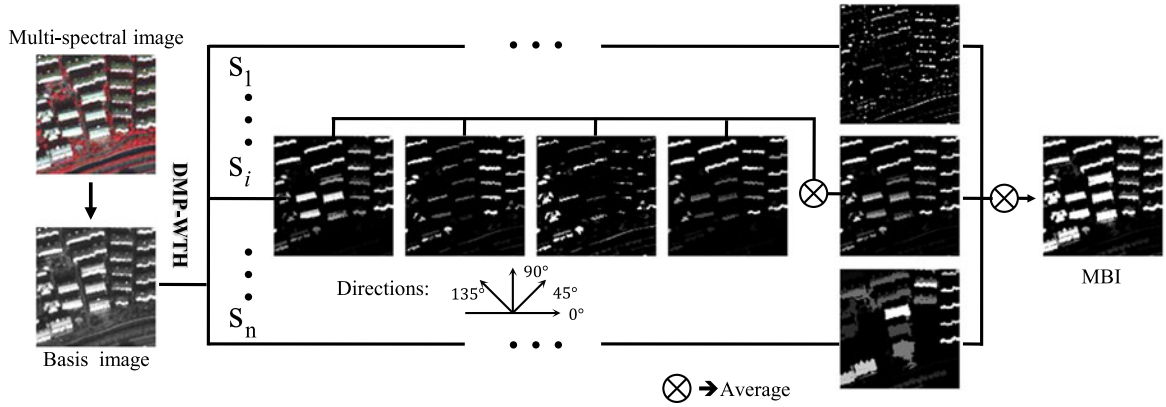


Fig. 1. Flowchart of MBI calculation for multispectral high-resolution remotely sensed imagery.

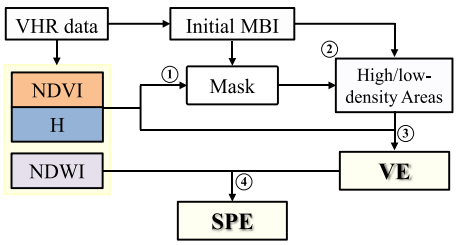


Fig. 2. Flowchart of SPE.

A. *Spectral Constraints*

It is difficult to directly apply spectral constraints to buildings due to their diverse spectral characteristics, but it is possible to impose the spectral constraints on the nonbuilding areas, and, therefore, refine the output of the original MBI, as illustrated in Fig. 2. In this section, we consider the NDVI, the hue component (H), and the NDWI of the high-resolution imagery

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{R}}{\mathrm{NIR} + \mathrm{R}} \tag{2}$$

$$\mathrm{NDWI} = \frac{\mathrm{G} - \mathrm{NIR}}{\mathrm{G} + \mathrm{NIR}} \tag{3}$$

$$\mathrm{H}_1 = \begin{cases} (\mathrm{G}-\mathrm{B})/\mathrm{C} \times 60 & \text{if } \mathrm{R} = \max(\mathrm{R},\mathrm{G},\mathrm{B}) \\ 120 + (\mathrm{B}-\mathrm{R})/\mathrm{C} \times 60 & \text{if } \mathrm{G} = \max(\mathrm{R},\mathrm{G},\mathrm{B}) \\ 240 + (\mathrm{R}-\mathrm{G})/\mathrm{C} \times 60 & \text{if } \mathrm{B} = \max(\mathrm{R},\mathrm{G},\mathrm{B}) \end{cases}$$

$$\textbf{IF } \mathrm{H}_1 < 0, \textbf{THEN } \mathrm{H} \leftarrow 360 + \mathrm{H}_1; \textbf{ELSE } \mathrm{H} \leftarrow \mathrm{H}_1 \tag{4}$$

where NIR, R, G, and B represent the reflectance values in the near-infrared, red, green, and blue channels, respectively, and C is a constant that is equal to $\max(\mathrm{R},\mathrm{G},\mathrm{B}) - \min(\mathrm{R},\mathrm{G},\mathrm{B})$.

1) Vegetation/Soil Elimination (VE): Bright vegetation and soil constitute the main source of commission errors for building detection, and VE is still a big challenge for the current building detectors. In this study, the vegetation index and the HSV (hue, saturation, value) color space are integrated to reduce the false alarms from bright vegetation and soil. In an ideal case, a high built-up density area with a lower probability of vegetation presence should be allocated a higher NDVI threshold, and vice versa. In view of this, a dual-threshold filter for VE can be useful. Meanwhile, some low building density areas will be overestimated by the initial MBI, due to the false alarms. To deal with this problem, a reasonable building density estimation alleviating the overestimation is required before the dual-threshold processing. That is, aiming at pruning the potential false alarms following the density estimation, a rough low NDVI threshold in conjunction with the hue constraint is first utilized to produce a temporary MBI-based mask for high-/low-density area separation (see ①in Fig. 2). As this rough step increases the gap between the high- and low-density areas, the density distribution for every pixel is recorded by locally averaging the temporary MBI-based mask in an $N \times N$ window. Otsu's method [29] is then used to extract the low-density area from the initial MBI (see ②in Fig. 2), by use of the refined density distribution. Finally, a low NDVI threshold (that is, v_{low}) in conjunction with the hue constraint is used to screen the vegetation/soil false alarms in the low building density area; while a high NDVI threshold (that is, v_{high}) in conjunction with the hue constraint is used in the high-density area (see ③in Fig. 2). It is notable that although the first rough NDVI threshold may remove some buildings as well, the goal of the density separation is easy to achieve. For simplicity, we can set the value of the rough NDVI threshold to the same as v_{low}, without degrading the performance. The dual-threshold step can be expressed as

$$\textbf{IF } \mathrm{NDVI}(x_{\mathrm{high}}) \geq v_{\mathrm{high}} \textbf{ AND } h_{\min} < \mathrm{H}(x_{\mathrm{high}}) < h_{\max}, \textbf{ THEN } \mathrm{MBI}(x_{\mathrm{high}}) = 0$$

$$\textbf{IF } \mathrm{NDVI}(x_{\mathrm{low}}) \geq v_{\mathrm{low}} \textbf{ AND } h_{\min} < \mathrm{H}(x_{\mathrm{low}}) < h_{\max}, \textbf{ THEN } \mathrm{MBI}(x_{\mathrm{low}}) = 0$$

where v_{high} and v_{low} denote the vegetation index threshold in the high-density and low-density regions, respectively. $\mathrm{NDVI}(.)$, $\mathrm{H}(.)$, and $\mathrm{MBI}(.)$ denote the specific feature values, respectively. $h_{\min}$ and $h_{\max}$ are fixed values representing the green and yellow components in the HSV space, corresponding to the spectral characteristics of vegetation and soil, respectively. Please note that some buildings (e.g., blue roofs) that show high NDVI values (see Fig. 3(b)) can be incorrectly removed when only considering the NDVI. In this study, therefore, both the NDVI and hue are jointly used to remove the bright vegetation

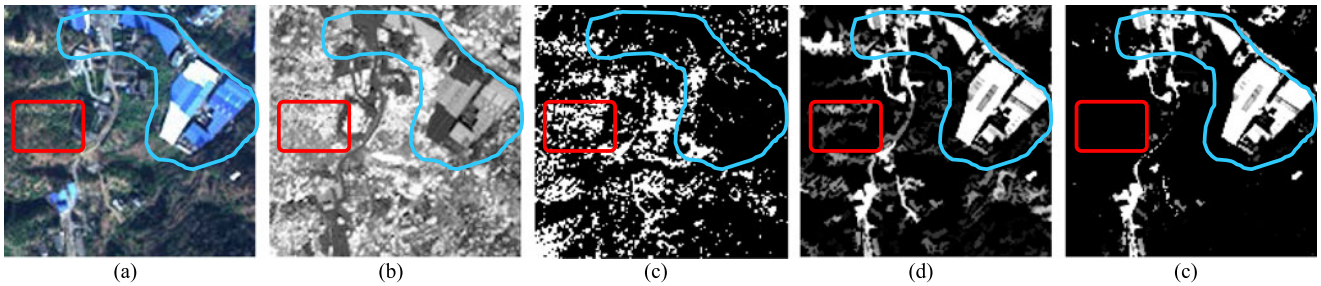


Fig. 3. Example of VE: (a) RGB; (b) NDVI; (c) binary hue image, where the white pixels are within the range of $h_{\min}$ and $h_{\max}$ in the H component; (d) original MBI; and (e) MBI feature after VE.

