This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Urban Building Density Estimation From High-Resolution Imagery Using Multiple Features and Support Vector Regression

Tao Zhang, Xin Huang, *Senior Member, IEEE*, Dawei Wen, and Jiayi Li, *Member, IEEE*

***Abstract*—Building density is considered as an environmental parameter of great significance for urban management, and it can also help us to understand how cities function. However, in developing countries, building density information is often lacking or incomplete in a large number of cities. To address this problem, this paper presents a method for the quantitative estimation of building density using a support vector regression model to establish the relationship between building density and the features extracted from the image, including spectral, morphological, and textural features. The importance and relevance of the different features is investigated by a feature selection method named recursive feature elimination. Tests performed on three representative megacities in China confirm that the proposed method achieves satisfactory results for building density estimation in terms of the low root-mean-square errors (∼0.05) and visual effect. It is also found that the combination of feature sets outperforms the single feature set, indicating the need for simultaneous consideration of all the categories of features. The knowledge of the most informative features provides an insight into selecting the most effective parameters for building density mapping. The investigation of the transferability between geographical regions confirms that the self-training achieves the better performance compared to the transfer training. The results with regard to the regression accuracy, feature selection, and the comparison with other methods indicate the effectiveness of the multifeature approach when applied to building density estimation.**

***Index Terms*—Building density, feature selection, mathematical morphology, remote sensing, support vector regression (SVR), texture analysis.**

I. Introduction

As THE world population and economy grow, the high rate of urbanization is resulting in vast urban expansion and some severe dilemmas, such as urban congestion, air pollution, and urban heat island [1], [2]. In this context, the geostatistical parameters of a city, e.g., urban extent, human settlements, and urban building density [3]–[5], have become essential for urban planning and environmental management. Building density, which indicates the ratio of the coverage of the buildings' footprints to the size of the area of interest [6], [7], is a crucial urban environmental parameter, which is more effective at evaluating disaster risk than building/nonbuilding dichotomy [6], [8]. Thus, building density has been the subject of great attention from many stakeholders, such as urban planners and real estate agents. However, in developing countries, building density information is often incomplete or unavailable in many cities [9], especially in the newly developed urban areas. Moreover, traditionally, building density information production relies on field investigation and manual delineation, which is extremely time consuming, and is not conducive to data updating. In view of this, the efficient generation and renewal of building density information is urgently needed in the related research community. Fortunately, high-resolution remote sensing data, as an economical and commercially available information source, have the promising potential to deal with such a task, considering their excellent ability to describe detailed urban land-cover types.

In the related studies, active remote sensing data have been utilized for building density mapping. The light detection and ranging remote sensing technology, which is able to provide accurate surface elevation measurement and directly extract building footprints for precise calculation of building density, is limited by the extraordinary cost of large-scale data collection and processing [9], [10]. Synthetic aperture radar images can be used to estimate building density by considering their polarization information [6] or textural features [11], but the complex scattering mechanisms in urban areas may reduce the estimation accuracy [12], and the building density parameter may not be fully characterized when just using textural information. With the improvement of remote sensors, optical very high-resolution (VHR) remotely sensed imagery can provide us with rich geospatial details, which opens up a new possibility for building density estimation. To data, most of the previous studies have focused on urban area extraction [13]–[17] or building detection [18], [19], leading to dichotomous classification, with no quantitative or detailed building density information. One simple method for building density estimation is based on building detection. However, this procedure is completely

Manuscript received October 28, 2016; revised February 9, 2017; accepted February 9, 2017. This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFB0501403, in part by the China Science Fund for Excellent Young Scholars under Grant 41522110, and in part by the Foundation for the Author of National Excellent Doctoral Dissertation of PR China under Grant 201348. *(Corresponding authors: Xin Huang and Dawei Wen.)*

T. Zhang and D. Wen are with theState Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: zhangtao437@163.com; daweiwen@whu.edu.cn).

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).

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.2017.2669217

1939-1404 © 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

dependent on the accuracy of the building extraction, and the accurate detection of buildings remains a big challenge since a number of urban structures can exhibit similar spectral properties, e.g., roads, soil, and other man-made urban facilities [20]. Moreover, due to the viewing angle of the imaging sensor, it is often very difficult to obtain the real footprints of buildings in urban areas from remote sensing imagery for accurate building density estimation. As an alternative to the approaches based on building detection, we simultaneously consider the spectral, morphological, and textural features to build a regression model to directly and accurately estimate the building density in this study. The features obtained from the VHR imagery can provide a sophisticated description of buildings and building clusters, which leads to more accurate building density estimation.

The image spectrum, which serves as the fundamental attribute of the image, has the ability to record the reflectance of objects in different spectral bands. In general, buildings exhibit relatively high signals in the visible bands [20], while vegetation shows high response values in the near-infrared band [21]. Therefore, spectral features, including multispectral bands as well as spectral differences [22], have the potential to represent both buildings and nonbuilding areas for building density estimation. Mathematical morphology, based on set theory, has been successful in dealing with a number of problems in remote sensing (e.g., classification and segmentation) [23]. Morphological features are able to effectively capture the basic geometric and structural characteristic of local buildings due to the fact that buildings are generally bright structures and cast shadows producing high local contrast [24]. The cooccurrence relationship of buildings and shadows can be characterized by differential morphological profiles (DMPs) [25], owing to their ability to highlight both bright and dark structures. Accordingly, DMPs have a great potential for building density estimation. Texture, which is the term used to characterize the tonal variation in an image, has played an important role in image processing and interpretation. In this study, we consider textural measures derived from the gray-level cooccurrence matrix (GLCM) [26], which has been demonstrated to be very efficient for analyzing the textural information of remote sensing imagery [27]. Textural features are capable of describing building clusters at a macrolevel, since buildings and their complex surroundings, e.g., roads, open spaces, and trees, produce high spatial heterogeneity [27], which implies that textural measures could also be utilized to delineate building density. Overall, these three feature sets (i.e., spectral, morphological, and textural features) are jointly considered, resulting in an integrated characterization of building density.

In this study, the relationship between building density and the multiple features extracted from the remote sensing imagery is investigated. Support vector regression (SVR) is employed to estimate the continuous parameter of building density in this paper due to its great ability to handle complex and nonlinear data distributions in high-dimensional feature spaces [28]. Considering that a large number of features may lead to information redundancy, both feature selection and the analysis of the relevance of the features are essential when encountering such input spaces. In this study, the importance and relevance of the features is assessed by a feature selection method named "recursive feature elimination" (RFE) [29], a feature selection algorithm designed for SVR, which calculates the changes in the decision function as a criterion for the feature ranking, and performs feature selection in the process of training [30]. Thus, RFE tends to find features better suited to the predetermined SVR model. Both the validity and transferability of the regression model are ascertained on representative images, and the proposed method is also compared with other building density estimation approaches.

The main contributions of this study lie in the following aspects: 1) building density, one of the most important urban environmental parameters, is effectively estimated using remote sensing imagery, which can help us to understand how cities function; 2) we propose a new framework, which simultaneously considers the spectral, morphological, and textural features of human settlements to comprehensively describe urban building density; and 3) the impact of the features is extensively analyzed, giving an insight into the most informative and efficient features on building density estimation.

The remainder of this paper is organized as follows. Section II specifically describes the building density estimation method. The datasets and experimental results are presented in Section III and discussed in Section IV, followed by the conclusion in Section V.

II. Methodology

The framework of the proposed method for building density estimation is illustrated in Fig. 1. First, three categories of features, including spectral, morphological, and textural features, are derived from high-resolution remote sensing imagery. It is expected that their combination can effectively indicate the building density. Subsequently, the regression model is trained and built by measuring the relationship between these features and reference building density values. Finally, the regression model is used to predict the building density of the whole image.

A. Input Features for Building Density Estimation

In this section, three categories of features (see Table I) are taken into account to describe building density: 1) the spectral features are used to record the reflectance of buildings and nonbuildings; 2) the morphological features are employed to delineate local building structures; and 3) the textural features are considered to describe the tonal variations of building clusters. Subsequently, the SVR model is used to build the relationship between these feature sets and the building density. Note that this framework is designed for images containing visible and near-infrared bands.

1) Spectral Feature: The spectral feature is the fundamental property of remotely sensed imagery. In general, buildings appear as bright structures with high reflectance in the visible bands [27], and vegetation (nonbuilding) presents correspondingly high values in the near-infrared band due to its biophysical property [21]. Furthermore, the spectral difference [22] can provide us with enhanced information about the targets (e.g., the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

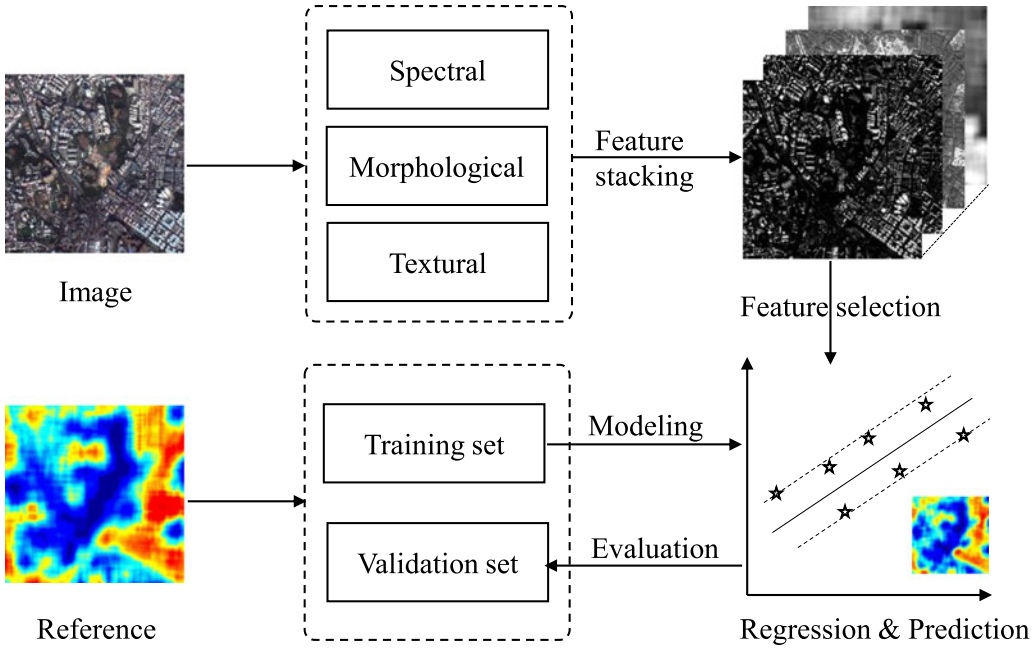


Fig. 1. Flowchart of the proposed method for building density estimation.

