

Article

Urban Built-Up Area Extraction from Landsat TM/ETM+ Images Using Spectral Information and Multivariate Texture

Jun Zhang ¹, Peijun Li ^{1,*} and Jinfei Wang ²

¹ Institute of Remote Sensing and GIS, School of Earth and Space Sciences, and Beijing Key Lab of Spatial Information Integration and 3S Application, Peking University, Beijing 100871, China; E-Mail: junzhang@pku.edu.cn

² Department of Geography, University of Western Ontario, 1151 Richmond Street, London, ON N6A 3K7, Canada; E-Mail: jfwang@uwo.ca

* Author to whom correspondence should be addressed; E-Mail: pjli@pku.edu.cn; Tel./Fax: +86-10-6275-7364.

Received: 1 May 2014; in revised form: 5 July 2014 / Accepted: 30 July 2014 /

Published: 6 August 2014

Abstract: Urban built-up area information is required by various applications. However, urban built-up area extraction using moderate resolution satellite data, such as Landsat series data, is still a challenging task due to significant intra-urban heterogeneity and spectral confusion with other land cover types. In this paper, a new method that combines spectral information and multivariate texture is proposed. The multivariate textures are separately extracted from multispectral data using a multivariate variogram with different distance measures, *i.e.*, Euclidean, Mahalanobis and spectral angle distances. The multivariate textures and the spectral bands are then combined for urban built-up area extraction. Because the urban built-up area is the only target class, a one-class classifier, one-class support vector machine, is used. For comparison, the classical gray-level co-occurrence matrix (GLCM) is also used to extract image texture. The proposed method was evaluated using bi-temporal Landsat TM/ETM+ data of two megacity areas in China. Results demonstrated that the proposed method outperformed the use of spectral information alone and the joint use of the spectral information and the GLCM texture. In particular, the inclusion of multivariate variogram textures with spectral angle distance achieved the best results. The proposed method provides an effective way of extracting urban built-up areas from Landsat series images and could be applicable to other applications.

Keywords: urban built-up area; multivariate texture; OCSVM; Landsat

1. Introduction

Over the past three decades, urban areas in China have been expanding at an unprecedented pace, due to significant economic development. As urban development initiatives encroach on previously un-developed land, the cumulative impact has resulted in the loss of valuable agricultural land, wetland and forest and also an increase in energy consumption and greenhouse gas emission [1]. In order to assess the spatial and temporal nature of urbanization and land-cover change occurring across urban landscapes, city planners, decision-makers and researchers require timely and accurate information.

Remote sensing technologies provide a reliable source for urban land cover/land use data acquisition. In particular, the United States' NASA Landsat satellite data series (e.g., MSS, TM and ETM+) have been widely used for mapping urban extent and monitoring urban growth [2–7], due to the sensors' capacity for synoptic view, repeat coverage over large areas and the availability of historical archive imagery [8]. Landsat sensors provide some advantages for the purposes of urban land mapping and change detection in terms of efficiency, as a single image can provide a synoptic view of an area of interest. In comparison to expensive higher-resolution sensors, the comparatively low-resolution nature of the Landsat TM/ETM+ sensor (30 m × 30 m) avoids complications from sparse coverage, limited scene availability and lack of data prior to 2000 when monitoring change for multiple periods. However, mapping urban areas using Landsat TM/ETM+ data remains a complex challenge [9], as there are few thematically pure urban pixels due to the mixture of manmade and vegetative land cover components that comprise urban areas. Complicating matters further, urban areas often display heterogeneous spectral characteristics and significant spectral confusion with other land cover classes: for example, barren land and asphalt concrete share similar spectral characteristics and, as a result, can be readily confused.

Over the past few decades, various methods for urban built-up area extraction have been developed using Landsat TM/ETM+ data. Generally, there are three main categories of urban built-up area extraction methods. First, spectral indexes (*i.e.*, calculated from the spectral information of remotely sensed images) or spatial indexes (e.g., texture), such as the built-up area index or the impervious surface related index, have been proposed and used to extract urban areas [10–16]. However, most existing spectral/spatial indexes mainly focus on the urban class itself, without full consideration of other land cover types; as such, these methods do not successfully address the confusion between urban area and other land cover types. In the second category, combined spectral data and spatial information, such as the combination of road density and spectral information or the combination of spectral data and texture information, are used in the extraction of urban built-up area [2,12,17–19]. However, these existing studies employing single-band texture information did not consider spatial variations between bands [20,21]. In the third category, multi-sensor data, such as Landsat TM/ETM+ and SAR (synthetic aperture radar) data are combined in urban built-up area extraction [4,7,22–27]. Due to the difficulties encountered in synthesizing different data sources, as well as the

time-consuming co-registration requirements for multi-sensor images, the methods of multi-sensor data combination have not been used widely.

Image texture is an important source of spatial information that can be derived from images [20]. Texture data, such as classical one-band texture (e.g., gray-level co-occurrence matrix (GLCM) texture), provide additional information for land cover mapping and have been widely used in diverse applications, including urban built-up area extraction [2,12,14,28–30]. However, since each band of a multispectral image generates band-specific texture data, the resultant texture bands have varying capabilities to discriminate between land cover types. The inclusion of one-band texture may not take full advantage of the available texture information contained in the multispectral image. In contrast, multivariate texture (or multi-band texture), e.g., measured by multivariate variogram, extracted simultaneously from all available spectral bands of a multispectral image [31], identifies the spatial variation of multiple bands, thus avoiding the need to select a single band as required in the traditional one-band texture method [21]. The multivariate texture data provide useful spatial information for discriminating between land cover types [31] and have been used in land cover classification, lithological mapping and urban damage assessment [31–34]. In previous studies, Euclidean distance and Mahalanobis distance were used in multivariate variogram texture computation and showed similar performances in land cover classification [31]. However, these two distance measures are sensitive to illumination variations [35], which could be produced by shadow and other factors in rugged terrain [36] and urban area. In contrast, the spectral angle distance [37–39] is insensitive to illumination and albedo effect. As such, the use of spectral angle distance in a multivariate variogram may provide an alternative method for multivariate texture extraction. Thus, the main objective of this study is to combine multivariate variogram texture and spectral data for urban built-up area extraction. The second objective is to investigate if the use of different distance measures in a multivariate variogram for multivariate texture extraction affects the performance of multivariate texture in combined classification for urban area mapping. The third objective is to evaluate the ability of this method in using historical archive Landsat imagery to the application of built-up area extraction

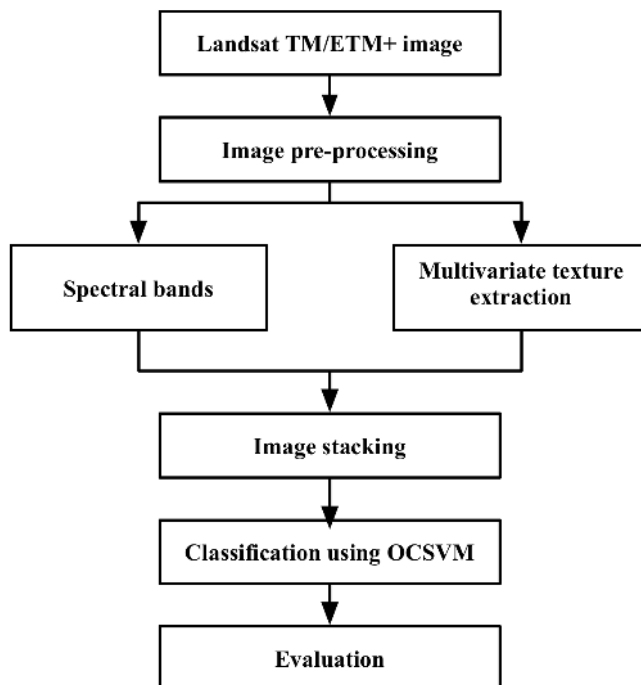
The term “urban area” is widely used in the literature and refers to the spatial extent of urbanized areas on a regional scale, but this is a fuzzy and inconsistent definition [7]. Since urban areas are mainly dominated by built-up lands with impervious surfaces, in this study, urban built-up area is used to represent the area directly occupied by a particular physical structure, including buildings, streets and impervious surfaces, all man-made structures, *i.e.*, a “settlement mask” [7,24].

