A Multidirectional and Multiscale Morphological Index for Automatic Building Extraction from Multispectral GeoEye-1 Imagery

Xin Huang and Liangpei Zhang

Abstract

This study proposes a novel morphological building index (MBI) for automatic building extraction from high-resolution remotely sensed imagery. The basic idea of MBI is to build a relationship between the implicit characteristics of buildings (e.g., brightness, size, and contrast) and the properties of morphological operators (e.g., reconstruction, granulometry, and directionality). Buildings are extracted by performing a threshold on the MBI feature image. Subsequently, the shape features, such as area and length-width ratio, are used to refine the binary building map. In order to validate the proposed algorithm, a comparative study was performed between MBI, a recently developed texture-derived built-up presence index (PanTex), and the widely used object-based approach. Experiments were conducted on a multispectral GeoEye-1 image, covering a study area of 5.5 km by 5.3 km in Hongshan district of Wuhan, central China. In experiments, MBI achieved satisfactory results and outperformed other algorithms in terms of both accuracies and visual inspection. The effects of parameters of MBI were also analyzed in detail, including directions, sizes and the binaryzation threshold.

Introduction

The precise location and identification of building features is one of the key information sources for urban planning, population estimation, landscape analysis, and environment surveying. Nowadays, the commercial high-resolution satellite images with multispectral channels became available, which provides more potential for automatic and accurate building detection. However, although the high-resolution data contain rich information in the spatial domain, it is generally agreed that the increase of spatial resolution does not necessarily signify the increase of interpretation accuracy. This phenomenon can be attributed to the high intra-class variance and low inter-class variance for the spectral statistics of high-resolution images, leading to the reduction of separability between spectrally similar classes in the spectral feature space. Therefore, recent advances in the high-resolution image processing focus on the spatial and structural feature extraction, e.g., pixel shape index (Huang *et al.*, 2007), differential morphological profiles (Pesaresi and Benediktsson, 2001), wavelet-based textures (Bian, 2003), and gray-level co-occurrence matrix (GLCM) textures (Pacifici *et al.*, 2009). The spatial information can be used to complement the spectral feature space and enhance the separability of classes with similar spectral responses. However, most of the aforementioned algorithms refer to the supervised machine learning. As a result, a large number of training samples are needed to precisely model the feature distribution of a class of interest.

In recent years, some sophisticated algorithms for building extraction have been presented. Lee *et al.* (2003) used ECHO and ISODATA classifiers to provide an approximate location of buildings, and a fine extraction was then carried out through a squaring approach based on the Hough transformation. Jin and Davis (2005) presented an automated building extraction strategy that simultaneously exploited structural, contextual, and spectral information. The final building extraction was done by integrating the results of the three different information sources. Lu *et al.* (2006) integrated the stereo image pairs, segmentation of classified image, and the level-set based shape model in a Dempster-Shafer algorithm for automatic building detection. Gamba *et al.* (2007) used neural networks and Markov random field for classification of boundary and non-boundary pixels, respectively, and the final result was obtained based on a decision fusion.

At the same time, the object-based analysis (OBA) approach has received much attention for building extraction from high-resolution data. Tian and Chen (2007) studied the optimization of multiscale segmentation for building recognition. Huang and Zhang (2008) proposed an adaptive mean-shift segmentation procedure for discrimination of urban structures with similar spectral attributes (e.g., buildings, roads, trails). However, some researches revealed that the major limitation of OBA was the inaccurate spatial relation between objects and segments

The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University Wuhan 430079, China (huang_whu@163.com).

Photogrammetric Engineering & Remote Sensing
Vol. 77, No. 7, July 2011, pp. 721–732.

0099-1112/11/7707–0721/$3.00/0

(Lhomme *et al.*, 2009), i.e., the segmentation results can not accurately describe the position, shape and context of objects. Moreover, the inaccurate segmentation (e.g., over- or under-segmentation) directly leads to inaccurate description of shape features. It should be admitted that it is still a difficult task to adaptively generate the optimal segmentation results for the OBA.

In this context, the objective of this research is to construct a feature index that can be directly applied to the building detection without any training or segmentation process. This kind of building index is simple, stable, and accurate for supporting massive processing on the high volume of current high-resolution satellite data. Several studies working towards the building index have been reported. Pesaresi *et al.* (2008) proposed a procedure for the calculation of a texture-derived built-up presence index, namely PanTex, based on a fuzzy composition of anisotropic co-occurrence measures. The basic idea of the PanTex is based on the fact that buildings cast a shadow that is producing high local contrast. Therefore, the PanTex was calculated using the Contrast measure of the GLCM for all the considered directions and displacements. Lhomme *et al.* (2009) proposed a semi-automatic building identification approach by a so-called "Discrimination by Ratio of Variance" (DRV). The definition of DRV is due to the evidence that there are high and low variances in the edge and body of buildings, respectively.

By summarizing the characteristics of both PanTex and DRV, it can be found that the construction of a building index is based on a low spectral variation corresponding to the building body and a high spectral variation corresponding to the building periphery. Based on the previous researches, in this study, we propose a novel morphological building index (MBI) for automatic building extraction by considering the following characteristics.

- Contrast and Brightness: The relatively high reflectance of roofs and the spatially adjacent shadows lead to a high local contrast, therefore, the brightness is used as the initial input for building detection, where bright means brighter than the surrounding features.
- Size: Sizes for most of urban buildings are within a scale range.
- Directionality: Buildings are more isotropic than roads.
- Shape: Shape features (e.g., rectangularity, length-width ratio) can be used as constrained conditions of buildings.

The proposed MBI considers the above characteristics of buildings by integrating the multiscale and multidirectional morphological operators, which is detailed in the next section. Subsequently, the PanTex and the object-based approach used in the comparative study are briefly described. The experimental section presents the validation of MBI algorithm and analysis of its parameters. The final section contains concluding remarks and prospects for future work.

Morphological Building Index (MBI)

Mathematical morphology is an effective tool for extracting image components that are useful in the representation and description of region features. Opening and closing are two commonly used operators, used to remove bright (opening) or dark (closing) details. Morphological operators are applied to an image with a set of a known shape, called a structural element (SE). The processing results depend on the interaction between SE and the size of structures or features in the image. Details of the mathematical morphology can be found in Soille (2003). Some key properties of the morphological transformation that are applied to the building extraction in this study are summarized as follows.

- Reconstruction: The reconstruction filter is an important class of morphological filters that have been proven to be very useful for image processing since they do not introduce discontinuities, and therefore, preserve the shapes observed in input images (Pesaresi and Benediktsson, 2001).
- Granulometry: It describes the sizes and scales of objects in an image. Granulometrices have been introduced in remote sensing image classification of urban areas (Pesaresi and Benediktsson, 2001; Benediktsson *et al.*, 2003). The multiscale morphological features, such as morphological profiles (MPs) and differential morphological profiles (DMPs), are built based on the operators with increasing size of SE. (Pesaresi and Benediktsson, 2001).
- Directionality: Most of the existing morphological approaches referred to the disk-shaped SE (Pesaresi and Benediktsson, 2001; Benediktsson *et al.*, 2005; Chanussot *et al.*, 2006; Huang and Zhang, 2009). However, the disk SE does not consider the directional information, which is essential for discrimination between spectrally similar objects such as buildings and roads since buildings are isotropic but roads are relatively anisotropic.