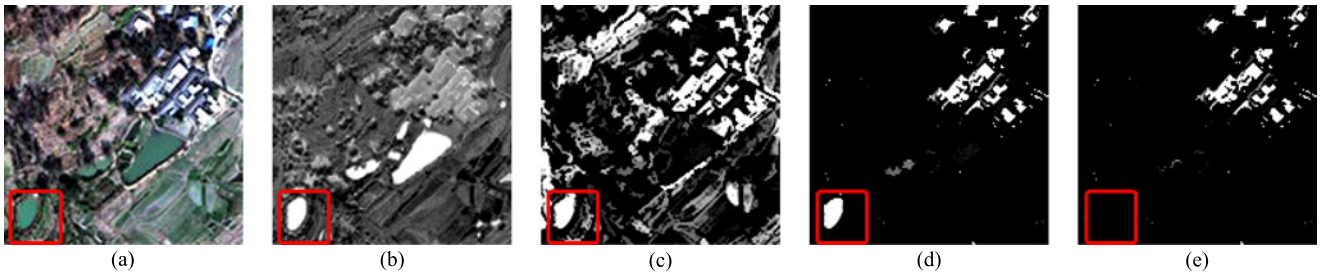


Fig. 4. Example of water body elimination: (a) RGB; (b) NDWI; (c) MBI; (d) MBI with VE; and (e) MBI with SPE constraint (including VE and WE).

and soil. As demonstrated in Fig. 3, by simultaneously considering NDVI and hue, the false alarms caused by the bright vegetation and soil can be effectively removed, and the blue buildings with high NDVI values can also be preserved.

2) Wetland Elimination (WE): As a robust water index, the NDWI can be useful for refining the morphological building detection (see ④in Fig. 2). The rule of WE can be expressed as follows:

$$\mathbf{IF}\ \mathrm{NDWI}\,(x) \geq w, \mathbf{THEN}\ \mathrm{MBI}\,(x) = 0$$

where w and $\mathrm{NDWI}(.)$ denote the water index threshold and the NDWI feature value, respectively. As can be seen in Fig. 4, a water body (left bottom), which shows a high NDVI response due to eutrophication, can be removed with the aid of the NDWI constraint.

After the spectral filtering, we then binarize the modified MBI feature to obtain an initial building map, as an input to the following analysis.

B. Shadow Verification and Shape Constraints

1) Shadow Elimination: Shadow can serve as a constraint for building extraction to remove false alarms, such as soil, roads, parking lots, and open areas. In this paper, we utilize the histogram thresholding approach [30] for shadow detection (see Fig. 5(b)), where the NIR band is considered due to its superior separability for shadow. Subsequently, the resulting shadow feature image is dilated the NIR band is considered due to its superior separability for shadow. Subsequently, the resulting shadow feature image is dilated the NIR band is considered due to its superior separability for shadow. Subsequently, the resulting shadow feature image is dilated using a linear SE in the opposite direction to the solar illumination angle. This constraint is used to test whether or not the candidate buildings are spatially adjacent to the shadow they cast. If there is an overlap, then the detected buildings can be retained; otherwise, the detected buildings are removed. The rule of SHD can be expressed as follows:

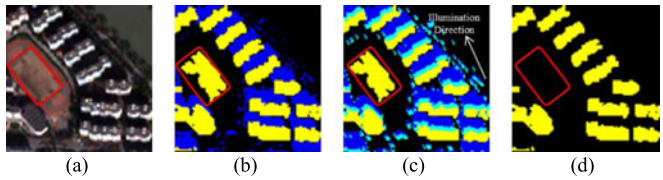


Fig. 5. Example of shadow verification: (a) RGB; (b) initial result for buildings and shadow, where the yellow and blue pixels represent buildings and shadow, respectively; (c) dilation of shadow, where the cyan pixels represent the dilated shadow areas; and (d) the result after shadow verification.

$$\mathbf{IF}\ \mathrm{SPE}\,(x) \cup \mathrm{dilate}\,(S\,(x)) = \emptyset,$$
$$\mathbf{THEN}\ x\ \mathrm{is\ identified\ as\ a\ non-building\ structure}$$

where SPE(x) represents the refined building detection result of pixel x after the spectral constraint, $S(.)$ is the shadow map, and dilate(.) denotes the morphological dilation operator. As shown in Fig. 5, buildings that are adjacent to the dilated shadow are confirmed and retained, but a playing field (circled by the red border) without surrounding shadow can be successfully filtered out.

2) Shape Elimination: Shape features, such as the area and length–width ratio, can be used as constrained conditions for buildings. On the basis of the building map generated by the connected component analysis, the area is calculated by counting

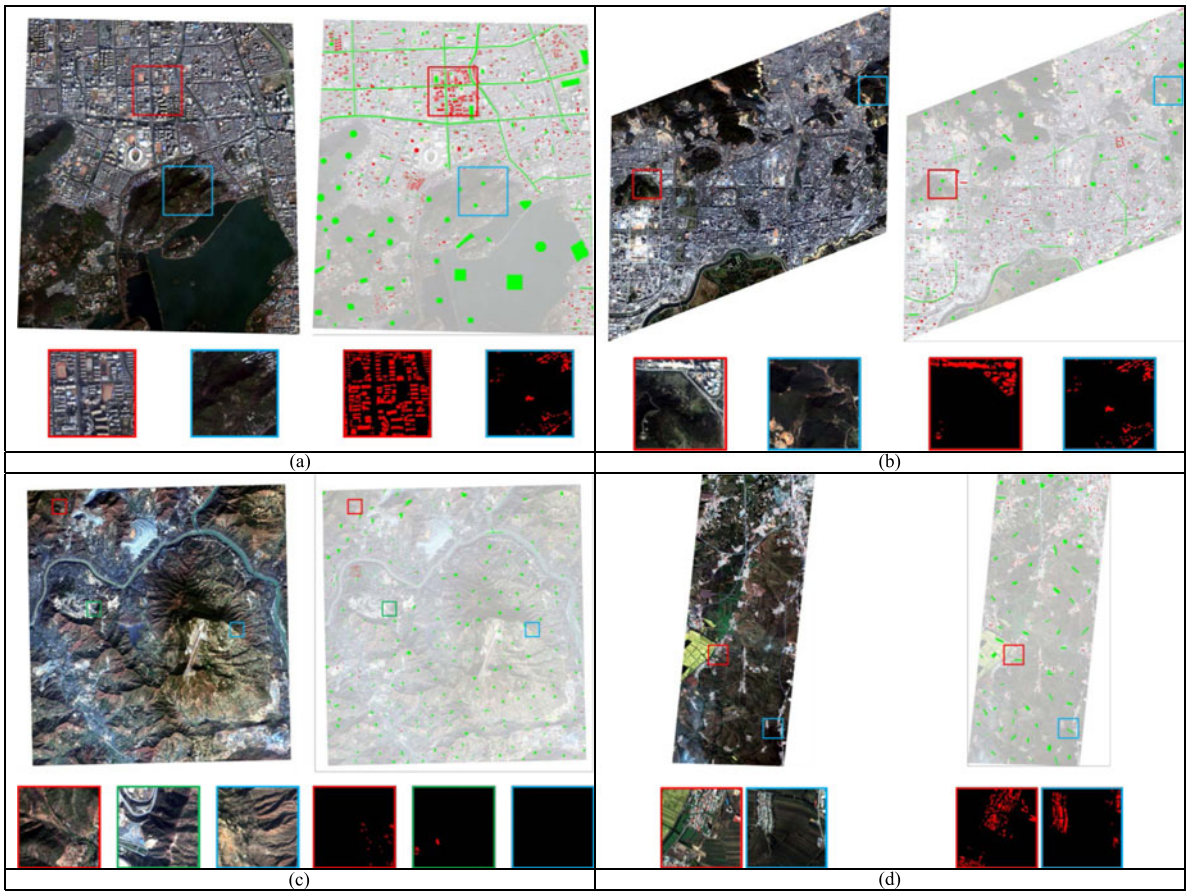


Fig. 6. Four selected test datasets: (a) the Hangzhou image and test patches #1 and #2; (b) the Shenzhen image and test patches #3 and #4; (c) the Panzhihua image and test patches #5 to #7; and (d) the Harbin image and test patches #8 and #9.