2. Methods

As mentioned previously, the main purpose of this study is to combine multi-band spectral data and multivariate texture information for the extraction of urban built-up areas using the Landsat series data. The main idea can be summarized as follows (Figure 1). First, multivariate variogram texture [31] is extracted from a Landsat TM/ETM+ multispectral image (Section 2.1), and then, the obtained multivariate texture is combined with the original multispectral image as an additional band. The combined spectral and texture images are classified to extract urban built-up area (Section 2.2). A one-class classifier, one-class support vector machine (OCSVM), is used in the classification

(Section 2.2). To demonstrate the effectiveness of the proposed method, the image texture measured by the classical GLCM is also used in the extraction of the built-up area for comparison.

Figure 1. Flowchart of the proposed method. OCSVM, one-class support vector machine.



2.1. Multivariate Texture Calculation

In this study, multivariate texture was measured using the multivariate variogram [31,40]. The multivariate variogram summarizes the spatial variations of multiple variables (e.g., spectral bands of a multispectral image) [40].

From a perspective of geostatistics, the digital number (DN) of a remote sensing image is considered as a regionalized variable, characterized by both random and spatial correlation aspects [41]. For a multispectral image that can be modeled as multivariate data, each pixel can be viewed as a vector of p spectral bands (variables):

$$DN(x)=[dn_1(x),dn_2(x),...,dn_k(x),...,dn_p(x)]_{1 \times p} \tag{1}$$

where $DN(x)$ is a row vector of pixel x and $dn_k(x)$ is the pixel value of the k -th band. The multivariate variogram for a multispectral image is defined as [31,40]:

$$\gamma_m(\mathbf{h})=\frac{1}{2}E[(DN(x)-DN(x+\mathbf{h}))\mathbf{M}(DN(x)-DN(x+\mathbf{h}))^T] \tag{2}$$

where $\gamma_m(\mathbf{h})$ represents half of the mathematical expectation $E[\bullet]$ of the increments of pixel vector pair values at distance \mathbf{h} , \mathbf{h} is the distance vector, $DN(x)$ and $DN(x + \mathbf{h})$ are pixel vectors at positions x and $x + \mathbf{h}$, respectively, the superscript T denotes the transpose of the matrix and \mathbf{M} is a symmetric positive-definite matrix defining the relations between the bands, such as the identity matrix (Euclidean) and the variance-covariance matrix (Mahalanobis) [40].

The multivariate variogram with Euclidean distance is expressed as:

$$\gamma_{m-Eu}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} \sum_{k=1}^p \{dn_k(x_i) - dn_k(x_i + \mathbf{h})\}^2 \quad (3)$$

where $N(\mathbf{h})$ is the number of distant pairs with lag \mathbf{h} and p is the number of bands.

The multivariate variogram with Mahalanobis distance is defined as:

$$\gamma_{m-Ma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} \left\{ \begin{bmatrix} dn_1(x_i) - dn_1(x_i + \mathbf{h}) \\ \vdots \\ dn_p(x_i) - dn_p(x_i + \mathbf{h}) \end{bmatrix}^T \begin{bmatrix} v_{11} & \cdots & v_{1p} \\ \vdots & \ddots & \vdots \\ v_{p1} & \cdots & v_{pp} \end{bmatrix} \begin{bmatrix} dn_1(x_i) - dn_1(x_i + \mathbf{h}) \\ \vdots \\ dn_p(x_i) - dn_p(x_i + \mathbf{h}) \end{bmatrix} \right\} \quad (4)$$

where the matrix containing element v_{ij} ($i = 1, \dots, p; j = 1, \dots, p$) is the covariance matrix of the multispectral image. In this study, the covariance matrix is calculated for the whole image (*i.e.*, global covariance).

In addition to the Euclidean distance and the Mahalanobis distance, the spectral angle distance [37,42], which has not been used for the extraction of multivariate texture in existing studies, is also used in this paper. The multivariate variogram with the spectral angle distance is expressed as:

$$\gamma_{m-Sa}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} \cos^{-1} \left\{ \frac{\sum_{k=1}^p dn_k(x_i) \cdot dn_k(x_i + \mathbf{h})}{\left[\sum_{k=1}^p (dn_k(x_i))^2 \right]^{1/2} \cdot \left[\sum_{k=1}^p (dn_k(x_i + \mathbf{h}))^2 \right]^{1/2}} \right\} \quad (5)$$

where $N(\mathbf{h})$ is the number of distant pairs with lag \mathbf{h} and p is the number of bands.

The multivariate texture was calculated in a local neighborhood (*i.e.*, moving window) from all multispectral bands using Equations (3)–(5), respectively. There are two parameters to be considered when calculating multivariate texture: the window size and the lag distance (including distance and direction) [31]. The obtained multivariate variogram value (for a specified moving window and a lag distance) is assigned to the central pixel of the window. The selection of an appropriate window size is usually done by a trial and error procedure. Generally, multiple lag distances can be used in the calculation of a multivariate variogram. However, it was found from the existing studies that the lag distance of one pixel is widely used for geostatistical texture computation, because it best describes the radiometric differences in the immediate neighborhood of the central pixel [21]. However, other lag distances can also provide useful information [43]. Usually, the lag distance can be measured in individual directions or in multiple directions within a moving window. In most previous studies, the omnidirectional texture that is produced by averaging the texture values from several directions (*e.g.*, NS, EW, NE-SW, NW-SE) was used [21,31]. However, omnidirectional texture usually shows severe edge effect, which is a well-known problem associated with the use of texture measures [12,34,44]. Thus, in this study, the minimum texture, *i.e.*, the minimum of the texture values for different directions, is used, since it is less sensitive to the edge effect [12]. It should be mentioned that since only one lag distance value was used in the multivariate variogram in this study, the multivariate texture obtained is actually the multivariate semivariance texture.

2.2. Urban Built-Up Area Extraction Using Combined Spectral and Texture Information

The obtained multivariate variogram texture and original multispectral image are combined for urban built-up area extraction. The combination is done by a vector stack method, *i.e.*, the derived multivariate variogram texture image is combined with spectral data as an additional band. For example, if there are six spectral bands of the original image and an extracted multivariate variogram texture band, the combined data will have seven bands. In this paper, multivariate textures measured using the multivariate variogram with three different distances (*i.e.*, Euclidean distance, Mahalanobis distance and spectral angle distance) are separately combined with spectral information. Thus, three data combinations are produced. In addition, to fully validate the performance of the proposed method, two other data combinations are used for comparison: (1) spectral information alone; and (2) spectral information and the one-band GLCM texture feature. There are five kinds of data combinations in total.

Since the spectral and texture features generally have different data ranges, the classification process may be dominated by data sources that display a larger scale of variation [45]. With such different data ranges, the classification directly using the original spectral features and the textural features may not achieve a satisfactory classification result. Therefore, prior to classification, both spectral and texture features are first normalized into the same range [0, 1]. The normalization can be formulated as:

$$d'_{x_i} = \frac{d_{x_i} - d_{\min}}{d_{\max} - d_{\min}} \quad (6)$$

where d_{x_i} denotes the original value of the pixel x_i , d_{\max} and d_{\min} represent the maximum and minimum values of that band and d'_{x_i} is the normalized value of the pixel x_i .

The combined spectral and texture images are classified using a supervised classifier to extract urban built-up area. Since the urban built-up area is the only class of interest (or target class), instead of using traditional multi-class classification methods, a one-class classifier, OCSVM, is adopted in this study.

The OCSVM is a recently developed one-class classifier, which is a special type of classical binary support vector machine (SVM). The OCSVM classifier only requires training data from one class (the target class) and can especially focus on the class. In the training process, only samples from the target class are used, and there is no information about outlier objects. The boundary between the target class and others has to be estimated from data in the only available target class. Thus, the task is to define a boundary around the target class; the boundary encircles as many target examples as possible and minimizes the chance of accepting outlier objects [33,46]. Furthermore, the OCSVM classifier is adopted in this study since it has also been successfully applied in specific land cover type mapping [47], change detection of one specific land cover class (e.g., urban expansion and urban building damage detection) [33] and ecological modeling [48].

2.3. Accuracy Assessment

A common method for accuracy assessment is the use of a confusion matrix. For each classification result of different data combinations, the accuracy assessment includes the producer's accuracy (PA),

user's accuracy (UA), overall accuracy (OA) and kappa coefficient. The kappa coefficient is a measure of the overall statistical agreement of a matrix and takes non-diagonal elements into account. Kappa analysis has been recognized as a powerful technique for analyzing a single error matrix and comparing different confusion matrices [49,50]. In this study, McNemar's test was used to evaluate the statistical significance of the difference between two classification results using the same set of reference data [51]. If the test value, $|Z|$, is greater than 1.96, the two classification results produce significantly different accuracies at a 95 percent confidence interval.

