

Article

# Representation of Block-Based Image Features in a Multi-Scale Framework for Built-Up Area Detection

Zhongwen Hu <sup>1,2</sup>, Qingquan Li <sup>2,\*</sup>, Qian Zhang <sup>3</sup> and Guofeng Wu <sup>2</sup>

<sup>1</sup> College of Information Engineering, Shenzhen University, Shenzhen 518060, China; zwhoo@szu.edu.cn

<sup>2</sup> Key Laboratory for Geo-Environmental Monitoring of Coastal Zone of the National Administration of Surveying, Mapping and Geoinformation & Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China; guofeng.wu@szu.edu.cn

<sup>3</sup> National Key Laboratory of Science and Technology on Multi-spectral Information Processing, School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China; hangfanzq@163.com

\* Correspondence: liqq@szu.edu.cn; Tel.: +86-0755-2653-6101

Academic Editors: Janet Nichol, Guoqing Zhou, Norman Kerle and Prasad S. Thenkabail

Received: 24 August 2015; Accepted: 2 February 2016; Published: 18 February 2016

**Abstract:** The accurate extraction and mapping of built-up areas play an important role in many social, economic, and environmental studies. In this paper, we propose a novel approach for built-up area detection from high spatial resolution remote sensing images, using a block-based multi-scale feature representation framework. First, an image is divided into small blocks, in which the spectral, textural, and structural features are extracted and represented using a multi-scale framework; a set of refined Harris corner points is then used to select blocks as training samples; finally, a built-up index image is obtained by minimizing the normalized spectral, textural, and structural distances to the training samples, and a built-up area map is obtained by thresholding the index image. Experiments confirm that the proposed approach is effective for high-resolution optical and synthetic aperture radar images, with different scenes and different spatial resolutions.

**Keywords:** block-based feature extraction; multi-scale feature representation; scale-space theory; built-up area detection; Harris corner point; high spatial resolution image

---

## 1. Introduction

Mapping landscapes is an important task for remote sensing applications. In recent years, mapping the distribution, growth, and characteristics of built-up areas has attracted increasing attention because it can provide important information for many applications. In recent decades, urbanization has accelerated globally and has caused a series of social and environmental problems, especially in developing countries. The accurate extraction and mapping of built-up areas plays an important role in many social, economic, and environmental studies, such as urban and transport planning [1,2], environmental protection [3], and assessment, rescue, and rebuilding efforts in disaster zones [4].

Much effort has been devoted to detecting and mapping built-up areas using remote sensing images. Due to their broad coverage, coarse resolution images such as the Moderate Resolution Imaging Spectroradiometer (MODIS) and Medium Resolution Imaging Spectrometer (MERIS) have been widely used for the global and regional mapping of built-up areas in global ecological, environmental, and climatological research [5]. With the development of imaging techniques, satellite images of a medium spatial resolution, such as the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+), are now often used for this task [6,7], as are China's Beijing-1 [8]

satellite data. These images have medium spatial resolutions and moderate coverage, which makes them suitable for regional built-up area mapping and related applications. The spatial resolution of remote sensing images has rapidly increased, and many satellites can now offer images with a spatial resolution range of 0.5–5.0 m and even better, such as Geoeye-1 (0.41 m) and WorldView-4 (0.3 m). These images can be employed to accurately map built-up areas at a much finer scale. For example, Pesaresi *et al.* [9] reported promising first results for accurate global human settlement mapping using SPOT, IKONOS, QuickBird, CBERS-2B, RapidEye, and GeoEye images. These maps can provide us with useful information for urban and transportation planning, as well as other social and ecological applications. The choice of high-, medium-, or coarse-resolution imagery is often application-dependent and is subject to several constraints and trade-offs inherent to all remote sensing data [5].

The methods for detecting built-up areas from images with different spatial resolutions are various. The urban landscape is a complex one, consisting of different land cover types, such as trees, lawns, and man-made objects. In medium- and coarse-resolution remotely sensed images, many different land cover types might be mixed in one pixel [10]; thus, it is difficult to model the textural and structural features. However, utilizing the specific spectral features of built-up areas, a set of normalized difference built-up indexes (NDBIs) has been designed for built-up area mapping [11–13]. With the increase in the spatial resolution, the textural and structural patterns of built-up areas become clearer, enabling them to be used for built-up area mapping [14–18]. However, the detailed patterns of built-up areas may be significantly different, according to the different types of roofs, trees, and shadows. Thus, it is more challenging to accurately detect and map built-up areas using high spatial resolution images than with medium- and coarse-resolution images.

Many methods have been applied to achieve this aim, and they can be categorized into three types: traditional classification-based methods, geometric feature based methods, and texture-based methods.

The traditional classification-based methods mainly focus on accurately modeling spectral, textural, and spatial patterns, and the built-up areas are obtained through a classification process. It is, however, very difficult to model complex built-up area patterns in high-resolution images using a single pixel, and thus segment- and block-based approaches are more frequently used for this purpose. It is reasonable to model urban structures using a segment or a region when the image is properly segmented. Many segment- or object-based classification approaches can be found in the literature [19–24]. However, the traditional segmentation algorithms are usually designed to segment an image into homogenous regions, in which the structure of the built-up area patterns may be missed. Thus, complex models, such as Markov random fields (MRFs) [24], are commonly employed to model the patterns, considering a region and its neighborhood. The block-based approaches are more effective at modeling complex textural and structural patterns than the object-based approaches [9]. However, the blocks should be no less than 50 m to contain enough information of the built-up area [14,17], which might result in very coarse boundaries for the detected areas.

Built-up areas are compound structures containing many man-made objects that have salient geometric features, such as dense corner points. Especially in high spatial resolution images, the corners of houses are distinct characteristics of built-up areas. Thus, many methods focus on the utilization of geometric features for built-up area detection. Local feature points, such as scale-invariant feature transform (SIFT) [25] and Harris corner points [26], are frequently used. In Sirmaçek's works [15,27], SIFT feature points were used as low-level clues for built-up areas, and in some other works [16,28–30], Harris corner points were used. The distribution density of local feature points has been found to be closely related to the presence of built-up areas [16]. In Shi's work [18], local feature points and edge structures were employed for accurate built-up area detection. These local structures are good clues for built-up area detection. However, these types of feature points and lines can be significantly affected by the textures of the trees and other objects in very high spatial resolution images.

The texture patterns of built-up areas are very different from those of other objects in high spatial resolution images. Many texture-based methods have been proposed to detect built-up areas using

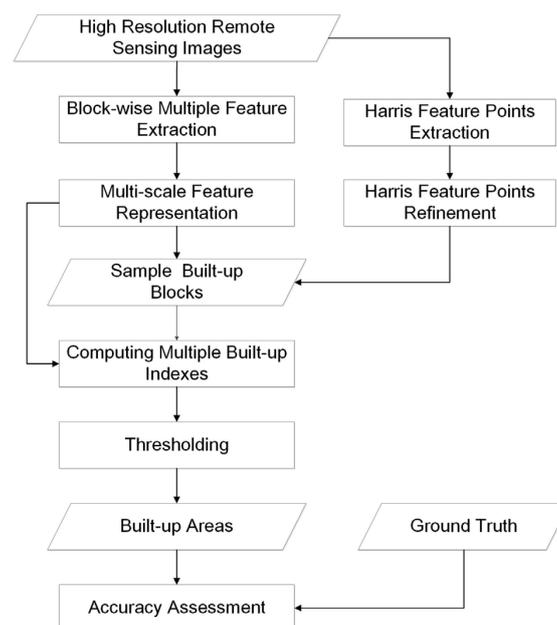
optical and synthetic aperture radar (SAR) images [14,31–33]. One of the most famous methods is the “Pantex” procedure [14], which proposes a texture-based built-up area presence index. It employs gray-level co-occurrence matrix (GLCM) features to extract a rotation-invariant index describing the presence of built-up areas. This method has been successfully used for global human settlement mapping [9]. A built-up area saliency index (BASI) has also been proposed [33], which employs the non-sampled contourlet transform (NSCT) [34] to describe the textures and measure the saliency of built-up areas. These texture-based indexes are effective over areas with regular textures. However, buildings are often sparsely distributed and mixed with trees in very high spatial resolution images and do not show regular textural patterns, and thus these types of methods may fail to detect built-up areas.

An important issue that should be addressed is that the micro-patterns of built-up areas under different resolutions may be different. This can significantly affect the performance of these methods. Moreover, the microstructures obtained at fine scales in very high spatial resolution images are not always effective in describing built-up areas. A possible solution to overcome this limitation is multi-scale analysis of the image, such as the discrete field of image descriptors (DFID) [9] and multi-scale image segmentation and feature representation [35].

In this paper, we present a novel approach for detecting built-up areas from high-resolution optical and SAR images as follows: (1) A block-based method is used, in which the spectral, textural, and structural features are represented in a multi-scale framework; (2) A set of refined Harris corner points are used as training samples; (3) A new built-up index is then obtained by computing, normalizing, and minimizing the feature distances to the training samples. The basic concept of multi-scale feature representation is that compound structures, which are difficult to model at a micro scale, show more stable statistical features at a macro scale [9].

