

Article

A Generalized Image Scene Decomposition-Based System for Supervised Classification of Very High Resolution Remote Sensing Imagery

ZhiYong Lv ^{1,*}, Haiqing He ^{2,3}, Jón Atli Benediktsson ⁴ and Hong Huang ⁵

¹ School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China

² Key Laboratory of Watershed Ecology and Geographical Environment Monitoring, National Administration of Surveying, Mapping and Geoinformation, Nanchang 330013, China; hyhqing@163.com

³ School of Geomatics, East China University of Technology, Nanchang 330013, China

⁴ Faculty of Electrical and Computer Engineering, University of Iceland, Reykjavik IS 107, Iceland; benedikt@hi.is

⁵ Key Laboratory of Optoelectronic Technology and Systems of the Education Ministry of China, Chongqing University, Chongqing 400044, China; hhuang.cqu@gmail.com

* Correspondence: Lvzhiyong_fly@hotmail.com; Tel.: +86-158-2967-3435

Academic Editors: Yuei-An Liou, Chyi-Tyi Lee, Yuriy Kuleshov, Jean-Pierre Barriot, Chung-Ru Ho, Guoqing Zhou, Xiaofeng Li and Prasad S. Thenkabail

Received: 25 July 2016; Accepted: 26 September 2016; Published: 30 September 2016

Abstract: Very high resolution (VHR) remote sensing images are widely used for land cover classification. However, to the best of our knowledge, few approaches have been shown to improve classification accuracies through image scene decomposition. In this paper, a simple yet powerful observational scene scale decomposition (OSSD)-based system is proposed for the classification of VHR images. Different from the traditional methods, the OSSD-based system aims to improve the classification performance by decomposing the complexity of an image's content. First, an image scene is divided into sub-image blocks through segmentation to decompose the image content. Subsequently, each sub-image block is classified respectively, or each block is processed firstly through an image filter or spectral-spatial feature extraction method, and then each processed segment is taken as the feature input of a classifier. Finally, classified sub-maps are fused together for accuracy evaluation. The effectiveness of our proposed approach was investigated through experiments performed on different images with different supervised classifiers, namely, support vector machine, k-nearest neighbor, naive Bayes classifier, and maximum likelihood classifier. Compared with the accuracy achieved without OSSD processing, the accuracy of each classifier improved significantly, and our proposed approach shows outstanding performance in terms of classification accuracy.

Keywords: observational scene; image decomposition; very high resolution; image classification

1. Introduction

Very high resolution (VHR) remote sensing images are images with very high spatial resolution. Higher spatial resolution remote sensing images not only exhibit better visual performance but also allow us to acquire a large amount of detailed ground information in urban areas. Thus, the potential applications of remote sensing, such as land cover mapping and monitoring [1], human activity analysis [2], and tree species classification [3], have increased. Numerous practical applications based on VHR remote sensing depend on classification results. However, given the limitation of remote sensing technology, VHR images are often coupled with poor radiometry, with less than three or five spectral bands. Therefore, this characteristic (i.e., a high spatial resolution but poor spectral

radiometry) does not necessarily reflect higher classification accuracies when the images are applied in practice [4–6].

Spatial features based on pixel-wise analysis are usually employed to improve the performance of VHR image classification. Zhang et al. proposed the pixel shape index to describe the spatial information around the central pixel [7]. Fauvel et al. presented a spatial-spectral classification method using SVMs and morphological profile (MP) [8]. In addition, morphological profiles (MPs) with different structural elements were proposed for structural feature extraction and classification of VHR images [9]. In previous studies [4,10], the spatial feature is combined not only with the spectral feature but is also often integrated with others. Many previous studies have demonstrated that high classification accuracies can be obtained by exploring the spatial features of VHR images in a pixel-wise manner. However, spatial-spectral feature-based methods depend on the performance of the spatial feature extraction algorithm and are usually data dependent. Any methods cannot be labeled as either a good or bad classification method [5], because although the high spatial resolution provides the possibility of detecting and distinguishing various objects with an improved visual effect, the classification of VHR images usually suffers from uncertainty regarding the spectral information, because the increase of the intra-class variance and decrease of the inter-class variance lead to a decrease of the separability in the spectral domain, particularly for the spectrally similar classes. In other words, as different objects may present similar spectral values in a VHR image, it is a difficult task to extract reasonable spatial features in practical application.

Object-based techniques have been proposed for VHR image classification to solve the above problems, which are brought about by the higher resolution and lower spectral information of the VHR images [11]. Object-based approaches usually start with segmentation to generate the image object, and an image object is formed by a group of pixels that are spectrally similar and spatially contiguous [12–14]. Recently, object-based classification methods for VHR remote sensing images have been studied extensively in practical applications [14–18]. For example, object-based spatial feature extraction has been proposed. This method aims to explore the relationship between each image object and its surrounding objects in the spectral and spatial domains. Zhong et al. defined an object-oriented conditional random field classification system for VHR remote sensing images, and the system utilized conditional random fields to extract contextual information in an objective manner [19]. Zhang et al. proposed an object correlative index that can describe the shape of a target in an objective manner [20]. Aside from the two aforementioned analyses, other powerful tools that maximize the spatial information of VHR image have been proposed for VHR images classification. These methods include graph theory [21], semantic clustering algorithm [22], and fuzzy rule-based approach [23]. As the analysis scale increases from a single pixel to an object, more features (such as texture, shape, and length) can be obtained than when a single pixel is used, and salt-and-pepper noise has usually been alleviated in the classification map of a VHR remote sensing image. Comparisons between pixel-wise and object-oriented approaches were conducted. The object-based approach demonstrated certain advantages in terms of classification results [24–27]; for instance, that it can smooth the noise in the classified map and that more image features (e.g., the texture and shape of the object) can be utilized. However, several limitations of the object-based approach to VHR image classification relative to a pixel-based approach were found. For example, although some scale estimate approaches have been proposed [28–31], it is still difficult to determine the optimal segmental scale parameters [32,33].

Aside from pixel-wise and object-based VHR image classification methods, a multi-scale strategy is an effective approach toward improving the classifying performance of VHR images. Bruzzone et al. revealed a multi-level context-based system for VHR classification. The system aimed to obtain accurate and reliable maps both by preserving geometrical details in images and properly considering multilevel spatial context information [34]. Huang et al. revealed a multi-scale feature fusion approach for the classification of VHR satellite imagery [35,36]. Johnson et al. proposed a competitive multi-scale object-based approach for land cover classification by using VHR remote sensing image [37]. All the

aforementioned works have verified that multi-scale feature extraction or multi-level system can effectively increase classification accuracy. Scale influence was also crucial in the classification of VHR images [38]. However, lots of papers focused on spectral–spatial feature extraction algorithms, object-based, or multi-scale based approaches for VHR image classification. To our knowledge, few techniques specifically focus on the influence of the image scene scale on the classification of VHR images. Moreover, from a theoretical viewpoint, classification accuracy can be improved by increasing the number of representative training samples for a supervised classifier, and this can be found in the existing literature [39,40]. Therefore, on the contrary, it is interesting to investigate how to improve the classification accuracies through decomposing the complexity of the image content in the case of pre-defined training samples.

In this paper, we attempt to improve the classification accuracy through reducing the complexity of image content, and propose a generalized system based on observational scene scale decomposition (OSSD) to improve the performance of VHR image classification. Unlike other methods presented in the literature, the proposed OSSD-based system is general and data-independent.

The rest of the paper is organized as follows: Section 2 presents a detailed description of the proposed OSSD-based system for classification of VHR imagery. Section 3 presents the dataset used for experiments and reports the experimental results. Finally, Section 4 provides a discussion of the proposed approach and concludes the paper.

2. The Proposed OSSD-Based Classification Method

In theory, the boundary among classes is more distinct, and a classifier can separate them more easily. However, because of the high spatial resolution and insufficient spectral information in a VHR image, intra-classes may have different spectral values, and the same spectral value may indicate different classes. Thus, determining the boundary between classes is difficult. Furthermore, spatial arrangement and spectral heterogeneity on the earth's surface are complex. In general, a larger geographic area may contain more classes. By contrast, a smaller area may cover fewer classes. Considerably complex image content will make classification difficult. Thus, an interesting task is to investigate whether observational scale decomposition can allow a supervised classifier to achieve a better boundary and obtain a higher accuracy than without any observational scene scale decomposition.