3. Study Areas and Data

Two megacity areas in China, Beijing and Tianjin, were selected as study areas for the evaluation of the proposed method. Beijing, the capital city of China, is located in northern China. Beijing has experienced a rapid urbanization in the past three decades, resulting in increased infrastructural and housing construction and urban expansion. The land cover types in the area include water, farmland, grassland, woodland, built-up area and bare land.

Bi-temporal Landsat TM/ETM+ data that were acquired in September 1995, and in July 2005, were used. Only six reflective multispectral bands (Bands 1–5, 7) with 30-m spatial resolution were used for each image. The image is free of clouds and of a high quality. No radiometric correction was conducted, since land cover classification does not gain from radiometric correction [52]. Bi-temporal images were co-registered. The image size finally used in the study is 1600×1860 pixels, covering an area of approximately $48 \text{ km} \times 55.8 \text{ km}$ (Figure 2). The training samples for the target class (built-up area) were visually selected from bi-temporal images. For accuracy assessment, 500 randomly distributed pixels were first generated for both built-up area and non-built-up area classes. After removing the pixels that were located near class boundaries and difficult in the determination of their class attributes and that were near the training pixels, a small polygonal area (cluster) (approximately 80–110 pixels each) around each remaining pixel was then selected as a test area. The training and test samples for the bi-temporal images of Beijing are shown in Table 1.

Figure 2. Landsat TM images of the Beijing study area (7, 4 and 2 bands as R, G and B) acquired in 1995 (a) and 2005 (b).

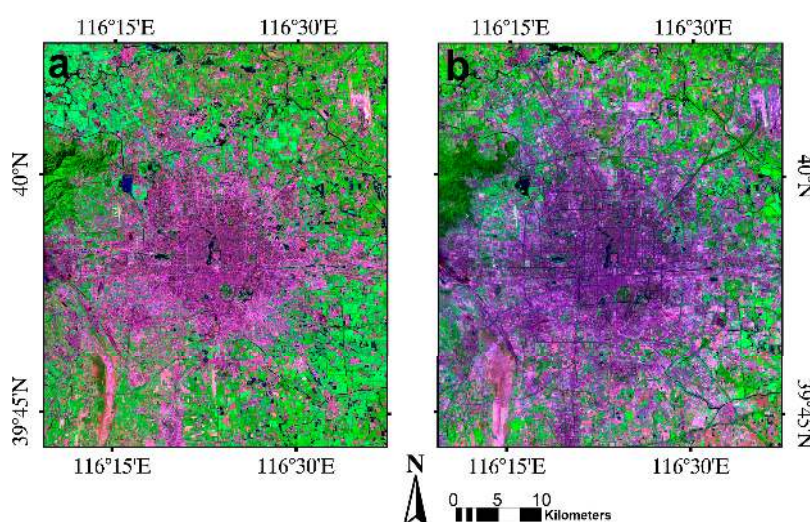
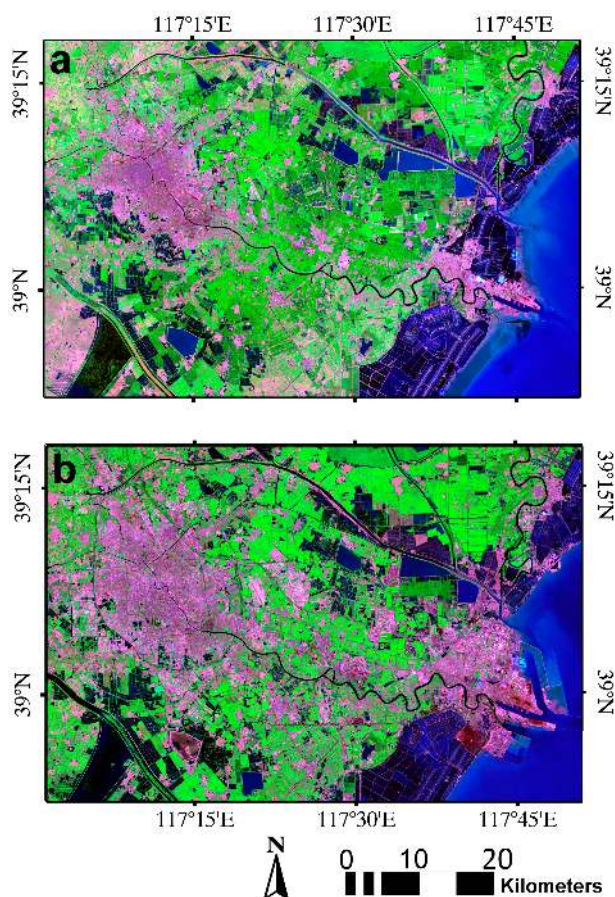


Table 1. Training and test samples used for the two study areas (pixels).

	Training Samples		Test Samples	
	Built-Up		Built-Up	Non-Built-Up
Beijing				
1995 image	17,703		18,861	25,608
2005 image	13,624		17,608	22,522
Tianjin				
1992 image	7717		10,220	18,547
2006 image	7154		13,332	17,280

The other study area, Tianjin, is also located in northern China. Tianjin is the third largest city and the second port city of China, which has experienced rapid urban and industrial development over twenty years. The land cover types in the area are the same as those in Beijing area. Two-temporal Landsat TM/ETM+ data acquired in July 1992, and in July 2006, were selected. Only six reflective multispectral bands (Bands 1–5, 7) with 30-m spatial resolution were used. The image is free of clouds and of a good quality. After image registration, the image size used in the study is 2400×1600 pixels, covering an area of approximately $72 \text{ km} \times 48 \text{ km}$ (Figure 3). The training and test samples were selected following the same procedure used for Beijing images. The training and test samples for bi-temporal images of Tianjin are shown in Table 1.

Figure 3. Landsat TM image of the Tianjin urban area (7, 4 and 2 bands as R, G and B) acquired in 1992 (a) and 2006 (b).

4. Results and Discussion

4.1. Built-Up Area Extraction Results in Beijing

The multivariate variogram texture was first extracted. In order to choose an appropriate window size for the final texture computation and subsequent combined classification, several window sizes, *i.e.*, 3×3 , 5×5 , 7×7 , 9×9 pixels, with a lag distance of one pixel, were respectively used to calculate image textures, which were compared in terms of the accuracy of the combined classification. Considering that spatial variations for these two areas are similar, the Landsat TM image of 2005 was used to estimate the optimal window size. The classification accuracies produced using spectral data and multivariate variogram texture data with different window sizes are shown in Table 2.

Table 2. Classification accuracies using spectral data and multivariate variogram texture data extracted with different window sizes for the 2005 image of the Beijing area (all in %).

Data Combinations	OA	Kappa	Built-Up Class		Non-Built-Up Class	
			PA	UA	PA	UA
Spectral + MV _{3_1} (Eu)	90.90	81.61	91.97	87.85	90.06	93.48
Spectral + MV _{5_1} (Eu)	91.23	82.23	91.15	89.10	91.28	92.96
Spectral + MV _{7_1} (Eu)	91.68	83.05	89.07	91.72	93.72	91.64
Spectral + MV _{9_1} (Eu)	91.58	82.84	88.57	91.95	93.94	91.31
Spectral + MV _{3_1} (Ma)	90.70	81.04	87.43	91.02	93.26	90.47
Spectral + MV _{5_1} (Ma)	90.52	80.84	91.27	87.64	89.93	92.95
Spectral + MV _{7_1} (Ma)	91.56	82.91	91.36	89.62	91.72	93.14
Spectral + MV _{9_1} (Ma)	91.68	83.11	90.54	90.50	92.57	92.60
Spectral + MV _{3_1} (Sa)	91.78	83.43	93.87	88.16	90.15	94.95
Spectral + MV _{5_1} (Sa)	92.15	84.14	93.08	89.46	91.43	94.42
Spectral + MV _{7_1} (Sa)	92.35	84.49	91.87	90.81	92.73	93.58
Spectral + MV _{9_1} (Sa)	92.31	84.43	92.54	90.19	92.13	94.04

OA: overall accuracy; PA: producer's accuracy; UA: user's accuracy; spectral: six TM bands; MV_{w_1}: multivariate variogram texture with a window size of $w \times w$ pixels and a lag distance of one pixel; Eu: Euclidean distance; Ma: Mahalanobis distance; Sa: Spectral angle distance.