TABLE I
FEATURES USED IN THE MODEL

Feature type	Feature name	Feature description
Spectral	Blue band	Blue band of the image
	Green band	Green band of the image
	Red band	Red band of the image
	NIR band	Near-infrared band of the image
	n(NIR−Red)	$\mathrm{n(NIR-Red)} = \frac{\mathrm{NIR-Red}}{\mathrm{NIR+Red}}$
	n(NIR−Green)	$\mathrm{n(NIR-Green)} = \frac{\mathrm{NIR-Green}}{\mathrm{NIR+Green}}$
	n(NIR−Blue)	$\mathrm{n(NIR-Blue)} = \frac{\mathrm{NIR-Blue}}{\mathrm{NIR+Blue}}$
	n(Red−Green)	$\mathrm{n(Red-Green)} = \frac{\mathrm{Red-Green}}{\mathrm{Red+Green}}$
	n(Red−Blue)	$\mathrm{n(Red-Blue)} = \frac{\mathrm{Red-Blue}}{\mathrm{Red+Blue}}$
	n(Green−Blue)	$\mathrm{n(Green-Blue)} = \frac{\mathrm{Green-Blue}}{\mathrm{Green+Blue}}$
Morphological	$\mathrm{DMP_{WTH}}$	$\mathrm{DMP_{WTH}} = \{\mathrm{DMP_{WTH}}(s,\mathrm{dir}) : s_{\min} \le s \le s_{\max}, \mathrm{dir} \in D\}$
		$\mathrm{DMP_{WTH}}(s,\mathrm{dir}) = \lvert\mathrm{WTH}(s+\Delta s,\mathrm{dir}) - \mathrm{WTH})s,\mathrm{dir})\rvert$
	$\mathrm{DMP_{BTH}}$	$\mathrm{DMP_{BTH}}(s,\mathrm{dir}) = \lvert\mathrm{BTH}(s+\Delta s,\mathrm{dir}) - \mathrm{BTH})s,\mathrm{dir})\rvert$
		$\mathrm{DMP_{BTH}} = \{\mathrm{DMP_{BTH}}(s,\mathrm{dir}) : s_{\min} \le s \le s_{\max}, \mathrm{dir} \in D\}$
Textural	Mean	Average of gray level
	Variance	Gray-level variance
	Homogeneity	$\mathrm{Homogeneity} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{p(i,j)}{1+(i-j)^2}$
	Contrast	$\mathrm{Contrast} = \sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j)-(i-j)^2$
	Dissimilarity	$\mathrm{Dissimilarity} = \sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j)\cdot\lvert i-j\rvert$
	Entropy	$\mathrm{Entropy} = -\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j)\cdot\log(p(i,j))$
	Second moment	$\mathrm{Secondmoment} = \sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j)^2$
	Correlation	$\mathrm{Correlation} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{(i\cdot j)\cdot p(i,j)-\mu_i\cdot\mu_j}{\sigma_i\cdot\sigma_j}$

vegetation information can be enhanced by the difference between the near-infrared band and the red band). Thus, the multispectral bands as well as the spectral differences are employed in this study since they have the ability to characterize both building and nonbuilding areas for building density estimation.

2) Morphological Feature: Mathematical morphology has been demonstrated to be an effective tool for the extraction of the geometrical structures of objects [24], [30]. Opening and closing are two commonly used operators, removing bright (compared to the surroundings) and dark structures of an image,

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

respectively. Top-hat transformation, which is defined as the difference between an original image and its morphological opening or closing, can be used to highlight the bright (white top-hat) or dark (black top-hat) structures of an image.

The morphological feature is able to delineate the spectral–structural characteristics of buildings (e.g., size, contrast) since buildings are generally bright structures of a certain size, and the spatially adjacent shadows lead to high local contrast [20]. DMPs of the white and black top-hat transformation have the ability to represent bright and dark structures, respectively, and can be used to describe the cooccurrence relationship between buildings and their cast shadows. Consequently, DMPs of the white and black top-hat transformation are considered due to their great potential for building density estimation.

3) Textural Feature: Texture refers to the visual effect produced by tonal variations over relatively small areas [31]. With the improvement of the spatial resolution of remote sensing images, spectral features have been found to be inadequate for the discrimination of spectrally similar classes in high-resolution imagery, such as buildings, roads, and parking lots. In this context, the addition of the textural feature provides supplementary information for the images. Thus, texture analysis is becoming increasingly important in remote sensing image processing and interpretation [31].

Urban areas are composed of diverse scene elements (e.g., buildings, roads, and open areas) with complex spatial patterns. A variety of materials (such as concrete, asphalt, plastic, glass, metal, etc.) are used to build houses, commercial buildings, and recreational areas. Buildings and their surroundings, such as trees, roads, public infrastructures, and open spaces, lead to tonal variations in remote sensing images. The spatial heterogeneity, as exhibited by urban areas and building clusters, can be exactly characterized by texture information at the macrolevel. Thus, textural features are taken into account since they have great potential to describe building density. In this study, two first-order and six second-order textural features are derived from the GLCM [26], which is a popular statistical algorithm for texture measurement. The textural measures are computed with the window size corresponding to the spatial scale used to calculate the building density.

Some representative features for a typical area with various land-use types (e.g., park, residential area) are presented in Fig. 2. It can be been seen that n(NIR−Red) (see the feature description in Table I) denotes the vegetation component, and vegetation is regarded as nonbuilding areas with low building density. $\mathrm{DMP}_{\mathrm{WTH}}$ mainly represents the buildings with high reflectance. The relatively dark open spaces between buildings and shaded areas are highlighted in the $\mathrm{DMP}_{\mathrm{BTH}}$ feature image. Thus, the local structures of individual buildings can be described with the morphological features. *Homogeneity* shows the low signals but *Contrast* has high response values for building areas with a relatively high density, which implies that building clusters exhibit heterogeneous textures in terms of visual effect. Summing up, the features of $\mathrm{DMP}_{\mathrm{BTH}}$, $\mathrm{DMP}_{\mathrm{WTH}}$, and *Contrast* present a positive association with building density, while n(NIR−Red) and *Homogeneity* are inversely correlated with building density. The integration of these features, therefore, has a great potential to achieve a better delineation of building density.

In Table I, WTH and BTH are the white and black top hat by reconstruction of the base image; *s* and dir represent the scale and direction of a linear structural element (SE), respectively; *D* denotes the set of directionality of the linear SE; *i* and *j* are the gray levels in the windows; *p*(*i,j*) represents the (*i,j*)th entry of the GLCM; *N* is the number of gray levels of the image; and μ and σ are the mean and standard deviation, respectively. The maximum of the visible bands for each pixel is computed as the base image for derivation of the morphological and textural features [27].

B. Support Vector Regression and Recursive Feature Elimination

In this paper, we estimate the building density through the multiple features derived from remotely sensed imagery. Regression models have previously been applied to predict continuous parameters in environmental studies [2]. In general, a regression model can be briefly written as

$$y = f(\boldsymbol{x}) \tag{1}$$

where $\boldsymbol{x}$ and y indicate the predictor variable and the associated dependent variable, respectively; and $f(\cdot)$ denotes the regression model, which is built and trained by measuring the relationship between the two kinds of variables. Specifically, in this study, the relationship between the reference building density (dependent variable) and multiple features (predictor variables) extracted from the images is established.

In this study, SVR [32] is employed to predict the building density due to its remarkable ability to deal with complex datasets, such as a nonlinear data distribution in a high-dimensional space. The main idea behind SVR is to estimate the dependence between pairs of predictor variables and a dependent variable by fitting an optimal approximating hyperplane to a set of training samples. The hyperplane is defined by an optimal linear function when it minimizes a cost function considering a maximized margin that encloses the samples, and a minimized error of approximation. The approximation is controlled by a ε-insensitive loss function that directly influences the width of the margin. The error of approximation, i.e., the samples located outside of the margin, is measured using slack variables. SVR shows a superior performance by implicitly mapping the original predictor variables into a higher dimensional feature space, wherein the new data distribution enables a better fitting of a linear combination. The standard form of SVR is

$$\begin{aligned} &\text{Minimize} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_i \left(\xi_i + {\xi_i}^*\right) \\ &\text{subject to } \begin{cases} k\left(\mathbf{w}, \boldsymbol{x}_i\right) + b - y_i \le \varepsilon + \xi_i \\ y_i - k\left(\mathbf{w}, \boldsymbol{x}_i\right) - b \le \varepsilon + {\xi_i}^* \\ \xi_i, {\xi_i}^* \ge 0 \end{cases} \end{aligned} \tag{2}$$

where $\|\cdot\|$, $(\cdot,\cdot)$, and $k(\cdot,\cdot)$ represent the Euclidean norm, the inner product, and the kernel mapping, respectively; $(\mathbf{w}, b)$ represents the regression coefficient and the bias term; and ε measures the precision of the approximation term. For SVR, *C*