## 2. Proposed Method

The proposed method is composed of four main steps (Figure 1): (1) Harris corner points are detected and refined; (2) the image is subdivided into small blocks, and their spectral, textural, and structural features are extracted and represented in a multi-scale framework; (3) spectral, textural, and structural built-up indexes are obtained using a supervised approach, and a built-up index image is obtained by normalizing and minimizing the normalized spectral, textural, and structural distances; and (4) a built-up area map is obtained by threshold segmentation of the index image.



**Figure 1.** The flowchart of the proposed method.

## 2.1. Feature Extraction and Multi-Scale Representation

### 2.1.1. Block-Based Feature Extraction

Compared with the traditional pixel- and object-based approaches, the block-based approaches have inherent advantages in describing built-up areas, including: (1) it is easier to model compound urban structures using a block rather than a single pixel; (2) the object-based approaches largely depend on image segmentation, which is often not effectively solved due to the complexity of ground objects [10]; and (3) it is more accurate to model built-up structures by aggregating several image object/region instances in one block, and it allows computationally efficient multi-scale pattern analysis of the image. Therefore, many studies have applied a block cell as the basic unit in built-up area detection. The following descriptors are derived to describe the blocks accurately and comprehensively:

- *Spectral Histogram*: Built-up areas cannot be accurately modeled using only spectral means due to their complex mixtures of buildings, trees, grass, and shadows, and thus the color histogram is used. The color histogram records the distribution of the spectral signatures and is one of the most commonly used descriptors. The number of bins depends on the quantization of the color space. In our work, the color space is quantized into 32 bins, and thus the histogram has 32 bins for each image channel.
- *RILBP/LC Histogram*: The textural features of a block can be modeled from two aspects: texture pattern and intensity. In this paper, the texture pattern and intensity are modeled using rotation-invariant local binary patterns (RILBP) and local contrast (LC) [36], respectively. RILBP/LC describes the textural features using the joint distribution histogram of the RILBP and LC. There are 10 patterns for the RILBP, and we quantize the LC to eight bins. Thus, a joint distribution histogram of  $10 \times 8$  bins is employed to describe the textural features.
- *HOG Histogram*: The structural features of built-up areas are one of their essential differences from other areas. In this paper, the histogram of oriented gradients (HOG) [37] is used to describe the structural features of the blocks. The HOG measures the structural features using the statistics of local orientation and intensity obtained from the image gradient. It is recorded using an orientation histogram, which is weighted by the intensity. Differing from the approach presented in [37], the HOG descriptor is obtained by computing the orientation histogram in a single block. The number of orientations is quantified to 12 bins, and thus the feature vector has 12 bins for each block.
- *Corner Response*: The value of the corner function describes the response of an image structure to a corner, which is a specific structure of man-made objects. There are many corners in built-up areas, such as the corners of buildings and road intersections. Thus, the response to the corner function should be high over these buildings and intersections. In this paper, the Harris corner function [26] is employed, and its response is anisotropic. The responses are computed pixel by pixel, and then the maximum in each block is used to describe the block.

### 2.1.2. Multi-Scale Feature Representation

Weizman [17] noted that a block should be no less than  $11 \times 11$  pixels to capture the basic textural and structural patterns of built-up areas. Zhong [38] used a block with  $16 \times 16$  pixels. Pesaresi [14] considered that the minimum detectable settlement structure was composed of at least two buildings, certain open spaces, and roads around the buildings, with an estimated minimum footprint of approximately 50 m. However, this is too coarse for accurate built-up area mapping using very high spatial resolution images. Accurate boundaries can be obtained using smaller blocks. However, if an image is segmented into too small blocks, it might not be robust enough to capture the micro-patterns of the urban structure.

To overcome the abovementioned limitations, and motivated by scale-space theory [39], we propose a block-based multi-scale framework for feature extraction and representation (Figure 2).

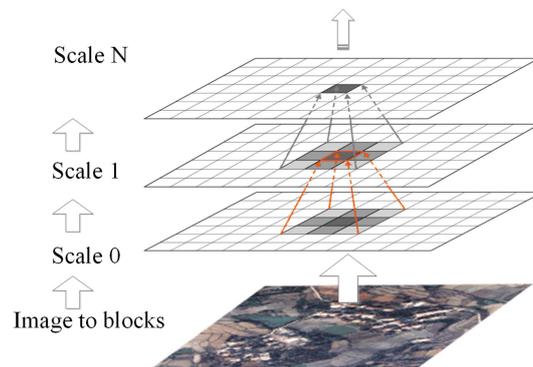
Scale-space theory is similar to the DFID [9], but its implementation is different. The implementation includes two main steps: (1) an image is divided into small blocks (e.g.,  $6 \times 6$  pixels), in which the spectral, textural, and structural features are extracted; and (2) the feature vectors are represented in the scale-space framework to capture the macro patterns. It is expected that the boundaries will be smoother using the smaller blocks and that the multi-scale representation can make the features more stable.

With the scale-space feature representation, the feature vectors of the blocks can be convolved with a Gaussian kernel (Equation (1)), which is the only possible scale-space kernel [25]:

$$G(dx, dy, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{dx^2 + dy^2}{2\sigma^2}}, \quad (1)$$

where  $dx$  and  $dy$  denote the horizontal and vertical distances to the center of the filtering window, and  $\sigma$  is the standard deviation of the Gaussian kernel. The standard deviation  $\sigma = 1.6$  and the corresponding radius of the filter window  $r = 5$  are used. In our experiments, they provide optimal performance in most of the cases. They might not be the optimal ones in some cases, but they provide close to optimal performance.

Each dimension of the features is convolved with those of its neighbors using the Gaussian kernel. In other words, the features of a block are Gaussian weighted by its surrounding blocks. Thus, the features of a block describe not only the block itself but also the surrounding blocks. For the multi-scale representation purpose, the linear scale-space convolution is executed iteratively, and the number of iterations is defined as the scale parameter. It is expected that by adopting the multi-scale representation, a small block can capture the macro patterns of a built-up area.



**Figure 2.** Block-based feature extraction and multi-scale representation.

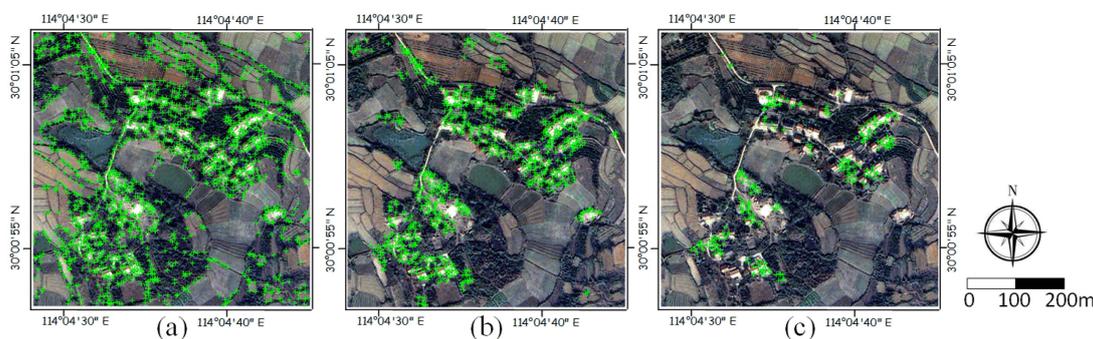
## 2.2. Multiple Built-Up Indexes (MBIs)

Some blocks can be labeled as built-up areas by estimating the density of the corner points from the refined corner points in built-up areas [27]. However, many urban areas might be missed when the density is very low. Thus, a novel supervised built-up index is proposed in this paper to solve this limitation. First, we utilized a set of refined Harris corner points to automatically select a set of training samples; second, four built-up indexes were obtained by calculating and normalizing the average distance to these training samples.

### 2.2.1. Automatic Corner Point Detection

Built-up areas are mixtures of buildings, roads, grasslands, and trees, in which the buildings and roads are typical man-made objects. The salient features of man-made objects are geometrical features such as straight lines and corners. Many studies have proven that the density of the corner points is a good clue for urban area extraction from high-resolution optical images [15,16,30]. However, many corner points may also exist over cropland and trees in very high spatial resolution images (Figure 3a),

which can be refined by eliminating some of the points by estimating the distribution density [16] (Figure 3b,c).



**Figure 3.** Harris corner point detection and refinement. (a) Original Harris corner points; (b) Refined results ( $R = 25$ ,  $N = 10$ ); (c) Refined results ( $R = 25$ ,  $N = 15$ ).

Harris corner points are applied in this paper. The distribution density of the corner points is estimated by placing a circle on a corner point, and the number of corner points in the circle is used as the density. Two parameters control the refinement of the corner points: the radius of the circle ( $R$ ) and the threshold ( $N$ ). If the number of corner points is less than  $N$ , the corresponding point should be eliminated. Finally, a set of refined points can be obtained (Figure 3b,c). As shown in the figures, although there are many corner points over the woodland areas, most of the distribution densities are less than the threshold, and thus most of these corner points are eliminated. The retained corner points can be used as reliable indicators of built-up areas.