From Table 2, it was found that the window size of 7×7 pixels used for texture calculation generally achieved the highest kappa coefficient among all window sizes. Hence, the window size of 7×7 pixels with a lag distance of one pixel was finally used in the multivariate variogram texture calculation.

In addition, eight GLCM textures were also calculated, namely mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation. As in many existing studies, the red band and near-infrared (NIR) band were used to compute eight GLCM textures [38,53]. In the process of GLCM texture calculation, a window size of 7×7 pixels and a one pixel offset for all directions were computed; the minimum of the texture values for different directions was finally used. After visual inspection and quantitative comparison, the dissimilarity texture computed from the NIR band

was used in subsequent analysis, since it achieved the highest accuracy among all combined spectral and GLCM texture classifications (Table 3).

Table 3. Classification accuracies using multispectral data and selected gray-level co-occurrence matrix (GLCM) textures for a 2005 image of the Beijing area (all in %).

Data Combinations	OA	Kappa	Built-Up Class	
			PA	UA
Spectral + Var	88.75	77.47	93.71	82.88
Spectral + Con	89.74	79.37	92.95	85.05
Spectral + SM	89.56	79.03	93.11	84.64
Spectral + Hom	90.12	80.15	93.71	85.23
Spectral + Dis	90.64	80.96	88.62	89.90
Spectral + Dis (Red)	89.09	77.86	87.90	87.32

OA: overall accuracy; PA: producer's accuracy; UA: user's accuracy; spectral: six TM bands; V: variance texture; C: contrast texture; SM: second moment texture; Hom: homogeneity texture; Dis: dissimilarity texture; Dis (Red): dissimilarity texture calculated from the red band. Textures that do not mention the band were calculated from the NIR band.

Table 4. Classification accuracies using multispectral data and texture features from Beijing images (all in %).

Data Combinations	OA	Kappa	Built-Up Class		Non-Built-Up Class	
			PA	UA	PA	UA
1995						
Spectral	89.20	77.96	88.29	86.53	89.87	91.25
Spectral + GLCM_Dis	89.73	78.98	87.92	87.88	91.07	91.10
Spectral + MV_Eu	89.82	79.11	87.08	88.70	91.83	90.61
Spectral + MV_Ma	89.90	79.29	87.48	88.56	91.68	90.86
Spectral + MV_Sa	91.72	83.03	90.04	90.39	92.95	92.68
2005						
Spectral	88.07	75.82	87.16	85.87	88.79	89.84
Spectral + GLCM_Dis	90.64	80.96	88.62	89.90	92.21	91.20
Spectral + MV_Eu	91.68	83.05	89.07	91.72	93.72	91.64
Spectral + MV_Ma	91.56	82.91	91.36	89.62	91.72	93.14
Spectral + MV_Sa	92.35	84.49	91.87	90.81	92.73	93.58

OA: overall accuracy; PA: producer's accuracy; UA: user's accuracy; spectral: spectral data; GLCM_Dis: dissimilarity texture from GLCM; MV_Eu: multivariate variogram texture with Euclidean distance; MV_Ma: multivariate variogram texture with Mahalanobis distance; MV_Sa: multivariate variogram texture with spectral angle distance.

Table 4 shows classification accuracies using different data combinations from the Beijing images. From the table, the classification using spectral information alone produced the lowest accuracies, with kappa coefficients of 75.82% and 77.96% for the 1995 and 2005 images, respectively. By adding textures, either one-band texture (GLCM dissimilarity) or multivariate textures, the classification accuracies were highly improved compared to the use of spectral information alone. For example, by adding one-band texture (GLCM dissimilarity), the increases in the kappa coefficient were 5.14% and

1.02% for the 1995 and 2005 images, respectively. Furthermore, by adding multivariate textures, the accuracies of combined classification were higher than that of classification including one-band texture. In particular, the inclusion of multivariate texture with spectral angle distance in combined classification achieved the highest accuracy among the results of adding one of three multivariate textures. For example, compared to the results using spectral information alone, the increases in the kappa coefficient were 5.07% and 8.67% for the 1995 and 2005 images, respectively; while compared to the results using spectral information and one-band texture, the increases in the kappa coefficient were 2.07% and 5.51% for the 1995 and 2005 images, respectively (Table 4).

Table 5 shows $|Z|$ values from McNemar's test for Beijing images, which were used to quantify the statistical significance of the difference between the two classification results. From Table 5, it is clear for two images that, compared with the classification results using spectral information alone, all classification results using combined spectral information and texture (GLCM or multivariate variogram) achieved significantly higher accuracies (at the 95% confidence level). Furthermore, the classification result using spectral information and multivariate variogram texture with spectral angle distance generated significantly higher accuracy than all other combinations of spectral data and texture (both GLCM and multivariate variograms with Euclidean and Mahalanobis distances).

Table 5. The $|Z|$ values from McNemar's test for the comparison of kappa coefficients from different classification results of Beijing images and corresponding statistical significances (at the 95% confidence level).

	Spectral	Spectral + GLCM_Dis	Spectral + MV_Eu	Spectral + MV_Ma	Spectra + MV_Sa
1995					
Spectral	\	S	S	S	S
Spectral + GLCM_Dis	2.4386	\	NS	NS	S
Spectral + MV_Eu	2.8265	0.3880	\	NS	S
Spectral + MV_Ma	3.2054	0.7671	0.3791	\	S
Spectral + MV_Sa	12.1263	9.7105	9.3136	8.9359	\
2005					
Spectral	\	S	S	S	S
Spectral + GLCM_Dis	11.1329	\	S	S	S
Spectral + MV_Eu	16.0409	4.9499	\	NS	S
Spectral + MV_Ma	15.4989	4.4018	0.5487	\	S
Spectral + MV_Sa	19.3730	8.3273	3.3851	3.9334	\

S: significant; NS: not significant.

Figures 4 and 5 show the three selected classification results for the 1995 and 2005 images of the Beijing area, respectively. For the 1995 image, it is seen that the results using different data combinations have a similar distribution of the built-up area over the whole scene (Figure 4). For the 2005 image, one obvious difference among these results is shown in the lower left region of three images: there is an area that was labeled as different classes by different methods (Figure 5). For two sets of results, differences among each set of results usually happened in detailed local regions, such as in rural settlement areas, harvested farmland area and built-up area that has unusual spectral

information corresponding to surrounding built-up areas. Thus, some local places in each image have been selected for comparison (Figure 6).

Figure 4. Classification results of the 1995 Beijing image using spectral information alone (a); using spectral information and one-band texture (b); and using spectral information and multivariate texture with spectral angle distance (c).

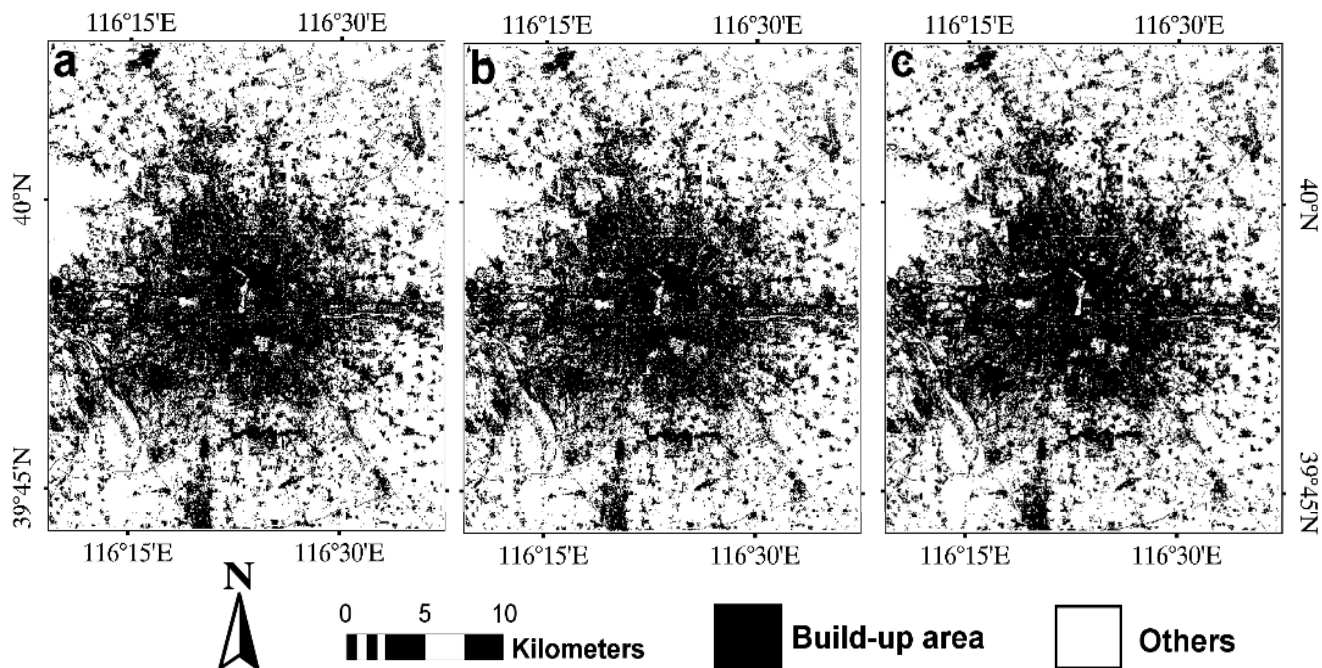


Figure 5. Classification results of the 2005 Beijing image using spectral information alone (a); using spectral information and one-band texture (b); and using spectral information and multivariate texture with spectral angle distance (c).