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

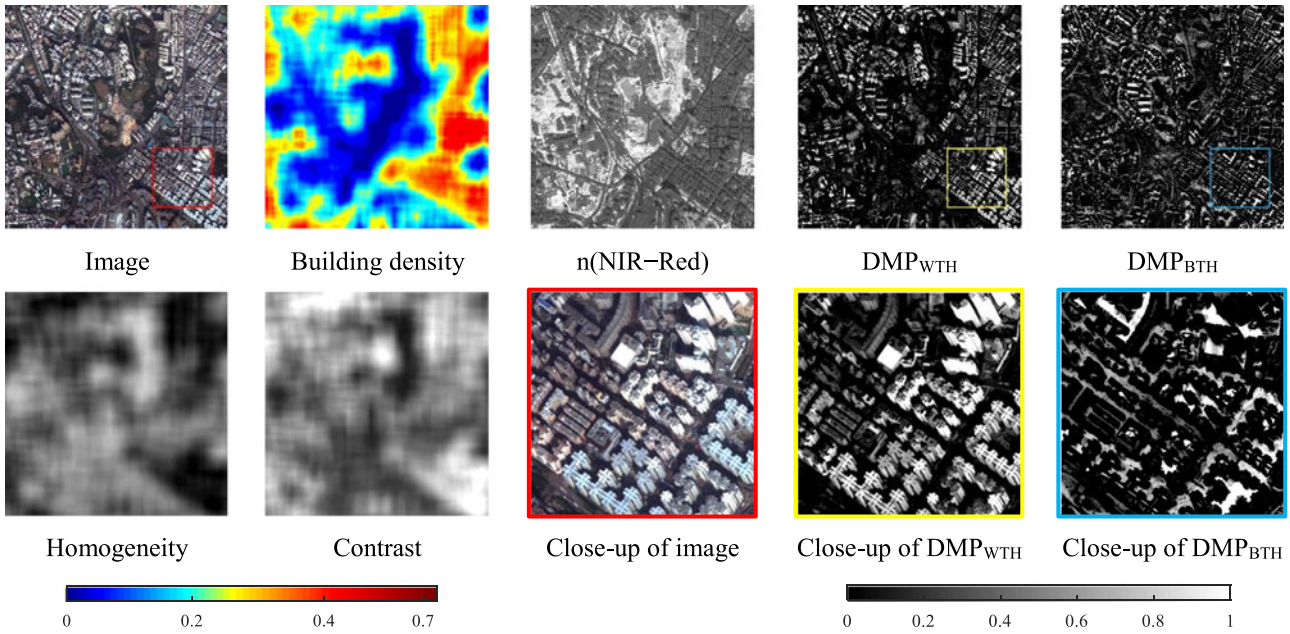


Fig. 2. Example of building density and some representative features for a typical area (see Table I for the complete feature description).

determines the tradeoff between the flatness of the regression model and the amount to which deviations larger than ε are tolerated, ξ_i and ξ_i^* refer to the ith slack variables used to relax the error of the approximation. In this study, SVR is carried out with an open-source software package, LIBSVM [33], in which the parameter settings, including the width of the Gaussian kernel, as well as C and ε, are determined by a grid search in conjunction with internal tenfold cross validation [34].

Considering that a large number of features may result in information redundancy, feature ranking and selection should be taken into account when dealing with such datasets. In the proposed framework, RFE, a feature selection algorithm designed for SVR, is able to find features that are better suited to the predetermined SVR model. Thus, it is applied to assess the importance of the multiple features whose selection criterion is based on an analysis of the smallest change in cost function [29]. In [29], the W^2 quantity, a measure of the predictive ability of the model, was proposed as follows:

$$W^2 = \|\mathbf{w}\|^2 = \sum_{i,j} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) k(x_i, x_j) \quad (3)$$

where α and α^* are Lagrange multipliers.

By using this property and assuming that the set of support vectors remains unchanged when eliminating the less informative features, it is possible to compute $W_{(-t)}^2$ for the feature subsets minus the considered feature t without retraining the model. Successively, the resulting ranking criterion is

$$DW_t^2 = \left| W^2 - W_{(-t)}^2 \right|. \quad (4)$$

The input corresponding to the smallest difference DW_t^2 is removed. The procedure can be iterated to use RFE to generate a feature ranking list [29].

III. Experiments and Results

A. Datasets

In this study, optical high-resolution remote sensing images derived from GeoEye-1 (GE-1) and WorldView-2 (WV-2) satellites were used for the validation. Specifically, three test images over three representative megacities in China were considered, as shown in Fig. 3. The data have a spatial resolution of 2 m, with image sizes of 6433 × 5409, 6327 × 4156, and 3288 × 3396 pixels for Shenzhen (WV-2), Wuhan (GE-1), and Hong Kong (WV-2), respectively. These cities present a wide variety of locations, economic development levels, and urban landscapes. The dense urban villages [35] in Wuhan and Shenzhen, as well as the high-rise buildings in Hong Kong, pose great challenges to the accurate estimation of building density.

In this study, the real building footprints for the building density estimation were obtained from Map World [36]. As the first official free mapping service of China, Map World has gradually become the main source of geographic information databases, aiming to offer the most authoritative and comprehensive mapping services to the public. Compared to other open-source maps, e.g., OpenStreetMap [37], Map World can provide more detailed building coverage information for Chinese cities. The reference building density was calculated as follows:

$$\text{Building density} = \frac{S_{\text{coverage}}}{S_{\text{land}}} \quad (5)$$

where S_{coverage} represents the building coverage provided by the Map World data, and S_{land} is the area considered. It can be seen that the determination of S_{land} has an impact on the building density estimation. Hence, in this experiment, the spatial unit S_{land} was set as different sizes of square window, with edge lengths of 100, 200, 300 m, etc.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

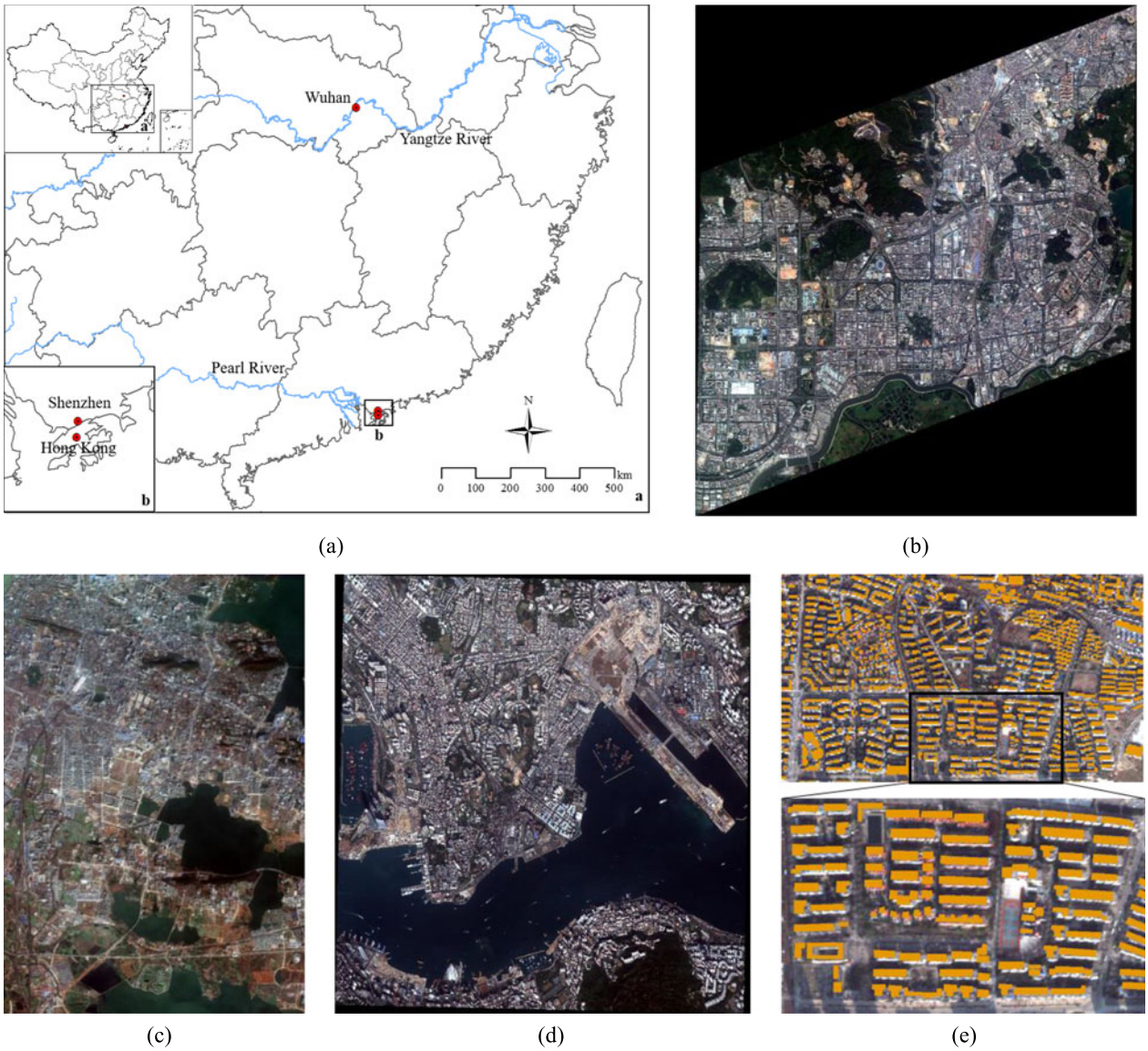


Fig. 3. Study areas and datasets. (a) Overview of the study areas. (b) Shenzhen. (c) Wuhan. (d) Hong Kong. (e) Example of building footprints obtained from Map World.

B. General Results

Thanks to the reference data obtained from Map World for the study areas, there were sufficient samples available for model training and validation. The reference building density was calculated at various spatial scales. Taking into account the spatial unit of the building density estimation, the training samples were randomly selected with a distance constraint [38] to ensure the spatial independence according to the spatial scale used to estimate the building density. For instance, when the building density was to be estimated with a spatial scale of 200 m, each training sample selected had to be at least 200 m away from the other ones. In this case, the number of training samples varied with the spatial scale. The other samples, accounting for the overwhelming majority (>99%) of the total samples available, were used for the validation. The regression model was generated by training samples in each image and tested with a validation dataset of the same image. The root-mean-square error (RMSE) [2] was used to quantitatively assess the estimation accuracy, where a lower RMSE indicates a better performance. In the following, the results of the spatial scale of 200 m are presented unless noted otherwise, and the effect of the spatial scale on building density estimation is discussed in Section IV-A.