### 2.2.2. Calculation of the Multiple Built-Up Indexes

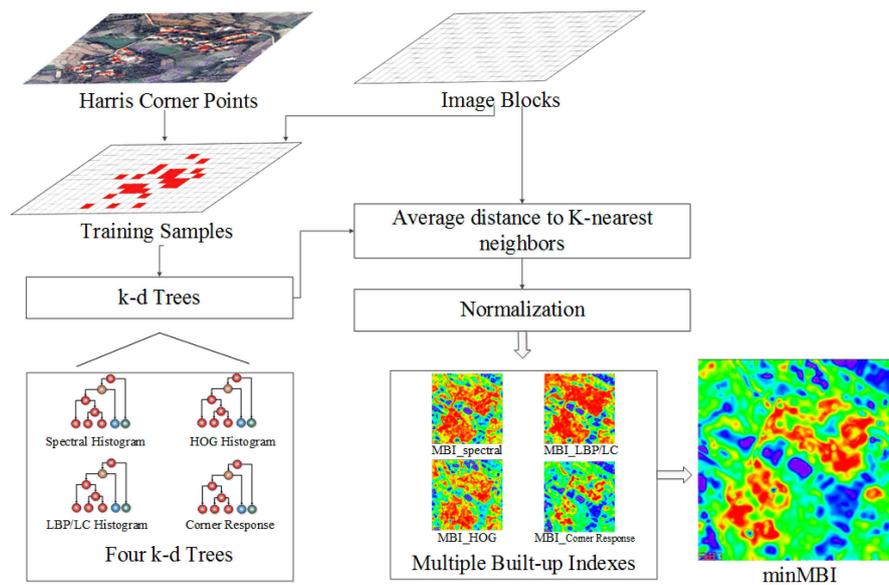
For a given block  $x$  and a set of training blocks, the distance from  $x$  to the nearest sample reflects the probability of being a built-up area. However, the samples extracted using the Harris corner points are not always completely convincing because they can be affected by the image texture and noise (e.g., the speckle noise of a SAR image). To alleviate such a negative impact, the average distance to the  $K$ -nearest samples is used. This process can be efficiently implemented using the  $k$ -dimensional tree ( $k$ -d tree) structure, which is a space-partitioning data structure for organizing points in a  $k$ -dimensional space and for multidimensional search (e.g., nearest neighbor search). An illustration is provided in Figure 4.

There are five steps in this method: (1) a set of blocks is selected as training samples, covering at least one refined Harris corner point; (2) all the samples are organized using four  $k$ -d trees, considering the spectral, textural, and structural features, and the corner response; (3) for each block of the image, its  $K$ -nearest training samples are obtained by the nearest neighbor search in the  $k$ -d trees, and the average distance to these samples is obtained; (4) the MBIs are obtained by normalizing the spectral, textural, and structural distances; and (5) a built-up index is obtained by calculating the minimum of the four indexes. The distance between two blocks is calculated using the Euclidean distance of the two feature vectors.

Because the blocks are described using four descriptors, four types of distances are obtained and used for generating the final built-up index. Thus, the four different built-up indexes are derived by normalizing the distances:

$$MBIs(i, j) = \frac{d_{\max} - d(i, j)}{d_{\max} - d_{\min}}, \quad (2)$$

where  $MBIs$  represent the multiple built-up indexes;  $(i, j)$  denotes the row and column of a block;  $d$  is the feature distance; and  $d_{\max}$  and  $d_{\min}$  are the maximum and minimum of distances, respectively. The MBIs range from 0 to 1, with a high value indicating a high probability of being a built-up area.

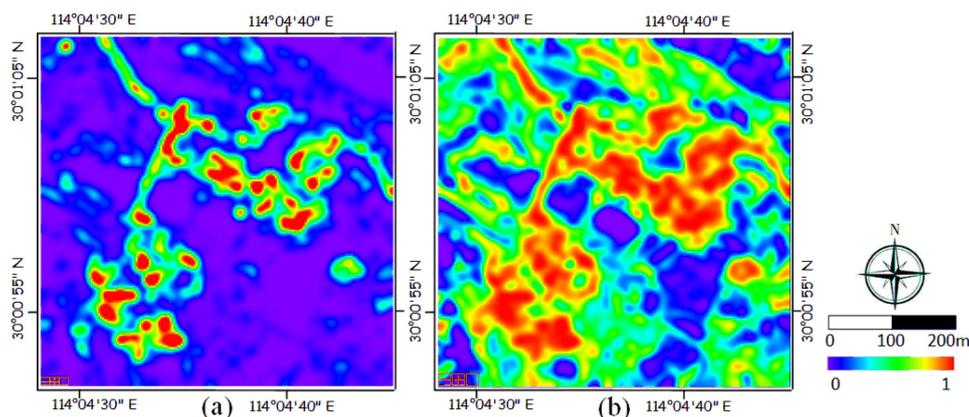


**Figure 4.** The flowchart for calculating the multiple built-up indexes.

It is worth noting that the intensity of the pixel-wise corner response is very high near corner points and is otherwise very low. Thus, even in urban areas, there will only be a small number of pixels with a very high response, as shown in Figure 5a. To better indicate the presence of built-up areas, a non-linear stretch of the distance of the corner response is applied before normalization (Equation (3)):

$$d_{CR'} = (d_{CR})^\beta, \tag{3}$$

where  $d_{CR}$  and  $d_{CR'}$  are the original and stretched distances, respectively; and  $\beta$  is a parameter ranging from 0 to 1. It is expected that the very high response to corner points can be suppressed by the non-linear stretch, and thus the intensity over built-up areas can be balanced. An example of a built-up index image obtained using the stretched distances is shown in Figure 5b, in which the indexes are smoothed and balanced over built-up areas.



**Figure 5.** The built-up index image obtained by normalizing the distance image of the corner responses. (a) The index image obtained using the original distances of the corner responses; (b) The index image obtained using the stretched distances of the corner responses ( $\beta = 0.1$ ).

### 2.2.3. The Minimum of the MBIs

It is assumed that the four MBIs should be significantly high when the corresponding block belongs to an urban area. Thus, we derive a single built-up index by computing the minimum of the four MBIs:

$$\min MBI(i, j) = \min \{ MBI_{s_t}(i, j), t = 1, 2, 3, 4 \}, \quad (4)$$

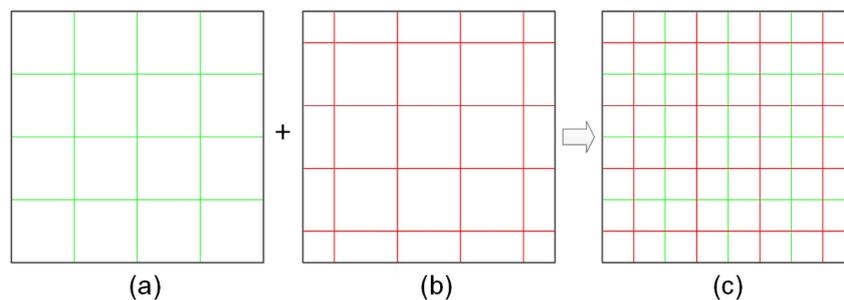
where  $MBI_{s_t}$  ( $t = 1, 2, 3, 4$ ) is the  $t$ -th MBI and  $(i, j)$  denotes the row and column of the block, respectively.

Finally, the boundaries of the built-up areas are derived by thresholding the minMBI image manually or automatically. Many automatic methods can be used to obtain the threshold, such as Otsu's method [40] and the Kittler and Illingworth minimum-error thresholding algorithm (K&I) [41]. To alleviate the effect of the different threshold methods, we used a trial and error approach. To overcome the serrated boundaries, the output vector result was further processed using a vector smoothing algorithm.

### 2.3. Block Offset and Data Fusion

One drawback of the proposed block-based approach is the serrated boundaries, especially when the block size is very large. However, this drawback can be overcome by dividing the image with an offset and fusing the indexes obtained with/without the offset. Two index images can be obtained by dividing the image into blocks with/without an offset ( $w/2$ ), and then the final index image is averaged by these two index images. An illustration of this approach is provided in Figure 6.

There are two main advantages to this approach: (1) by fusing the built-up indexes obtained with/without offset, the block size can be reduced to half the original size (Figure 6), and thus the spatial accuracy of the boundary can be significantly improved; and (2) the computational load can be reduced by 50% compared to dividing the image into smaller blocks ( $w/2$ ).