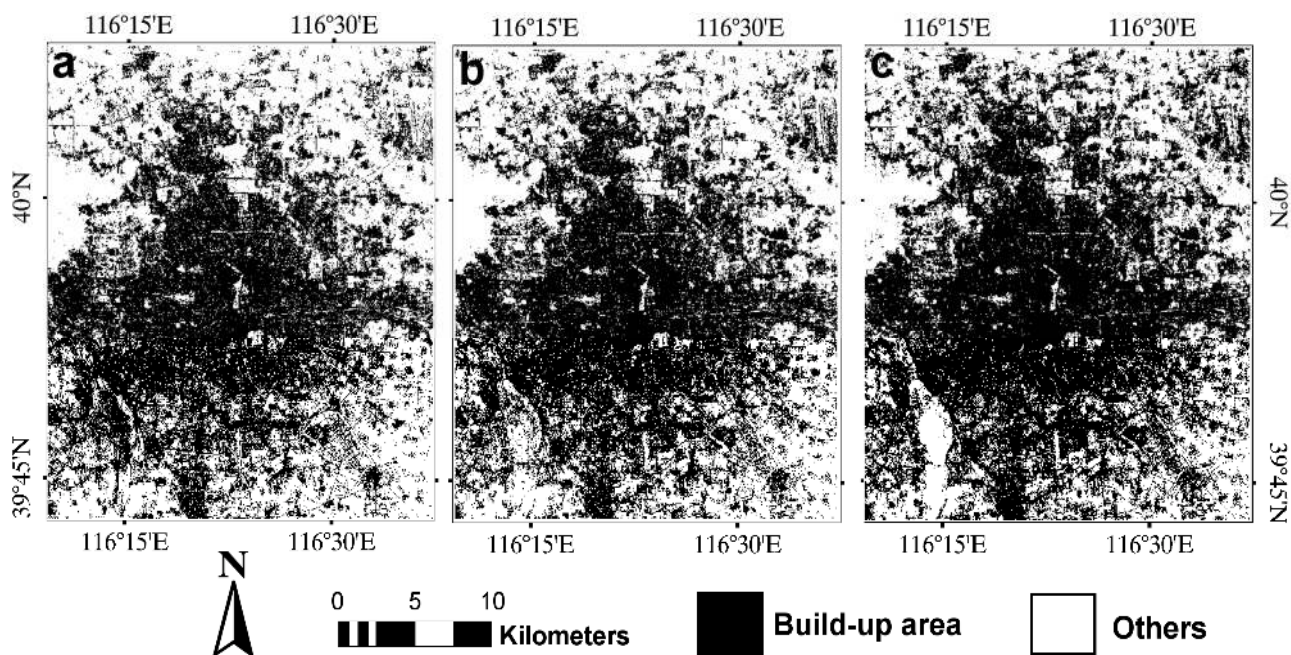
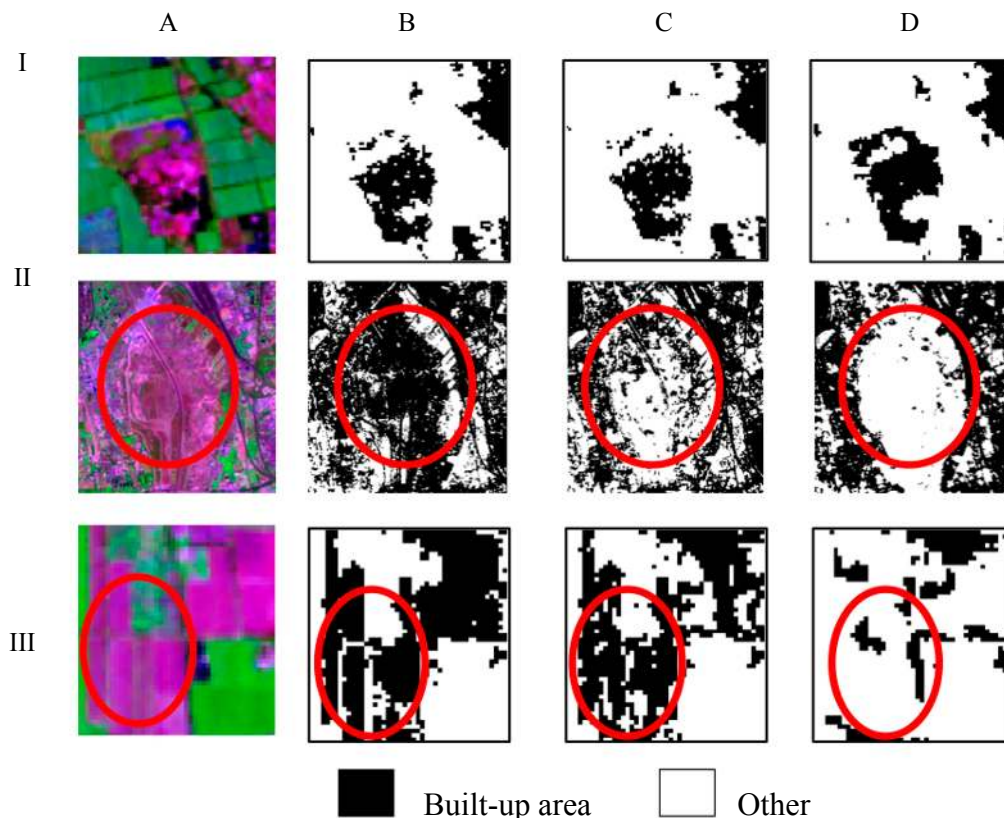


Figure 6. Portions of Landsat TM images (Bands 7, 4 and 2 as R, G and B) (A); I-A: built-up area (mostly in the middle, red color) and surrounding vegetated farmland; II-A: bare land (mostly in the middle, red color) and neighboring urban built-up area; III-A: harvested farmland (red color) and vegetation; corresponding classification results of different data combinations in Beijing area. (B) Spectral data alone; (C) spectral data and GLCM texture; (D) spectral data and variogram texture with spectral angle distance.



From Figure 6, by using spectral data alone and using combined spectral data and GLCM texture, the urban built-up areas were partially recognized (e.g., I-B, I-C of Figure 6). However, when spectral data and multivariate variogram texture were jointly used, the built-up area was more accurately extracted. When spectral data were used alone, some areas of bare land and harvested land were misclassified as urban built-up area (e.g., II-B, III-B of Figure 6). When GLCM texture was included in classification, the misclassification error was slightly reduced (e.g., II-C, III-C of Figure 6). However, when multivariate variogram texture was included in classification, the misclassification error was significantly reduced (e.g., II-D, III-D of Figure 6).

In summary, the proposed method outperformed the use of spectral information alone and the joint use of spectral information and GLCM texture in bi-temporal images of the Beijing area. Specifically, the proposed method significantly reduced the confusion between built-up area and bare land and the confusion between built-up area and harvested farmland. These results demonstrate the effectiveness of the proposed methods in urban built-up area extraction.