Table II lists the training sample size and the accuracy of the estimated building density. "All cities" denotes the overall results of all the test sites. The small model errors (i.e., RMSE) indicate that satisfactory accuracies are obtained for all the test data. In the meantime, the performances of the estimated building densities derived by various combinations of feature sets are also presented in Fig. 4. For simplification, in the figure, "Spectral + Textural" means joint feeding of these two types of features into the regression model. In general, all the cases show similar tendencies, in that the combination of all the predictor

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE II
TRAINING SAMPLE SIZE AND ACCURACY OF THE TEST CITIES AT THE SPATIAL SCALE OF 200 M USING ALL FEATURES AND SELECTED FEATURES BY THE RFE METHOD (SEE SECTION IV-D)

City	# Training sample	RMSE	
		All features	Selected features
Shenzhen	1814	0.0514	0.0533
Wuhan	1886	0.0526	0.0544
Hong Kong	590	0.0551	0.0592
All cities	4290	0.0525	0.0546

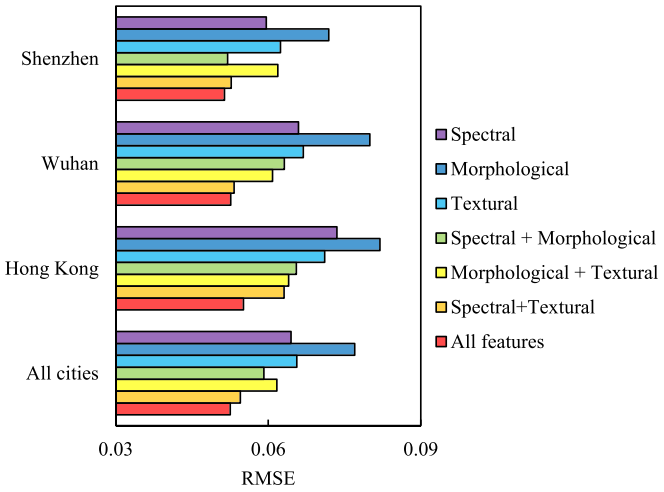


Fig. 4. Regression performance of building density estimation for different input feature sets.

features, i.e., "Spectral + Morphological + Textural," achieves the best regression performance, indicating that the simultaneous consideration of local building structures and building cluster textures, as well as the spectral information, is the most accurate way to estimate building density. Specifically, when focusing on a single feature category, the spectral and textural features show a very close performance, and they both are superior to the morphological features. This is probably due to the fact that the former two feature sets consist of predictors, such as visible bands, n(NIR−Red), *Homogeneity*, and *Contrast* (see Fig. 2), denoting both positive and negative associations with building density, which leads to a more comprehensive characterization of the building density. In addition, it can be seen that the performance is significantly improved when two feature categories are integrated. In fact, the combination of the textural and spectral features (which ranks second among all the combinations) achieves a satisfactory accuracy, which is close to the result of "All features."

The visual effects of the estimated building density for the three study areas are further illustrated in Fig. 5. Note that the large water bodies (e.g., lake and sea) are masked out in black since these areas have no relevance to building density. In general, it is noticeable that the estimated building density is in close accordance with the reference data. Low-density areas within the cities are mainly related to mountains, parks, squares, and harbors [see Fig. 6(a) and (b)]. On the other hand, the high-density areas are usually associated with central business districts, urban villages [35], and traditional residential clusters [see Fig. 6(c) and (d)]. In Wuhan, the high-density areas concentrate in the top left part of the image, which corresponds to the old town of the city with intensive low-rise buildings [see Fig. 6(e)]. As for Hong Kong, it can be clearly seen that a large continuous area in the left side of the image, which is a famous commercial district, Mong Kok [see Fig. 6(f)], presents high-density values. The densely built areas in Shenzhen show a relatively dispersed distribution compared to Wuhan and Hong Kong. On the whole, through a visual inspection, the areas with low-/high-density values are visibly apparent and uniform in both the reference data and estimated results.

IV. DISCUSSIONS

A. *Effect of the Spatial Scale on Building Density Estimation*

The window size for calculating building density is an important parameter [8], which impacts both the prediction accuracy and mapping detail. In this experiment, building densities with different spatial scales (e.g., 100, 200, 300 m, etc.) were generated. We investigated the effect of the spatial scale in terms of both accuracy and detail. For accuracy, the RMSEs between the reference data and estimated building density at the different spatial scales were evaluated. For the map detail, according to [3], mutual information, one measure of information theory, was employed, which is expressed as

$$I(X,Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \tag{6}$$

where X and Y are two random variables; $p(x, y)$ is the joint probability distribution function of X and Y; and $p(x)$ and $p(y)$ are the marginal probability distribution functions of X and Y, respectively. This index is able to denote how much detail is lost with the increase of the spatial scale, where a lower value indicates less information preservation. Since there are two variables involved in the calculation of mutual information, we need to choose a base image as a comparison to all the other images. According to [3], the density at a 50-m scale was considered as an alternative base image to compute the mutual information. The interval of building density was set as 0.01. In this case, $p(x)$ or $p(y)$ denotes the marginal probability of density values of one image falling into each interval (e.g., 0−0.01, 0.01−0.02), and $p(x, y)$ represents the joint probability of density values of two images falling into certain intervals.

As expected, the RMSE values tend to decline with the growth of the spatial size [see Fig. 7(a)], but the results lose more information at the larger spatial scales [see Fig. 7(b)]. Both measures show the largest changes between the spatial scales of 100 and 200 m. An apparent difficulty when deriving an appropriate spatial scale is to balance the reduction of the model error and the preservation of spatial details. According to [3], by referring to the two quantitative criteria (i.e., RMSE and mutual information), in this study, 200 m can be considered as

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

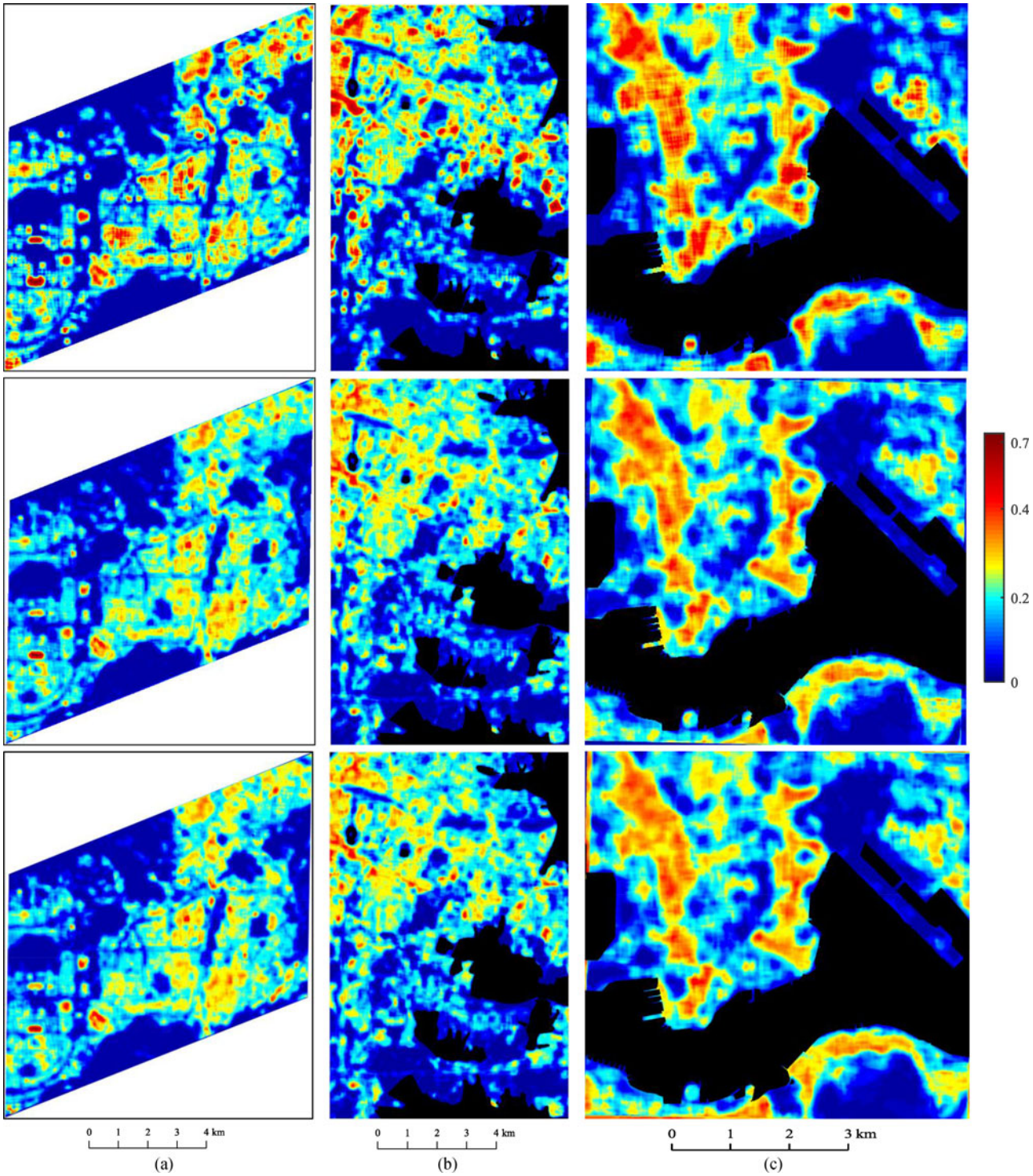


Fig. 5. Reference building density (first row), estimated building density using all features (second row), and estimated building density using selected features by the RFE method (third row, see Section IV-D) for: (a) Shenzhen, (b) Wuhan, and (c) Kong Hong. The spatial scale of the building density is 200 m.

a reasonable spatial scale for all the study cities for building density mapping, since it achieves low RMSEs (∼0.05) and preserves more detailed information. This selection of the spatial scale, which is an acceptable tradeoff between the prediction accuracy and mapping detail, is very close to the result reported in [3].