To verify our deduction, a primary example with different sized images is shown in Figure 1, where the image block is extracted from an aerial image that is acquired by an ADS-80 sensor with three false-color bands and high spatial resolution. The image content becomes increasingly complex as the image size increases. In this comparison, bands 1 and 2 are the input features. Based on the pixels' distribution in the feature space, the classification region is predicted by using a support vector machine (SVM) classifier with a Radial basis function (RBF) and a five-fold cross-validation. Data points are related to each pixel of the upper-left image. The training sample is set to ensure the fairness of each prediction. In Figure 1, region "A" matches with class "C1," region "B" matches with class "C2," and so on. It can be found that the distribution of each class is greater mixed with the increase of the image size, meaning that more points are misclassified compared with the related region.

Therefore, exploring the relationship between image content decomposition and its classification accuracies is valuable and significant. To achieve this purpose, an image will be decomposed into sub-image blocks, each of which is classified. The flowchart of the proposed OSSD-based classification system is shown in Figure 2. Three main blocks are contained. First, the whole image (I) is segmented by chessboard segmentation, and the image scene is decomposed according to the segmental vector. Second, each sub-image block ($S_1, S_2, S_3, \dots, S_n$) can be classified directly, or each sub-image block is processed through a filter, e.g., a median filter, to remove noise, and the filtered result is given as $(S'_1, S'_2, S'_3, \dots, S'_n)$. Then, each band of the sub-image block is taken as the input feature for a classifier, and the classification result is $(R_1, R_2, R_3, \dots, R_n)$. Finally, the accuracy is evaluated based on the classification map, which is fused using each sub-image classification. In Figure 2, n is the number of sub-image blocks.

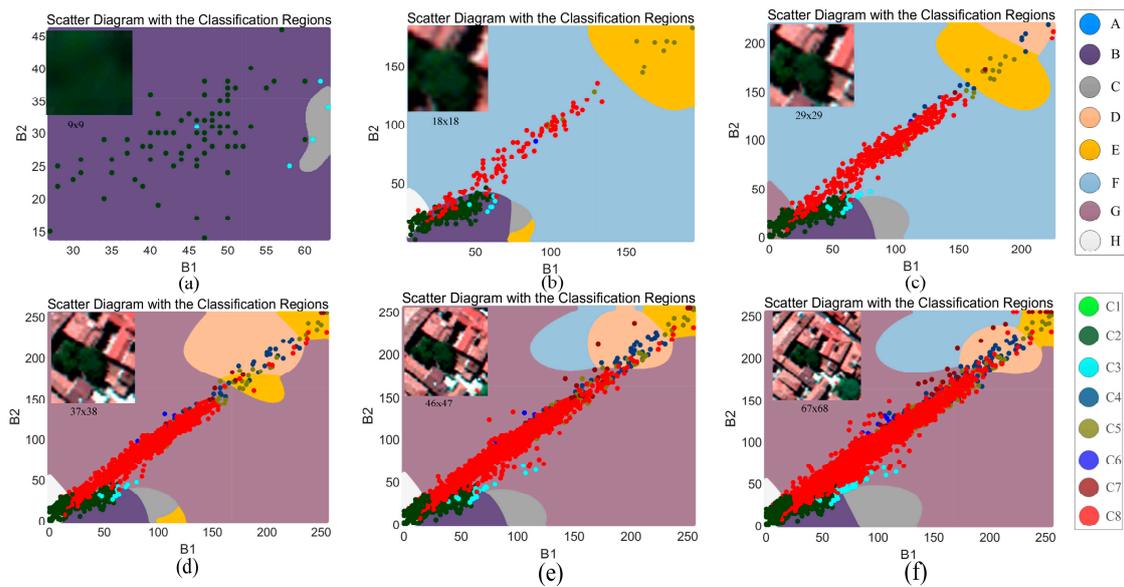


Figure 1. A demonstration to show the relationship between image content and its classified region using SVM, with an increase of candidate image spatial scale. From (a) to (f), the image size is 9×9 , 18×18 , 29×29 , 37×38 , 46×47 , 67×68 pixels, respectively.

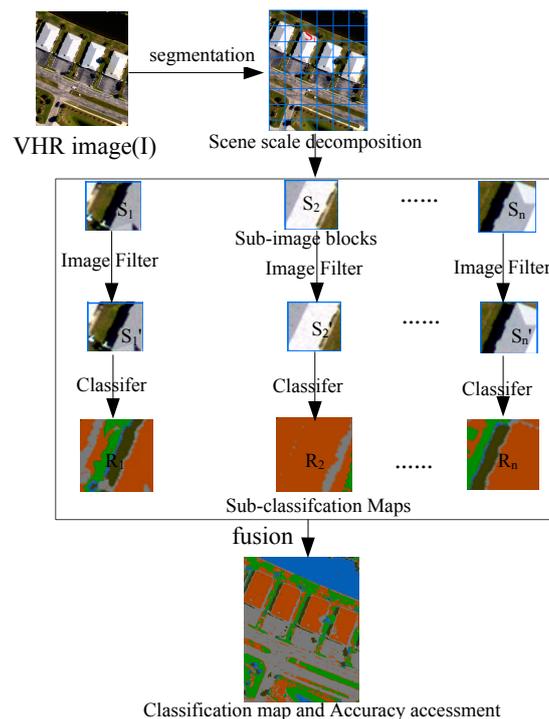


Figure 2. Flowchart of the proposed system based on OSSD coupled with image filter.

2.1. Scene Decomposition Using Chessboard Segmentation

OSSD is adopted to limit spatial range and reduce the image complexity for classification. Observational scene scale is defined as the range of a sub-image block, as shown in Figure 3. In theory, a smaller observation scene scale corresponds to lower complexity of a sub-image block, because more targets will be covered for classification as the observational scene scale increases. Therefore, the observational scene scale potentially affects the classification of VHR remote sensing images.

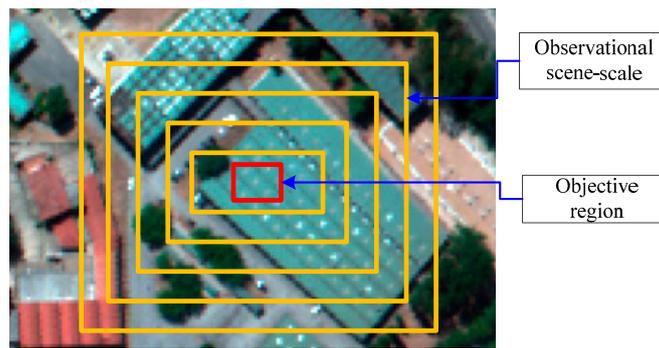


Figure 3. Definition of observational scene scale.

In this study, OSSD is achieved by chessboard segmentation, which is embedded in eCognition 6.8 business software. Chessboard segmentation decomposes the scene into equal squares of a given size. The chessboard segmentation algorithm produces simple square objects. Thus, this algorithm is often used to subdivide images. Here, chessboard segmentation can satisfy the requirements of the scene decomposition of our proposed system. A parameter called object size is important for chessboard segmentation and is symbolized by O . Object size is defined as the size of the square grid in pixels, and variables are rounded to the nearest integer.

2.2. Sub-Image Block Processing with Image Filter or Feature Extraction

While a scene is being decomposed by the chessboard segmentation method, an image scene is split into numerous sub-image blocks. In this step, each entire sub-image block will be processed by a median filter to remove noise. The median filter algorithm replaces the pixel value with the median of the neighboring pixels. The pattern of neighbors is called the “window,” which slides over the entire sub-image block pixel by pixel, and the size of the window is defined as r . As shown in Figure 4, the bands (b_1, b_2 and b_3) of each sub-image block are processed by median filter, and the result of this process (b'_1, b'_2 and b'_3) is fused together as S'_1 .