TABLE I
DETAILS OF THE DATASETS CONSIDERED IN THIS RESEARCH

Test image	Satellite sensor	Spatial resolution	Spatial size	Major coverage	Major error sources
Hangzhou	WorldView-2	2.0	2454×2750	Medium-density urban	roads, bright soil, open areas
Shenzhen	QuickBird	2.4	4507×5360	highly dense urban	roads, bright soil, open areas
Panzhihua	QuickBird	2.4	8805×9449	mountainous	roads, bright soil, open areas, paddy fields
Harbin	QuickBird	2.4	7310×3554	agricultural, rural	bright soil, roads

the total number of pixels in an object, and the length–width ratio is measured as the ratio between the maximum and minimum lengths of an object. The aim of the shape constraints is to alleviate the commission errors by removing small errors, e.g., narrow and elongated roads. The rule of SHA can be expressed as

$$\textbf{IF}\ \mathrm{Count}\,(x) < c\ \ \textbf{OR}\ \mathrm{LWR}\,(x) > r$$
$$\textbf{THEN}\ x\ \text{is identified as a nonbuilding structure}$$

where $\mathrm{Count}(.)$ indicates the number of pixels in an object and c denotes its threshold. $\mathrm{LWR}(.)$ represents the length–width ratio and r is its threshold.

IV. EXPERIMENTS AND DISCUSSION

A. Datasets

We selected four high-resolution datasets covering the cities of Hangzhou, Shenzhen, Panzhihua, and Harbin to validate the performance of the proposed framework. Located in different regions of China, these study areas represent a variety of typical scenes, such as urban, rural, mountainous, and agricultural areas. The study images, as well as a set of representative test patches (for detailed analysis and comparison), are shown in Fig. 6, and the details of these datasets are listed in Table I.

B. Experimental Settings

To investigate the effectiveness of the proposed framework, the original MBI [23] and DMP-SVM [24] were used as comparison methods (see Section IV-C1). First, please note that the

TABLE II
PARAMETERS OF THE PROPOSED FRAMEWORK FOR THE STUDY DATASETS

	N (m)	v_{min}	v_{max}	w	h_{min}	h_{max}	r	c	b
Hangzhou	120	0.12	0.2	0.5	20	140	7	9	0.12
Shenzhen	120	0.12	0.2	0.5	20	140	7	4	0.12
Panzhihua	120	0.12	0.2	0.5	20	140	7	4	0.20
Harbin	120	0.12	0.2	0.5	20	140	7	4	0.25

threshold values for the MBI and the proposed framework are different due to their different motivations. The threshold of the MBI is aimed at maximizing the discrimination between building and nonbuilding structures, but the proposed framework prefers a smaller threshold value, so as to keep more building candidates before importing the spectral, shadow, and shape filters. Meanwhile, DMP-SVM utilizes multiscale and multidirectional DMPs as features and support vector machine (SVM) [31] as the classifier in a binary classification framework to discriminate between buildings and nonbuildings. In this case, 300 building pixels and 300 background pixels were randomly and manually delineated from each image as training samples.

To further explore the role of the shadow verification of the proposed framework, two recent shadow-based detectors, kindly implemented by the original authors, were used for a detailed comparison (see Section IV-C2). The first shadow-based detector is GrabCut [18], which models the directional spatial relationship between buildings and their shadows using a new fuzzy landscape generation approach, and the buildings are then detected by GrabCut with the help of the detected shadow and landscape regions. The second detector is multilabel partitioning [19], which improves the GrabCut [18] approach by the use of the two-level graph theory approach. Both of these approaches were designed for the automatic detection of buildings in complex environments, and they can perform well in challenging scenes. Due to the availability of the metadata of the test site, a relatively large area in Shenzhen was tested (see Section IV-C2). For a fair comparison, the results of both the benchmark methods achieved by the optimal parameter settings are reported, and the suggested parameter settings for the proposed framework in all the test datasets are listed in Table II (the hue images were rescaled to [0,255]). Concerning the NIR band is considered due to its superior separability for shadow. Subsequently, the resulting shadow feature image is dilated the NIR band is considered due to its superior separability for shadow. Subsequently, the resulting shadow feature image is dilated the parameter selection, several issues should be noted. First, appropriate ranges of parameters for the proposed framework are suggested in this study. Second, most of the parameters can be kept the same for different datasets. The parameter sensitivity is further analyzed in Section IV-C. All the experimental steps, except for SVM, which was implemented with the help of the LibSVM package using C++, were performed using MATLAB R2015b on a workstation with a single i7-3770K 2.80 GHz processer and 64.0 Gb of RAM.

To quantitatively evaluate the building detection results, a large number of manually delineated representative reference samples were uniformly distributed in the images. The reference samples were generated by field investigation and careful image interpretation of the study areas. The following metrics were used to assess the building detection accuracy: completeness (Com), correctness (Cor), and quality (Q). Referring to as the "detection rate," completeness is the percentage of correctly detected buildings to the total number of reference buildings. Correctness is the percentage of correctly detected buildings to the total number of detected buildings [32]. An ideal detection result should simultaneously have high completeness and correctness. The quality of the results provides a tradeoff metric that balances completeness and correctness [33].

C. Experimental Results

1) General Results and Analysis of the Representative Test Patches: The building detection results are shown in Fig. 7, where the white and black pixels represent the detected buildings and background, respectively. The quantitative accuracies are listed in Table III, where the best results for each quality index are labeled in bold, and the suboptimal results for each quality index are underlined. In general, the MBI shows the worst performance and the proposed framework obtains the best performance for all cases, in terms of both the visual interpretation and quantitative assessment. In the four test images, building extraction for Panzhihua and Harbin, corresponding to mountainous and rural/agricultural areas, respectively, was challenging, but the proposed framework still presents promising performances in these areas, i.e., the quality indices are $\sim$82%. The detection results obtained by the proposed framework, MBI, and DMP-SVM are further discussed as follows.

MBI performs well in the urban areas, but it is subject to overestimation (false alarms) in the nonurban areas, due to the complexity of the challenging scenes, especially in Panzhihua and Harbin. According to Table III, it is apparent that the proposed framework is superior to the original MBI in both completeness and correctness. For instance, the increase of completeness, correctness, and quality are 0.55%, 7.91%, and 6.48% in Hangzhou, and 5.75%, 18.75%, and 18.84% in Panzhihua, respectively. The significant improvements in correctness indicate that a number of false alarms can be successfully removed by the use of the filters in the proposed framework. Meanwhile, compared to the original MBI, a smaller binary threshold for the proposed framework preserves more potential buildings, and, hence, ensures a higher completeness in the experiments.

Benefiting from the supervised machine learning, the accuracy of DMP-SVM is better than the original MBI, and the completeness is close to that achieved by the proposed framework. However, similar to the MBI, DMP-SVM does not solve the problem of false alarms, resulting in a lower correctness (see Table III). Furthermore, DMP-SVM requires intensive manual labor since additional training samples need to be manually collected to precisely model the features of the buildings. To further compare the performance of the proposed framework and DMP-SVM, the computational efficiency (computation time) is reported in Table III. Despite ignoring the time cost for the selection of the training samples, the proposed framework can

Fig. 7. Building detection results: (a) original RGB images; (b) building maps with the result of the MBI; (c) building maps with the result of DMP-SVM; and (d) building maps with the result of the proposed framework. The green color represents the boundary between the high-density and low-density areas.

TABLE III
REFERENCE INFORMATION, ACCURACY (%), AND COMPUTATION TIMES (S) OF THE STUDY DATASETS

	# Test samples		MBI				DMP-SVM				The proposed framework			
Test image	Buildings	Background	Com	Cor	Q	Time	Com	Cor	Q	Time	Com	Cor	Q	Time
Hangzhou	121671	296094	85.86	87.39	76.40	**7.8**	86.10	88.51	77.44	28.3	**86.41**	**95.30**	**82.88**	12.9
Shenzhen	134979	244670	89.11	90.65	81.61	**34.0**	89.32	91.07	82.12	105.0	**93.84**	**92.42**	**87.14**	43.4
Panzhihua	112780	916871	83.65	72.74	63.69	**140.6**	87.81	73.91	67.03	348.5	**89.40**	**91.49**	**82.53**	174.1
Harbin	77317	365455	86.74	80.25	71.48	**44.5**	86.45	83.78	74.06	130.2	**89.58**	**96.07**	**86.41**	52.6