**Figure 6.** An illustration of block offset and data fusion. (a) Offset = 0; (b) Offset =  $w/2$ ; (c) Fused blocks.

### 2.4. Parameter Settings

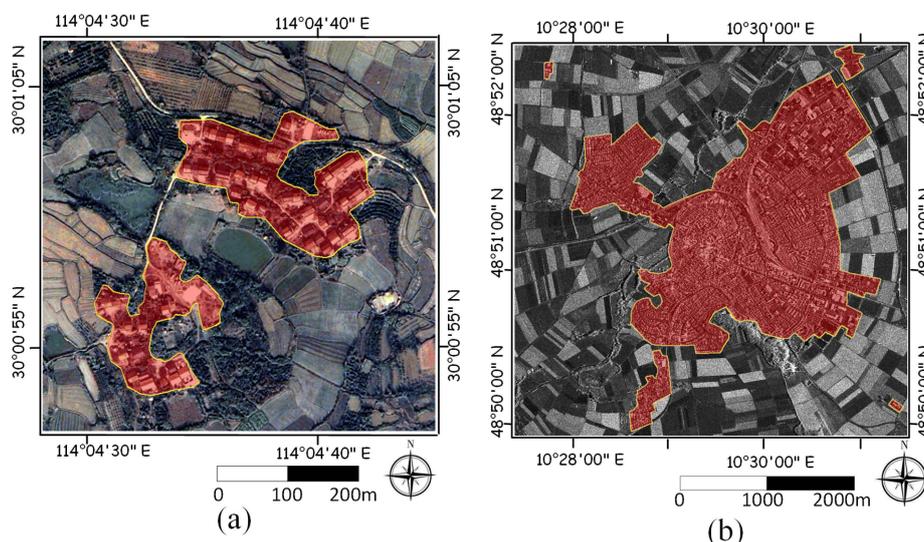
Several parameters affect the performance of the proposed method: the parameters  $R$  and  $N$  in the Harris corner point refinement, the size of the blocks, the scale parameter, and the parameter  $K$  for the nearest neighbor searching.

Generally,  $R = 25$ ,  $N = 15$ , and  $K = 10$  can be used for most images. The most important parameters are the size of the blocks and the scale parameter. In the Pantex procedure, the size of the texture window is estimated using the spatial resolution of the input image. Motivated by this idea, we automatically estimate the size of the block and the scale parameter. As presented by Pesaresi [14], a detectable-target minimum settlement structure is approximately 50 m on the ground. However, as the blocks are represented in the multi-scale framework, the block size can be much smaller than 50 m. Thus, in this paper, we simply estimate the parameters using the criterion  $s \times w \times r = 50$ , where  $s$ ,  $w$ , and  $r$  are the scale parameter, the block size, and the spatial resolution, respectively. In this paper,  $s = 3$  is used for most of the optical images with a 1–3 m/pixel spatial resolution. Thus,  $s = 3$  and  $w = 24$  (or alternatively,  $s = 5$  and  $w = 14$ ) can be used for QuickBird images with a 0.6 m spatial resolution, and  $s = 3$  and  $w = 8$  can be used for ZY-3 images with a 2.1 m spatial resolution. However, the parameters can be tuned according to the patterns of the built-up areas. To make the spectral histograms non-sparse, the minimum block size is  $6 \times 6$  pixels.

### 3. Experiments

#### 3.1. Experiments and Dataset Description

A number of different high-resolution images were used to evaluate the performance of the proposed method. First, a multi-spectral QuickBird image ( $750 \times 750$  pixels, 0.6 m/pixel) covering a suburban area near the city of Wuhan, Hubei province, China, was used to evaluate the performance on an optical image. The test image was obtained by fusing the 0.6 m/pixel panchromatic image and the 2.4 m/pixel multi-spectral image. Second, a TerraSAR image ( $1506 \times 1506$  pixels, 3 m/pixel) covering Nördlinger Ries, Germany, was used to demonstrate the effectiveness on a radar image. Quantitative and qualitative comparisons with two state-of-the-art methods were also conducted using these images. The test images and reference built-up areas are shown in Figure 7. Third, a panchromatic ZY-3 image ( $5000 \times 5000$  pixels, 2.1 m/pixel) from Hebei province, China, was used to demonstrate the performance over a large area, and the visual comparisons with the Pantex and BASI procedures were carried out. Finally, a set of high spatial resolution images was used to show the effectiveness on images with different scenes and different spatial resolutions.



**Figure 7.** The experimental data set and the corresponding reference built-up areas. (a) Multi-spectral QuickBird image with 0.6 m/pixel; (b) TerraSAR image with 3 m/pixel.

#### 3.2. Evaluation Metrics

Three metrics are employed for quantitatively evaluating the proposed method: the precision ( $P$ ), the recall ( $R$ ), and the F-measure ( $F$ ). The  $P$  and  $R$  metrics evaluate the correctness and completeness, respectively. They are defined as:

$$P = \frac{TP}{TP + FP} \quad (5)$$

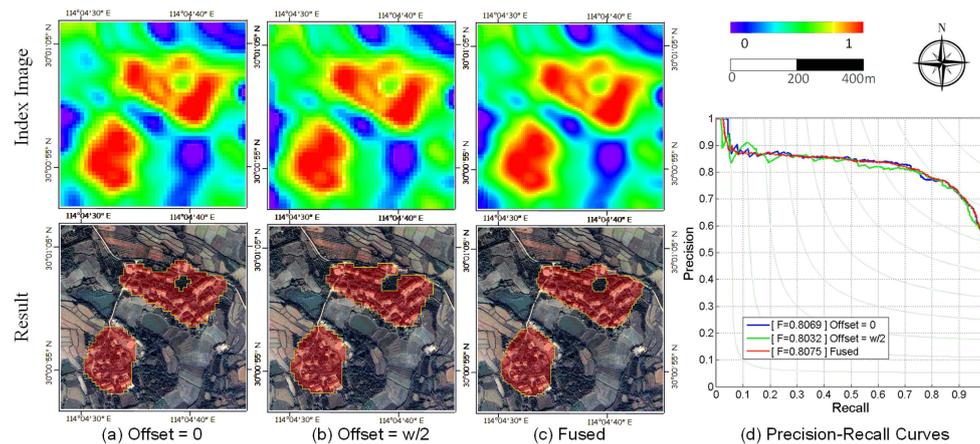
$$R = \frac{TP}{TP + FN} \quad (6)$$

$$F = 2 \frac{P \cdot R}{P + R} \quad (7)$$

where  $TP$ ,  $FP$ , and  $FN$  are the numbers of true positive, false positive, and false negative detected pixels, respectively.

### 3.3. Effectiveness of the Block Offset and Data Fusion Procedure

First, the multi-spectral QuickBird image was used to evaluate the performance of the block offset and data fusion procedures. The parameters used in this experiment were  $R = 25$ ,  $N = 15$ ,  $K = 10$ ,  $w = 16$ , and  $s = 5$ . The image was first divided into blocks of  $16 \times 16$  pixels without offset, and a built-up index image was obtained (Figure 8a). The image was then divided into blocks with an offset of eight pixels, and a new built-up index image was obtained (Figure 8b). Finally, an index image (Figure 8c) was obtained by averaging and normalizing these two index images. The final results were obtained using the same threshold.



**Figure 8.** The built-up indexes and final results obtained with (a) offset = 0; (b) offset =  $w/2$ ; (c) the fused result; and (d) the precision-recall curves. The boundaries of the built-up areas are not smoothed.

Because the size of the blocks is very large, the mosaic effect is obvious in the index images. Thus, the final boundaries are very rough and serrated (Figure 8a,b), and cannot easily be smoothed. However, the mosaic effect is greatly suppressed in the fused index image. Furthermore, the boundaries of the final results are smoother and can be better processed at a later stage. The reason for this is that the block size can be reduced to  $w/2$  by the fusion procedure. As a result, the fused index image provides a smoother precision-recall curve and better F-measure values (Figure 8d).

### 3.4. Effectiveness of the Multi-Scale Feature Representation

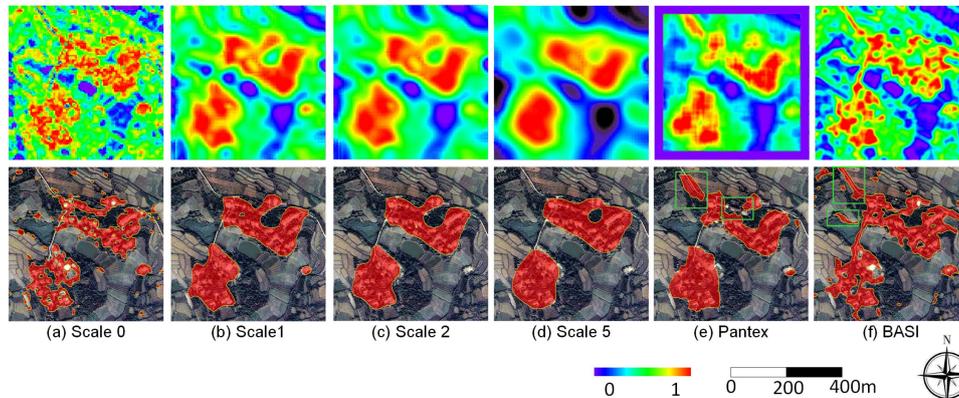
#### 3.4.1. Qualitative Evaluation

The impact of the scale parameter of the proposed method was evaluated using the multi-spectral QuickBird image. The parameters for the Harris corner point refinement were  $R = 25$  and  $N = 15$ . The test image was first divided into blocks of  $16 \times 16$  pixels, corresponding to approximately 9 m on the ground. The set of built-up indexes was obtained using  $s = 0$  to 5. Finally, the built-up areas were extracted by thresholding the indexes manually, and then the boundaries were smoothed through post-processing. To better illustrate the different performances, the results with the best F-measure values are presented. The index images and final results are shown in Figure 9a–d.