B. Correlations Between Single Features and Building Density

The correlations between single features and the reference building density, quantitatively assessed by correlation coefficients [6], are shown in Fig. 8. Among the spectral features, four features containing the near-infrared band, indicating vegetation components, are inversely correlated to building density

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

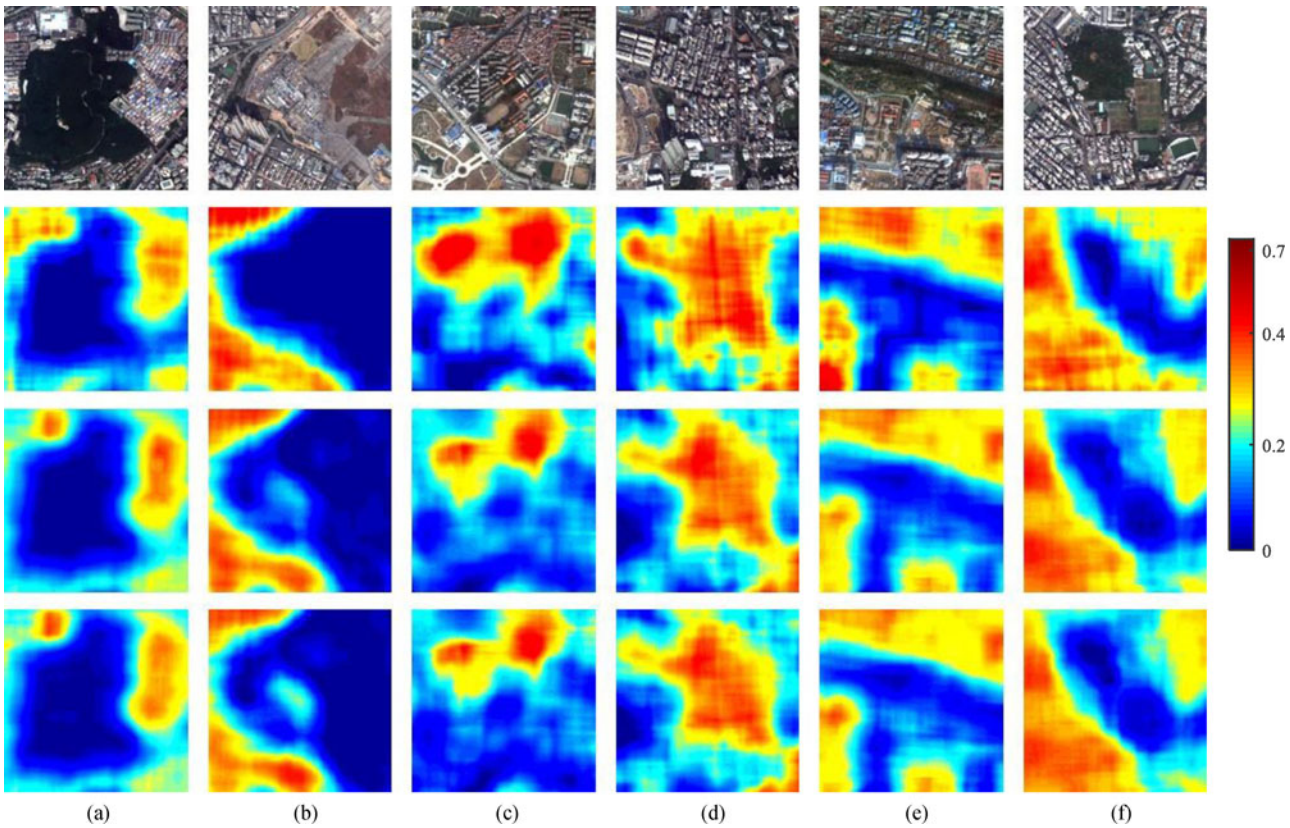


Fig. 6. Examples of different urban landscapes. True-color image (first row). Reference building density (second row). Estimated building density using all features (third row). Estimated building density using selected features by the RFE method (last row, see Section IV-D). The spatial scale of building density is 200 m.

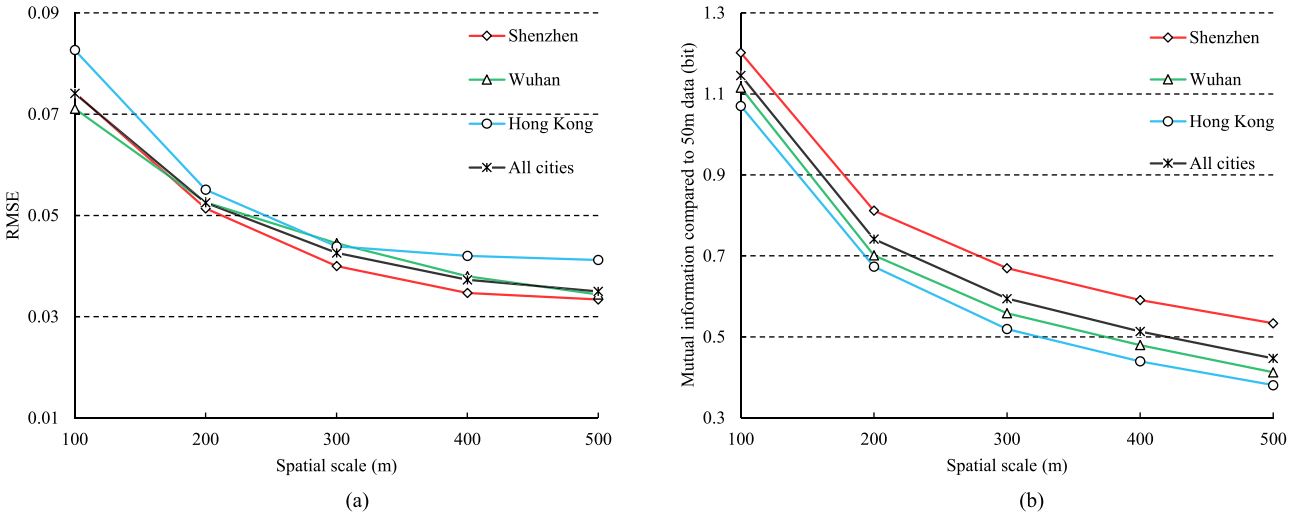


Fig. 7. Evaluation of the building density estimation with different spatial scales: (a) RMSE between the reference data and estimated building density, and (b) mutual information compared to the 50-m spatial scale.

because of their representativeness for nonbuilding areas, while the other features present a positive correlation with building density, since buildings often have a relatively high reflectance in the visible bands. With regard to the morphological features, $\mathrm{DMP}_{\mathrm{WTH}}$ shows larger values, suggesting that the bright structures are more correlated with building density than the dark structures represented by $\mathrm{DMP}_{\mathrm{BTH}}$. $\mathrm{DMP}_{\mathrm{WTH}}$ is perceived as a direct descriptor for local building structures, and $\mathrm{DMP}_{\mathrm{BTH}}$ is used to characterize the shadow, which is an indirect approach to imply the existence of buildings. Accordingly, the performance of $\mathrm{DMP}_{\mathrm{WTH}}$ is more efficient than that of $\mathrm{DMP}_{\mathrm{BTH}}$ with respect to describing building density. As for the textural features,

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

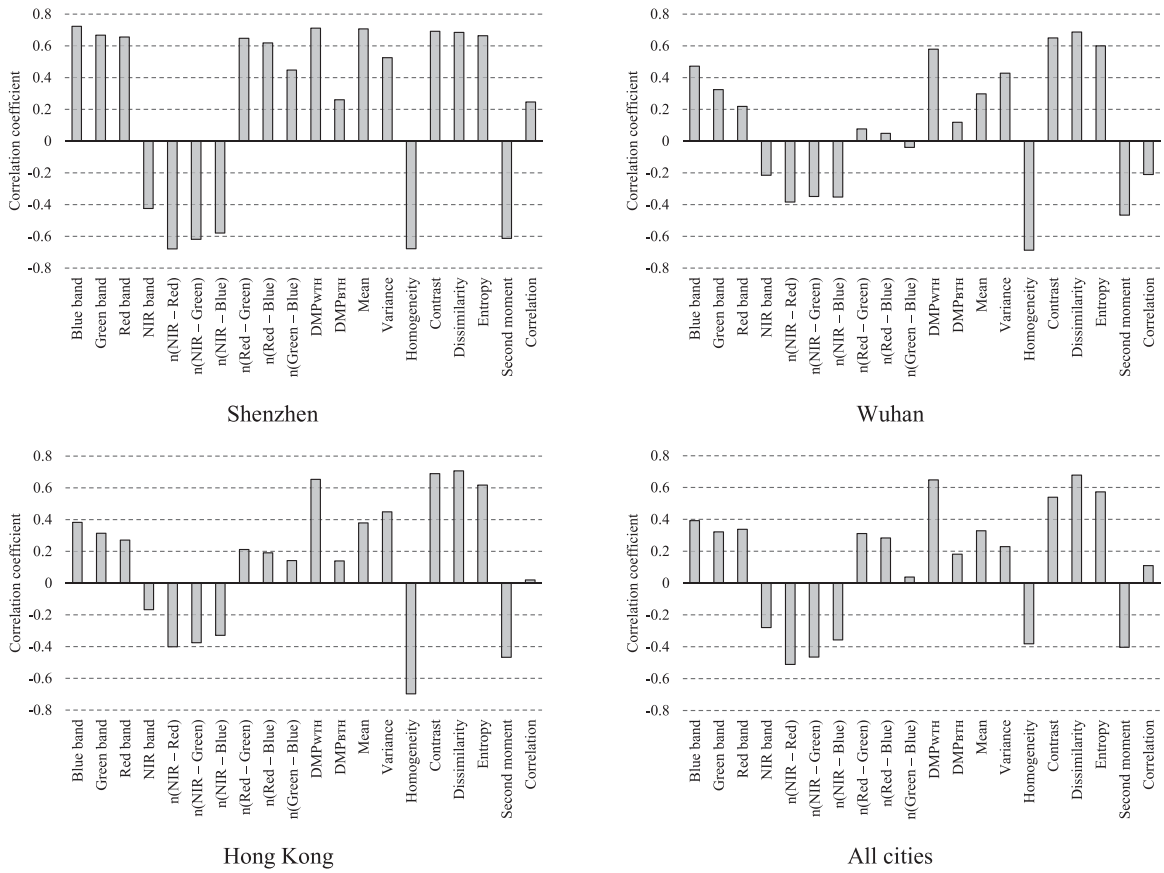


Fig. 8. Correlation coefficients between single features and building density.

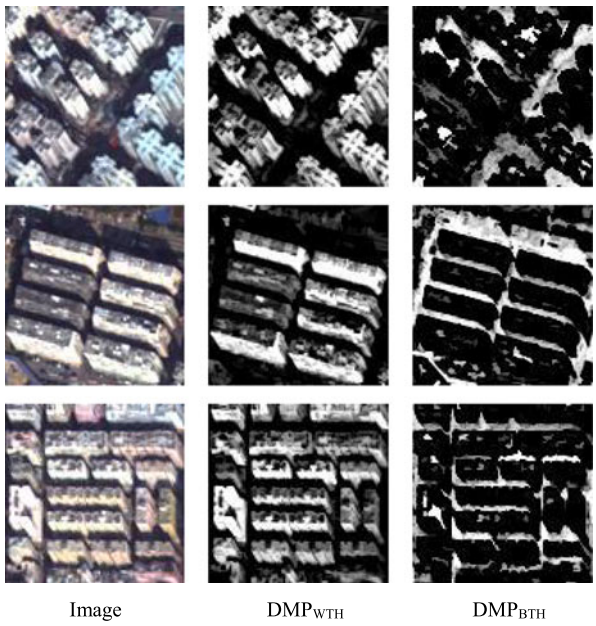


Fig. 9. Some typical examples for morphological features used in this study.