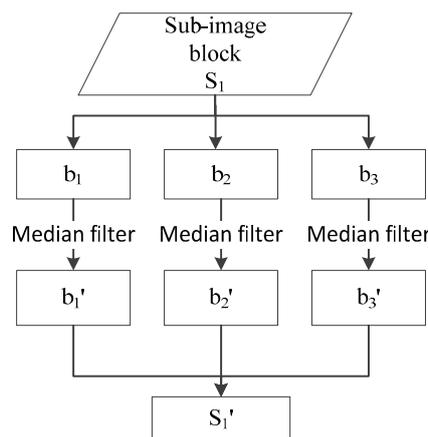


Figure 4. Example of each band of a sub-image block (S_1) filtered using a median filter.

Notably, the processing method for a sub-image block is not limited to a median filter. In general, a spatial feature extraction approach, such as a morphology profile (MP) [41,42], can also be used. MP is an effective and classical approach for extracting the spatial information of a VHR image; morphological opening and closing operators are used in order to isolate bright (opening) and dark (closing) structures in the images. Thus, in the following experiments, the proposed system coupled with the MP will also be investigated extensively.

2.3. Sub-Image Block Classification Using a Supervised Classifier

In this section, a training set was selected randomly from the whole image scene in a manual manner and a supervised classifier is adopted to achieve each sub-image block classification. On one hand, from the view of visual cognition, the complexity of the image scene is reduced by the OSSD. On the other hand, from the view of the previous demonstration (Figure 1), for a pre-given training sample and a supervised model (such as an SVM classifier), it can be found that it is more difficult to classify a larger sub-image block than that of a smaller sub-image block. In theory, it is difficult to evaluate the ability of a supervised classifier for different size testing image-blocks with a pre-given training sample. The classification accuracies not only relate to the distribution of testing datasets and training sample, they also have a strong correlation with the separability of the data and the ability of classifier. Therefore, we try to investigate the effectiveness and feasibility of the proposed OSSD-based approach by experimental comparisons.

3. Experimental Section

In this section, two VHR remote sensing images are used to validate the effectiveness and adaptability of the proposed system for improving the classification performance. Three parts are designed to achieve this purpose. In the first part, the study area is well described. In the second part, two experiments are set up, and the parameter settings are detailed. Finally, the results of the two experiments are given.

3.1. Datasets

Two real VHR remote sensing images are used to assess the effectiveness and adaptability of the proposed system.

The first image is an aerial image, which is acquired by an airborne ADS80 sensor, as shown in Figure 5. The relative flying height is approximately 3000 m, and the spatial resolution is 0.32 m. This aerial image is used to compare the accuracy obtained by the proposed OSSD-based classification system with that of without OSSD. The image is classified into five classes, namely, water, shade, grass, road, and building.

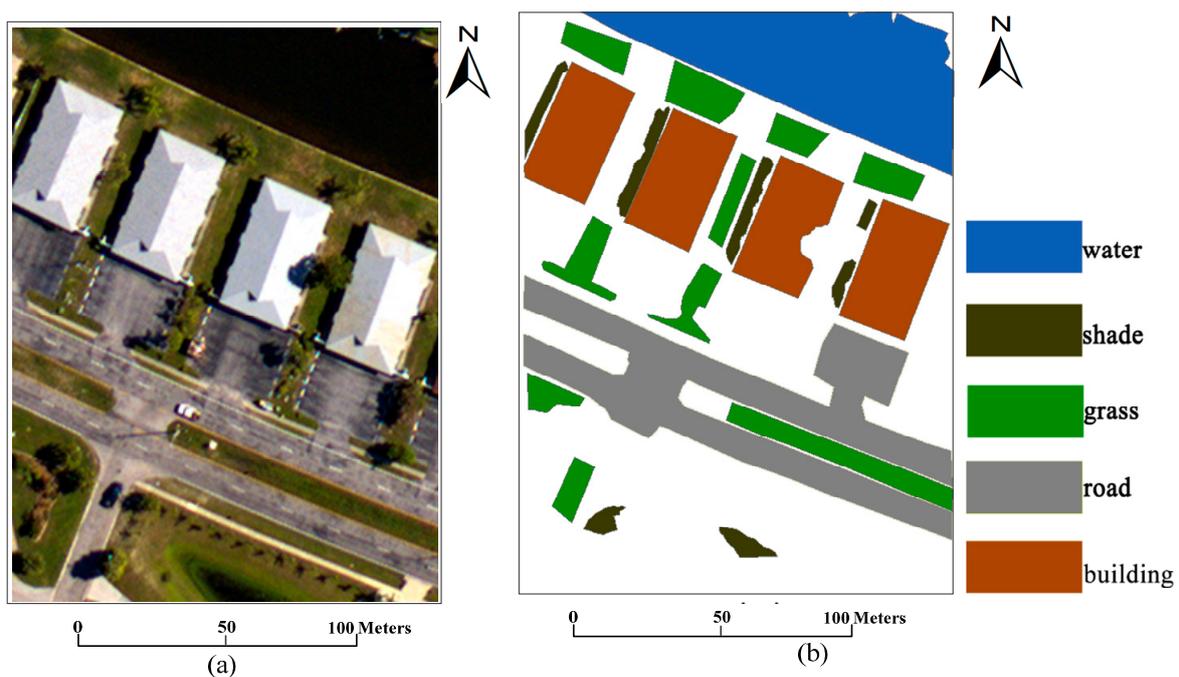


Figure 5. First study area: (a) aerial data; (b) ground reference data.

The second dataset is for an image acquired by the ROSIS-03 sensor, which covers the area of Pavia University, north Italy. The original dataset is 610×340 pixels with 115 spectral bands, with a spatial resolution of 1.3 m. A false color composite of the image with channels 60, 27, and 17 for red, green, and blue, respectively, is adopted here for our experiment, and the false color image is shown in Figure 6a. The following nine classes of interest were considered: trees, asphalt, bitumen, gravel, painted metal sheets, shadows, bricks, meadows, and soil. The ground reference map is shown in Figure 6b.

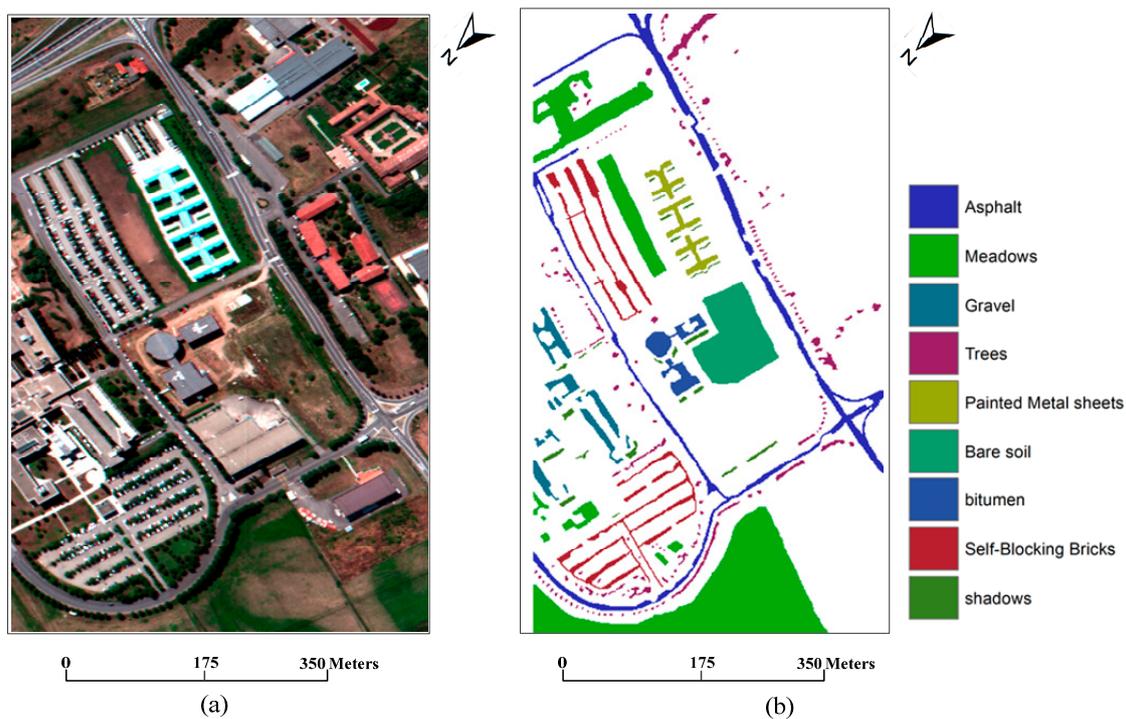


Figure 6. Second study area: (a) scene of Pavia University acquired by ROSIS-03; (b) ground reference of Pavia University.