The multi-scale feature representation plays an important role in the proposed approach and can significantly improve the performance (Figure 9a,b). The index image seems to be polluted by noise; there are many “holes” over the settlements, and many other objects are falsely detected (Figure 9a). Such a result is caused by the fact that a block of  $16 \times 16$  pixels cannot describe the macro patterns of a built-up area at the scale of 0 because the blocks covering different parts of the built-up areas may have very different features.

Through representing the blocks in the multi-scale framework, the features become more stable (Figure 9b–d), and the indexes over built-up areas are significantly higher than those of other objects. Most built-up areas are correctly detected, and with the increase in the scale parameter from 1 to 5,

the indexes are gradually smoothed and the “noise” is eliminated. Such a result can be explained by the fact that with the increase in the scale parameter, a block covers a larger area on the ground. Thus, the block describes the macro patterns and is not sensitive to the local “noise.” Moreover, because the index image is smoothed, the boundaries are also smoothed.



**Figure 9.** The built-up indexes and the corresponding final results obtained at different scales with the Pantex and BASI procedures. The final results are those with the best F-measure values. (a–d) The built-up indexes and final results of the proposed method with  $s = 0, 1, 2$  and  $5$  respectively; (e,f) The built-up indexes and final results of the Pantex and BASI.

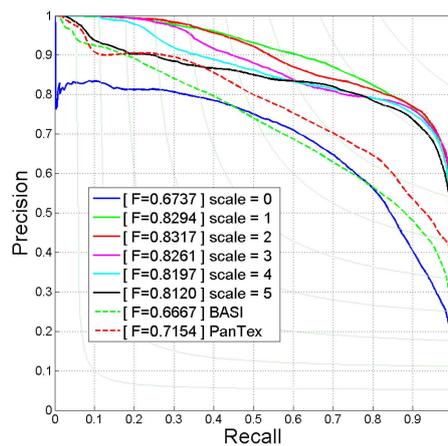
Using the QuickBird image, the proposed method was compared with two similar techniques: Pantex [14] and BASI [33]. The window size for the Pantex procedure was  $83 \times 83$  pixels (corresponding to approximately 50 m on the ground). The BASI index was computed in eight directions and three scales, and the size of the window used to compute the texture energy was 83.

The indexes and the final results of Pantex and BASI are presented in Figure 9e,f. The proposed method derives similar indexes to the Pantex procedure. However, the proposed method provides more reliable results over the area marked in Figure 9e. Moreover, the computational load of the Pantex procedure is much higher than the proposed method under the parameter setting  $w = 83$ . One way to reduce the computational load is to downsample the image to a coarser resolution. However, the small residential areas, as shown in Figure 9, may be undetectable in an image of a coarser spatial resolution. The BASI index does not perform as well as the proposed method and Pantex, and some cropland areas, especially those near roads, are falsely detected. Moreover, there are many “holes” over settlements.

### 3.4.2. Quantitative Evaluation

The selection of the threshold can significantly affect the final result. Thus, to evaluate the impact of different thresholds, a set of thresholds ranging from 0 to 1 was applied, and a set of corresponding precisions, recalls, and F-measure values was obtained, which can be plotted as a  $P$ - $R$  curve.

The  $P$ - $R$  curves of all the methods are presented in Figure 10. The  $P$ - $R$  curves of  $s = 1$  to  $5$  are much better than that of  $s = 0$ . Such a result can be explained by the fact that a block of  $16 \times 16$  pixels cannot describe the macro patterns of a built-up area at the scale of 0; thus, the “holes” and “spots” (Figure 9a) significantly affect the quantitative metrics. However, the multi-scale feature representation framework makes a block at a larger scale to capture the macro patterns of the built-up areas by aggregating the features of the neighbors. There is also another tendency that the maximum F-measure value increases from  $s = 0$  to  $2$ , but decreases when  $s > 2$ . Such a result can be explained by the fact that with the increase in the scale parameter ( $s = 0, 1, 2$ ), the multi-scale representation makes the block more stable to describe the built-up structures. As a result, the quantitative metrics increase. However, the boundaries of built-up areas might be over-smoothed (Figure 9d) at higher scales ( $s > 2$ ) for the test image, which results in a decrease in the quantitative metrics.

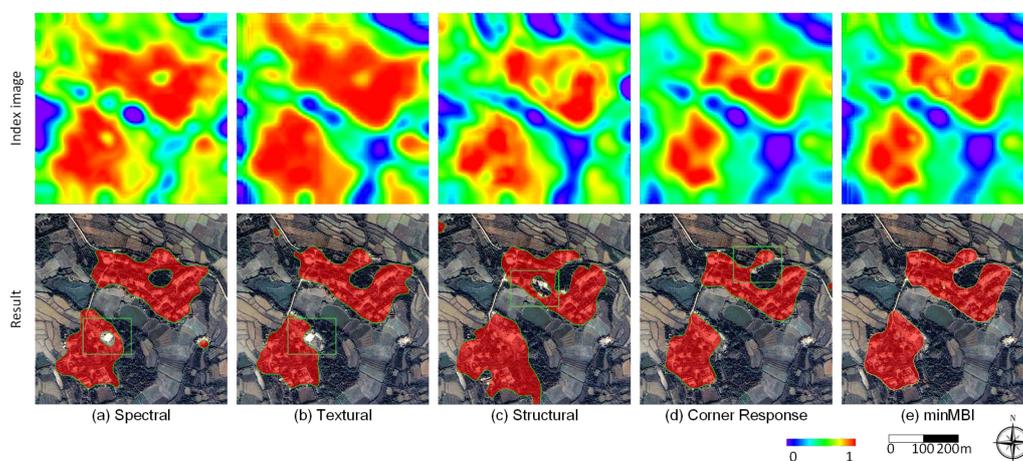


**Figure 10.** Quantitative evaluation of the performances at different scales.

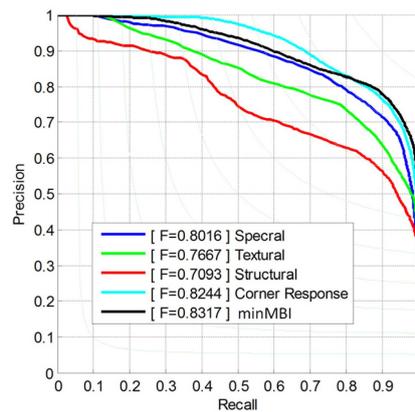
Figure 10 shows that the proposed method achieves much better quantitative scores ( $F > 0.8$ ) than Pantex ( $F = 0.7154$ ) and BASI ( $F = 0.6667$ ) on the test site. Such a result can be explained by the fact that both Pantex and BASI utilize only textural features, which can be affected by the image details in high-resolution images. Moreover, in the 0.6 m/pixel test image, it is very difficult to model the textural patterns of the built-up areas at the original resolution using the GLCM and NSCT descriptors. Thus, the Pantex and BASI indexes are unstable in the test image. Finally, some other objects are falsely detected, and some built-up areas are missed (Figure 9e,f). As a result, the best F-measure values are much lower than those of the proposed method. A possible solution is downsampling the image to a moderate resolution, where the textural patterns can be better described. However, as mentioned before, some small settlements may be undetectable in a coarser-resolution image, and the boundaries might be coarse.

### 3.5. Effectiveness of the Multiple Features

The multi-spectral QuickBird image was used to evaluate the performance of different image features. The parameters used in this experiment were  $R = 25$ ,  $N = 15$ ,  $K = 10$ ,  $w = 16$ , and  $s = 2$ . Four index images were obtained using spectral, textural, structural, and corner response features alone (Figure 11a–d), and the minMBI image was obtained using the method described in Equation (4), respectively (Figure 11e). The corresponding quantitative evaluation metrics are shown in Figure 12.



**Figure 11.** Built-up indexes obtained using different image features and the corresponding results. (a–d) The index image and the final results using spectral, textural, structural and corner response, respectively; (e) The proposed minMBI index image and the final result.



**Figure 12.** Quantitative evaluation of the performances using different image features.

As shown in the first row of Figure 11, all these index images can be used to indicate the presence of built-up area. However, some differences exist. Some trees around the built-up areas have very high values using only spectral or textural features alone (Figure 11a,b). The reason is that a large number of trees over the built-up areas share similar textural and spectral patterns with trees outside built-up areas; however, their structural features and corner response are very different. In the structural index image, some built-up areas have low values compared with other built-up areas. The index image of corner response is very similar to that of the proposed minMBI, which indicates that the corner response contributes a lot to the final minMBI image.

The second row of Figure 11 presents the corresponding final results of best F-measures. Although most of the built-up areas are detectable from a single feature, there exist some limitations, as marked by the rectangles in Figure 11. The result obtained using the proposed minMBI is much better than those obtained using only one feature (Figure 11e). Moreover, the quantitative precision-recall curves in Figure 12 show that a better F-measure is obtained by using the proposed minMBI.

The experiment has shown that a combination of multiple features is significantly better than using only one feature.