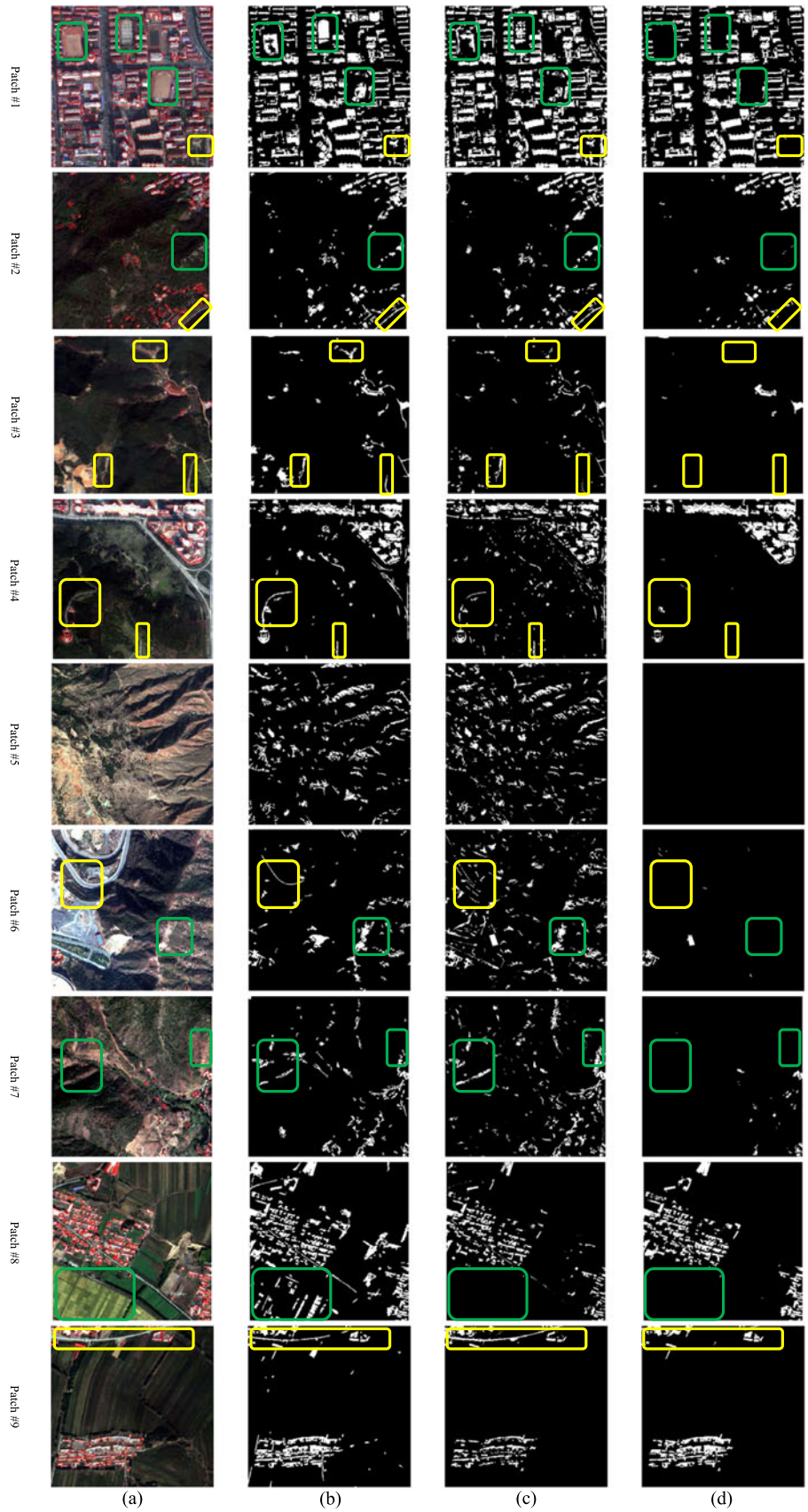


Fig. 8. Building detection results and the ground reference for the representative test patches: (a) original RGB images with ground reference (with red boundary); (b) the results of the MBI; (c) the results of DMP-SVM; and (d) the results of the proposed framework.

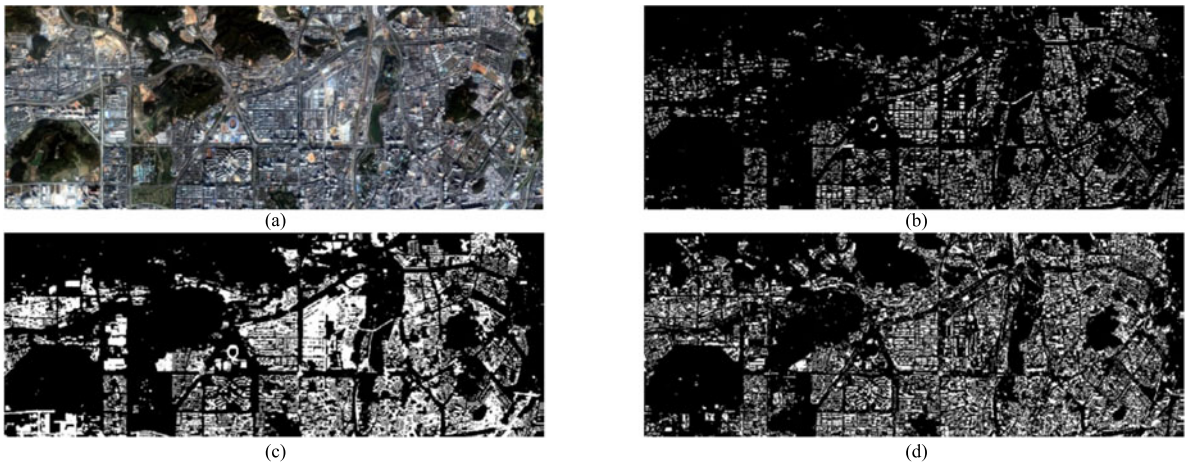


Fig. 9. Building detection results: (a) original RGB image; (b) result of GrabCut [18]; (c) result of multilabel partitioning [19]; and (d) result of the proposed framework.

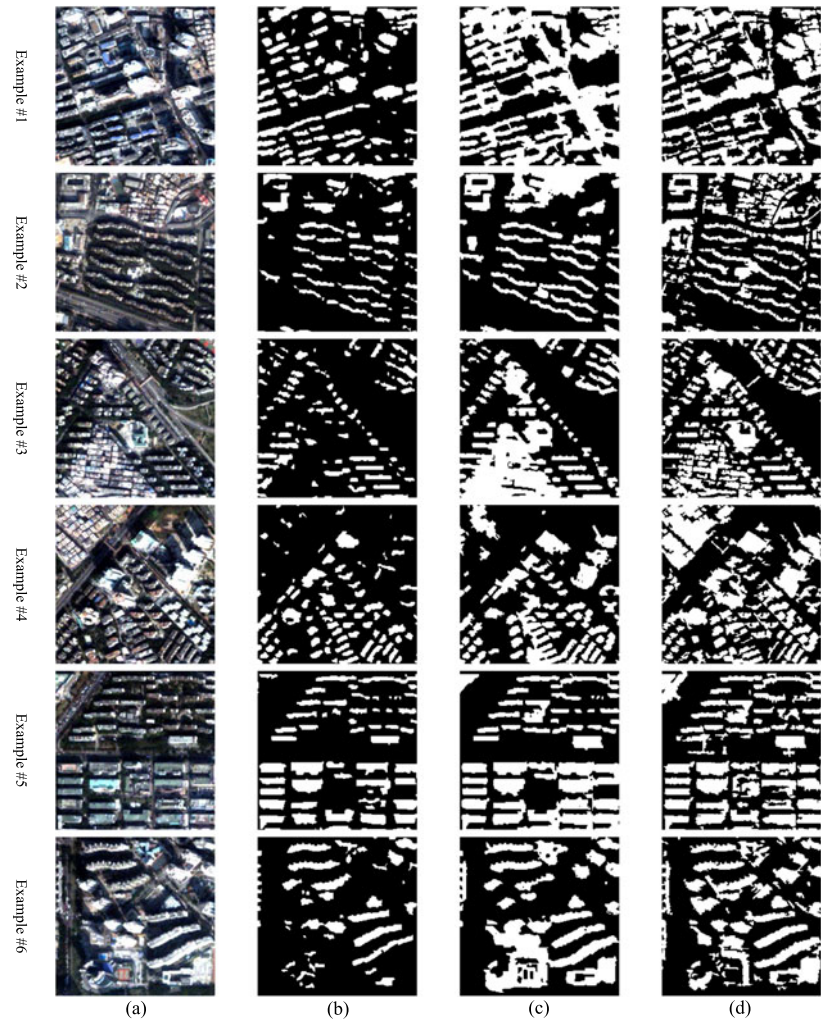


Fig. 10. Building detection results and the ground reference (with red boundary) for the representative examples: (a) original RGB images; (b) results of GrabCut [18]; (c) results of multilabel partitioning [19]; and (d) results of the proposed framework.