For both datasets, classification is challenging because of the very high spatial resolution, which is 1.3 m or higher. Only three bands are used as the false color image for each experiment. In addition, as shown in the scene of each image, roads, buildings, water, and shade may be confused with one another. Hence, uncertainties may be engendered during land cover classification.

3.2. Experimental Setup and Parameter Settings

On the basis of the two VHR remote sensing images shown above, the accuracy and adaptability of the proposed OSSD-based classification system are investigated through the following two experimental setups.

The first experiment aims to evaluate the classification accuracies obtained by the proposed OSSD-based system and that of without OSSD processing. Training samples and reference data are collected for the following comparisons. The number of training samples and the reference data are shown in Table 1. The following four tests are designed: (1) the original test, in which the aerial image is classified without any image processing; (2) a median filter (MedF) test, in which the image is classified on the basis of median filter processing with a 3×3 window; (3) an OSSD-based test, in which the image is decomposed by chessboard segmentation, and each sub-image block is classified respectively by the supervised classifier; and (4) the MedF + OSSD test, in which the image is classified through the proposed OSSD-based system. In all the tests, the object size (O) is equal to 90 for the chessboard

algorithm. Then, to ensure a fair comparison, the median filter with a 3×3 window was also used to process each sub-image block. The above tests were performed based on the same training and reference pixels to ensure fairness.

Table 1. Number of training samples and reference data for the aerial image.

Class	Training Pixels	Test Pixels
Water	84	15,543
Shade	33	2930
Grass	100	8094
Road	102	16,879
Building	91	17,441

Adaptability is investigated with four different supervised classifiers, namely, K-nearest neighbor (KNN) [43], naive Bayes classifier (NBC) [44], maximum likelihood classifier (MLC) [45], and SVM [46], in the above comparison tests. KNN is a non-parametric method used for classification and regression. NBC is a simple probabilistic classifier based on the use of Bayes' theorem with strong (naive) independence assumptions between the features. MLC depends on the maximum likelihood estimation for each class. The SVM classifier with RBF kernel function and parameters is set by five-fold cross-validation.

OSSD coupled with the spatial feature extraction-MP approach is also investigated in this experiment. A structuring element with a disk shape and a size of 5×5 is used to extract structural information in the experiment. The image was first processed by OSSD with a parameter $O = 70$. Then, each sub-image block is processed by MP instead of a median filter. The processed result is used for classification.

The second experiment: This experiment has two objectives. First, similar to the first experiment, it aims to test the effectiveness of the proposed system. Second, it aims to explore the relationship between parameters and classification accuracies. In this experiment, for the study area of the Pavia University image, training samples were selected randomly, and the reference data are detailed in Table 2. To achieve these objectives, this experiment is designed as follows:

(1) Classification based on the proposed OSSD-based system is compared with that without OSSD. In addition, the proposed system is also compared with OSSD without any image processing.

(2) The sensitivity of the parameters, including the object size (O) and the radius (r) of the median filter, is investigated. The experimental approach incorporated one varied parameter, and the other parameters are constant. The details are shown in the experimental results section.

Table 2. Number of training samples and reference data for the Pavia University image.

Class	Training Sample	Test Pixels
Asphalt	96	6631
Meadows	100	18,649
Gravel	45	2099
Trees	46	3064
Painted metal	46	1345
Bare soil	97	5029
Bitumen	24	1330
Self-blocking bricks	51	3682
Shadows	36	847

3.3. Experimental Results

Experiments on the aerial image: The classification results of the four tests for the aerial image are compared in Figure 7. The results of each test are detailed as follows.

In the original test, the aerial image was directly classified by KNN, MLC, NBC, and SVM. As shown in the first column of Figure 6, salt-and-pepper noise is observed in each classified map, and the highest accuracy is obtained by KNN. The overall accuracy (OA) for KNN is 84.1%, whereas the kappa coefficient (Ka) is 0.79.

In the MedF-test, the aerial image was classified on the basis of median filter processing. The radius (r) is given as 3×3 . Compared with the accuracy achieved in the original test, the accuracy of each classifier is improved by approximately 3%, and the noise in the classification map was more or less eliminated.

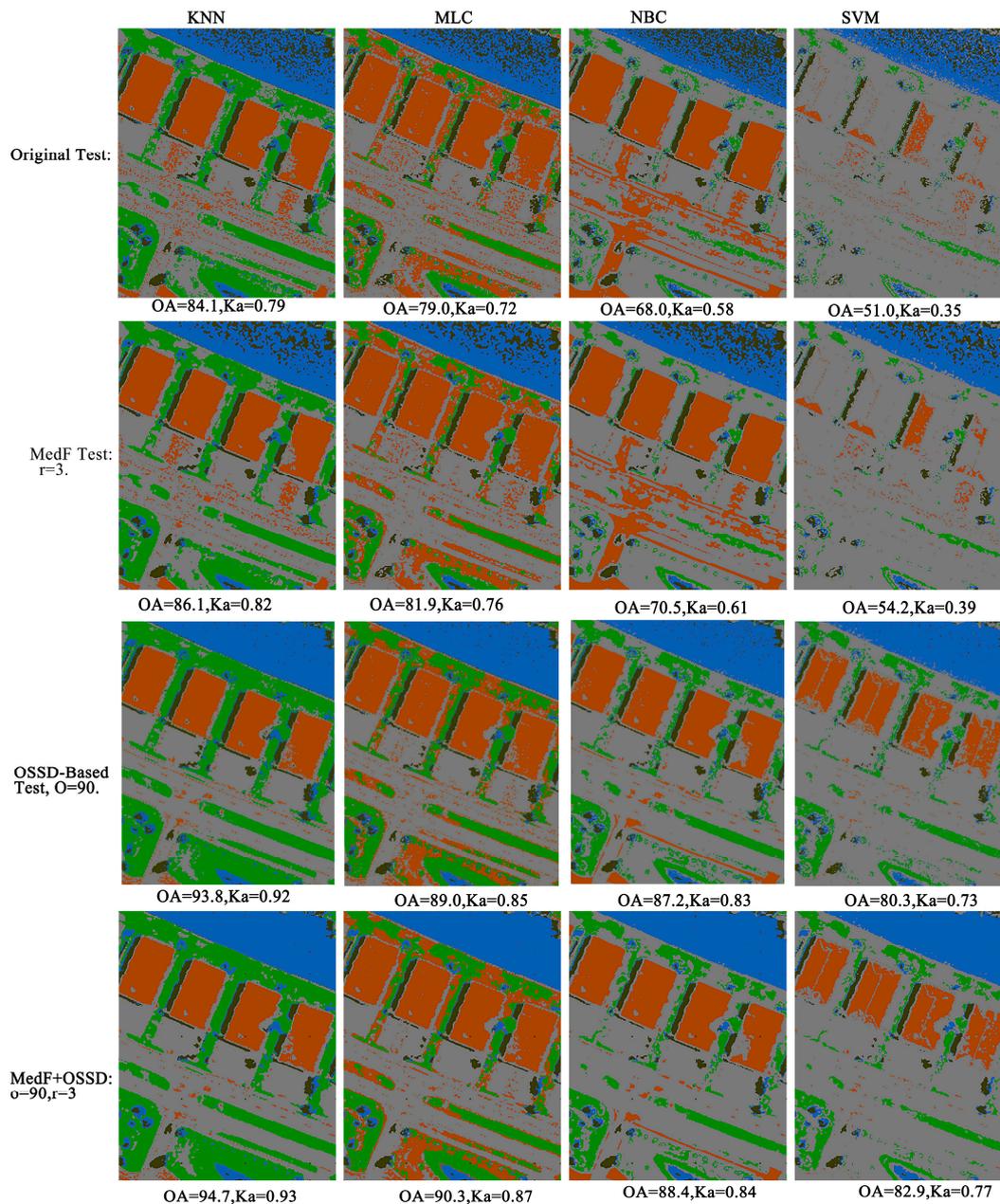


Figure 7. Classification results (aerial image); **The first column, the second column, the third column and the fourth column** are the classification maps for KNN, MLC, NBC, and SVM, respectively. Original test: image without any processing; MedF Test: the image is processed only with the median filter, $r = 3 \times 3$; OSSD-based test: the image is decomposed with $O = 90$; OSSD + MedF Test: the image is classified by the proposed system, with parameters $O = 90$ and $r = 3 \times 3$. The value of OA is given in percentage.