In this paper, the MBI is defined by describing the characteristics of buildings (e.g., brightness, size, contrast, directionality, and shape) based on the aforementioned morphological transformation. Calculation of MBI is detailed as the following steps:

Step 1: Calculation of Brightness
The maximum of multispectral bands for each pixel is restored as the brightness value

$$b(x) = \max_{1 \le k \le K} (M_k(x)) \tag{1}$$

where $M_k(x)$ indicates the spectral value of pixel x at band k, and $b(x)$ represents its brightness. K is the total number of multispectral bands. The maximums of multispectral bands correspond to the features with high reflectance (candidate buildings), moreover, this step also reduces the number of spectral bands that are processed in the subsequent steps.

Step 2: Top-hat by Reconstruction
Top-hat is defined as the difference between an original image and its morphological opening. The opening of the brightness image b with a structural element s is defined based on two basic morphological operators (erosion ε and dilation δ):

$$\gamma^s(b) = \delta^s(\varepsilon^s(b)), \tag{2}$$

where s is the size of the structural element, and its unit is the number of pixels. The top-hat (TH) transformation is then defined as

$$\mathrm{TH}^s(b) = b - \gamma^s(b). \tag{3}$$

It should be noted that not all structures from the original image will be recovered when opening and top-hat filters are applied (Benediktsson *et al.*, 2005). Consequently, a reconstruction filter is used since this family of filters has

proven to have a better shape preservation than classical morphological filters (Pesaresi and Benediktsson, 2001; Benediktsson *et al.*, 2005). Accordingly, in this study, the top-hat transformation is performed in a reconstruction manner (Soille, 2003):

$$\mathrm{THR}^{s}(b) = b - \gamma^{s}_{\mathrm{RE}}(b) \tag{4}$$

where THR and γ_{RE} are top-hat by reconstruction and opening by reconstruction (OBR), respectively. It should be noted that the top-hat in this study refers in particular to the white top-hat transformation, which is chosen for building extraction because it is able to detect bright structures that have a size equal or less than a given scale parameter (length of SE) and remove other dark structures. The THR values reflect the difference of brightness between structures and their neighborhoods within the region of the structural element. Therefore, information regarding the contrast of the structures is contained in the THR features.

Step 3: Directional THR

In Equation 4, the commonly used disk-shaped SE is isotropic and does not contain directional information. Therefore, we propose to use linear SE (Soille and Talbot, 2001) for calculation of MBI since it is effective in extracting multidirectional and anisotropic features. The multidirectional information of THR is integrated using an average operator:

$$\overline{\mathrm{THR}}^{s}(b) = \underset{dir}{\mathrm{mean}}\left(\mathrm{THR}^{s.dir}(b)\right) \tag{5}$$

where *dir* represents the directionality of structural element. In Equation 5, the building candidates give larger feature values than other land-cover classes because buildings are relatively isotropic and have high top-hat values in all the directions.

Step 4: Granulometry by Top-hat

The multiscale THR is considered in this study because buildings in high-resolution images show complex spatial patterns with different sizes, shape, heights, and areas. The multiscale THR is built on the differential morphological profiles (Pesaresi and Benediktsson, 2001):

$$\begin{cases} \mathrm{THR}_{\mathrm{DMP}} = \left\{\mathrm{THR}^{s^{\min}}_{\mathrm{DMP}}, \ldots \mathrm{THR}^{s}_{\mathrm{DMP}}, \ldots \mathrm{THR}^{s^{\max}}_{\mathrm{DMP}}\right\} \\ \mathrm{THR}^{s}_{\mathrm{DMP}} = \left|\overline{\mathrm{THR}}^{s+\Delta s}(b) - \overline{\mathrm{THR}}^{s}(b)\right| \\ s^{\min} \le s \le s^{\max} \end{cases} \tag{6}$$

where Δs is the interval of the granulometry.

Step 5: Morphological Building Index (MBI)

In the above steps, the following characteristics of structures are addressed: brightness (Step 1), contrast (Step 2), directionality (Step 3), and size (Step 4). Therefore, MBI is defined based on these implicit features of buildings that are contained in the $\mathrm{THR}_{\mathrm{DMP}}$ histogram:

$$\mathrm{MBI} = \underset{s}{\mathrm{mean}}\ (\mathrm{THR}_{\mathrm{DMP}}). \tag{7}$$

The calculation of MBI is based on the observation that building structures have larger values in the $\mathrm{THR}_{\mathrm{DMP}}$ histogram, and hence a larger value of MBI means higher possibility to be a building structure.

Step 6: Postprocessing of MBI

The initial result of the building map is obtained by simply setting a threshold:

IF $\mathrm{MBI}(x) \ge$ t, THEN $\mathrm{map}_1(x) = 1$;
ELSE $\mathrm{map}_1(x) = 0$,

where $\mathrm{MBI}(x)$ and $\mathrm{map}_1(x)$ indicate the value of MBI and the initial label for pixel x, respectively. In order to refine the building map, the vegetation index and the shape attributes are then used to suppress noise from the initial result:

Rule 1: IF $\mathrm{map}_1(x) = 0$; THEN $\mathrm{map}(x) = 0$,
Rule 2: IF $\mathrm{map}_1(x) = 1$ AND ($\mathrm{NDVI}(x) \ge t1$)
OR $\mathrm{ratio}(x) \ge t2$ OR $\mathrm{area}(x) \le t3$)
THEN $\mathrm{map}(x) = 0$,

where $\mathrm{NDVI}(x)$ is the normalized difference vegetation index of pixel x, and $\mathrm{map}(x)$ denotes the final building map after the postprocessing. The shape attributes length-width ratio and area are computed based on the connected component analysis (Gonzalez and Woods, 2002) of the binary building map. The area is the number of pixels in an object, and the ratio is measured by comparing the maximum and minimum lengths of an object. The objective of the postprocessing is to reduce the commission errors by simultaneously removing the bright vegetation ($\mathrm{NDVI}(x) \ge t1$), narrow and elongated roads ($\mathrm{ratio}(x) \ge t2$), and small noises ($\mathrm{area}(x) \le t3$).

A sample test is presented here in order to demonstrate the processing steps of the proposed procedure (from Figure 1 to Figure 4). Figure 1a shows the brightness feature of the sample image, and Figure 1b is the ground truth map of buildings. Figure 2 presents the graphical examples of multidirectional and multiscale features used to calculate the MBI. The histograms of the top-hat DMP for the three land-cover classes are compared in Figure 3, where the DMP histograms are computed based on 6,148, 5,440, and 945 samples for buildings, backgrounds, and roads, respectively. These samples are manually chosen from the test image. For instance, 6,148 pixels of buildings are generated from the ground truth map (Figure 1b), and each pixel has a DMP histogram with 20-dimensional morphological channels (4 directions × 5 scales). The DMP

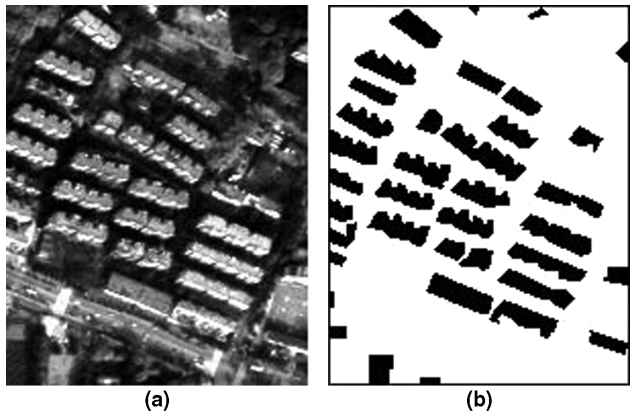


Figure 1. A test sample of GeoEye-1 image in the study area: (a) the brightness image, and (b) the ground truth map of buildings.