TABLE IV
ACCURACY (%) OF THE BUILDING DETECTION RESULTS FOR EACH STEP OF THE PROPOSED FRAMEWORK

	Initial MBI			SPE			SPE + SHD			SPE + SHD + SHA		
Test image	Com	Cor	Q	Com	Cor	Q	Com	Cor	Q	Com	Cor	Q
Hangzhou	**87.90**	85.12	76.19	86.63	91.81	80.41	86.42	94.70	82.43	86.41	**95.30**	**82.88**
Shenzhen	**94.64**	81.19	77.62	94.20	87.95	83.42	94.17	90.95	86.10	93.84	**92.42**	**87.14**
Panzhihua	**90.91**	49.50	47.17	89.89	86.48	78.82	89.48	89.47	80.96	89.40	**91.49**	**82.53**
Harbin	**90.19**	70.37	65.37	89.67	93.52	84.42	89.58	95.51	85.96	89.58	**96.07**	**86.41**

TABLE V
Q INDEX (%) OF THE BUILDING DETECTION RESULTS FOR THE SPECTRAL, SHADOW, AND SHAPE CONSTRAINTS

Test image	Initial MBI	SPE	SHD	SHA	SPE + SHA + SHD	SPE + SHD + SHA
Hangzhou	76.19	80.41	80.19	76.97	82.75	**82.88**
Shenzhen	77.62	83.42	81.48	78.95	86.86	**87.14**
Panzhihua	47.17	78.82	58.36	49.12	81.47	**82.53**
Harbin	65.37	84.42	81.86	66.00	85.97	**86.41**

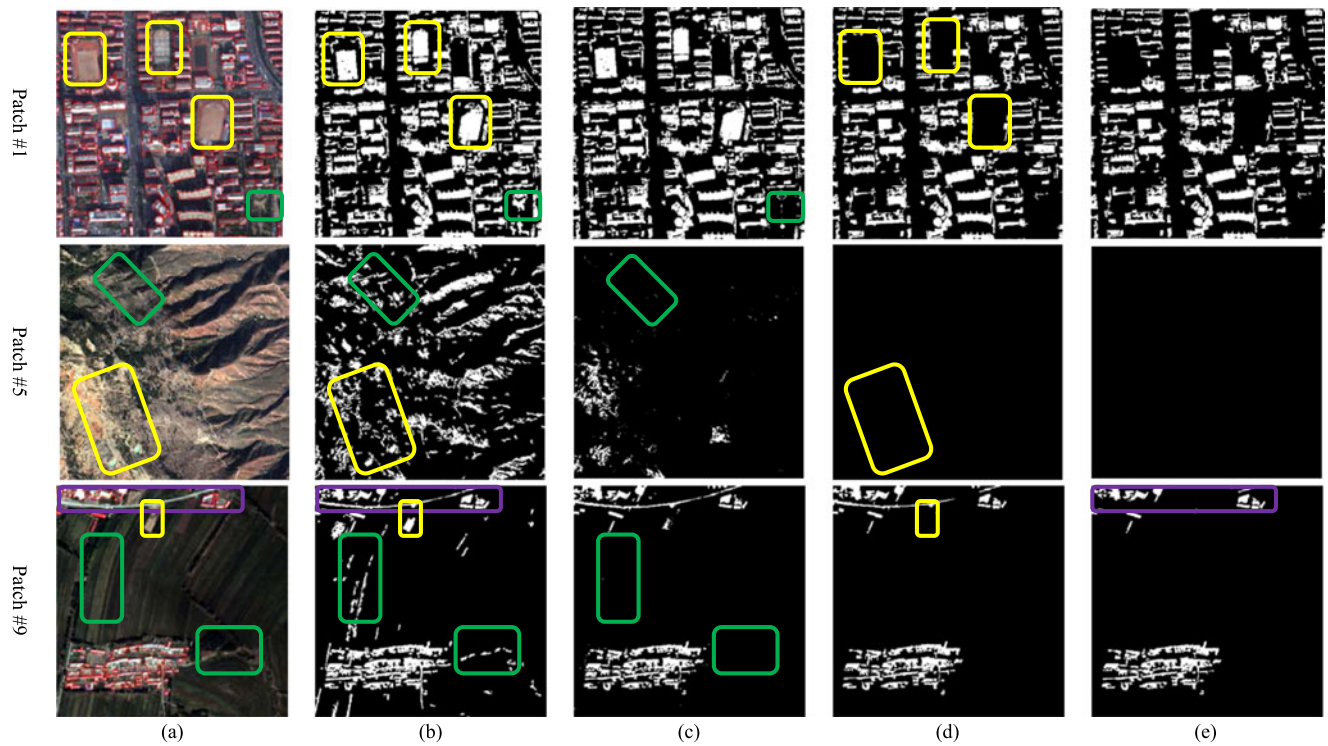


Fig. 11. Building detection results and the ground reference for patches #1, #5, and #9: (a) original RGB images with ground reference (with red boundary); (b) results of the initial MBI; (c) results of the initial MBI + spectral constraint; (d) results of the initial MBI + spectral constraint + shadow verification; and (e) results of the initial MBI + spectral constraint + shadow verification + shape constraint. The red, yellow, and purple regions emphasize the performances of the spectral, shadow, and shape constraints, respectively.

be carried out much more rapidly than DMP-SVM. The above experimental results confirm that the proposed framework is efficient for building detection from high-resolution imagery.

The results of the set of representative test patches (see Fig. 6) are further compared and analyzed in this section. The original image, the reference data, and the results of the MBI, DMP-SVM, and the proposed framework are shown in Fig. 8, where the ground reference is delineated with a red boundary in the original image. As can be seen in the figure, the proposed framework is the most robust and promising detector in the various challenging scenes. In these test patches, it can be clearly seen that the MBI and the DMP-SVM are subject to many false alarms, especially in the nonurban scenes. Each example patch is analyzed in detail in the following.

In patch #1, which is a typical urban area, all the detectors show similar completeness but different correctness, as the MBI and DMP-SVM wrongly detect a number of false alarms, such as playing fields (green rectangles), open areas, and bare soil (yellow rectangle). In patches #2, #3, and #4, where the landscapes mostly consist of natural green space, both the MBI and DMP-SVM are subject to false alarms, such as bare soil (green rectangles) and roads (yellow rectangles). As for the mountainous areas, e.g., patches #5, #6, and #7, a large quantity of bare soil, open areas, and paddy fields are incorrectly identified, corresponding to the relatively low correctness values reported in Table III. In particular, patch #5, where no buildings exist, was used to test the commission errors of the different building detectors. In this case, the proposed framework achieves the best results, but the MBI and DMP-SVM lead to a series of commission errors, with a number of bright soil and vegetation areas incorrectly detected. Patches #8 and #9 are associated with agricultural areas, where there is confusion between the buildings/dry farmland in patch #8 (green rectangle) and the elongated roads/buildings (yellow rectangle) in patch #9, due to their similar spectral reflectance. All in all, the representative test patches show that the proposed framework is able to produce more robust and desirable results in various image scenes than the two comparative methods.

2) Comparison With Recent Shadow Detection-Based Detectors: In order to compare the proposed framework with the existing shadow detection-based approaches, the building detection maps of the test image (a relatively large area in the Shenzhen test image) obtained with the proposed framework, GrabCut [18], and multilabel partitioning [19] are shown in Fig. 9, and the according representative patches are displayed in Fig. 10 for further analysis. GrabCut [18] and multilabel partitioning [19] are shadow-based methods, which tend to be limited by the shadow detection results, although the multilabel partitioning [19] is less sensitive, which can be attributed to the introduction of a two-level graph partitioning framework. According to Fig. 9, the proposed framework provides comparable or even slightly better results than the multilabel partitioning [19], and both methods perform better than GrabCut [18], since it misses more buildings. Specifically, the proposed framework performs better than the two comparative methods in some local regions because the shadow detection result may not be very good in these regions, but the multilabel partitioning [19] acquires better building shape preservation and higher completeness in some local regions. As can be seen in Fig. 10, all of the approaches perform well since they detect most of the buildings in these representative patches. GrabCut [18] is subject to the most severe omission problem, and the multilabel partitioning [19] also misses some buildings due to the difficulty of precisely detecting the shadow in such a complex and large image, while the proposed framework shows less sensitivity because the shadow is only utilized to verify the detected building candidates. However, some dark buildings are still left out of the results of the proposed framework, which is one of the drawbacks of the original MBI.