Further, in the OSSD + MedF test, the proposed OSSD-based system is applied to classify the aerial image. The results indicate that the classification performance achieved by the proposed system is evidently improved for each supervised classifier. The highest accuracy was achieved by the proposed system in terms of the four comparisons with different classifiers. Furthermore, the accuracy of each specific class based on the classifier KNN was also considered in the comparison, as shown in Table 3. Notably, the classification accuracies of water and grass are obviously improved. The classification accuracy of roads increased from 84.3% to 96.3%.

Finally, another test was adopted for MP. This test investigated the effectiveness of the proposed OSSD coupled with a spatial feature extraction MP, and a disk-shaped structural element with a size of 5×5 is employed for MP. The results are shown in Table 4. The comparison among each result of the original test, MP, and MP + OSSD test indicates that the OA, average accuracy (AA), and Ka are improved when the proposed OSSD is employed for the KNN supervised classifier.

Table 3. Classification accuracies (aerial image) in percentage for the KNN classifier and the four tests.

Class	Original Test	MedF	OSSD	OSSD + MedF
Water	76.1	80.1	93.9	97.8
Shade	80.6	81.0	73.3	74.5
Grass	74.7	75.7	88.6	90.4
Road	84.3	87.1	95.2	96.3
Building	95.8	96.1	95.4	95.7
OA	84.1	84.0	93.8	94.7
AA	82.3	86.1	89.9	91.0
Ka	0.79	0.82	0.92	0.93

Table 4. Aerial image classification accuracies for the test using only the MP feature and OSSD integrated with MP based on KNN.

Class	MP	OSSD + MP
Water	45.1	97.4
Shade	76.7	30.9
Grass	87.6	95.8
Road	98.1	98.1
Building	96.0	94.2
OA	81.6	93.3
AA	80.7	83.3
Ka	0.764	0.91

Experiments on the Pavia University Image: In this experiment, the effectiveness of the proposed system was evaluated by comparing the different classification methods for the Pavia University image. Then, the different values for the parameters of the proposed system were tested to detect their influence on the system. The number of training and test pixels in each class is listed in Table 2, and the available reference data are shown in Figure 6b. Details about this experiment are given below.

Classification maps acquired by different methods are grouped into four tests and, similar to the first experiment, four supervised classifiers are used, namely, KNN, MLC, NBC, and SVM. As shown in Figure 8, the original test obtains classification without any scene decomposition or image processing; the MedF test provides the result of the Pavia University image filtered by a median filter with $r = 3$. Compared with the original classification, the image scene is decomposed with $O = 70$, and the accuracy is improved from 50.3% to 60.6% for KNN. Finally, the OSSD + MedF test shows the results of our proposed approach with parameter settings of $O = 70$ and $r = 3$.

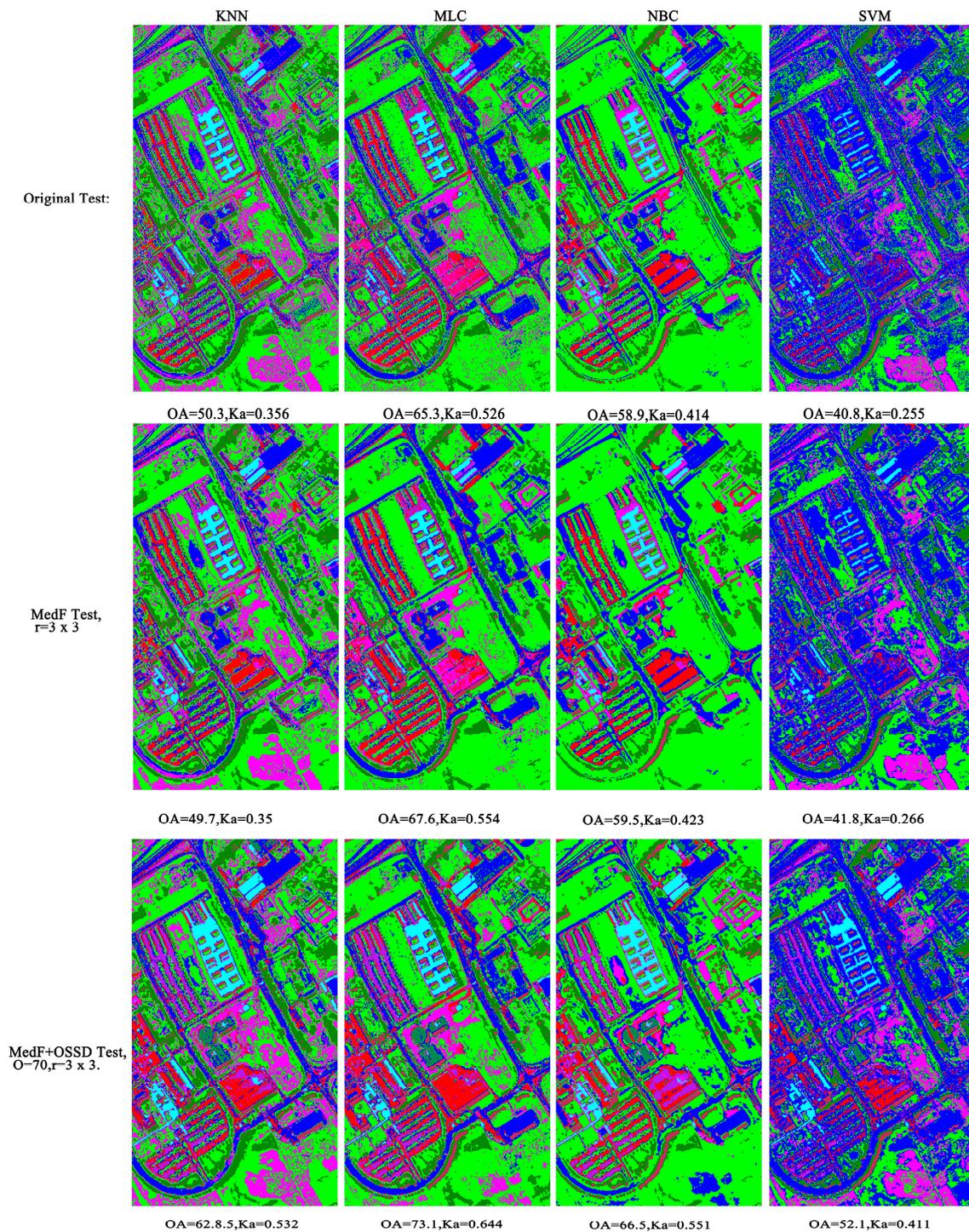


Figure 8. Classification results (Pavia University false color image); **The first column, the second column, the third column and the fourth column** are the classification maps of KNN, MLC, NBC, and SVM, respectively. Original test: image without any processing; MedF test: the image processed with a median filter, $r = 3 \times 3$; OSSD-Based Test, the original image is decomposed with $O = 70$; OSSD + MedF test: the image is classified by the proposed system, with the parameters $O = 70$ and $r = 3 \times 3$. The value of OA is given in percentage.

Specific classes of the four tests according to the KNN classifier were also assessed, as shown in Table 5. This table, original test and MedF test demonstrated similar accuracies, and the proposed approach for OSSD-based or OSSD + MedF test improved the classification accuracy of most

classes. In particular, bitumen and self-blocking bricks were improved by approximately 75% and 51%, respectively. This result was obtained because, compared with their neighboring targets, these classes were more easily identified in a limited spatial range of a single sub-image block by using OSSD processing.