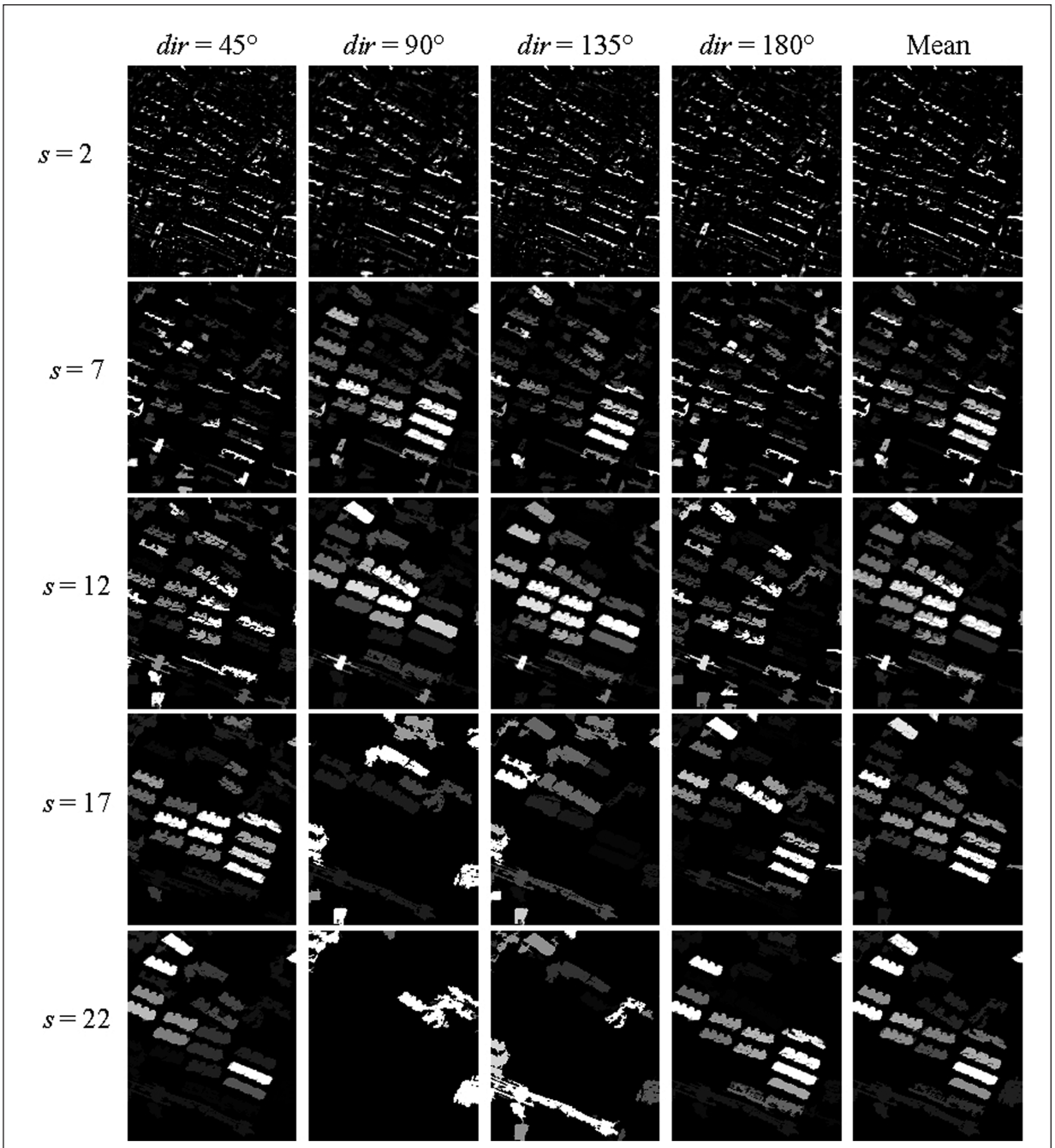


Figure 2. The multidirectional and multiscale morphological profiles of the THR transformation.

of buildings is then calculated by averaging the feature values of all the building samples for each channel. In Figure 3, SE(2,45) represents s = 2 and dir = 45 degrees. It can be seen that the three classes show strikingly different characteristics. Due to the isotropy of buildings, 16 out of the 20-dimensional morphological channels have feature values larger than 5. With respect to the roads, their DMP values are much smaller than buildings, and only three channels have a large value (greater than 5). Furthermore, due to the directionality of roads, most of their large values are with *dir* = 135 degrees (e.g., SE(12,135), SE(17,135), and SE(22,135)), corresponding to the actual direction of roads in the test image.

Figure 4 presents the graphical results of the automatic building extraction. Figure 4a shows the MBI feature image. Figure 4b and 4c are the building maps without and with the post-processing, respectively. It can be seen that the postprocessing is able to remove small noise in the background and narrow structures in the roads. In order to show the superiority of the anisotropic structural element (linear) over the isotropic one (disk), Figure 4d presents the building map resulted from the disk SE. It can be found that the disk structure is not effective for discrimination between buildings and roads since the multidirectional information is not taken into account.

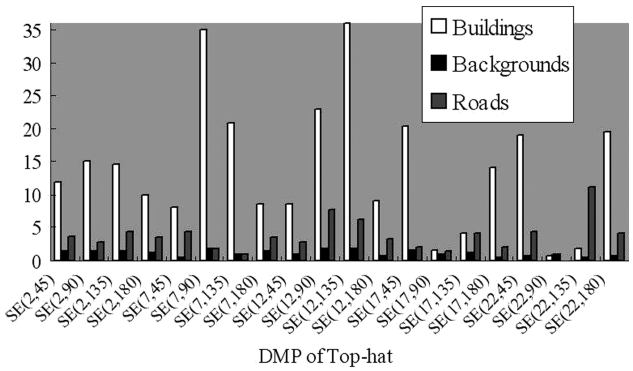


Figure 3. The multiscale and multidirectional histograms of the top-hat DMP for buildings, backgrounds, and the roads.