D. Experimental Analysis

1) Constraints of the Proposed Framework: To clearly illustrate the specific roles of the different constraints in the proposed framework, the results of the filters are displayed step by step in Table IV and Fig. 11. The accuracy indices in the table show that, when compared with the original MBI, the correctness and quality scores are raised significantly, but the completeness scores only decreases slightly when the SPE, SHD, and SHA constraints are gradually considered. In this regard, it can be stated that the proposed framework successfully removes a large number of false alarms, and at the same time hardly misses any correct buildings. Specifically, the spectral constraints effectively remove the bare soil in patch #1 (urban area), the soil and paddy fields in patch #5 (mountainous area), and the dry farmland and road in patch #9 (agricultural area). The shadow constraint (SHD) effectively filters out the playing fields and open areas in patch #1 and the bright soil in patch #5. Finally, the shape constraint can clear up the residual dry farmland and roads in patch #9.

Furthermore, the effects of the three filters are individually displayed in Table V, columns 2–4, which list the quality scores (Q) of the detection results associated with the specific filter (SPE, SHD, and SHA). Among these filters, with regard to the Q scores (see Table V), the spectral constraint achieves the best results in terms of improving the original MBI. The shape constraint presents only a slight improvement over the initial MBI, as it is only aimed at removing small noise points and elongated objects. Moreover, the order of the three filters was investigated, and two different combinations are compared in last two columns of Table V. This shows that the most appropriate order for the filters of the proposed framework is spectral, shadow, and shape in order.

2) Parameters of the Proposed Framework: In this section, suggestions for the parameter settings are given and the sensitivity of some of the critical parameters is analyzed.

a) Window Size N (m): N is used to calculate the average MBI value of the pixels in an $N \times N$ window size, which is used to separate high MBI areas from low MBI areas. In order to investigate the impact of N, the completeness, correctness, and quality scores of the building detection results with different window sizes are displayed in Fig. 12. N is not sensitive to these test images as their results are relatively stable when the value of N is varied from 60 to 300. However, the performances with 120 and 180 are better than the performances with the other sizes, while the former size has a lower computational cost.

b) Spectral Constraint: The dual thresholds of the NDVI (v_{high} and v_{low}) are applied to eliminate vegetation and soil in different degrees for high MBI areas and low MBI areas. In Fig. 13, the completeness, correctness, and quality of the building detection results with different combinations of v_{high} and v_{low} are shown in order to analyze the impact of the thresholds of the NDVI. The Hangzhou, Shenzhen, Panzhihua, and Harbin datasets achieve the highest quality score when $v_{\mathrm{low}} = 0.12$ and $v_{\mathrm{high}} = 0.24, v_{\mathrm{low}} = 0.12$ and $v_{\mathrm{high}} = 0.2, v_{\mathrm{low}} = 0.08$ and $v_{\mathrm{high}} = 0.2$, and $v_{\mathrm{low}} = 0.12$ and

TABLE VI
ACCURACY (%) OF THE BUILDING DETECTION RESULTS FOR THE AUTOMATIC SELECTION OF THRESHOLDS

	MBI			NDVI (Otsu's)			NDWI (Otsu's)			Binary threshold (Otsu's)			The proposed framework		
Test image	Com	Cor	Q	Com	Cor	Q	Com	Cor	Q	Com	Cor	Q	Com	Cor	Q
Hangzhou	85.86	87.39	76.40	*84.18*	96.06	81.37	*58.58*	97.82	57.82	*62.02*	**98.84**	*61.57*	**86.41**	95.30	**82.88**
Shenzhen	89.11	90.65	81.61	**94.21**	*87.71*	83.22	*25.80*	91.41	25.19	*84.52*	**97.60**	82.80	93.84	92.42	**87.14**
Panzhihua	83.65	72.74	63.69	**89.83**	87.18	79.35	*15.48*	77.11	14.80	*82.59*	**92.10**	77.13	89.40	91.49	**82.53**
Harbin	86.74	80.25	71.48	**90.27**	93.59	85.01	*49.27*	87.70	46.09	*82.73*	**96.93**	80.61	89.58	96.07	**86.41**

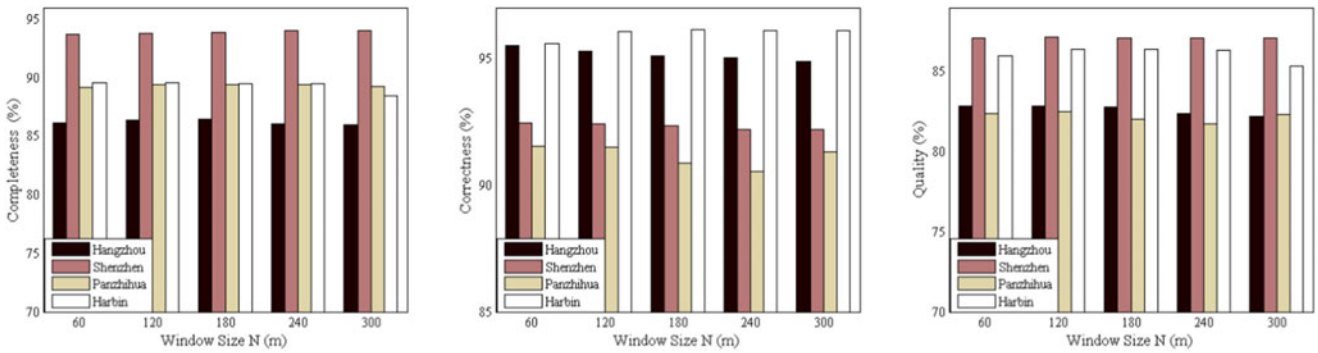


Fig. 12. Correlation between the window size *N* and the building detection accuracies: (a) completeness; (b) correctness; and (c) quality.

$v_{\mathrm{high}} = 0.24$, respectively. That is, the optimal value in each subfigure belongs to the nondiagonal ones, which confirms the superiority of the dual-threshold setting. According to the authors' experience, the suggested thresholds of v_{high} and v_{low} are in the range [0.1, 0.16] and [0.2, 0.3], respectively, where the quality scores are satisfactory and stable.

Thresholds of the NDWI (w) and hue component (h_{min} and h_{max}): The NDWI is aimed at filtering out the commission errors of water features, and its threshold *w* was fixed as 0.5 in this study. However, the threshold may need to be changed when the NDWI is applied to different high-resolution remote-sensing images. Although it may be difficult to select a threshold to distinguish water from other targets, it is easy to select a relatively large threshold in order to filter out the bright water and retain the buildings. The thresholds of the hue component are used to highlight the green and yellow components in the color space. Therefore, h_{min} and h_{max} can be safely set as 20 and 140, respectively (the hue image is rescaled to [0, 255]).

c) Binary Threshold (b): The binary threshold *b* is the threshold used to extract the initial building structures from the original MBI. It is suggested that it is set as a small value in order to keep more building candidates for the subsequent spectral, shadow, and shape filters. The building detection accuracies for the proposed framework with regard to parameter *b* in all four datasets are displayed in Fig. 14. I can be said that this parameter should be fine-tuned according to the different test datasets. For instance, an appropriate range of *b* is between 0.04 and 0.12 for the Hangzhou dataset, but the parameter is not sensitive to the other three test images as their results are relatively stable when its value is changed from 0.04 to 0.28. In general, however, the preferred threshold can be selected in terms of the Q score to balance the completeness and correctness.

d) Shape Constraint: Thresholdof the area (c): The area is used to remove small noise points. Considering that most small buildings have an area of approximately 5×5 m [13], in this study, the area threshold was fixed as $c = 4$ for the QuickBird images and $c = 9$ for the WV-2 images.

Threshold of the length–width ratio (*r*): The length–width ratio is used to remove narrow and elongated nonbuilding structures. The correlation between the threshold of the length–width ratio and the building detection accuracies is illustrated in Fig. 15. It can be seen in Fig. 15 that a small value of r can filter out some buildings, leading to decrease in the completeness (see Fig. 15(a)). The correctness index shows an improvement as the value of *r* is increased from 4 to 12 (see Fig. 15(b)), illustrating that the shape constraint is able to reduce the false alarms. In general, a suitable range for threshold *r* is from 7 to 10 in terms of the Q score reported in Fig. 15(c).