Table 5. Classification accuracies (Pavia University image) in percent (%) for the KNN classifier in the three tests.

Class	Original	MedF	OSSD	MedF + OSSD
Asphalt	61.38	61.7	77.1	81.38
Meadows	59.69	57.8	55.9	57.13
Gravel	30.3	28.9	68.5	77.32
Trees	62.89	63.71	59.6	61.22
Painted metal sheets	99.1	99.63	99.6	99.78
Bare soil	37.1	38.06	47.2	49.35
Bitumen	1.2	0.9	69.5	75.04
Self-blocking bricks	0.16	0.02	52.6	51.68
Shadows	58.08	59.56	59.2	63.46
OA	50.3	49.7	60.6	62.8
AA	45.5	45.6	65.5	68.5
Ka	0.356	0.35	0.50	0.53

4. Discussion

Discussion of Experiment 1: Compared with the results of raw classification, as shown in Figure 9, while the original image is decomposed and each sub-image block is classified respectively, it can be found that the classification accuracy is notably improved by the decomposition of the image scene.

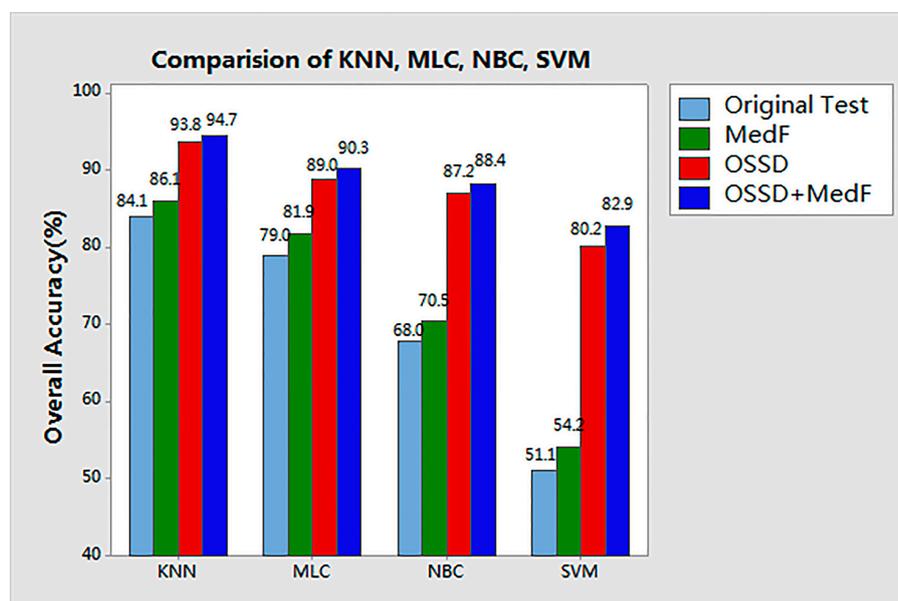


Figure 9. Aerial image accuracy comparisons of OA (%) based on different classification methods and supervised classifiers.

Notably, in the first experimental results, ground objects with a large scale, such as water, grass, and roads, can be identified with a higher accuracy when OSSD is integrated with MP. However, some smaller and regular-shaped ground objects, such as shade and buildings, exhibited poorer accuracy because their spatial features (such as shape and structure) may be destroyed when OSSD processing is adopted.

Discussion of Experiment 2: A comparison among the results of the four tests on the Pavia University image, as shown in Figure 10, shows that the classification performance of each supervised classifier in terms of accuracy is improved by the proposed OSSD-based system. The OA of the KNN classifier in particular exhibited an improvement of approximately 12.4%, while OSSD is adopted to decompose the complexity of the image content. MLC, NBC, and SVM all achieved an obvious improvement in terms of OA when OSSD is employed.

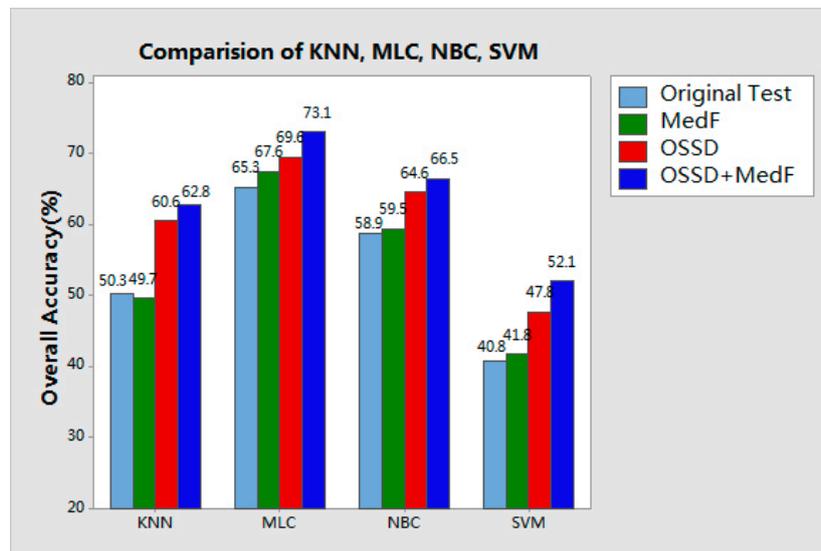


Figure 10. Pavia University image accuracy comparisons of OA (%) based on different classification methods and supervised classifiers.

Furthermore, the different values for the two parameters, namely, object size (O) and filter radius (r), of the proposed system were investigated in this experiment. The training set (Table 2) and the four classifiers (KNN, MLC, NBC, and SVM) are also employed for this evaluation.

As discussed above, object size (O) is a parameter of chessboard segmentation. This parameter also indicates the observational scene scale of the proposed system as prepared for the classifier. As mentioned in methodology section A, when the value of O increases, the larger observational scene scale image block will be extracted from the scene of the image through chessboard segmentation. Therefore, the value of O is larger, and fewer sub-image blocks will be produced through chessboard segmentation. The classification results with different O values are shown in Figure 11. At a fixed median filter radius of $r = 3 \times 3$, the OA curve grows almost to the horizontal level as O increases from 10 to 20. Moreover, the classification accuracies of each classifier also increase. However, the curve does not increase with the increment of O . In fact, the curve gradually reaches a stable trend and remains above 62% (KNN), 72% (MLC), 66% (NBC), and 52% (SVM), despite the wide range of O from 30 to 100. This test indicates that, if the value of O is small, then to a certain extent the proposed OSSD-based system cannot work effectively. This result occurs because if a smaller scene covers a single class, then it cannot provide sufficient information for distinguishing between classes.

Next, the parameter r indicates the window size for the median filter processing of every sub-image block. The relationship between OA and r is shown in Figure 12. This figure shows that the value of O was fixed at 90. When the value of r was varied from 1 to 5, the accuracy gradually escalated to the maximum value. Then, OA declined successively to the minimum when the value of r increased from 5 to 15. With the increase in window size, sufficient spatial information is considered, and the processing effect of the median filter reaches the optimum. However, an overly large r value will excessively remove the noise and blur the edge between different objects. Classification map comparisons based on the different values of r also reveal this fact, as seen in Figure 13.

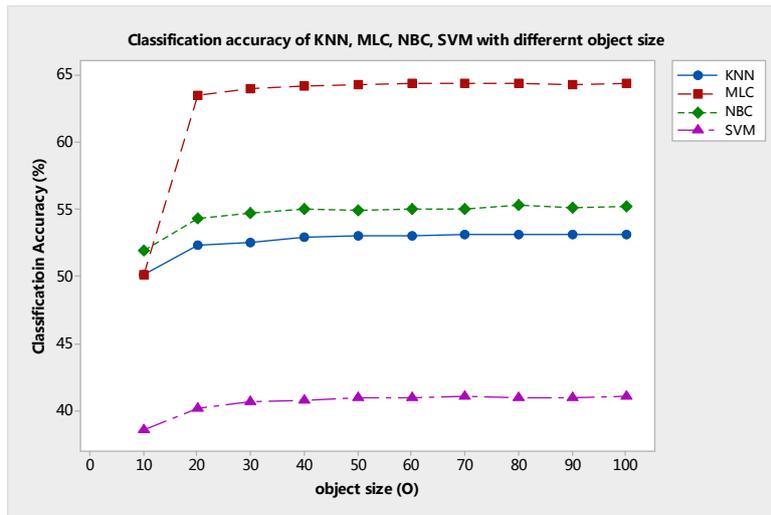


Figure 11. The relationship between OA and object size (O).