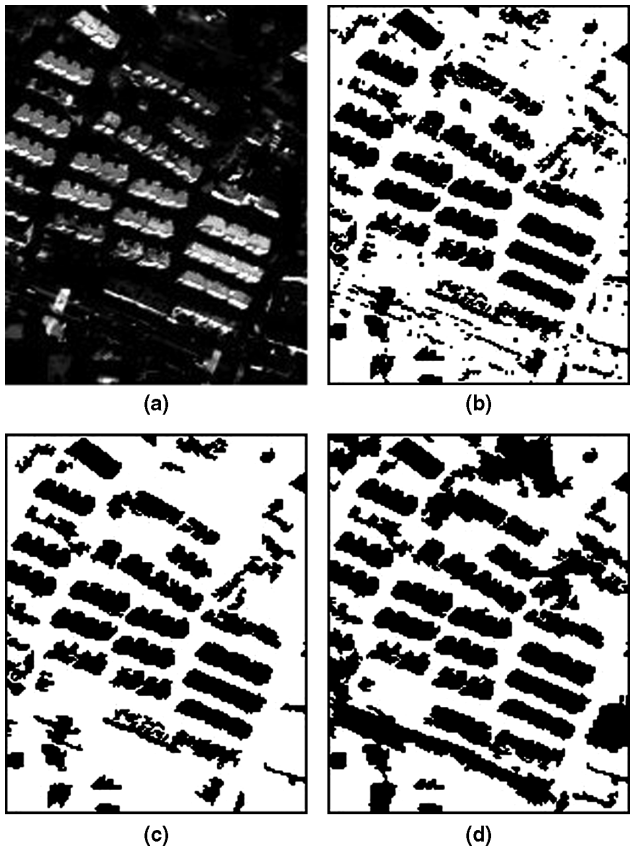


Figure 4. A sample test case for the procedure of automatic building extraction: (a) the MBI feature image, (b) and (c) the MBI-based results for automatic building extraction without and with the post-processing, respectively, and (d) the map resulted from an isotropic (disk) SE.

Comparisons

In order to evaluate the effectiveness of MBI, a recently developed procedure for the calculation of a texture-derived built-up presence index (Pesaresi *et al.*, 2008), namely PanTex, is carried out for comparison in this study. In addition, a multilevel object-based approach based on the well-known eCognition® platform is also implemented in the comparative study.

PanTex Index

PanTex is a fuzzy rule-based composition of anisotropic textural co-occurrence measures by the gray-level co-occurrence matrix (GLCM). The basic idea of the PanTex is to capture a basic structural characteristic of buildings, namely the fact that buildings cast a shadow and hence produce high local contrast. Therefore, the contrast measure of GLCM is used to represent the structural characteristics of built-up areas. PanTex is then calculated by integration of different directions and displacements of the contrast. Although the spatial resolution of 5 m was considered sufficient for discrimination of built-up areas using the PanTex (Pesaresi *et al.*, 2008), in this study, the index is calculated from the multispectral GeoEye-1 dataset with 2 m spatial resolution. Moreover, in experiments, different window sizes of the PanTex are discussed so as to adapt the index to the considered dataset. In addition, the same post-processing algorithm is used for the PanTex.

Object-based Approach

The basic idea of the object-based analysis is to group the spatially adjacent pixels into spectrally homogeneous objects, and then conduct classification on objects as the minimum processing units. The reason why the object-based approach has been widely employed for high-resolution image processing lies in that it is able to generate multi-features, such as spectral, textural, shape and contextual information, in different scales for each segment. In this study, the object-based building extraction is implemented on the commercial software eCognition®. Considering the complexity of the spatial patterns of buildings, two segmentation levels are generated by defining two scale parameters in the segmentation process (18 and 35). The scale parameter controls the size of segments, i.e., a small scale parameter will give rise to a small object size on average, while a large value will lead to a large object size on average. The procedure of multilevel object-based building extraction in this study is presented in Table 1.

All the four multispectral bands of GeoEye-1 image are used as the input for segmentation. Based on the boundaries of segments, several shape attributes are exploited for building extraction:

1. Length-width ratio is used to remove long and narrow objects (roads).
2. Shape index calculated by $\frac{\text{perimeter}}{4 \cdot \sqrt{\text{area}}}$ aims to describe the elongation of the objects. Road objects have larger perimeters and relatively small areas, hence they show larger shape

TABLE 1. THE MULTILEVEL OBJECT-BASED BUILDING EXTRACTION APPROACH BASED ON THE ECOGNITION® PLATFORM

Scale	Object of Interest	Decision Rule
35	Bare Soil	140<brightness<180; NDVI<0.04; area>2990; 1.1<length/width<4.5; shape index<2.5
	Roads	length/width>3.8; area>1200; NDVI<0.01
18	Vegetation	NDVI>0.054
	Buildings	brightness>136; rectangular fit >0.64; shape index<2.6; 1.16<length/width<5.8

index. Buildings with similar areas as roads should have smaller perimeters and hence show smaller shape index.

3. The rectangular fit is defined based on the creation of a rectangle with the same area as the considered object. It is measured by comparing the number of pixels inside the rectangle and the total number of pixels for the considered object. Buildings are potential to have large values of rectangular fit due to their rectangularity.

At the large scale (scale parameter = 35), some large objects such as bare soil and roads are detected and then removed from the image. At the small scale (scale parameter = 18), the vegetation is removed and the shape attributes are used to identify buildings. All the parameters are chosen manually according to the characteristics of the images.

Dataset and Study Area

The GeoEye-1 image used in this study was acquired on 22 December 2009, with four multispectral bands of 2.0 m spatial resolution. The study site covers Hongshan District of Wuhan in central China. As shown in the Figure 5, the study area is located besides the famous East Lake forming a special urban-lake landscape. The area measures approximately 5.5 km by 5.3 km and encompasses several land-cover types including forest, grassland, bare soil, buildings, roads, lakes. The satellite image shown in Figure 5 is the brightness image of the multispectral GeoEye-1 data. Four sub-regions with detailed field surveying are chosen for validation and comparison of different building extraction algorithms:

- Region A, Hongshan Square: a subway is currently under construction.
- Region B, Lake landscape: the fruit lake and the east lake are seated, respectively, at the left and right side of the bridge in the image.
- Region C, T-bridge: this region is characterized by the heavy traffic and the dense residential area.
- Region D, Wuhan University: it shows a regular pattern of buildings with rich green space.

TABLE 2. NUMBERS OF THE GROUND TRUTH SAMPLES FOR THE FOUR VALIDATION AREAS

	Validation Areas	No. of Samples	
		Buildings	Backgrounds
A	Hongshan Square	11,198	11,051
B	Lake landscape	14,332	14,068
C	T-bridge	15,650	15,301
D	Wuhan University	12,428	14,690

Based on the field surveying, over one hundred thousand samples of ground truth were selected for assessment of the algorithms. The numbers of the ground truth samples are listed in Table 2, where the backgrounds indicate the non-building areas.

Experiments

The experimental section includes two subsections: at first the proposed MBI algorithm was assessed based on the selected four validation areas, and at the same time a comparative study was conducted between the MBI, PanTex, and the object-based approach; second, the sensitivity of the MBI parameters, such as length and direction of the linear SE, setup of thresholds, was analyzed. Four statistical measures were used to evaluate the results of building extraction (Congalton and Green, 2009). Omission error (OE) and commission error (CE) denote the fractions of samples that are wrongly identified as backgrounds and that are wrongly classified as buildings,

Figure 5. The study area of Hongshan district of Wuhan in central China. The image is the brightness feature of multispectral GeoEye-1 data. The sub-regions (a), (b), (c), and (d) are the validation areas with detailed ground surveying.

respectively. Overall accuracy (OA) and Kappa coefficient for the building-background classification are computed based on the confusion matrix (Congalton, 1991).