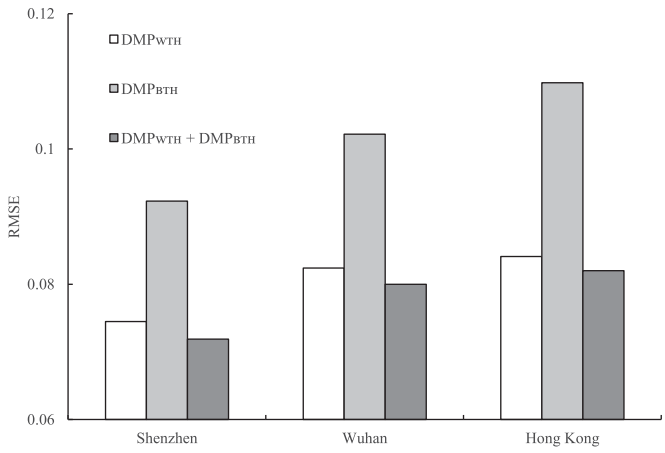


Fig. 10. Accuracies of building density regression for various combinations of morphological features.

most of the measures present a relatively high correlation with building density, wherein the heterogeneous textural measures (e.g., *Contrast, Dissimilarity*) are positively associated with the building density, while the homogeneous textural features (e.g., *Homogeneity, Second moment*) present negative associations. The different buildings, which are constructed with complex and diverse materials, as well as their complicated surroundings (e.g., roads, open spaces), lead to high regional tonal variations in remote sensing images. Consequently, building clusters exhibit evident heterogeneity in terms of visual effect.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

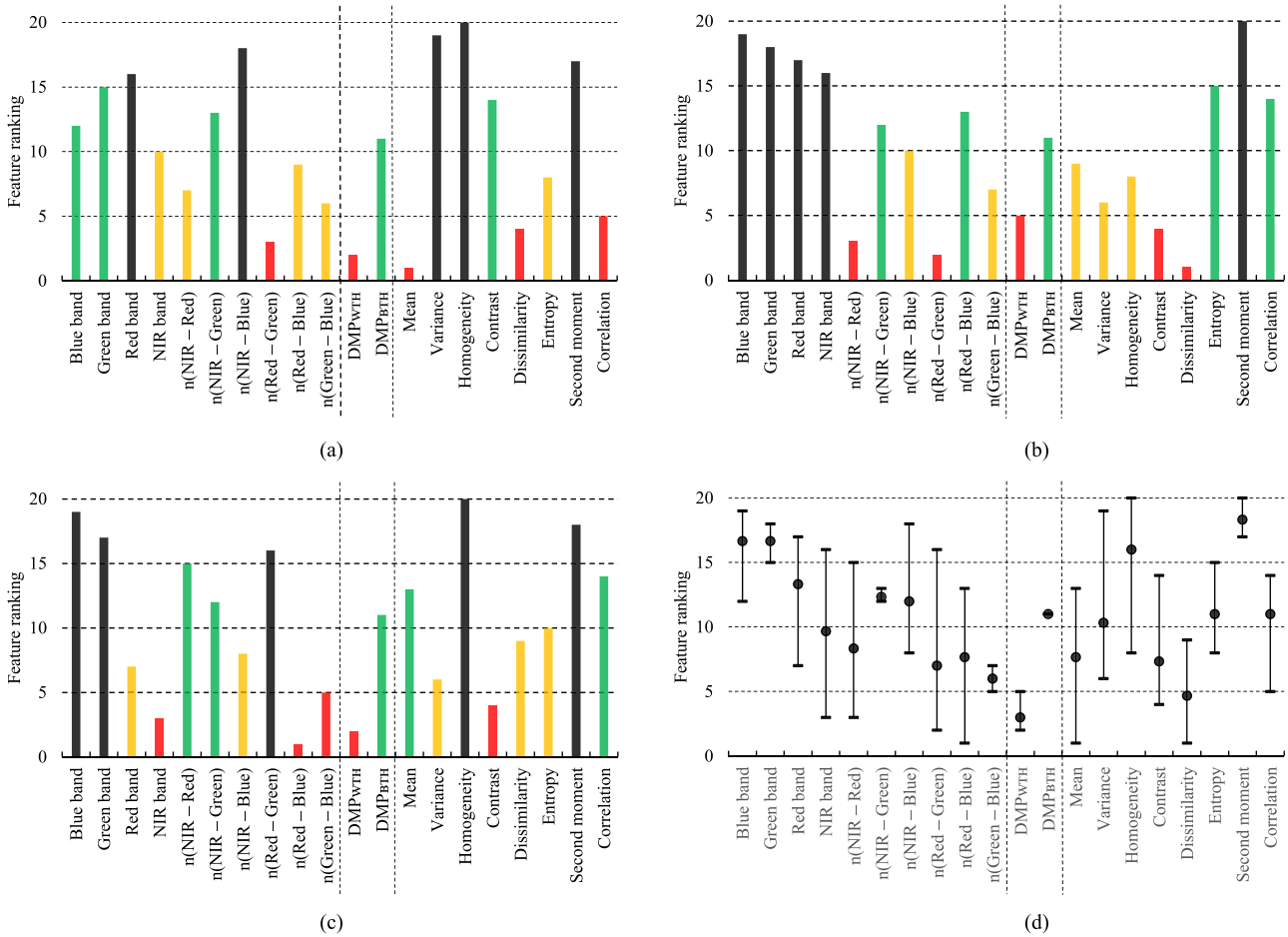


Fig. 11. Feature ranking by the RFE algorithm for: (a) Shenzhen, (b) Wuhan, and (c) Hong Kong. The red bars represent features ranking 1−5, the orange bars represent features ranking 6−10, the green bars represent features ranking 11−15, and the black bars represent features ranking 16−20. (d) Overall feature ranking result for the three cities, where the upper horizontal line denotes the maximum, the lower horizontal line denotes the minimum, and the black dot denotes the average value.

TABLE III
TOP TEN FEATURES IN TERMS OF AVERAGE RANKING FOR THE THREE CITIES

Overall ranking	Feature
1	$\mathrm{DMP_{WTH}}$
2	Dissimilarity
3	n(Green−Blue)
4	n(Red−Green)
5	Contrast
6	Mean
7	n(Red−Blue)
8	n(NIR−Red)
9	NIR band
10	Variance

C. Consideration of Shadows in High-Resolution Imagery

The shadows in high-resolution imagery should be taken into consideration when estimating building density. The morphological features used in this study have ability to highlight both bright ($\mathrm{DMP_{WTH}}$) and dark ($\mathrm{DMP_{BTH}}$) structures, which can help us to characterize the cooccurrence relationship of buildings and shadows. $\mathrm{DMP_{WTH}}$ mainly indicates the buildings with high spectral reflectance, and the relative dark areas (e.g., shadows) are highlighted in the $\mathrm{DMP_{BTH}}$ feature image. Some typical examples are presented in Fig. 9. It can be observed that in high-resolution imagery, the shadows can also indicate the presence of buildings. In Fig. 10, it can be seen with the addition of $\mathrm{DMP_{BTH}}$, the estimated building density achieves a more accurate performance (smaller RMSE) compared to only using $\mathrm{DMP_{WTH}}$ as an input feature for regression. As a summary, dark structures (e.g., shadows) characterized by $\mathrm{DMP_{BTH}}$ can help us to estimate building density, and, hence, we choose $\mathrm{DMP_{BTH}}$ as a member of candidate feature pool for subsequent building density regression in this study.

D. Feature Ranking and Feature Subset Selection

The importance of the input variables for the building density estimation was measured by the RFE algorithm described in Section II. The sequential feature ranking acquired by RFE is shown in Fig. 11. The features ranked from 16 to 20 (i.e., the first five features removed by RFE) belong to the spectral and textural feature sets for each city, which implies that a strong redundancy may exist in these two categories of features.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

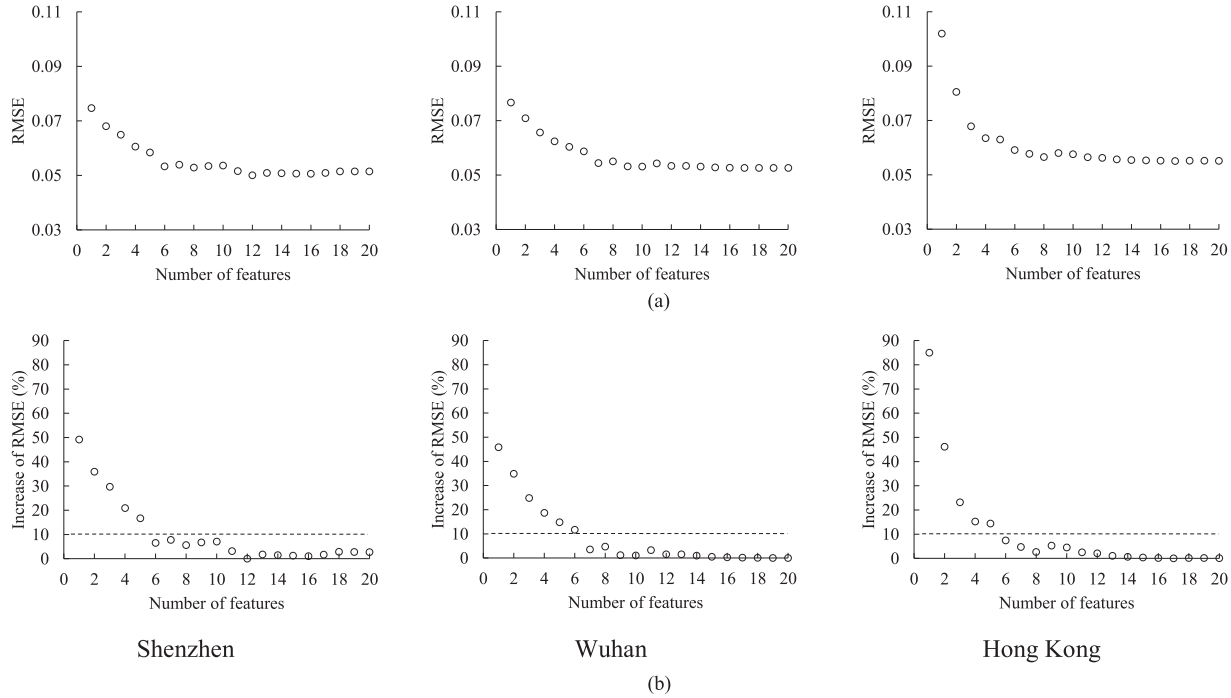


Fig. 12. Accuracy of the estimated building density obtained using feature subsets composed of different numbers of features. (a) RMSE. (b) Percentage increase of RMSE.