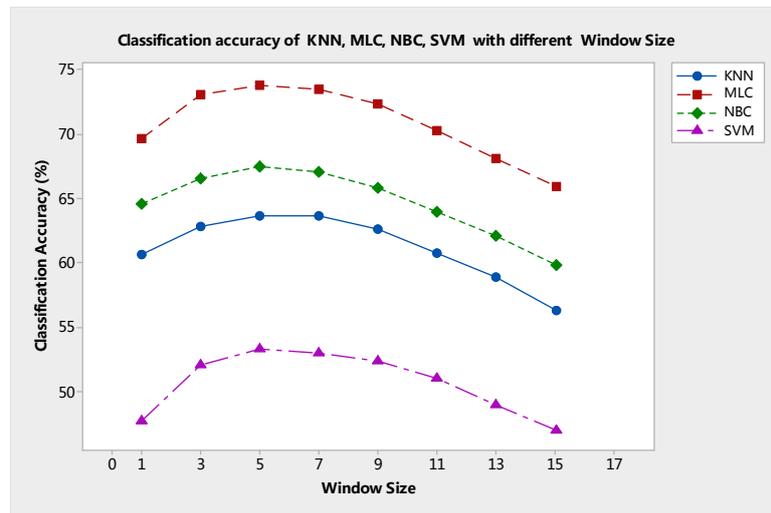


Figure 12. The relationship between OA and filter window size (r).

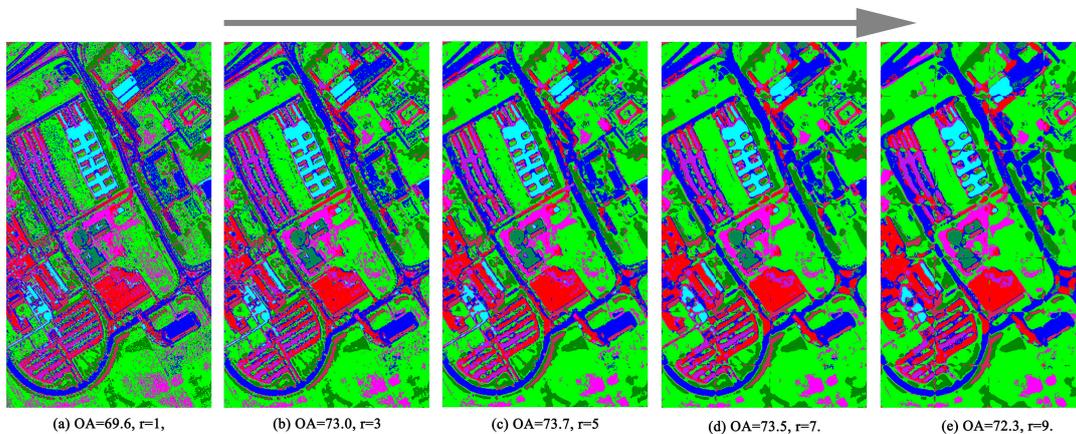


Figure 13. Classification maps based on the proposed system using MLC classifier. The fixed object size of OSSD-based processing is 90, and the value of window size for median filter (a–e) varies from 1 to 9.

Based on the two experimental discussions, it can be seen that the SVM classifier exhibits lower performance compared to the kNN ones in the proposed OSSD method, especially in the Pavia University image scene. This is owing to two aspects: (i) the training size of the Pavia University image is very small, with the training size being only 541 pixels, thus only 1.26% of the total ground truth 42,676 pixels; (ii) when the given image is decomposed into smaller sub-blocks, the number of training/testing samples contained by them is very small, which leads to insufficient learning of the respective classifiers. Therefore, one limitation of the proposed OSSD method is that accuracy depends on the decomposition scale (object size “*O*”) and each obtained block should contain a few training samples for proper learning of the respective classifiers.

5. Conclusions

In this paper, the relationship between the decomposition of image scene and classification accuracies has been investigated, and a generalized OSSD-based system has been proposed. The proposed system utilized OSSD to limit the spatial range of an image and reduce the complexity of the image content to improve the classification performance. The accuracies of the supervised classifier can be improved by the OSSD-based framework. Experiments were carried out on two real VHR remote sensing images. The results of the experiments showed the effectiveness and feasibility of the proposed OSSD method in improving the classification performance. Moreover, the proposed method has demonstrated several advantages, as is shown in the following:

(1) Our proposed OSSD-based system provides a generalized approach to improving the performance of classifiers. Furthermore, it can be generally applied to many supervised classifiers, because the proposed system presents no a priori constraints and is also data independent. Therefore, our proposed approach has more potential applications in VHR imagery processing and analysis.

(2) With respect to the experimental results, it reveals the fact that land cover classification accuracy can be improved through image scene decomposition. In addition, when OSSD is employed or coupled with an image filter or a feature extraction approach, an overall accuracy and kappa coefficient higher than that without OSSD processing can also be obtained. Therefore, the proposed OSSD-based system proves that the observational scene scale needs to be considered in the classification of VHR images.

With the development of remote sensing technology, a large amount of VHR images can be acquired conveniently, with land cover classification referring to the surface cover on the ground, whether it be vegetation, urban infrastructure, water, bare soil, or other. Identifying, delineating, and mapping land cover is important for global monitoring studies, resource management, and planning activities. Therefore, a generalized and simple classification system plays an important role in practical application.

This study provides a simple and effective strategy to improve VHR image classification accuracy. From the current experiments, it has been noted that the number and size of sub-blocks should be determined according to the class content variability in the terrain. A segmentation method should be chosen to delineate the blocks spatially. However, the class content of a sub-block is uncertain and unknown before classification. In future research, an extensive study of this will be conducted. In addition, the theory of the proposed OSSD-based system will be investigated extensively in other remote sensing applications and image scenes, such as for SAR images or forest image scenes.

Acknowledgments: The authors would like to thank the editor-in-chief, the anonymous associate editor, and the reviewers for their insightful comments and suggestions. This work was supported by the Geographic National Condition Monitoring Engineering Research Center of Sichuan Province (GC201515), a project of the China Postdoctoral Science Foundation (2015M572658XB), and the Visiting Scholar Foundation of Key Laboratory of Optoelectronic Technology and Systems (Chongqing University), Ministry of Education. The National Natural Science Foundation of China (41401526) and the Open Research Fund of Key Laboratory of Watershed Ecology and Geographical Environment Monitoring, NASG (WE2015003) also supported this work.

Author Contributions: ZhiYong Lv was primarily responsible for the original idea and experimental design. HaiQing He provides the fund supporting this research and publication. In addition, HaiQing He also contributed to the experiments. Jón Atli Benediktsson provided important suggestions for improving the paper’s quality. Hong Huang contributed to the experimental analysis and revised the paper.