4.2. Built-Up Area Extraction Results in Tianjin Area

Table 6 shows classification results using different data combinations in the Tianjin area. From the Table 6, urban built-up area extraction using spectral information alone produced the lowest accuracies. By adding one-band GLCM texture, the accuracy of the combined classification was slightly improved compared to the use of spectral information alone. However, by adding multivariate variogram textures, the classification accuracies were significantly higher than the use of spectral data alone and the inclusion of one-band GLCM texture. In particular, the inclusion of multivariate variogram texture with spectral angle distance in classification achieved the highest accuracy among the results of using one of three multivariate variogram textures. For example, compared to the results using spectral information and GLCM texture, the increases in the kappa coefficient for the 1992 and 2006 images were 8.30% and 7.55%, respectively.

Table 6. Classification accuracies using multispectral data and texture features of the Tianjin area (all in %).

Data Combinations	OA	Kappa	Built-Up Class		Non-Built-Up Class	
			PA	UA	PA	UA
1992						
Spectral	86.54	70.64	81.24	80.94	89.46	89.64
Spectral + GLCM_Dis	86.86	71.37	81.92	81.25	89.58	89.99
Spectral + MV_Eu	89.31	76.70	85.23	84.76	91.56	91.84
Spectral + MV_Ma	88.20	74.15	82.58	83.95	91.30	90.49
Spectral + MV_Sa	90.72	79.67	86.19	87.50	93.21	92.45
2006						
Spectral	85.21	69.92	82.97	83.06	86.94	86.87
Spectral + GLCM_Dis	86.07	71.60	83.00	84.70	88.44	87.08
Spectral + MV_Eu	86.72	72.96	84.32	85.05	88.56	87.98
Spectral + MV_Ma	86.64	72.83	84.66	84.67	88.17	88.17
Spectral + MV_Sa	89.54	78.65	86.49	89.16	91.89	89.81

OA: overall accuracy; PA: producer's accuracy; UA: user's accuracy; spectral: spectral data; GLCM_Dis: dissimilarity texture from GLCM; MV_Eu: multivariate variogram texture with Euclidean distance; MV_Ma: multivariate variogram texture with Mahalanobis distance; MV_Sa: multivariate variogram texture with spectral angle distance.

Table 7 shows $|Z|$ values from McNemar's test for the Tianjin images, which were used to quantify the statistical significance of the difference between two classification results. From the table, it is similar to the Beijing area, *i.e.*, compared with the classification results using spectral information alone, all classification results using combined spectral information and texture (GLCM or multivariate variogram) achieved significantly higher accuracies (at the 95% confidence level). Furthermore, the classification result using spectral information and multivariate variogram texture with spectral angle distance generated significantly higher accuracy than all other combinations of spectral data and texture (both GLCM and multivariate variograms with Euclidean and Mahalanobis distances).

Table 7. The $|Z|$ values from McNemar’s test for the comparison of kappa coefficients from different classification results of the Tianjin images and corresponding statistical significances (at the 95% confidence level).

	Spectral	Spectral + GLCM_Dis	Spectral + MV_Eu	Spectral + MV_Ma	Spectral + MV_Sa
1992					
Spectral	\	NS	S	S	S
Spectral + GLCM_Dis	1.0517	\	S	S	S
Spectral + MV_Eu	9.5622	8.5150	\	S	S
Spectral + MV_Ma	5.6076	4.5573	3.9662	\	S
Spectral + MV_Sa	14.8610	13.8211	5.3433	9.2973	\
2006					
Spectral	\	S	S	S	S
Spectral + GLCM_Dis	2.7942	\	S	NS	S
Spectral + MV_Eu	4.9731	2.1802	\	NS	S
Spectral + MV_Ma	4.7187	1.9256	0.2547	\	S
Spectral + MV_Sa	15.0591	12.2892	10.1222	10.3755	\

S: significant; NS: not significant.

Figures 7 and 8 show the four selected classification results for the 1992 and 2006 images, respectively. From these figures, it is clear that the salt-and-pepper appearance was reduced in the classification results produced by the proposed methods (Figures 7c,d and 8c), compared with the results using spectral data alone (Figures 7a and 8a) and the results including GLCM texture (Figures 7b and 8b), especially in the lower right region of these images.

Figure 7. Classification results of the 1992 Tianjin image using spectral information alone (a); using spectral information and one-band texture (b); using spectral information and multivariate variogram texture with spectral angle distance (c).

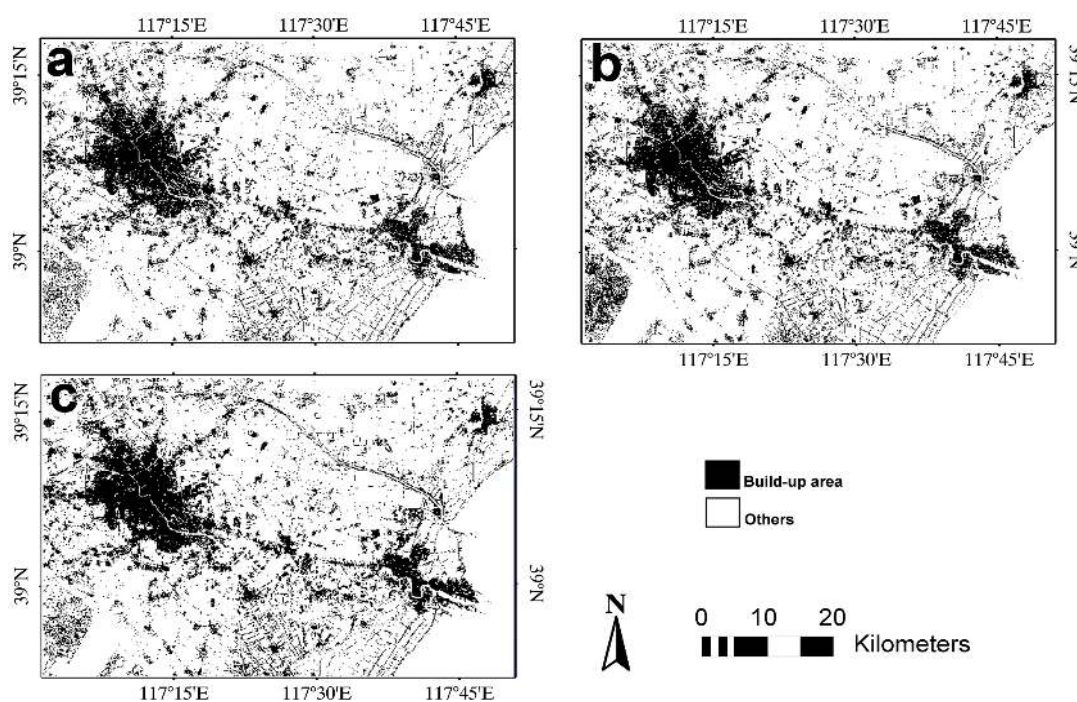


Figure 8. Classification results of the 2006 Tianjin image using spectral information alone (a); using spectral information and one-band texture (b); using spectral information and multivariate variogram texture with spectral angle distance (c).