TABLE IV
ACCURACY OF THE TEST CITIES WITH DIFFERENT FEATURE SELECTION METHODS

City	SVR-CFS	LASSO
Shenzhen	0.0556	0.0569
Wuhan	0.0565	0.0594
Hong Kong	0.0639	0.0654

The top five features for each study area all come from different feature sets, indicating the necessity for the simultaneous involvement of spectral, morphological, and textural features. The overall feature contribution result (i.e., the maximum, minimum, and mean values of the feature ranking) is also presented in Fig. 11(d). The importance of features varies in the different study areas, but some features remain relatively stable, such as $\mathrm{DMP}_{\mathrm{WTH}}$ and $\mathrm{DMP}_{\mathrm{BTH}}$. The visible bands and the homogeneous textures (e.g., *Homogeneity*, *Second moment*) make the least contribution to building density estimation, which is probably due to the presence of other similar features, such as the spectral difference in visible bands and n(NIR−Red). The top ten features according to the average ranking of the three cities are reported in Table III. $\mathrm{DMP}_{\mathrm{WTH}}$, as a single feature, is of the most importance due to its great ability to directly describe local building structures. The heterogeneous textural measures, such as *Dissimilarity*, *Contrast*, and *Variance,* play very important roles in the regression model owing to their ability to characterize the heterogeneity of building clusters. Spectral features, in the form of spectral differences between visible bands as well as the NIR band, also make great contributions to building density estimation since they can characterize the reflectance of both buildings and nonbuildings.

According to the feature ranking sequence, the accuracies (i.e., RMSE) using the feature subsets of different numbers of features are shown in Fig. 12(a). It is clear that the RMSE values tend to decline sharply at first and then become stable after several informative features are involved in the regression, which implies that information redundancy exists in our model, and a satisfactory accuracy for building density estimation with fewer but more important features could be achieved. A smaller feature subset selection was obtained based on the following:

$$\mathrm{IncRMSE} = \frac{\mathrm{RMSE} - \mathrm{RMSE}_{\min}}{\mathrm{RMSE}_{\min}} \times 100\% \qquad (7)$$

where $\mathrm{RMSE}_{\min}$ is the minimal RMSE value, and $\mathrm{IncRMSE}$ is the percentage increase compared with $\mathrm{RMSE}_{\min}$.

In this study, the percentage growth of RMSE was set as less than 10%, which is a slight degradation that does not significantly affect the performance [39]. Under this circumstance, the feature set with the smallest size was selected as an acceptable approach for the reduction of feature redundancy [see Fig. 12(b)]. Based on such a criterion, the number of remaining features only accounts for about one-third of the initial dimensionality of all the feature sets [see Fig. 12(b)]. By keeping only six or seven features, the performances of the selected features are very close to those of all features in terms of both quantitative comparison and visual effect (see Table II and Fig. 5), which indicates the effectiveness of the feature selection using the RFE method.

As comparisons, one of the filter-based methods named "correlation-based feature selection" (CFS) [40] and another

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

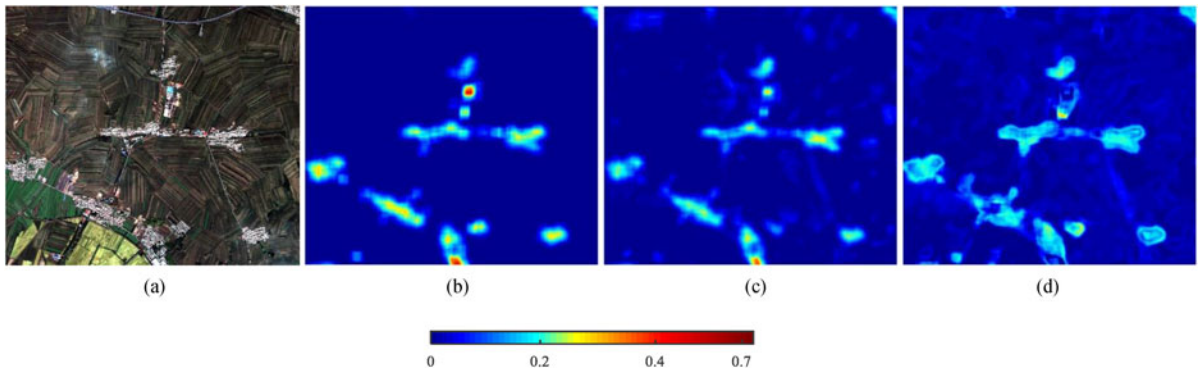


Fig. 13. Results for a rural area of Harbin. (a) True-color image. (b) Reference building density. (c) Estimated building density using self-training. (d) Estimated building density using transfer training (model trained in Wuhan).

TABLE V
ACCURACY OF BUILDING DENSITY ESTIMATION USING SELF-TRAINING (DIAGONAL LINE) AND TRANSFER TRAINING (OFF-DIAGONAL ELEMENTS)

Training \ Test	Shenzhen	Wuhan	Hong Kong	Harbin
Shenzhen	0.0587	0.1077	0.0768	0.1289
Wuhan	0.0878	0.0580	0.0776	0.0307
Hong Kong	0.0723	0.0839	0.0646	0.0881
Harbin	0.1628	0.1071	0.1085	0.0166

regression analysis method called "least absolute shrinkage and selection operator" (LASSO) [41] were also used. CFS can be considered as a preprocessing step and is independent of the choice of predictor. It aims to find the features that are highly correlated with the dependent variable and uncorrelated with each other. LASSO performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability of the regression model. More details can be found in [40] and [41]. The regression accuracies obtained with CFS and LASSO are presented in Table IV, which are inferior to those obtained with SVR-RFE. (see Table II for comparison). The results show that SVR-RFE is an effective regression and feature selection method in our case.

E. Trained Model for Transfer Between Geographical Regions

To investigate the geographical transferability between the test cities, the regression model was trained based on one image and applied to the other test sites [42]. Since the correlations of some features with building density could be inconsistent in different cities, the features which have a relatively high relevance to building density in all the test sites were considered to have more potential transferability [43]. According to Fig. 8 (all cities), the features whose correlation coefficients in absolute values were larger than 0.4 [44], indicating a relatively consistent distribution across the cities, were selected to test the generalizability of the trained model for transfer between geographical regions.

TABLE VI
ACCURACY OF BUILDING DENSITY ESTIMATION BASED ON BUILDING DETECTION

City	MBI	MBI-pp
Shenzhen	0.0784	0.0755
Wuhan	0.0828	0.0796
Hong Kong	0.0774	0.0769

An additional QuickBird image of a rural area of Harbin with a spatial resolution of 2.4 m (see Fig. 13) was also used. For this image, we manually delineated the building roofs for the reference building density calculation as the low-rise buildings in the rural area of Harbin are not significantly affected by the viewing angle. The diagonal elements in Table V represent the RMSE values achieved by self-training (i.e., the regression model is trained based on one image and applied to the same test site). On the other hand, the off-diagonal elements record the accuracy of the transfer-training model. It is demonstrated that in all the cases, the self-training achieves the best results compared to the transfer training, which can be easily understood due to the differences (e.g., landscapes, sensors) between the test sites.

F. Comparison With Building Density Estimation Based on Building Detection

The proposed framework for building density estimation was also compared with approaches based on building detection. Two recently developed methods were used for the building de-

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

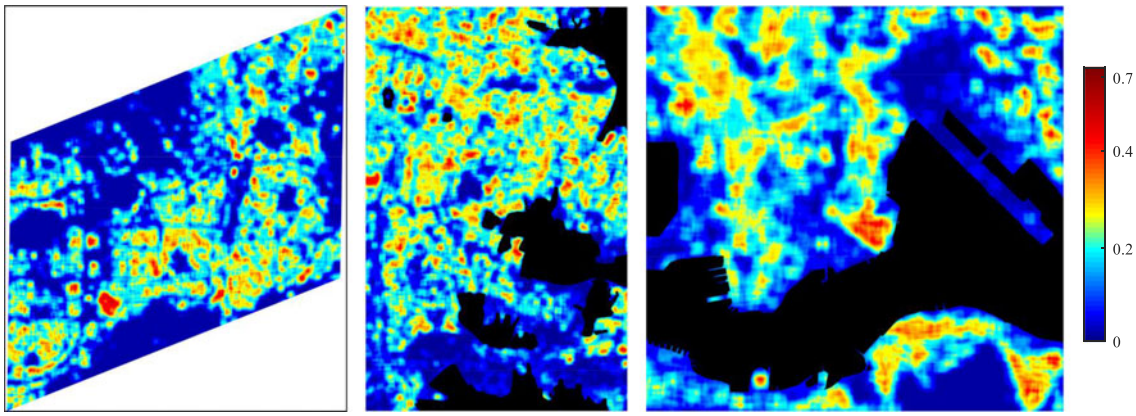


Fig. 14. Estimated building density based on building detection (MBI-pp).

tection: the original morphological building index (MBI) [20], [24] and the MBI combined with a postprocessing procedure (MBI-pp) [45]. MBI aims to describe the spectral–spatial characteristic of buildings (e.g., brightness, size, contrast, and directionality) based on a set of morphological operations, and MBI-pp improves the MBI detector by considering spectral, shadow, and shape constraints. MBI-pp has previously been compared with other state-of-the-art methods, such as multilabel partitioning [46] and GrabCut [47], which confirmed that MBI-pp can achieve a comparable or even better performance for building extraction. Therefore, in this study, as a comparison with the proposed framework, MBI-pp and the original MBI were used to detect the building structures and compute the building density.

The results of the estimated building density based on building detection are presented in Table VI and Fig. 14. The method proposed in this paper performs better (see Table II and Fig. 5) than the approaches based on MBI and MBI-pp, which are subject to complex threshold selection and are dependent on the accuracy of the building identification. Moreover, due to the viewing angle of the remote sensing images, both MBI and MBI-pp have difficulty in extracting the real building footprints in the urban areas, while the building density is defined as the building coverage ratio within an area of interest and the reference building inventories obtained from Map World also refer to the building footprints.