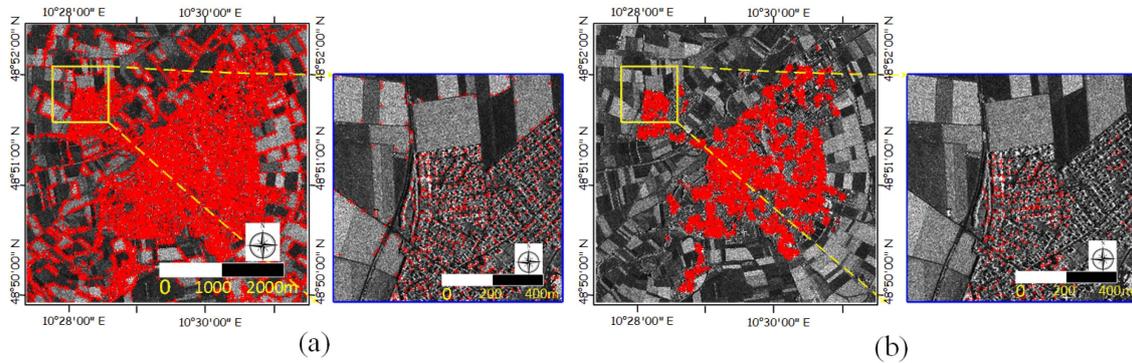
### 3.6. Effectiveness of SAR Images

The dense distribution of corner points is suitable not only for optical images but also for SAR images (Figure 13), although the imaging mechanisms are totally different. Such a result means that the local feature points are also good clues for detecting built-up areas from high-resolution SAR images.

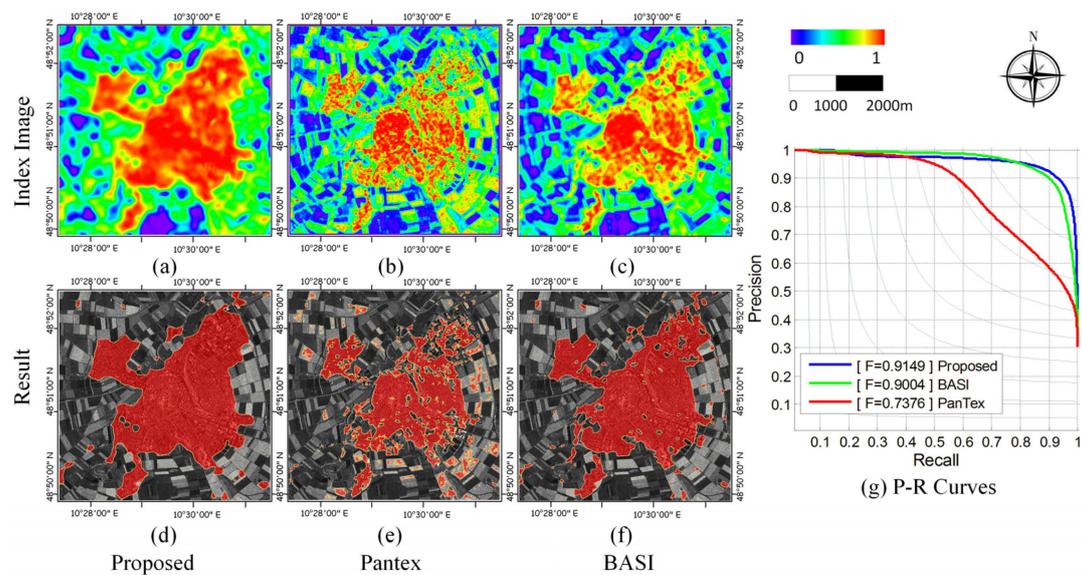
We evaluated the performance of the proposed method using a high-resolution SAR image. The block size of the proposed method was set as 6, and  $s = 5$ . The Pantex and BASI procedures were also used for comparison. The window size was set as 17 for Pantex. The BASI index was computed in eight directions and three scales, and the window size used to compute the texture energy was 17. The results using the TerraSAR image are shown in Figure 14, which demonstrates that the Pantex procedure works well over most of the built-up areas, but the indexes are not convincing over some cropland areas (Figure 14b,e), which could be caused by the serious speckle noise in the radar images. The Pantex procedure was greatly affected by the textural features, and many cropland areas are detected as built-up areas. The BASI procedure performs better than the Pantex procedure; however, it is also affected by the speckle noise. Moreover, there are also many “holes” over the built-up areas in Figure 14c,f. Such results can be explained by the fact that both the Pantex and BASI procedures are computed at a single scale and at the pixel level; thus, the index can be low, and holes can happen over flat areas (e.g., a square).

The proposed approach performs much better than Pantex and BASI from both the visual comparison (Figure 14d–f) and quantitative metrics (Figure 14g). The cropland and built-up areas can be easily distinguished. Such a result can be explained by the fact that (1) most of the features

are recorded using statistical histograms, which are more robust to regular speckle noise and (2) the block-based multi-scale feature representation makes the indexes robust to small flat areas over urban areas. Thus, the proposed method produces a better built-up area map (Figure 14a) and *P-R* curve (Figure 14g) than Pantex and BASI in this experiment.



**Figure 13.** Corner point detection and refinement using a high spatial resolution SAR image. (a) Original Harris corner points and the zoomed view in the rectangle area; (b) Refined Harris corner points and the zoomed view in the rectangle area.



**Figure 14.** Comparison with Pantex and BASI using a TerraSAR image. (a) Proposed method built-up index (b) Pantex built-up index; (c) BASI built-up index; (d–f) The final results using the proposed method, Pantex and BASI, respectively; (g) The *P-R* curves of Pantex, BASI, and the proposed method.

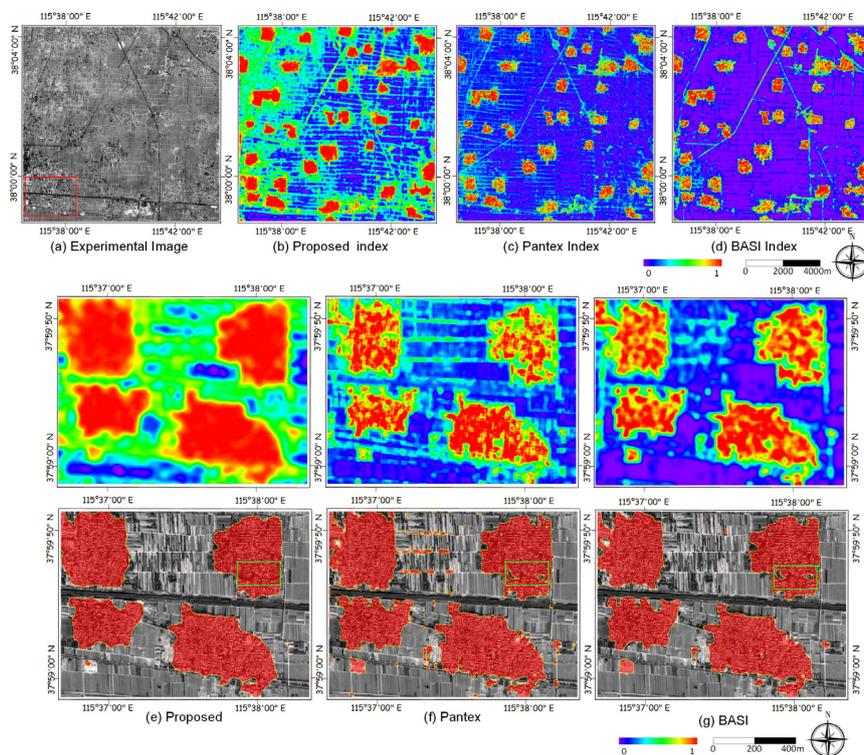
### 3.7. Effectiveness of ZY-3 Image

The ZY-3 image (2.1 m/pixel, 5000 × 5000 pixels) covers the center of the North China Plain. Most of the residential areas in this image are densely distributed. The parameters of the proposed method were set as  $s = 3$  and  $w = 8$ . The window size of Pantex was set as  $w = 23$ . The BASI index was computed in eight directions and three scales. The size of window used to compute the texture energy was 23. The test image and index images are presented in Figure 15a–d, and subsets of the details are presented in Figure 15e–g.

As shown in Figure 15a–d, the proposed index shows a different behavior than the Pantex and BASI procedures, especially over the areas on the left of the test image. This is mainly caused by the high contrast of the cropland areas.

Considering the detailed subsets in Figure 15e,f, the proposed index produces results that are much smoother over the settlement areas than Pantex and BASI. It also outperforms Pantex and BASI in handling “holes” over settlements, as marked in the green rectangles. This is mainly due to the multi-scale feature representation, which uses a block covering a larger area to capture the macro patterns of the built-up areas. One limitation of the proposed method that can be observed in the figures is that the boundaries are over-smoothed in some cases. This is caused by the over-smoothing of the details in the multi-scale feature representation.

It is also worth noting that the computational efficiency of the proposed method is much higher than that of Pantex and BASI. The proposed method takes only 215 s to process the ZY-3 image, while Pantex takes approximately 50 min and BASI takes more than 12 h. The reason for this is that the image is divided into small blocks, and thus the computational load is dramatically reduced. Moreover, the image descriptors are more efficient than the GLCM and NSCT texture descriptors.