3) Threshold Setting Approaches: As a typical automatic approach, Otsu's method, which partitions the whole data into two clusters (i.e., background and target) by measuring the variance, was used as a comparison. The thresholds associated with the NDVI, the NDWI, and the binary thresholds (*b*) were investigated. We manually fixed the other parameters as the corresponding optimal and focused on one specific parameter at a time. The quantitative accuracies are listed in Table VI, where the best results for each quality index are labeled in bold. In general, the superior Q scores in the last column and the

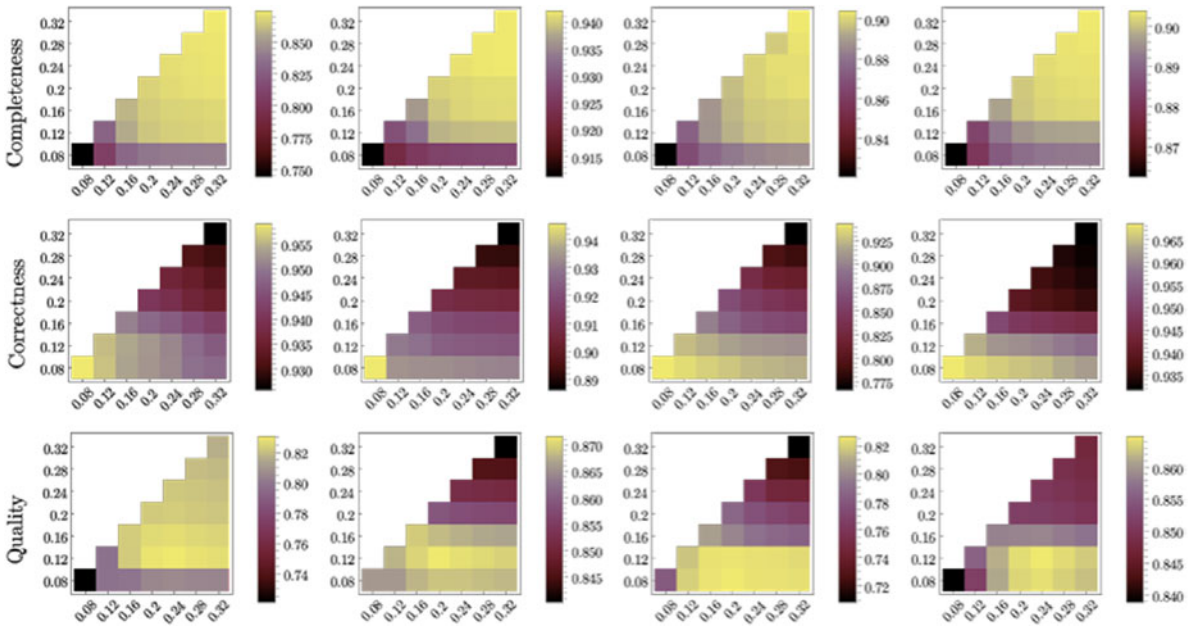


Fig. 13. Correlation between the thresholds of the NDVI and the building detection accuracies: the horizontal axis and vertical axis represent the values of v_{high} and v_{low}, respectively.

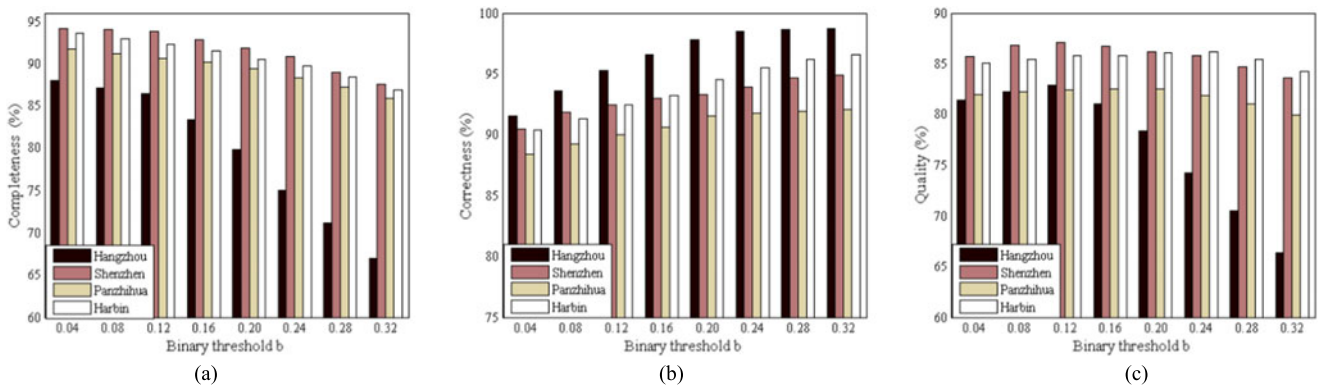


Fig. 14. Correlation between the binary threshold *b* and the building detection accuracies: (a) completeness, (b) correctness, and (c) quality.

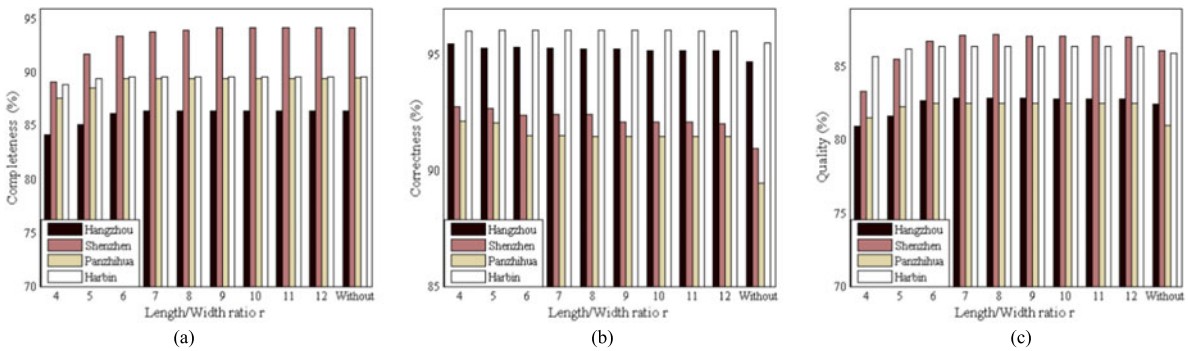


Fig. 15. Correlation between the threshold of the length–width ratio (*r*) and the building detection accuracies: (a) completeness; (b) correctness; and (c) quality.

inferior values associated with Otsu's method confirm the superior performance of the proposed framework.

For the first case, the thresholds of v_{high} and v_{low} were equivalent to the automatic NDVI threshold. The Q scores for the automatic NDVI threshold show a decrease in accuracy for all the datasets compared to the manually selected thresholds, although the completeness presents a tiny improvement in these datasets, except for the Hangzhou dataset. The results of the automatic NDWI threshold are very poor, which is related to the proportion of water in these images. However, it is a good idea to select a relatively large threshold in order to filter out the bright water and retain the buildings. The automatic binary threshold tends to be too large, and the completeness values are significantly lower while the correctness values are only slightly higher, thus causing a decrease in the quality. Overall, the automatic approach is sensitive to the amount of the land cover, especially when dealing with rare target cases. In view of this, using the suggested ranges for the parameters may be a better approach.

V. Conclusion

In this study, a new building detection postprocessing framework has been proposed, aiming at adapting and extending building detection to different and complex environments, such as urban, agricultural, mountainous, and rural areas. To this end, the newly proposed framework improves the recently developed MBI detector by additionally considering spectral, shadow, and shape constraints, which actually form a filtering system to deal with more challenging image scenes. Four high-resolution images acquired by commercial high-spatial-resolution sensors (e.g., QuickBird, WorldView-2) were used to validate the proposed framework, and the results were promising and robust as the quality scores of all the test images were higher than 80%.

The advantages of the proposed framework can be summarized as follows: 1) it is a robust and effective building detector for high-resolution remote-sensing imagery, by simultaneously integrating multiple constraints, including morphological, spectral, shadow, and shape features. 2) The proposed framework is able to achieve satisfactory performances for various challenging scenes (not only in urban areas); thus, showing its potential for building detection from large-area high-resolution remote-sensing imagery. 3) The appropriate range for the thresholds has been suggested based on a large number of tests. 4) The high computational efficiency and simple implementation of the proposed framework show its potential for not only building detection, but also for other applications, such as building density estimation, urban environmental monitoring, socioeconomic parameter analysis, etc.

Acknowledgment

The authors would like to thank Dr. A. O. Ok for providing the results of his approaches for comparison and the anonymous reviewers for their insightful and constructive comments, which significantly improved the quality of this paper.

References

[1] X. Jin and C. H. Davis, "Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information," *EURASIP J. Appl. Signal Process*, vol. 2005, pp. 2196–2206, 2005.

[2] D. Lu, H. Tian, G. Zhou, and H. Ge, "Regional mapping of human settlements in southeastern China with multisensor remotely sensed data," *Remote Sens. Environ.*, vol. 112, no. 9, pp. 3668–3679, 2008.