Validation of MBI

The parameters of MBI were listed as follows.

1. Directionality: In Step 3, Equation 5, eight directions were considered ($D = 8$ stands for the total number of the directions), and the values of dir range from 22.5 to 180 degrees with an interval of 22.5.
2. Size of SE: In Step 4, Equation 6, the sizes of the linear SE were set to ($2 \leq s \leq 22$, $\Delta s = 5$) according to the spatial resolution of the image and the sizes of buildings in the study area.
3. Threshold: In Step 6, the threshold for the initial binaryzation of MBI was set to $t = 2$. Its influence was further discussed in the sensitivity analysis.
4. Postprocessing: In Step 6, the parameters of shape attributes were defined as $t1 = 180$ (after a 0 to 255 normalization of NDVI), $t2 = 9.6$, and $t3 = 30$. These parameters were tuned within a small area and then extended to the whole study area. They were kept constant in this study.

The binaryzation threshold of PanTex was set to 40 based on a careful tuning on the test images. The same post-processing algorithm was also performed on the PanTex for a fair comparison. In addition, different sizes of the analysis window for PanTex were discussed.

The accuracies of building detection for different algorithms are shown in Table 3, where (a), (b), (c), and (d) represent the four validation regions, respectively. The first comment to the table is that the proposed MBI can give substantially more accurate results than the PanTex and the object-based approach. Compared with PanTex (13 × 13), the improvements of OA achieved by MBI are 20.9 percent, 23.6 percent, 7.5 percent, and 27.4 percent for the four areas, respectively, while the improvements of OA are, respectively, 22.8 percent, 17.1 percent, 16.7 percent, and 9.4 percent, compared with the eCognition®-based approach. The improvements of Kappa coefficient are also significant. Moreover, MBI gives the lowest omission and commission errors.

The influence for the postprocessing of MBI is also analyzed in Table 3. Results show that the post-processing can reduce the commission errors effectively by removing bright vegetation and small noise. Commission errors decrease by 3.8 percent, 4.2 percent, 5.4 percent, and 9.2 percent for the four validation areas, respectively. On the other hand, comparable even lower omission errors are achieved (Table 3b and 3c) when the postprocessing is performed. Therefore, it can be stated that the post-processing is effective in refining the initial result of MBI. In addition, Table 3 provides the information about the window sizes of PanTex. It is found that although the PanTex with a larger window size gives better accuracies for OA and Kappa, it leads to higher commission errors at the same time.

The extracted building maps of the four validation areas are shown in Figure 6 for a visual inspection. Rows 1 and 2 are the brightness and the MBI feature images, respectively. Row 3 indicates the manually delineated ground truth maps. Rows 4, 5, and 6 show the results of MBI, eCognition® and PanTex (13×13), respectively. The comments to the figure are summarized as follows.

1. The results provided by the object-based approach are acceptable, however, the shape attributes extracted based on the boundaries of objects are not fully effective for discrimination between bare soil, roads, and buildings due to the inaccurate segmentation (e.g., over- and under-segmentation). The inaccuracy of segmentation leads to the fact that it is difficult to choose suitable thresholds for identification of a building object. For instance, in the area of T-bridge, the bright bare soils in the center of the region were wrongly identified as buildings, at the same time some buildings had been filtered out when roads and other noise were removed.
2. Although PanTex is able to predict the presence of buildings, it fails to accurately locate their boundaries because the pre-defined window size may not favor all the building structures with different scales.
3. In general, the proposed MBI algorithm provides the most accurate results according to the ground truth. Roads and buildings are effectively separated due to the anisotropy of MBI, and buildings with different sizes are considered using the multiscale profiles.

TABLE 3. STATISTICAL ACCURACIES OF BUILDING EXTRACTION FOR MBI, PANTEX, AND OBJECT-BASED APPROACH (POST- = POSTPROCESSING)

(A) VALIDATION REGION = HONGSHAN SQUARE

Method		Accuracies			
		OE (%)	CE (%)	OA (%)	Kappa
MBI	Without post-	10.9	6.2	91.5	0.830
	With post-	11.4	2.4	93.1	0.862
PanTex	11 by 11	52.8	15.8	68.9	0.380
(With post-)	13 by 13	41.6	18.8	72.2	0.446
eCognition®	NA	40.1	23.8	70.3	0.407

(B) VALIDATION REGION = LAKE LANDSCAPE

Method		Accuracies			
		OE (%)	CE (%)	OA (%)	Kappa
MBI	Without post-	7.3	6.3	93.1	0.862
	With post-	6.9	2.1	95.5	0.910
PanTex	11 by 11	43.7	18.2	71.5	0.432
(With post-)	13 by 13	41.4	19.4	71.9	0.440
eCognition®	NA	38.8	6.0	78.4	0.569

(C) VALIDATION REGION = T-BRIDGE

Method		Accuracies			
		OE (%)	CE (%)	OA (%)	Kappa
MBI	Without post-	8.1	7.4	92.1	0.843
	With post-	7.6	2.0	95.1	0.903
PanTex	11 by 11	29.9	4.6	83.1	0.663
(With post-)	13 by 13	18.4	6.7	87.6	0.753
	17 by 17	13.5	7.1	89.0	0.780
eCognition®	NA	37.6	7.4	78.4	0.569

(D) VALIDATION REGION = WUHAN UNIVERSITY

Method		Accuracies			
		OE (%)	CE (%)	OA (%)	Kappa
MBI	Without post-	17.2	15.6	85.0	0.699
	With post-	18.2	6.4	89.0	0.777
PanTex	9 by 9	79.1	30.5	59.5	0.138
(With post-)	11 by 11	71.5	31.4	61.1	0.180
	13 by 13	67.5	33.0	61.6	0.196
eCognition®	NA	34.0	13.6	79.6	0.581