**Figure 15.** Experiments on the ZY-3 image. (a) The test image; (b) The proposed method; (c) The Pantex index; (d) The BASI index; (e–g) The details of the proposed method, Pantex, and BASI in the rectangle of (a).

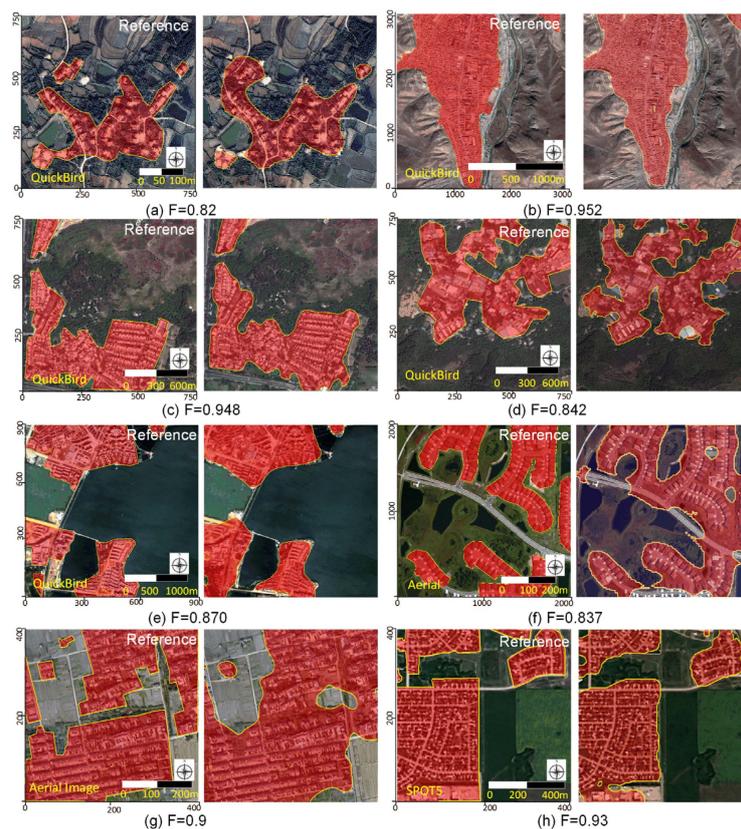
### 3.8. Effectiveness in Different Scenes and Different Spatial Resolutions

The objective of this paper is to develop a robust and effective method for the detection of built-up areas in different scenes. As there are many commercial remote sensing images with spatial resolutions ranging from 0.3 m/pixel to 10 m/pixel, the method also needs to be effective on images of different scenes and different spatial resolutions.

A set of high spatial resolution remote sensing images was used to show the applicability of the proposed method, including QuickBird (0.6 and 2.4 m/pixel), ZY-3 (2.1 m/pixel), SPOT-5 (2.5 m/pixel), and aerial images (0.32 and 1 m/pixel). The image scenes contain both mountainous and flat areas. The results are shown in Figure 16. Parameter settings of  $R = 25$ ,  $N = 10$ ,  $K = 10$ , and  $s = 3$  were used. The sizes of the blocks were estimated using the method proposed in Section 2.4. The results over the hilly and mountainous areas are presented in Figure 16a–d, and the results over the flat areas using different images are presented in Figure 16e–h. The F-measure values are reported in the captions.

As shown in Figure 16a,d, the proposed method performs well over hilly and mountainous areas ( $F = 0.82$  and  $0.842$ , respectively). The orchard and forest areas also do not greatly affect the results, despite having regular textural patterns. Although the houses show a scattered distribution in the orchards (Figure 14), and the textural patterns are not obviously different from those of other land-cover types, they can be easily distinguished by their structures and spectral signatures. Most of the houses are ultimately detected successfully. The image in Figure 16b covers a valley containing many ridge lines and a dry river. The built-up areas in this study area can be easily distinguished by their structures, although the spectral features are similar. The built-up area shown in Figure 16c is a typical crowded residential area, with regular texture and similar structures and spectral signatures. Thus, accurate boundaries of the residential areas can be obtained. The built-up area presented in Figure 16d is composed of small residential houses and large factory buildings, and the spatial and textural patterns are very different. However, both the residential houses and factory buildings have orthogonal corners. Thus, many Harris corner points can be detected, and both types of built-up area can be detected.

As shown in Figure 16e–h, the built-up areas have regular textures and structures. Moreover, they are clearly different from the surrounding scenes. Although the resolutions range from  $0.32$  m/pixel to  $2.5$  m/pixel, the proposed method successfully detects the built-up areas. These experiments demonstrate that the proposed method can be used for built-up area detection from different high spatial resolution images, including QuickBird, SPOT, ZY-3, RapidEye, and aerial images.



**Figure 16.** Built-up area detection using different images and different scenes with the proposed method. (a,b) Built-up area detection over hilly areas using QuickBird images with a resolution of  $0.6$  m/pixel;  $750 \times 750$  and  $2961 \times 3108$  pixels; (c,d) Built-up area detection over mountainous areas using QuickBird images with a resolution of  $2.4$  m/pixel,  $750 \times 750$  and  $750 \times 750$  pixels; (e–h) The detection results over flat areas using QuickBird ( $2.4$  m/pixel,  $900 \times 900$  pixels), aerial ( $0.32$  m/pixel,  $2001 \times 2001$  pixels, and  $1$  m/pixel,  $400 \times 400$  pixels), and SPOT-5 ( $2.5$  m/pixel,  $400 \times 400$  pixels) images, respectively.

## 4. Discussion

### 4.1. Selection of Training Samples

In this study, the built-up area samples were selected using the Harris corner points. Although the Harris corner points are very effective in detecting densely distributed settlements, they may fail in detecting industrial areas with very large and elongated buildings (Figure 17).



**Figure 17.** Built-up area detection over settlements and industrial areas. The Harris corner points are drawn in green.

Such a result is caused by the sparse distribution of the Harris corner points. There are no Harris corner points retained in the industrial area (green rectangle), and thus it cannot be detected. However, as presented in the figure, another salient feature of built-up areas is the dense distribution of short lines. Thus, a combination of corner points and short straight lines might be a possible solution to this problem and will be evaluated in our future work. Besides, some typical built-up samples could be selected manually to improve the robustness of the proposed method. Especially in some areas, the patterns of built-up areas are very similar. Some datasets containing the typical samples could be established and used.

### 4.2. Fusion of Multi-Scale Information

The multi-scale feature representation can significantly improve the robustness of the final results. However, the boundaries of the final results are blurred. One possible solution to this is fusing the information obtained at different scales, in which the coarser contours are obtained at coarser scales and are used as auxiliary data for detecting accurate boundaries at finer scales. However, the fusion of multi-scale information is still a challenging task.

### 4.3. Acceleration and Parallel Processing

In the proposed method, the whole image is processed. This results in a low computational efficiency for processing very large images. In our experiments, we found that a built-up area could be detected using only a subset of the image. Thus, if the image is properly split, different parts of the built-up areas could be detected separately, and a parallel implementation could be used to accelerate the process.

## 5. Conclusions

In this paper, we have proposed a block-based multi-scale framework for built-up area detection. The image is first divided into blocks, in which the spectral, textural, and structural features are extracted and represented in a multi-scale framework. A set of refined Harris corner points is then used as clues for selecting the training blocks, and a built-up index is obtained by minimizing the

normalized spectral, textural, and structural distances of the image blocks to the nearest  $K$  samples. Finally, the built-up areas are obtained by thresholding the index image.

The experiments showed the effectiveness and robustness of the proposed method in detecting built-up areas from high-resolution optical and SAR images. Furthermore, the method achieves pleasing results in different image scenes and using images of different spatial resolutions. Thus, the proposed method can be used for built-up area mapping or as a pre-processing for accurate building detection in very high spatial resolution images.

We conclude that the proposed method can be used for built-up area mapping, and it is robust with regard to optical and SAR images of different scenes and different spatial resolutions. Thus, the proposed method has the potential to be applied for regional and global built-up area mapping. It could also be used to produce a mask for accurate building detection and related applications.

**Acknowledgments:** The authors would like to thank the anonymous reviewers for their valuable comments. This study was jointly supported by grants from the National Natural Science Foundation of China under Grant No. 41501369, the Postdoctoral Science Foundation of China under Grant No. 2014M562206, the Geographic National Condition Monitoring Engineering Research Center of Sichuan Province under Grant No. GC201513, the Shenzhen Scientific Research and Development Funding Program under Grants No. ZDSY20121019111146499 and JSJG20121026111056204, and the Shenzhen Dedicated Funding of Strategic Emerging Industry Development Program under Grant No. JCYJ20121019111128765. We would also like to thank Google, the Spot Image Company, the German Aerospace Center (DLR), and EADS Astrium for providing freely available images for research purposes. We are also grateful to Yingjie Tian (Wuhan University) for providing the experimental results of BASI.