V. Conclusion

Building density is an essential urban environmental parameter and plays a key role in research into urban planning and human behavior. However, at the present time, building density information in China is either lacking or incomplete for a number of major cities. To address this problem, in this study, we have proposed a method for quantitative building density estimation using high-resolution satellite images (e.g., GeoEye-1, WorldView-2), by combining spectral, morphological, and textural features. The spectral features are able to characterize the reflectance of buildings and nonbuildings, the morphological features have ability to delineate local building structures, and the textural features are effective in describing building clusters. The experiments undertaken using data from three representative megacities of China (Shenzhen, Wuhan, and Hong Kong) achieved satisfactory results in terms of low RMSE values and visual results. The combination of different feature sets outperformed the single feature category, suggesting that the simultaneous consideration of all categories of features is necessary for building density estimation in different urban scenarios.

In the proposed framework, an SVR is used to construct the model of building density estimation due to its good generalization capability, particularly in high-dimensional spaces. In order to investigate the importance of the different features, the RFE algorithm is used to analyze the contribution of the different features to building density estimation. A feature subset selection based on RFE takes account of both regression accuracy and data redundancy reduction. The analysis of the relevance of the input features provides an insight into the most informative features, which may help other researchers to select more appropriate and efficient features for building density estimation. The investigation of the transferability between geographical regions also confirms that the self-training achieves the better results compared to the transfer training due to the differences (e.g., landscapes, sensors) between the test sites. A more in-depth understanding of urban building density, especially a better delineation for extremely high-density areas that may be closely associated with population aggregation and economic development, will be investigated in our future work.

Acknowledgment

The authors would like to thank anonymous reviewers for the insightful and constructive suggestions, which significantly improved the quality of this paper.

References

[1] H. Taubenböck, T. Esch, A. Felbier, M. Wiesner, A. Roth, and S. Dech, "Monitoring urbanization in mega cities from space," *Remote Sens. Environ.*, vol. 117, pp. 162–176, Feb. 2012.

[2] H. C. Ho, A. Knudby, P. Sirovyak, Y. Xu, M. Hodul, and S. B. Henderson, "Mapping maximum urban air temperature on hot summer days," *Remote Sens. Environ.*, vol. 154, pp. 38–45, Nov. 2014.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[3] J. Susaki, M. Kajimoto, and M. Kishimoto, "Urban density mapping of global megacities from polarimetric SAR images," *Remote Sens. Environ.*, vol. 155, pp. 334–348, Dec. 2014.

[4] H. Taubenbock, T. Esch, A. Felbier, A. Roth, and S. Dech, "Pattern-based accuracy assessment of an urban footprint classification using TerraSAR-X data," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 278–282, Mar. 2011.

[5] A. Krehl, S. Siedentop, H. Taubenböck, and M. Wurm, "A comprehensive view on urban spatial structure: Urban density patterns of German city regions," *ISPRS Int. J. Geoinf.*, vol. 5, no. 6, pp. 76–96, May 2016.

[6] M. Kajimoto and J. Susaki, "Urban density estimation from polarimetric SAR images based on a POA correction method," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 6, no. 3, pp. 1418–1429, Jun. 2013.

[7] H. Taubenböck *et al.*, "Delineation of central business districts in mega city regions using remotely sensed data," *Remote Sens. Environ.*, vol. 136, pp. 386–401, Sep. 2013.

[8] M. Pesaresi, M. Halkia, and G. K. Ouzounis, "Quantitative estimation of settlement density and limits based on textural measurements," in *Proc. 2011 Joint Urban Remote Sens. Event*, Munich, Germany, 2011, pp. 89–92.

[9] D. Gonzalez-Aguilera, E. Crespo-Matellan, D. Hernandez-Lopez, and P. Rodriguez-Gonzalvez, "Automated urban analysis based on LiDAR-derived building models," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 3, pp. 1844–1851, Mar. 2013.

[10] B. Yu, H. Liu, J. Wu, Y. Hu, and L. Zhang, "Automated derivation of urban building density information using airborne LiDAR data and object-based method," *Landscape Urban Plan.*, vol. 98, nos. 3/4, pp. 210–219, Dec. 2010.

[11] M. Schmidt, T. Esch, D. Klein, M. Thiel, and S. Decha, "Estimation of building density using TerraSAR-X-Data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Honolulu, HI, USA, 2010, pp. 1936–1939.

[12] Z. Zhu, C. E. Woodcock, J. Rogan, and J. Kellndorfer, "Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data," *Remote Sens. Environ.*, vol. 117, pp. 72–82, Feb. 2012.

[13] P. Zhong and R. Wang, "A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3978–3988, Dec. 2007.

[14] B. Sirmacek and C. Unsalan, "Urban area detection using local feature points and spatial voting," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 146–150, Jan. 2010.

[15] A. Kovács and T. Szirányi, "Improved Harris feature point set for orientation-sensitive urban-area detection in aerial images," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 796–800, Jul. 2013.

[16] S. Yu, M. Berthod, and G. Giraudon, "Toward robust analysis of satellite images using map information-application to urban area detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 4, pp. 1925–1939, Jul. 1999.

[17] C. Tao, Y. Tan, Z.-R. Zou, and J. Tian, "Unsupervised detection of built-up areas from multiple high-resolution remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1300–1304, Nov. 2013.

[18] B. Sirmacek and C. Unsalan, "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.

[19] A. Manno-Kovács and A. O. Ok, "Building detection from monocular VHR images by integrated urban area knowledge," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 10, pp. 2140–2144, Oct. 2015.

[20] X. Huang, "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, Jul. 2011.

[21] J. W. Rouse, Jr., R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS," in *Proc. 3rd Earth Resour. Technol. Satell. Symp.*, Washington, DC, USA, 1973, pp. 309–317.

[22] M. E. Martin and J. D. Aber, "High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes," *Ecol. Appl.*, vol. 7, no. 2, pp. 431–443, May 1997.

[23] X. Huang *et al.*, "Multiple morphological profiles from multicomponent-base images for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 7, no. 12, pp. 4653–4669, Dec. 2014.

[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] 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.

[26] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, May 1979.

[27] 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.

[28] X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 257–272, Jan. 2013.

[29] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Mach. Learn.*, vol. 46, no. 1, pp. 389–422, Jan. 2002.

[30] D. Tuia, F. Pacifici, M. Kanevski, and W. J. Emery, "Classification of very high spatial resolution imagery using mathematical morphology and support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3866–3879, Nov. 2009.

[31] F. Dell'Acqua and P. Gamba, "Texture-based characterization of urban environments on satellite SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 153–159, Jan. 2003.

[32] V. N. Vapnik and V. Vapnik, *Statistical Learning Theory*, vol. 1. New York, NY, USA: Wiley, 1998.

[33] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, Apr. 2011, Art. no. 27.

[34] H. Yang, Q. Du, and G. Chen, "Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 5, no. 2, pp. 544–554, Apr. 2012.

[35] X. Huang, H. Liu, and L. Zhang, "Spatiotemporal detection and analysis of urban villages in mega city regions of china using high-resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3639–3657, Jul. 2015.

[36] (2015, Jun. 10). MapWorld. [Online]. Available: http://www.tianditu.cn/

[37] M. Haklay and P. Weber, "Openstreetmap: User-generated street maps," *IEEE Pervasive Comput.*, vol. 7, no. 4, pp. 12–18, Oct.–Dec. 2008.

[38] X. Li, W. Li, A. Middel, S. L. Harlan, A. J. Brazel, and B. L. Turner, "Remote sensing of the surface urban heat island and land architecture in Phoenix, Arizona: Combined effects of land composition and configuration and cadastral–demographic–economic factors," *Remote Sens. Environ.*, vol. 174, pp. 233–243, Mar. 2016.

[39] M. Kuhn. (2012). Variable selection using the caret package. [Online]. Available: http://cran.cermin.lipi.go.id/web/packages/caret/vignettes/caretSelection.pdf

[40] M. A. Hall, "Correlation-based feature selection for machine learning," Ph.D. dissertation, Dept. Comput. Sci., Univ. Waikato, Hamilton, New Zealand, 1999.

[41] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. Roy. Stat. Soc.*, vol. 58, no. 1, pp. 267–288, 1996.

[42] M. Wurm, A. Schmitt, and H. Taubenböck, "Building types' classification using shape-based features and linear discriminant functions," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 9, no. 5, pp. 1901–1912, May 2016.

[43] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

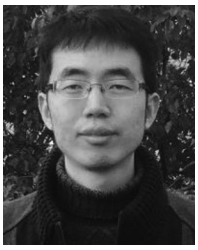
[44] F. López-Granados, M. Jurado-Expósito, J. M. Peña-Barragán, and L. García-Torres, "Using geostatistical and remote sensing approaches for mapping soil properties," *Eur. J. Agron.*, vol. 23, no. 3, pp. 279–289, Oct. 2005.

[45] X. Huang, W. Yuan, J. Li, and L. Zhang, "A new building extraction postprocessing framework for high-spatial-resolution remote-sensing imagery," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 10, no. 2, pp. 654–668, Feb. 2017.

[46] 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, Dec. 2013.

[47] 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.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



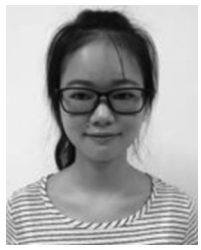
Tao Zhang received the B.S. degree in surveying and mapping from Central South University, Changsha, China, in 2014. He is currently working toward the Ph.D. degree in photogrammetry and remote sensing at the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China.

His research interests include high-resolution image processing and remote sensing applications.

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. He is the Founder and Director of the Institute of Remote Sensing Information Processing, School of Remote Sensing and Information Engineering, Wuhan University. He has published more than 80 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 received 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 the 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) and also the Lead Guest Editor of the special issue on Sparsity-Driven High-Dimensional Remote Sensing Image Processing and Analysis for the J*ournal of Applied Remote Sensing* (vol. 10, no. 4, October 2016). Since 2014, he has been an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. Since 2016, he has been an Associate Editor of the Photogrammetric Engineering and Remote Sensing.



Dawei Wen received the B.S. degree in surveying and mapping from Wuhan University, Wuhan, China, in 2013, where she is currently working toward the Ph.D. degree in photogrammetry and remote sensing at the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing.

Her research interests include change analysis of multitemporal remote sensing images 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*.