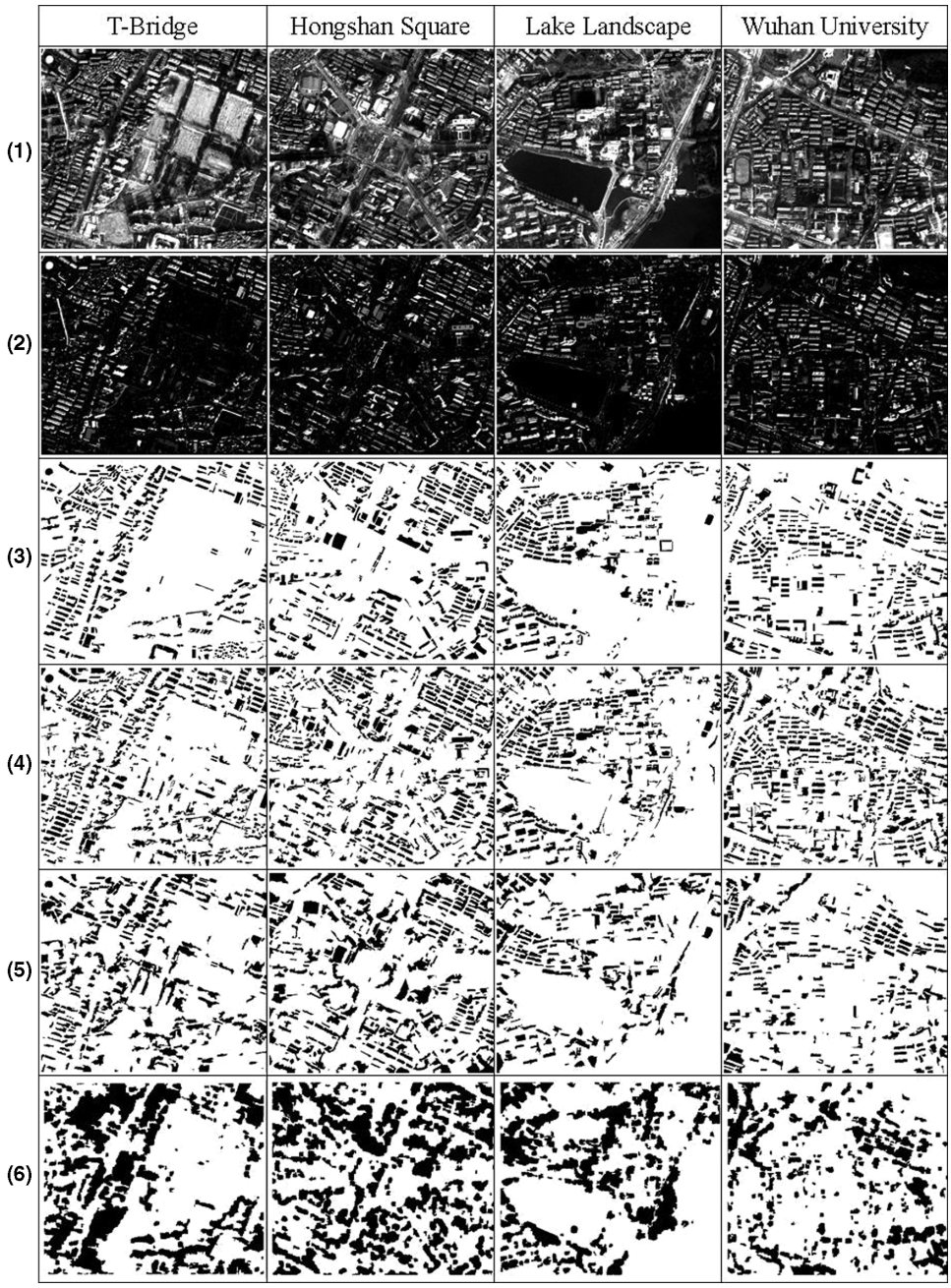
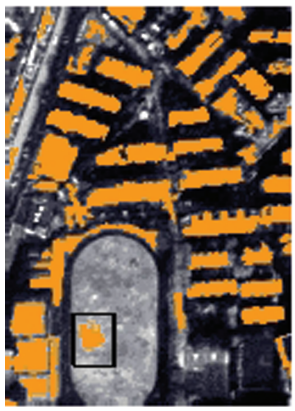
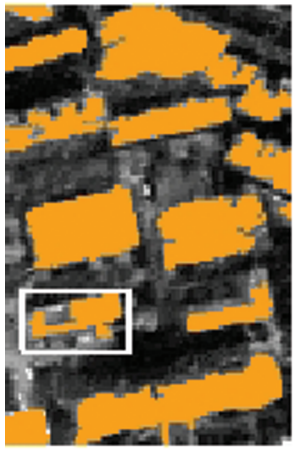


Figure 6. Results of building detection for the four validation areas (white = backgrounds and black = buildings). Rows 1, 2, and 3 represent brightness, MBI, and the manually delineated ground truth maps, respectively. Rows 4, 5, and 6 represent the building maps extracted by MBI, eCognition® and PanTex (13 × 13), respectively.

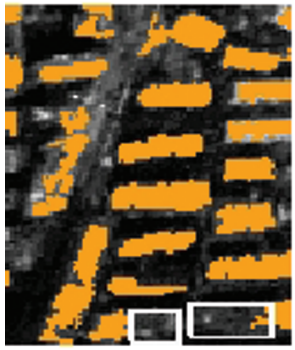
In order to give a comprehensive conclusion, in Plate 1 some images are taken as examples to show the drawbacks of MBI. Plate 1 – Case A shows a commission error, where an object of bare soil (in the playground) is wrongly identified as a building since they have similar brightness, structure and shape attributes. Plate 1– Case B is an omission error showing that a building is partially identified due to the heterogeneity of the roof. Plate 1 – Case C is also an omission error, where dark and shadowed roofs are partially recognized since the calculation of MBI depends on the locally bright structures. However, it should be kept in mind that it is not easy to automatically extract buildings from a complicated urban landscape in a high-resolution image, especially without any supervised machine learning. Therefore, MBI can be considered potential for accurate estimation of presence of buildings from large amount of high-resolution images. The results of the whole study area are shown in Figure 7.



Case (A) Commission Error: An object of bare soil is wrongly identified as a building structure.



Case (B) Omission Error: A roof is partially identified due to the heterogeneity of the roof.



Case (C) Omission Error: Dark and shadowed roofs are partially identified.

Plate 1. Three sample images used to show the uncertainties of MBI. The buildings extracted by MBI are overlaid on the brightness images. The extracted buildings are highlighted in orange.

Directionality of MBI

Directionality is a key parameter of MBI since multidirectional structural element (SE) is able to represent the anisotropy of objects and discriminate roads and buildings. In order to verify the effectiveness of the multidirectional SE, the isotropic disk-shaped SE was used for a comparison. The sizes of the disk SE were set to the same values as the linear one: ($2 \leq s \leq 22$, $\Delta s = 5$). Results are reported in Table 4, from which we can obtain two interesting observations about the directionality of MBI.

1. The results of the linear SE are much higher than those of the disk one. Due to the consideration of anisotropy, the linear SE is able to separate roads (narrow and elongated) and buildings (rectangular). An example has been shown in Figure 4. Most of the road structures are removed by the linear SE (Figure 4c), but are retained in the result of the disk SE (Figure 4d).
2. There is not significant difference between the four-directional ($D = 4$) and the eight-directional ($D = 8$) SE. This phenomenon signifies that the four-directional MBI is adequate for estimation of presence of buildings; moreover, it is able to reduce the computational cost and hence will benefit to process large amount of images.

Sizes of the Structural Element

This objective of this subsection is to analyze the effect of the sizes of SE, and give comments on how to select them. To this aim, different sizes of SE are used for calculation of MBI, and results are compared in Figure 8. The vertical axis represents the overall accuracy, and the horizontal one stands for the maximum size of SE (s^{max} in Equation 6). For instance, "$s^{max} = 27$" means that the structural elements of $s = (2, 7, 12, 17, 22, 27)$ are used for construction of morphological profiles. From the figure, it can be seen that the horizontal axis is divided into three sections:

1. $7 \leq s^{max} \leq 17$, where accuracies increase substantially.
2. $22 \leq s^{max} \leq 32$, where accuracies are stably high.
3. $s^{max} \geq 37$, where accuracies decrease slowly.