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

References

1. Moser, G.; Serpico, S.B.; Atli, J. Land-cover mapping by Markov modeling of spatial–contextual information in very-high-resolution remote sensing images. *Proc. IEEE* **2013**, *101*, 631–651. [[CrossRef](#)]
2. Longbotham, N.; Chaapel, C.; Bleiler, L.; Padwick, C.; Emery, W.J.; Pacifici, F. Very high resolution multiangle urban classification analysis. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1155–1170. [[CrossRef](#)]
3. Dalponte, M.; Bruzzone, L.; Gianelle, D. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sens. Environ.* **2012**, *123*, 258–270. [[CrossRef](#)]
4. Huang, X.; Zhang, L. An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 257–272. [[CrossRef](#)]
5. Wilkinson, G.G. Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 433–440. [[CrossRef](#)]
6. Gianinetto, M.; Rusmini, M.; Candiani, G.; Dalla Via, G.; Frassy, F.; Maianti, P.; Marchesi, A.; Nodari, F.R.; Dini, L. Hierarchical classification of complex landscape with VHR pan-sharpened satellite data and OBIA techniques. *Eur. J. Remote Sens.* **2014**, *47*, 229–250. [[CrossRef](#)]
7. Zhang, L.; Huang, X.; Huang, B.; Li, P. A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2950–2961. [[CrossRef](#)]
8. Fauvel, M.; Benediktsson, J.A.; Chanussot, J.; Sveinsson, J.R. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 3804–3814. [[CrossRef](#)]
9. Lv, Z.Y.; Zhang, P.; Benediktsson, J.A.; Shi, W.Z. Morphological profiles based on differently shaped structuring elements for classification of images with very high spatial resolution. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4644–4652. [[CrossRef](#)]
10. Zhang, L.; Zhang, L.; Tao, D.; Huang, X. On combining multiple features for hyperspectral remote sensing image classification. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 879–893. [[CrossRef](#)]
11. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; van der Meer, F.; van der Werff, H.; van Coillie, F.; et al. Geographic object-based image analysis—towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191. [[CrossRef](#)] [[PubMed](#)]
12. Baatz, M.; Schäpe, A. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. *Angew. Geogr. Informationsverarbeitung XII* **2000**, *58*, 12–23.
13. Tang, Y.; Zhang, L.; Huang, X. Object-oriented change detection based on the Kolmogorov–Smirnov test using high-resolution multispectral imagery. *Int. J. Remote Sens.* **2011**, *32*, 5719–5740. [[CrossRef](#)]
14. Han, N.; Du, H.; Zhou, G.; Sun, X.; Ge, H.; Xu, X. Object-based classification using SPOT-5 imagery for Moso bamboo forest mapping. *Int. J. Remote Sens.* **2014**, *35*, 1126–1142. [[CrossRef](#)]
15. Zhang, C.; Zhao, Y.; Zhang, D.; Zhao, N. Application and evaluation of object-oriented technology in high-resolution remote sensing image classification. In Proceedings of the International Conference on Control, Automation and Systems Engineering, Singapore, 30–31 July 2011; pp. 1–4.
16. Malinverni, E.S.; Tassetti, A.N.; Mancini, A.; Zingaretti, P.; Frontoni, E.; Bernardini, A. Hybrid object-based approach for land use/land cover mapping using high spatial resolution imagery. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 1025–1043. [[CrossRef](#)]
17. De Pinho, C.M.D.; Fonseca, L.M.G.; Korting, T.S.; De Almeida, C.M.; Kux, H.J.H. Land-cover classification of an intra-urban environment using high-resolution images and object-based image analysis. *Int. J. Remote Sens.* **2012**, *33*, 5973–5995. [[CrossRef](#)]
18. Rasi, R.; Beuchle, R.; Bodart, C.; Vollmar, M.; Seliger, R.; Achard, F. Automatic updating of an object-based tropical forest cover classification and change assessment. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 66–73. [[CrossRef](#)]

19. Zhong, Y.; Zhao, J.; Zhang, L. A hybrid object-oriented conditional random field classification framework for high spatial resolution remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7023–7037. [[CrossRef](#)]
20. Zhang, P.; Lv, Z.; Shi, W. Object-based spatial feature for classification of very high resolution remote sensing images. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1572–1576. [[CrossRef](#)]
21. Chen, X.; Fang, T.; Huo, H.; Li, D. Graph-based feature selection for object-oriented classification in VHR airborne imagery. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 353–365. [[CrossRef](#)]
22. Yi, W.; Tang, H.; Chen, Y. An object-oriented semantic clustering algorithm for high-resolution remote sensing images using the aspect model. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 522–526. [[CrossRef](#)]
23. Hamedianfar, A.; Shafri, H.Z.M. Development of fuzzy rule-based parameters for urban object-oriented classification using very high resolution imagery. *Geocarto Int.* **2014**, *29*, 268–292. [[CrossRef](#)]
24. Robertson, D.L.; King, D.J. Comparison of pixel-and object-based classification in land cover change mapping. *Int. J. Remote Sens.* **2011**, *32*, 1505–1529. [[CrossRef](#)]
25. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [[CrossRef](#)]
26. Myint, S.W.; Gober, P.; Brazel, A.; Grossman-Clarke, S.; Weng, Q. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* **2011**, *115*, 1145–1161. [[CrossRef](#)]
27. Tehrany, M.S.; Pradhan, B.; Jebuv, M.N. A comparative assessment between object and pixel-based classification approaches for land use/land cover mapping using SPOT 5 imagery. *Geocarto Int.* **2014**, *29*, 351–369. [[CrossRef](#)]
28. Drăguț, L.; Tiede, D.; Levick, S.R. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 859–871. [[CrossRef](#)]
29. Drăguț, L.; Csillik, O.; Eisank, C.; Tiede, D. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 119–127. [[CrossRef](#)] [[PubMed](#)]
30. Montagni, A.; Larsen, R.; Greve, M.H. Accuracy assessment measures for image segmentation goodness of the Land Parcel Identification System (LPIS) in Denmark. *Remote Sens. Lett.* **2013**, *4*, 946–955. [[CrossRef](#)]
31. Weng, Q.; Hofmann, P. Defining robustness measures for OBIA framework: A case study for detecting informal settlements. In *Global Urban Monitoring and Assessment through Earth Observation*; Weng, Q., Ed.; CRC Press: Boca Raton, FL, USA, 2014; pp. 303–324.
32. Liu, D.; Xia, F. Assessing object-based classification: Advantages and limitations. *Remote Sens. Lett.* **2010**, *1*, 187–194. [[CrossRef](#)]
33. Arvor, D.; Durieux, L.; Andrés, S.; Laporte, M.-A. Advances in geographic object-based image analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 125–137. [[CrossRef](#)]
34. Bruzzone, L.; Carlin, L. A multilevel context-based system for classification of very high spatial resolution images. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2587–2600. [[CrossRef](#)]
35. Huang, X.; Zhang, L.; Li, P. A multiscale feature fusion approach for classification of very high resolution satellite imagery based on wavelet transform. *Int. J. Remote Sens.* **2008**, *29*, 5923–5941. [[CrossRef](#)]
36. Huang, X.; Zhang, L. A multilevel decision fusion approach for urban mapping using very high-resolution multi/hyperspectral imagery. *Int. J. Remote Sens.* **2012**, *33*, 3354–3372. [[CrossRef](#)]
37. Johnson, B.A. High-resolution urban land-cover classification using a competitive multi-scale object-based approach. *Remote Sens. Lett.* **2013**, *4*, 131–140. [[CrossRef](#)]
38. Quattrochi, D.A.; Goodchild, M.F. *Scale in Remote Sensing and GIS*; CRC Press: Boca Raton, FL, USA, 1997.
39. Huang, X.; Guan, X.; Benediktsson, J.A.; Zhang, L.; Li, J.; Plaza, A.; Dalla Mura, M. Multiple Morphological Profiles From Multicomponent-Base Images for Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4653–4669. [[CrossRef](#)]
40. Shao, Y.; Lunetta, R.S. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. *ISPRS J. Photogramm. Remote Sens.* **2012**, *70*, 78–87. [[CrossRef](#)]

41. Huang, X.; Han, X.; Zhang, L.; Gong, J.; Liao, W.; Benediktsson, J.A. Generalized Differential Morphological Profiles for Remote Sensing Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1736–1751. [[CrossRef](#)]
42. Yıldırım, I.; Ersoy, O.K.; Yazgan, B. Improvement of classification accuracy in remote sensing using morphological filter. *Adv. Space Res.* **2005**, *36*, 1003–1006. [[CrossRef](#)]
43. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [[CrossRef](#)]
44. Zhang, M.L.; Peña, J.M.; Robles, V. Feature selection for multi-label naive Bayes classification. *Inf. Sci.* **2009**, *179*, 3218–3229. [[CrossRef](#)]
45. Strahler, A.H. The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sens. Environ.* **1980**, *10*, 135–163. [[CrossRef](#)]
46. Chang, C.C.; Lin, C.J. LIBSVM: A Library for Support Vector Machine. 2012. Available online: <http://www.csie.ntu.edu.tw/~cjlin/libsvm> (accessed on 25 September 2016).



© 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 Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).