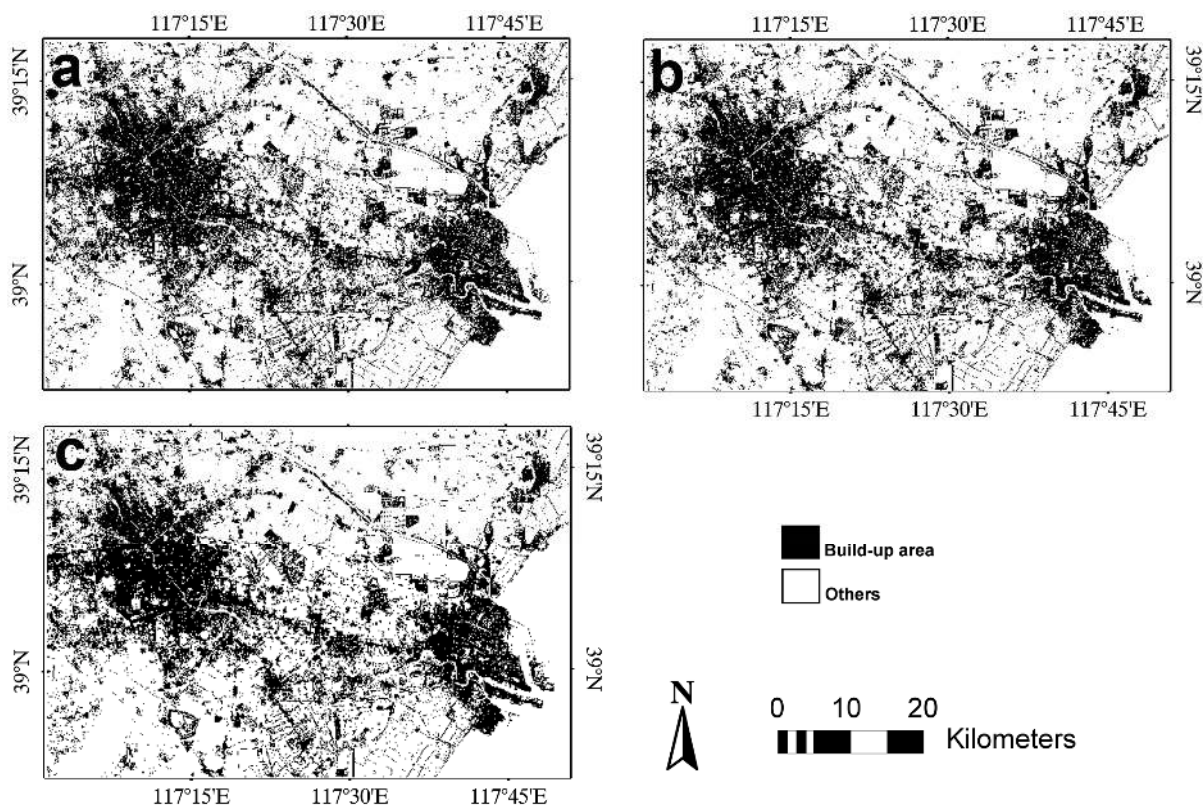
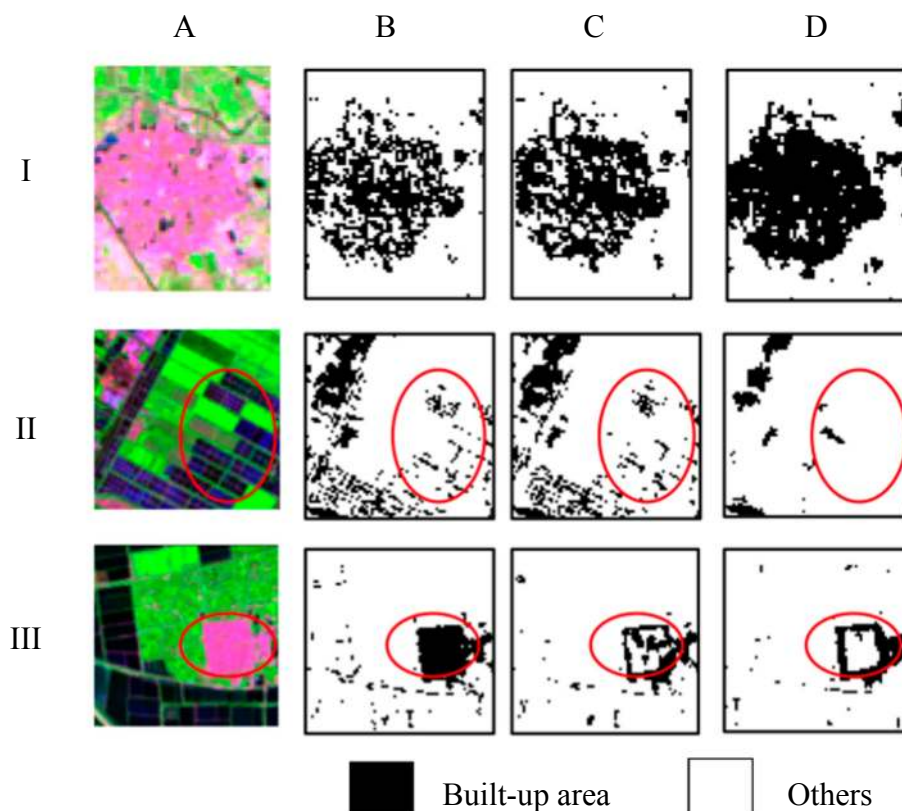


Figure 9 shows some portions of the classification results using different data combinations. From the figure, using spectral data alone or combining spectral data and GLCM texture, urban built-up areas were partially recognized (e.g., I-B, I-C of Figure 9). However, when spectral data and multivariate variogram texture were combined in the classification, the built-up area was more accurately extracted (e.g., I-D of Figure 9). When spectral data were used alone, some areas of water and harvested farmland were misclassified as urban built-up area (e.g., II-B, III-B of Figure 9). When GLCM texture was included in combined classification, the misclassification errors were slightly reduced (e.g., II-C, III-C of Figure 9). However, when multivariate variogram texture was included in classification, these misclassification errors were significantly reduced (e.g., II-D, III-D of Figure 9).

It is of note that there are some misclassification errors occurring at the boundaries between urban built-up area and other land cover classes in the classification result by the proposed method (e.g., III-D of Figure 9). These misclassification errors are mainly produced by the edge effect of the texture property [54]. In general, as in the first example (the Beijing case), the proposed method outperformed the use of spectral information alone and the joint use of spectral information and GLCM texture in bi-temporal images of the Tianjin area. Specifically, the proposed method significantly reduced the confusion between built-up area and water area and the confusion between built-up area and harvested farmland. These results demonstrated the effectiveness of the proposed methods in urban built-up area extraction.

Figure 9. Portions of Landsat TM images (Bands 7, 4 and 2 as R, G and B) (A); I-A: built-up area (mostly in the middle) and surrounding farmland; II-A: farmland and water area (blue color areas); III-A: harvested farmland (red color, in the middle) and vegetated farmland. Corresponding classification results using different data combinations in Tianjin area. (B) Spectral data alone; (C) spectral data and GLCM texture; (D) spectral data and multivariate variogram texture with spectral angle distance



5. Conclusions

Timely and accurate information about urban areas is very important for diverse applications. Due to significant spectral heterogeneity and spectral confusion with other land cover classes at 30-m resolution (e.g., Landsat TM/ETM+ data), the urban area extraction using spectral data from Landsat series data alone is a challenging task. This paper proposed a method that combines multivariate variogram texture and spectral data in urban area classification. The OCSVM classifier used in this study only requires the training samples from the target class (*i.e.*, urban built-up area), while the traditional multi-class classifiers require training samples from all land cover classes. The proposed method was evaluated using bi-temporal Landsat TM/ETM+ images from two megacity areas. The experimental results indicated that the inclusion of multivariate variogram texture outperformed the use of spectral data alone and the inclusion of classical GLCM texture in classification. In particular, the inclusion of multivariate variogram texture with spectral angle distance achieved increases in the kappa coefficient from 4% to 9% compared to all other data combinations. The increases in the kappa coefficient are statistically significant in terms of McNemar's test (at the 95% confidence level). The proposed method that employs multivariate texture and a one-class classifier (e.g., OCSVM) provides

an effective and efficient way of extracting urban built-up area. The method can be also applicable to other relevant applications.

Acknowledgements

This work was funded by National Science Foundation of China (Grant Number 41371329). We would like to thank anonymous reviewers for their constructive comments, which greatly improved the quality of our manuscript.

Author Contributions

The main idea of extracting built-up area by multivariate texture and spectral information was proposed by Peijun Li. The experiments were carried out by Jun Zhang, who also prepared the figures. Jinfei Wang provided much in-depth analysis of the results. This manuscript was written by Jun Zhang.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Chameides, W.L.; Kasibhatla, P.S.; Yienger, J.; Levy, H. Growth of continental-scale metro-agro-plexes, regional ozone pollution, and world food production. *Science* **1994**, *264*, 74–77.
2. Guindon, B.; Zhang, Y.; Dillabaugh, C. Landsat urban mapping based on a combined spectral-spatial methodology. *Remote Sens. Environ.* **2004**, *92*, 218–232.
3. Small, C. A global analysis of urban reflectance. *Int. J. Remote Sens.* **2005**, *26*, 661–681.
4. Griffiths, P.; Hostert, P.; Gruebner, O.; van der Linden, S. Mapping megacity growth with multi-sensor data. *Remote Sens. Environ.* **2010**, *114*, 426–439.
5. Sun, Z.; Guo, H.; Li, X.; Lu, L.; Du, X. Estimating urban impervious surfaces from Landsat-5 TM imagery using multilayer perceptron neural network and support vector machine. *J. Appl. Remote Sens.* **2011**, doi:10.1117/1.3539767.
6. Schneider, A. Monitoring land cover change in urban and pen-urban areas using dense time stacks of Landsat satellite data and a data mining approach. *Remote Sens. Environ.* **2012**, *124*, 689–704.
7. Taubenböck, H.; Esch, T.; Felbier, A.; Wiesner, M.; Roth, A.; Dech, S. Monitoring urbanization in mega cities from space. *Remote Sens. Environ.* **2012**, *117*, 162–176.
8. Wulder, M.A.; White, J.C.; Masek, J.G.; Dwyer, J.; Roy, D.P. Continuity of landsat observations: Short term considerations. *Remote Sens. Environ.* **2011**, *115*, 747–751.
9. Small, C.; Lu, J.W.T. Estimation and vicarious validation of urban vegetation abundance by spectral mixture analysis. *Remote Sens. Environ.* **2006**, *100*, 441–456.
10. Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594.
11. Xu, H. Extraction of urban built-up land features from landsat imagery using a thematic-oriented index combination technique. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 1381–1391.