According to the properties of the THR transformation defined in Equation 4, the selection of lengths of SE depends on the sizes of buildings in the image. The building sizes in the study area range from 6 m (e.g., informal settlements) to 60 m (e.g., large apartments), equivalent to 3 and 30 pixels ($s^{max} = 3$ and $s^{max} = 30$), respectively. Therefore, high accuracies are obtained with $22 \leq s^{max} \leq 32$ due to the correspondence to the actual sizes of buildings. Actually, from Figure 8, it can be observed that the overall accuracies are relatively robust to the sizes of SE, since the results are still acceptable when the size is very large.

Threshold of MBI

Figure 9 presents the effect of the threshold of MBI (t in the postprocessing). The validation area of T-bridge is used for this discussion. OE and CE are omission and commission errors, respectively. Overall accuracy (OA) and Kappa coefficient indicate the accuracies of buildings-backgrounds classification. Figure 9 shows that the most accurate results are obtained with $t = 1.5$, 2, and 3. Subsequently, accuracies decrease gradually when the values of t become larger. A large value of t signifies that more building candidates are removed in the initial binaryzation, resulting in a large omission error and small commission error. The figure also shows the proportion of building pixels to the whole test image. It should be noted that the ratio of building pixels to the whole image is 13.8 percent in the ground truth reference (50,189 building pixels in a 564×645 test image), which is approximately equivalent to the case of $t = 4$ (51,115 building pixels). Therefore, the reasonable range of t should be smaller than 4 ($t \leq 4$) in order to avoid the omission errors. In addition, it can be seen that $t = 0.1$ gives satisfactory results (OA = 93.2 percent and Kappa = 0.864), which shows that although all the structures with nonzero MBI values are detected as buildings, a high detection accuracy can be still achieved.

Conclusions

In this paper, a novel morphological building index (MBI) is proposed for automatic building extraction from

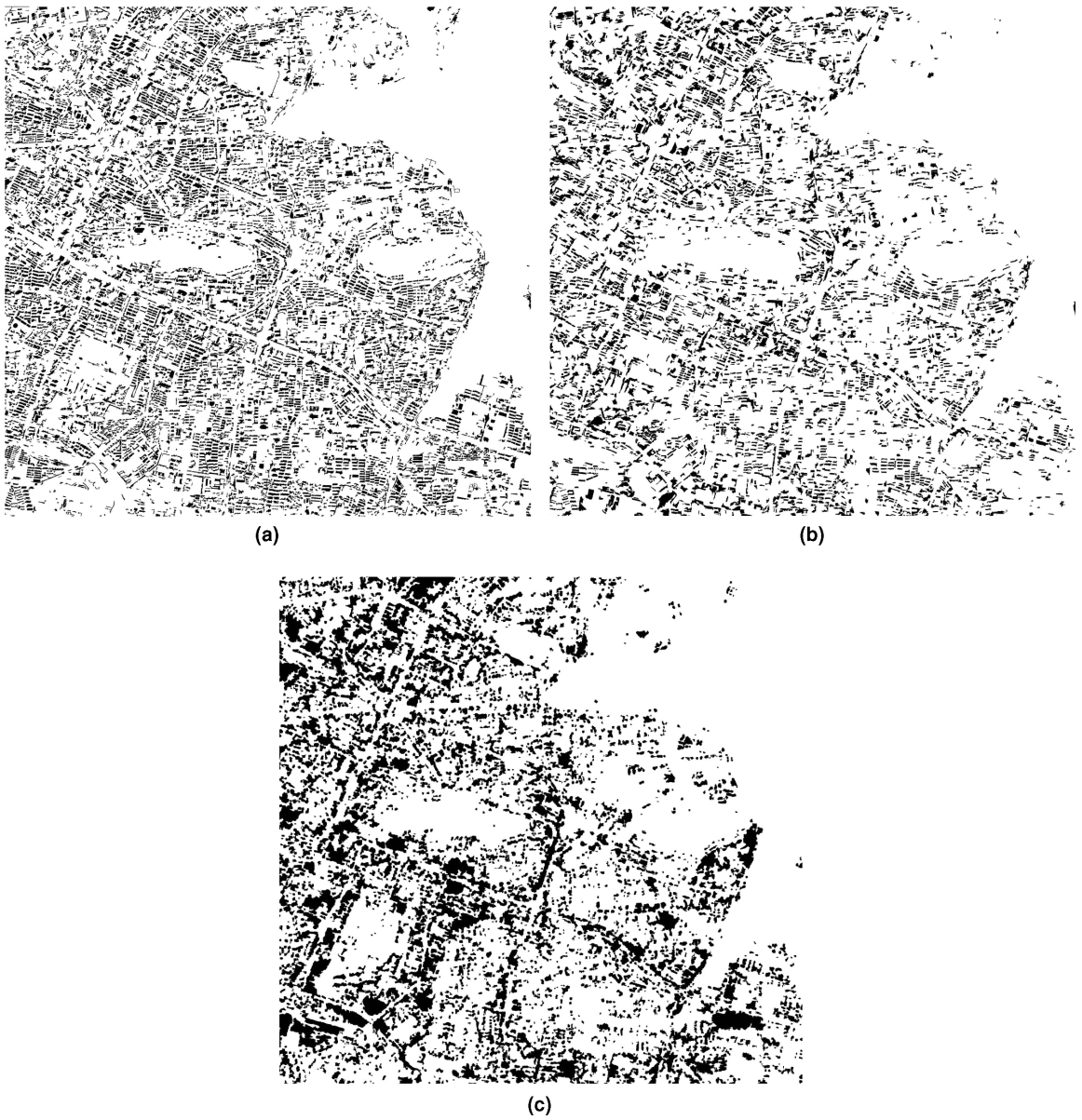


Figure 7. Building maps of the whole study area with results of: (a) MBI, (b) object-based approach, and (c) PanTex.

TABLE 4. COMPARISON OF THE LINEAR SE (ANISOTROPIC) AND THE DISK ONE (ISOTROPIC); D IS THE NUMBER OF DIRECTIONS OF THE LINEAR SE

Area	Structural Element	Accuracies			
		OE (%)	CE (%)	OA (%)	Kappa
Hongshan Square	Linear ($D = 4$)	11.8	2.2	93.0	0.860
	Linear ($D = 8$)	11.4	2.4	93.1	0.862
	disk	7.5	13.2	89.0	0.781
Lake landscape	Linear ($D = 4$)	6.3	1.9	95.8	0.917
	Linear ($D = 8$)	6.9	2.1	95.5	0.910
	disk	4.0	34.8	72.0	0.439
T-bridge	Linear ($D = 4$)	8.3	2.8	94.4	0.888
	Linear ($D = 8$)	7.6	2.0	95.1	0.903
	disk	6.0	23.6	82.2	0.643
Wuhan University	Linear ($D = 4$)	19.0	7.7	88.1	0.759
	Linear ($D = 8$)	18.2	6.4	89.0	0.777
	disk	7.4	35.8	72.9	0.472

high-resolution remotely sensed imagery. A multispectral GeoEye-1 image, covering a study area of 5.5 km by 5.3 km, in the Hongshan district of Wuhan (central China) was used in experiments. The contribution of this study lies in the following items:

1. The MBI is modeled by relating the implicit characteristics of buildings such as brightness, size, contrast, directionality and shape, to the properties of morphological transformation such as reconstruction, granulometry, and directionality.
2. A comparative study was made between the MBI, the PanTex (a recently proposed built-up index), and the object-based approach. Results verified that the MBI gave better results than other algorithms in terms of both statistical accuracies and visual inspection.
3. The experiments were conducted on the multispectral GeoEye-1 imagery. To our knowledge, few studies have been reported for information extraction from the GeoEye-1

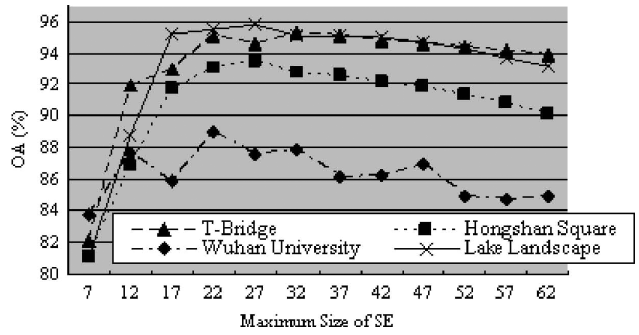


Figure 8. Relationship between the overall accuracies and the sizes of SE in the four validation areas.

imagery. Therefore, it is worthwhile evaluating the effectiveness of building extraction from this new data type.

The parameters of MBI were also analyzed in the experiments and conclusions were summarized as follows:

1. Directionality of MBI: Results verified that the anisotropic SE (linear) was able to separate roads and buildings and outperformed the isotropic SE (disk). It was also revealed that the number of directions for the linear SE ($D = 4$ and $D = 8$) did not affect the accuracies significantly.
2. Sizes of the SE: The experiment showed that high accuracies were obtained when the sizes of SE were consistent with the actual sizes of buildings in the image.
3. Binaryzation threshold of MBI: This parameter controls the balance of commission and omission errors. According to our analysis, a small value is suggested for the threshold t in order to avoid a high omission error.

In this paper, the drawbacks of MBI were discussed in Plate 1. The basic assumption of MBI, i.e., a building is a bright structure due to the high reflectance of roofs and has high contrast due to the adjacent shadows, is not always true, especially in a complicated urban landscape. However, it should be noted that it is not an easy task to automatically extract buildings from high-resolution images without any supervised learning. Therefore, it can be stated that MBI is effective for indication of buildings and will benefit to process large amount of images.

Future research includes its further validation using different high-resolution images. Moreover, MBI is potential to estimate sizes and directions of buildings based on the morphological profiles. The application of MBI to urban monitoring and landscape analysis is also planned.

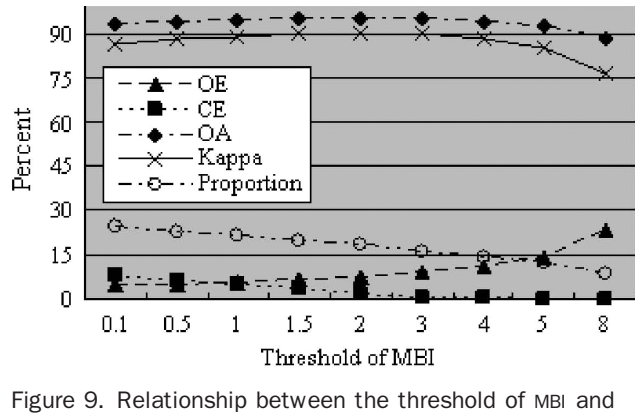


Figure 9. Relationship between the threshold of MBI and the results of building extraction.

Acknowledgments

This work was supported by the National Science Foundation of China under Grant No. 40930532, the Fundamental Research Funds for the Central Universities under Grant No. 3101016, and the LIESMARS Special Research Funding.

References

Benediktsson, J.A., M. Pesaresi, and K. Arnason, 2003. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations, *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):1940–1949.

Benediktsson, J.A., J.A. Palmason, and J. R. Sveinsson, 2005. Classification of hyperspectral data from urban areas based on extended morphological profiles, *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):480–491.

Bian, L., 2003. Retrieving urban objects using a wavelet transform approach, *Photogrammetric Engineering & Remote Sensing*, 69(2):133–141.

Chanussot, J., J.A. Benediktsson, and M. Fauvel, 2006. Classification of remote sensing images from urban areas using a fuzzy possibilistic model, *IEEE Geoscience and Remote Sensing Letters*, 3(1):40–44.

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

Congalton, R.G., and K. Green, 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Second edition, CRC Press, Boca Raton, Florida 208 p.

Gamba, P., F. Dell'Acqua, G. Lisini, and G. Trianni, 2007. Improved VHR urban mapping exploiting object boundaries, *IEEE Transactions on Geoscience and Remote Sensing*, 45(8):2676–2682.

Gonzalez, R.C., and R.E. Woods, 2002. *Digital Image Processing*, Second edition, Prentice Hall, Upper Saddle River, New Jersey, 793 p.

Huang, X., L. Zhang, and P. Li, 2007. Classification and extraction of spatial features in urban areas using high-resolution multispectral imagery, *IEEE Geoscience and Remote Sensing Letters*, 4(2):260–264.

Huang, X., and L. Zhang, 2008. An adaptive mean-shift analysis approach for object extraction and classification from urban hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 46(12):4173–4185.

Huang, X., and L. Zhang, 2009. A comparative study of spatial approaches for urban mapping using hyperspectral ROSIS images over Pavia City, northern Italy, *International Journal of Remote Sensing*, 30(12):3205–3221.

Jin, X., and C.H. Davis, 2005. Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information, *EURASIP Journal on Applied Signal Processing*, 14:2196–2206.

Lee, D.S., J. Shan, and J.S. Bethel, 2003. Class-guided building extraction from Ikonos imagery, *Photogrammetric Engineering & Remote Sensing*, 69(2):143–150.

Lhomme, S., D.C. He, C. Weber, and D. Morin, 2009. A new approach to building identification from very-high-spatial resolution images, *International Journal of Remote Sensing*, 30(5):1341–1354.

Lu, Y.H., J.C. Trinder, and K. Kubik, 2006. Automatic building detection using the Dempster-Shafer algorithm, *Photogrammetric Engineering & Remote Sensing*, 72(4):395–403.

Pacifici, F., M. Chini, and W.J. Emery, 2009. A neural network approach using multi-scale textural metrics from

very high-resolution panchromatic imagery for urban land-use classification, *Remote Sensing of Environment*, 113(6):1276–1292.

Pesaresi, M., and J.A. Benediktsson, 2001. A new approach for the morphological segmentation of high resolution satellite imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 39(2):309–320.

Pesaresi, M., A. Gerhardinger, and F. Kayitakire, 2008. A robust built-up area presence index by anisotropic rotation-invariant textural measure, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1(3):180–192.

Soille, P., 2003. *Morphological Image Analysis: Principles and Applications*, Springer-Verlag Berlin, Germany, 391 p.

Soille, P., and H. Talbot, 2001. Directional morphological filtering, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1313–1329.

Tian, J., and D.M. Chen, 2007. Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition, *International Journal of Remote Sensing*, 28(20):4625–4644.

(Received 11 June 2010; accepted 20 January 2011; final version 01 February 2011)