[3] Q. Weng, *Global Urban Monitoring and Assessment Through Earth Observation*. Boca Raton, FL, USA: CRC Press, 2014, p. 84.

[4] L. Bruzzone and L. Carlin, "A multilevel context-based system for classification of very high spatial resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2587–2600, Sep. 2006.

[5] S. W. Myint, N. S.-N. Lam, and J. M. Tyler, "Wavelets for urban spatial feature discrimination," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 7, pp. 803–812, 2004.

[6] J. A. Benediktsson, M. Pesaresi, and K. Amason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.

[7] J. Inglada, "Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features," *ISPRS J. Photogramm.*, vol. 62, no. 3, pp. 236–248, 2007.

[8] C. Senaras, M. Ozay, and F. T. Yarman Vural, "Building detection with decision fusion," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 6, no. 3, pp. 1295–1304, Jun. 2013.

[9] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.

[10] L. Bian, "Retrieving urban objects using a wavelet transform approach," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 2, pp. 133–141, 2003.

[11] X. Huang, L. Zhang, and P. Li, "Classification and extraction of spatial features in urban areas using high-resolution multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 260–264, Apr. 2007.

[12] M. Pesaresi, A. Gerhardinger, and F. Kayitakire, "A robust built-up area presence index by anisotropic rotation-invariant textural measure," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 1, no. 3, pp. 180–192, Sep. 2008.

[13] K. Stankov and D.-C. He, "Detection of buildings in multispectral very high spatial resolution images using the percentage occupancy hit-or-miss transform," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 7, no. 10, pp. 4069–4080, Oct. 2014.

[14] K. Karantzalos and N. Paragios, "Recognition-driven two-dimensional competing priors toward automatic and accurate building detection," *IEEE Trans Geosci. Remote Sens.*, vol. 47, no. 1, pp. 133–144, Jan. 2009.

[15] B. Sirmacek and C. Ünsalan, "Urban-area and building detection using SIFT keypoints and graph theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1156–1167, Apr. 2009.

[16] J. Wang, X. Yang, X. Qin, X. Ye, and Q. Qin, "An efficient approach for automatic rectangular building extraction from very high resolution optical satellite imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 487–491, Mar. 2015.

[17] C. Benedek, X. Descombes, and J. Zerubia, "Building detection in a single remotely sensed image with a point process of rectangles," in *Proc. Int. Conf. Pattern Recog.*, 2010, pp. 1417–1420.

[18] A. O. Ok, C. Senaras, and B. Yuksel, "Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery," *IEEE Trans Geosci. Remote Sens.*, vol. 51, no. 3, pp. 1701–1717, Mar. 2013.

[19] A. O. Ok, "Automated detection of buildings from single VHR multispectral images using shadow information and graph cuts," *ISPRS J. Photogramm.*, vol. 86, pp. 21–40, 2013.

[20] C. Senaras and F. T. Yarman Vural, "A self-supervised decision fusion framework for building detection," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 9, no. 5, pp. 1780–1791, May 2016.

[21] M. Pesaresi and A. Gerhardinger, "Improved textural built-up presence index for automatic recognition of human settlements in arid regions with scattered vegetation," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 4, no. 1, pp. 16–26, Mar. 2011.

[22] P. Gamba, M. Pesaresi, K. Molch, A. Gerhardinger, and G. Lisini, "Anisotropic rotation invariant built-up presence index: Applications to SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2008, pp. V-338–V-341.

[23] X. Huang and L. Zhang, "A multidirectional and multiscale morphological index for automatic building extraction from multispectral GeoEye-1 imagery," *Photogramm. Eng. Remote Sens.*, vol. 77, no. 7, pp. 721–732, 2011.
[24] X. Huang and L. Zhang, "Morphological building/shadow index for building extraction from high-resolution imagery over urban areas," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 5, no. 1, pp. 161–172, Feb. 2012.
[25] X. Huang, Q. Lu, and L. Zhang, "A multi-index learning approach for classification of high-resolution remotely sensed images over urban areas," *ISPRS J. Photogramm.*, vol. 90, pp. 36–48, 2014.
[26] Y. Tang, X. Huang, and L. Zhang, "Fault-tolerant building change detection from urban high-resolution remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1060–1064, Sep. 2013.
[27] X. Huang, L. Zhang, and T. Zhu, "Building change detection from multitemporal high-resolution remotely sensed images based on a morphological building index," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 7, no. 1, pp. 105–115, Jan. 2014.
[28] B.-C. Gao, "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 257–266, 1996.
[29] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, nos. 285/296, pp. 23–27, 1975.
[30] H. Song, B. Huang, and K. Zhang, "Shadow detection and reconstruction in high-resolution satellite images via morphological filtering and example-based learning," *IEEE Trans Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2545–2554, May 2014.
[31] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[32] W. Song and T. L. Haithcoat, "Development of comprehensive accuracy assessment indexes for building footprint extraction," *IEEE Trans Geosci. Remote Sens.*, vol. 43, no. 2, pp. 402–404, Feb. 2005.
[33] M. Rutzinger, F. Rottensteiner, and N. Pfeifer, "A comparison of evaluation techniques for building extraction from airborne laser scanning," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 2, no. 1, pp. 11–20, Mar. 2009.

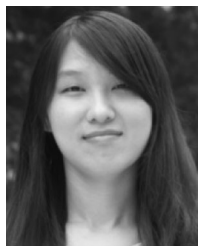
Xin Huang (M'13–SM'14) received the Ph.D. degree in photogrammetry and remote sensing from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China, in 2009.

He is currently a Full Professor with Wuhan University, where he teaches remote sensing, photogrammetry, image interpretation, etc., and where he is also the Founder and Director at the Institute of Remote Sensing Information Processing, School of Remote Sensing and Information Engineering. He has published more than 70 peer-reviewed articles in the international journals. His research interests include hyperspectral data analysis, high-resolution image processing, pattern recognition, and remote-sensing applications.

Prof. Huang was the recipient of the Top-Ten Academic Star of Wuhan University in 2009, the Boeing Award for the Best Paper in Image Analysis and Interpretation from the American Society for Photogrammetry and Remote Sensing in 2010, the New Century Excellent Talents in University from the Ministry of Education of China in 2011, the National Excellent Doctoral Dissertation Award of China in 2012, and the China National Science Fund for Excellent Young Scholars in 2015. In 2011, he was recognized by the IEEE Geoscience and Remote Sensing Society (GRSS) as the Best Reviewer of IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He was the winner of the IEEE GRSS 2014 Data Fusion Contest. He was the lead Guest Editor of the special issue on information extraction from high-spatial-resolution optical remotely sensed imagery for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (vol. 8, no.5, May 2015). Since 2014, he has been an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Wenliang Yuan received the B.S. degree from Wuhan University, Wuhan, China, in 2014, where he is currently working toward the M.S. degree at the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing

His current research interests include high-resolution building extraction, image processing, pattern recognition, machine learning, and remote-sensing applications.



Jiayi Li (S'12–M'15) received the B.S. degree from Central South University, Changsha, China, in 2011, and the Ph.D. degree from Wuhan University, Wuhan, China, in 2016.

She is currently an Associate Professor with the School of Remote Sensing and Information Engineering, Wuhan University. Her research interests include hyperspectral imagery, sparse representation, computation vision and pattern recognition, and remote-sensing images.

Dr. Li is a Reviewer of more than ten international journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, the IEEE SIGNAL PROCESSING LETTER, and the *International Journal of Remote Sensing*.

Liangpei Zhang (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, ChangSha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is currently the Head at the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. He is also a Chang-Jiang Scholar Chair Professor appointed by the Ministry of Education of China. He is currently a Principal Scientist for the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. He has more than 410 research papers. He is the holder of 15 patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is a Fellow of the Institution of Engineering and Technology, an Executive Member (Board of Governor) of the China National Committee of the International Geosphere-Biosphere Programme, and an Executive Member of the China Society of Image and Graphics. He regularly serves as a Cochair of the series SPIE Conferences on Multispectral Image Processing and Pattern Recognition, Conference on Asia Remote Sensing, and many other conferences. He edits several conference proceedings, issues, and Geoinformatics symposiums. He also serves as an Associate Editor of the *International Journal of Ambient Computing and Intelligence*, the *International Journal of Image and Graphics*, the *International Journal of Digital Multimedia Broadcasting*, the *Journal of Geo-Spatial Information Science*, the *Journal of Remote Sensing*, and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He received the best paper Boeing award in 2010 and best paper ERDAS award of the American Society of Photogrammetry and Remote Sensing in 2013.