12. Pesaresi, M.; Gerhardinger, A.; Kayitakire, F. A robust built-up area presence index by anisotropic rotation-invariant textural measure. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2008**, *1*, 180–192.
13. Xu, H. A new index for delineating built-up land features in satellite imagery. *Int. J. Remote Sens.* **2008**, *29*, 4269–4276.
14. Pesaresi, M.; Gerhardinger, A. Improved textural built-up presence index for automatic recognition of human settlements in arid regions with scattered vegetation. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2011**, *4*, 16–26.
15. Deng, C.; Wu, C. BCI: A biophysical composition index for remote sensing of urban environments. *Remote Sens. Environ.* **2012**, *127*, 247–259.
16. Weng, Q. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sens. Environ.* **2012**, *117*, 34–49.
17. Zhang, Q.; Wang, J.; Peng, X.; Gong, P.; Shi, P. Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data. *Int. J. Remote Sens.* **2002**, *23*, 3057–3078.
18. He, C.; Wei, A.; Shi, P.; Zhang, Q.; Zhao, Y. Detecting land-use/land-cover change in rural–urban fringe areas using extended change-vector analysis. *Int. J. Appl. Earth Observ. Geoinf.* **2011**, *13*, 572–585.
19. Lyons, M.B.; Phinn, S.R.; Roelfsema, C.M. Long term land cover and seagrass mapping using Landsat and object-based image analysis from 1972 to 2010 in the coastal environment of South East Queensland, Australia. *ISPRS J. Photogramm. Remote Sens.* **2012**, *71*, 34–46.
20. Haralick, R.M.; Karthikeyan, S.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybernet.* **1973**, *6*, 610–621.
21. Chica-Olmo, M.; Abarca-Hernandez, F. Computing geostatistical image texture for remotely sensed data classification. *Comput. Geosci.* **2000**, *26*, 373–383.
22. Wentz, E.A.; Stefanov, W.L.; Gries, C.; Hope, D. Land use and land cover mapping from diverse data sources for an arid urban environments. *Comput. Environ. Urban Syst.* **2006**, *30*, 320–346.
23. Leinenkugel, P.; Esch, T.; Kuenzer, C. Settlement detection and impervious surface estimation in the Mekong delta using optical and SAR remote sensing data. *Remote Sens. Environ.* **2011**, *115*, 3007–3019.
24. Bagan, H.; Yamagata, Y. Landsat analysis of urban growth: How Tokyo became the world's largest megacity during the last 40 years. *Remote Sens. Environ.* **2012**, *127*, 210–222.
25. Huang, B.; Zhang, H.; Yu, L. Improving Landsat ETM Plus urban area mapping via spatial and angular fusion with MISR multi-angle observations. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2012**, *5*, 101–109.
26. Zhu, Z.; Woodcock, C.E.; Rogan, J.; Kellndorfer, J. Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data. *Remote Sens. Environ.* **2012**, *117*, 72–82.
27. Kenduiywo, B.K.; Tolpekin, V.A.; Stein, A. Detection of built-up area in optical and synthetic aperture radar images using conditional random fields. *J. Appl. Remote Sens.* **2014**, doi:10.1117/1.JRS.8.083672.

28. Karathanassi, V.; Iossifidis, C.; Rokos, D. A texture-based classification method for classifying built areas according to their density. *Int. J. Remote Sens.* **2000**, *21*, 1807–1823.
29. Stefanov, W.L.; Ramsey, M.S.; Christensen, P.R. Monitoring urban land cover change: An expert system approach to land cover classification of semiarid to arid urban centers. *Remote Sens. Environ.* **2001**, *77*, 173–185.
30. Dekker, R.J. Texture analysis and classification of ERS SAR images for map updating of urban areas in the Netherlands. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1950–1958.
31. Li, P.; Cheng, T.; Guo, J. Multivariate image texture by multivariate variogram for multispectral image classification. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 147–157.
32. Li, P.; Yu, H.; Cheng, T. Lithologic mapping using ASTER imagery and multivariate texture. *Can. J. Remote Sens.* **2009**, *35*, S117–S125.
33. Li, P.; Xu, H.; Guo, J. Urban building damage detection from very high resolution imagery using OCSVM and spatial features. *Int. J. Remote Sens.* **2010**, *31*, 3393–3409.
34. Li, P.; Xu, H.; Song, B. A novel method for urban road damage detection using very high resolution satellite imagery and road map. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 1057–1066.
35. Sohn, Y.S.; Moran, E.; Gurri, F. Deforestation in north-central Yucatan (1985–1995): Mapping secondary succession of forest and agricultural land use in sotuta using the cosine of the angle concept. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 947–958.
36. Guo, L.J.; Moore, J.M. Cloud-shadow suppression technique for enhancement of airborne thematic mapper imagery. *Photogramm. Eng. Remote Sens.* **1993**, *59*, 1287–1291.
37. Kruse, F.A.; Lefkoff, A.B.; Boardman, J.W.; Heidebrecht, K.B.; Shapiro, A.T.; Barloon, P.J.; Goetz, A.F.H. The spectral image-processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.* **1993**, *44*, 145–163.
38. Lu, D.; Hetrick, S.; Moran, E. Land cover classification in a complex urban-rural landscape with Quickbird imagery. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 1159–1168.
39. Pouch, G.W.; Campagna, D.J. Hyperspherical direction cosine transformation for separation of spectral and illumination information in digital scanner data. *Photogramm. Eng. Remote Sens.* **1990**, *56*, 475–479.
40. Bourgault, G.; Marcotte, D. Multivariable variogram and its application to the linear-model of coregonalization. *Math. Geol.* **1991**, *23*, 899–928.
41. Ramstein, G.; Raffy, M. Analysis of the structure of radiometric remotely-sensed images. *Int. J. Remote Sens.* **1989**, *10*, 1049–1073.
42. Murphy, R.J.; Monteiro, S.T.; Schneider, S. Evaluating classification techniques for mapping vertical geology using field-based hyperspectral sensors. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3066–3080.
43. Lark, R.M. Geostatistical description of texture on an aerial photograph for discriminating classes of land cover. *Int. J. Remote Sens.* **1996**, *17*, 2115–2133.
44. Ferro, C.J.S.; Warner, T.A. Scale and texture in digital image classification. *Photogramm. Eng. Remote Sens.* **2002**, *68*, 51–63.
45. Mather, P.; Tso, B. *Classification Methods for Remotely Sensed Data*; CRC Press: Boca Raton, FL, USA, 2003.

46. Tax, D.M.; Duin, R.P. Combining one-class classifiers. In *Multiple Classifier Systems*; Springer-Berlin: Heidelberg, Germany, 2001; pp. 299–308.
47. Sanchez-Hernandez, C.; Boyd, D.S.; Foody, G.M. One-class classification for mapping a specific land-cover class: SVDD classification of fenland. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1061–1073.
48. Guo, Q.; Kelly, M.; Graham, C.H. Support vector machines for predicting distribution of Sudden Oak Death in California. *Ecol. Modell.* **2005**, *182*, 75–90.
49. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46.
50. Smits, P.C.; Dellepiane, S.G.; Schowengerdt, R.A. Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach. *Int. J. Remote Sens.* **1999**, *20*, 1461–1486.
51. Foody, G.M. Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 627–633.
52. Song, C.; Woodcock, C.E.; Seto, K.C.; Lenney, M.P.; Macomber, S.A. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sens. Environ.* **2001**, *75*, 230–244.
53. Marceau, D.J.; Howarth, P.J.; Dubois, J.M.M.; Gratton, D.J. Evaluation of the gray-level cooccurrence matrix-method for land-cover classification using SPOT imagery. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 513–519.
54. Maillard, P. Comparing texture analysis methods through classification. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 357–367.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).