**Author Contributions:** Zhongwen Hu developed the algorithm, carried out the experiments, and drafted the manuscript. Qingquan Li and Guofeng Wu provided suggestions and helped to improve the manuscript. Qian Zhang helped with the implementation of the algorithm.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

|        |   |
|--------|---|
| MODIS  | moderate resolution imaging spectroradiometer |
| MERIS  | medium resolution imaging spectrometer        |
| TM     | thematic mapper                               |
| ETM+   | enhanced thematic mapper plus                 |
| NDBI   | normalized difference built-up index          |
| SAR    | synthetic aperture radar                      |
| MRF    | Markov random field                           |
| SIFT   | scale-invariant feature transform             |
| GLCM   | gray-level co-occurrence matrix               |
| NSCT   | non-sampled contourlet transform              |
| DFID   | discrete field of image descriptors           |
| RILBP  | rotation-invariant local binary patterns      |
| LC     | local contrast                                |
| HOG    | histogram of oriented gradients               |
| BASI   | built-up areas saliency index                 |
| MBIs   | multiple built-up indexes                     |
| minMBI | minimum of the MBIs                           |

## References

1. Svensson, M.K.; Eliasson, I. Diurnal air temperatures in built-up areas in relation to urban planning. *Landsc. Urban Plann.* **2002**, *61*, 37–54. [[CrossRef](#)]

2. Kasanko, M.; Barredo, J.I.; Lavalle, C.; McCormick, N.; Demicheli, L.; Sagris, V.; Brezger, A. Are European cities becoming dispersed? A comparative analysis of 15 European urban areas. *Landsc. Urban Plann.* **2006**, *77*, 111–130. [[CrossRef](#)]
3. Pauleit, S.; Ennos, R.; Golding, Y. Modeling the environmental impacts of urban land use and land cover change—A study in Merseyside, UK. *Landsc. Urban Plann.* **2005**, *71*, 295–310. [[CrossRef](#)]
4. Pesaresi, M.; Gerhardinger, A.; Haag, F. Rapid damage assessment of built-up structures using VHR satellite data in tsunami-affected areas. *Int. J. Remote Sens.* **2007**, *28*, 3013–3036. [[CrossRef](#)]
5. Potere, D.; Schneider, A.; Angel, S.; Civco, D.L. Mapping urban areas on a global scale: which of the eight maps now available is more accurate? *Int. J. Remote Sens.* **2009**, *30*, 6531–6558. [[CrossRef](#)]
6. Tatem, A.J.; Noor, A.M.; Hay, S.I. Defining approaches to settlement mapping for public health management in Kenya using medium spatial resolution satellite imagery. *Remote Sens. Environ.* **2004**, *93*, 42–52. [[CrossRef](#)] [[PubMed](#)]
7. Lu, D.; Tian, H.; Zhou, G.; Ge, H. Regional mapping of human settlements in southeastern China with multisensor remotely sensed data. *Remote Sens. Environ.* **2008**, *112*, 3668–3679. [[CrossRef](#)]
8. Wang, L.; Gong, P.; Ying, Q.; Yang, Z.; Cheng, X.; Ran, Q. Settlement extraction in the North China Plain using Landsat and Beijing-1 multispectral data with an improved watershed segmentation algorithm. *Int. J. Remote Sens.* **2010**, *31*, 1411–1426. [[CrossRef](#)]
9. Pesaresi, M.; Huadong, G.; Blaes, X.; Ehrlich, D.; Ferri, S.; Gueguen, L.; Halkia, M.; Kauffmann, M.; Kemper, T.; Lu, L.; *et al.* A Global Human Settlement Layer from Optical HR/VHR RS Data: Concept and First Results. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 2102–2131. [[CrossRef](#)]
10. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16. [[CrossRef](#)]
11. 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. [[CrossRef](#)]
12. Xu, H. A new index for delineating built-up land features in satellite imagery. *Int. J. Remote Sens.* **2008**, *29*, 4269–4276. [[CrossRef](#)]
13. He, C.; Shi, P.; Xie, D.; Zhao, Y. Improving the normalized difference built-up index to map urban built-up areas using a semiautomatic segmentation approach. *Remote Sens. Lett.* **2010**, *1*, 213–221. [[CrossRef](#)]
14. 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 Obs. Remote Sens.* **2008**, *1*, 180–192. [[CrossRef](#)]
15. Sirmacek, B.; Unsalan, C. Urban-area and building detection using SIFT keypoints and graph theory. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1156–1167. [[CrossRef](#)]
16. Tao, C.; Tan, Y.; Zou, Z.R.; Tian, J. Unsupervised detection of built-up areas from multiple high-resolution remote sensing images. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1300–1304. [[CrossRef](#)]
17. Weizman, L.; Goldberger, J. Urban-area segmentation using visual words. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 388–392. [[CrossRef](#)]
18. Shi, H.; Chen, L.; Bi, F.K.; Chen, H.; Yu, Y. Accurate Urban Area Detection in Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2015**. in press.
19. Hofmann, P. Detecting Informal Settlements from IKONOS Image Data Using Methods of Object Oriented Image Analysis—An Example from Cape Town (South Africa). Available online: [http://www.ecognition.com/sites/default/files/395\\_hofmann.pdf](http://www.ecognition.com/sites/default/files/395_hofmann.pdf) (accessed on 24 August 2015).
20. Hofmann, P.; Strobl, J.; Blaschke, T.; Kux, H. Detecting informal settlements from QuickBird data in Rio de Janeiro using an object based approach. In *Object-Based Image Analysis*; Springer: London, UK, 2008; pp. 531–553.
21. Jacquin, A.; Misakova, L.; Gay, M. A hybrid object-based classification approach for mapping urban sprawl in periurban environment. *Landsc. Urban Plann.* **2008**, *84*, 152–165. [[CrossRef](#)]
22. Zhou, Y.; Wang, Y. Extraction of impervious surface areas from high spatial resolution imagery by multiple agent segmentation and classification. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 857–868. [[CrossRef](#)]
23. Hu, X.; Weng, Q. Impervious surface area extraction from IKONOS imagery using an object-based fuzzy method. *Geocarto Int.* **2011**, *26*, 3–20. [[CrossRef](#)]
24. Zheng, C.; Wang, L.; Zhao, H.; Chen, X. Urban area detection from high-spatial resolution remote sensing imagery using Markov random field-based region growing. *J. A. Remote Sens.* **2014**, *8*, 083566. [[CrossRef](#)]

25. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comp. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
26. Harris, C.; Stephens, M. A combined corner and edge detector. *Alvey Vis. Conf.* **1988**, *15*, 50.
27. Sirmaçek, B.; Ünsalan, C. Urban area detection using local feature points and spatial voting. *IEEE Geosci. Remote Sens. Lett.* **2010**, *7*, 146–150. [[CrossRef](#)]
28. Kovacs, A.; Szirányi, T. Improved Harris feature point set for orientation-sensitive urban-area detection in aerial images. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 796–800. [[CrossRef](#)]
29. Li, Y.; Tan, Y.; Deng, J.; Wen, Q.; Tian, J. Cauchy Graph Embedding Optimization for Built-Up Areas Detection From High-Resolution Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, in press. [[CrossRef](#)]
30. Gueguen, L.; Soille, P.; Pesaresi, M. A new built-up presence index based on density of corners. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012; pp. 5398–5401.
31. Gamba, P.; Aldrighi, M.; Stasolla, M. Robust extraction of urban area extents in HR and VHR SAR images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2011**, *4*, 27–34. [[CrossRef](#)]
32. Li, N.; Bruzzone, L.; Chen, Z.; Liu, F. A Novel Technique Based on the Combination of Labeled Co-Occurrence Matrix and Variogram for the Detection of Built-up Areas in High-Resolution SAR Images. *Remote Sens.* **2014**, *6*, 3857–3878. [[CrossRef](#)]
33. Shao, Z.; Tian, Y.; Shen, X. BASI: A new index to extract built-up areas from high-resolution remote sensing images by visual attention model. *Remote Sens. Lett.* **2014**, *5*, 305–314. [[CrossRef](#)]
34. Da Cunha, A.L.; Zhou, J.; Do, M.N. The nonsubsampling contourlet transform: Theory, design, and applications. *IEEE Trans. Image Process.* **2006**, *15*, 3089–3101. [[CrossRef](#)] [[PubMed](#)]
35. Gueguen, L. Classifying compound structures in satellite images: A compressed representation for fast queries. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1803–1818. [[CrossRef](#)]
36. Ojala, T.; Pietikainen, M.; Maenpää, T. Multiresolution gray-scale and rotation-invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [[CrossRef](#)]
37. Dalal, N.; Triggs, B. Histograms of Oriented Gradients for Human Detection. Available online: <https://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf> (accessed on 24 August 2015).
38. Zhong, P.; Wang, R. A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3978–3988. [[CrossRef](#)]
39. Lindeberg, T. *Scale-Space Theory in Computer Vision*; Springer: London, UK, 1993.
40. Otsu, N. A threshold selection method from gray-level histograms. *Automatica* **1975**, *11*, 23–27.
41. Kittler, J.; Illingworth, J. Minimum error thresholding. *Pattern Recogn.* **1986**, *19*, 41–47. [[CrossRef](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).