Article

Object-Oriented Change Detection Method Based on Spectral–Spatial–Saliency Change Information and Fuzzy Integral Decision Fusion for HR Remote Sensing Images

Chuting Ge [1], Haiyong Ding [1,*], Inigo Molina [2], Yongjian He [3] and Daifeng Peng [1]

[1] School of Remote Sensing & Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 20201235003@nuist.edu.cn (C.G.); daifeng@nuist.edu.cn (D.P.)
[2] ETSI Topografía, Geodesia y Cartografía, Campus Sur UPM, C/Mercator 2, Universidad Politécnica de Madrid, 28031 Madrid, Spain; inigo.molina@upm.es
[3] School of Geography Science, Nanjing University of Information Science and Technology, Nanjing 210044, China; 001529@nuist.edu.cn
* Correspondence: 002093@nuist.edu.cn

Abstract: Spectral features in remote sensing images are extensively utilized to detect land cover changes. However, detection noise appearing in the changing maps due to the abundant spatial details in the high-resolution images makes it difficult to acquire an accurate interpretation result. In this paper, an object-oriented change detection approach is proposed which integrates spectral–spatial–saliency change information and fuzzy integral decision fusion for high-resolution remote sensing images with the purpose of eliminating the impact of detection noise. First, to reduce the influence of feature uncertainty, spectral feature change is generated by three independent methods, and spatial change information is obtained by spatial feature set construction and the optimal feature selection strategy. Secondly, the saliency change map of bi-temporal images is obtained with the co-saliency detection method to complement the insufficiency of image features. Then, the image objects are acquired by multi-scale segmentation based on the staking images. Finally, different pixel-level image change information and the segmentation result are fused using the fuzzy integral decision theory to determine the object change probability. Three high-resolution remote sensing image datasets and three comparative experiments were carried out to evaluate the performance of the proposed algorithm. Spectral–spatial–saliency change information was found to play a major role in the change detection of high-resolution remote sensing images, and the fuzzy integral decision strategy was found to effectively obtain reliable changed objects to improve the accuracy and robustness of change detection.

Keywords: change detection; fuzzy integral decision fusion; co-saliency detection; spectral–spatial features; object-oriented method

1. Introduction

With the availability of increasingly high-resolution (HR) satellite images, remote sensing is extensively utilized in many fields, such as urban planning, forest fire monitoring, and vegetation phenology change [1–4]. As an important application of remote sensing technology, change detection, as well as its role in revealing changes in land cover, is now one of the critical research hotspots due to the close relationship between residents and their environment [5–8].

Change detection (CD) refers to the process of detecting changes in the land surface from bi-temporal, multi-temporal, and time series images acquired by different types of sensors [9–13]. As one of the most important applications of satellite images, CD plays a key role not only in finding change objects, but also in providing further insight into the process of the evolution of the land surface. Recently, due to the global rapid

urbanization process, CD is much more important because it provides accurate information on changes in the land cover, for example, damages caused by earthquakes and flooding, the extents of urban expansion, and the areas of forest fires [14–19]. Due to the rich spatial characteristics of high-resolution remote sensing images, there are obvious differences between different surface objects and certain spatial heterogeneity within the same object. This hinders applying conventional change detection or semantic detection methods, such as the difference method (DI), log ratio method, and change vector analysis method (CVA), to the change detection of high-resolution remote sensing images. Therefore, accurate and robust change detection or semantic detection approaches are still needed to meet these application requirements in order to obtain a better and deeper understanding of the change in land cover.

According to the gray value, used only as statistical information, the change detection results obtained are often incomplete, and there are several spurious changed areas. A variety of transform-based change detection algorithms were proposed, such as iterative slow feature analysis (ISFA) [20], iteratively reweighted multivariate alteration detection (IRMAD) [21], and principal component analysis (PCA) [22]. The errors caused by the illumination conditions and radiation differences prove the limitation of utilizing spectral information alone. In contrast, texture and structural features are more stable and are not affected by spectral differences. Therefore, the idea of merging multiple features for change detection is widely adopted.

Spectral, texture, structural features, and other changes are widely used in existing studies. Based on spectral characteristics, such as the spectral correlation mapper (SCM) [23], the spectral gradient difference (SGD) [23], the Kullback–Leiber divergence [24], and the neighborhood correlation image (NCI) are used for the change detection of remote sensing images. Based on texture features, for instance, the Markov random field (MRF) texture [25], grey level co-occurrence matrix (GLCM) [26,27], and wavelet based textural features [28] are used for the change detection and object extraction of remote sensing images with high spatial resolution. Based on structural features, such as extended morphological profiles (EMPs) [29], the rolling guide filter (RGF), histogram of the orientated gradient (HOG) [30], and channel characteristics of orientated gradients (CFOGs) [30], morphological attribute profiles (APs) [31] are used to detect the land use and land cover. In addition, other change information, such as the morphological building index (MBI) [32,33], the normalized difference vegetation index (NDVI) [34], and the modified normalized difference water index (MNDWI) [35] can supplement the shortcomings of the change results with the help of image features to optimize the final detection results. In addition, deep learning is receiving much attention in different computer vision research areas, including the analysis of remote sensing images change detection [36–39]. Deep features of pixels or objects are extracted through deep learning methods, such as the neural network of spatial–temporal attention [14], the transformer-based model [40], and fully convolutional two-stream architecture [41]. Meanwhile, to better aggregate contextual and detailed information from remote sensing images, some researchers introduced feature fusion networks for change detection [41–43]. It should be noted that deep learning-based methods require a certain number of labelled training samples. Unfortunately, there are often not enough training data that represent the real change information of land cover objects [39]. In summary, the ability to fuse multi-information features for change detection in order to obtain reliable land cover change results is very necessary for high-resolution remote sensing images.

It is noted that some of the algorithms were developed for medium or coarse spatial resolution multispectral images. The abundance of spectral features in the multispectral images makes it much easier to implement these methods to detect land cover changes. Meanwhile, with the availability of higher resolution or very high-resolution images, the requirement of developing CD algorithms for HR images becomes much more pressing. As noticed, the pixel-based CD algorithms developed for the medium spatial resolution multispectral images are not fully appropriate for HR images due to the heterogeneity of

a pixel-based semantic image, meaning it does not account for the spatial context of an image [22]. Two kinds of strategies are used to detect changes in HR images. The first strategy is to extract as many features as possible from multiple scale images in order to compensate for the scarcity of spectral features, so that the CD algorithms developed on the multispectral features can be used at pixel-level CD in HR images [44,45]. The second strategy is to develop object-based CD algorithms [46,47] by segmenting an HR image into many non-overlapping objects. More robust change detection results are obtained by generating and processing superpixels for optical and SAR images [48–52]. In addition, the organic combination of the above two strategies provide a new idea for high-resolution image change detection.

Compared with pixel-level methods, object-level change detection approaches can effectively integrate change information from remote sensing images and avoid the influence of salt and pepper noise. However, the detection accuracy depends on the quality of the segmentation results [53], so it is worth pondering how to choose the optimal segmentation scale. Moreover, compared to the cumbersome process of the direct object comparison method and the object classification post-comparison method, the idea of directly combining the segmentation result with the initial detection results, for example, the Dempster–Shafer fusion theory [23,54,55], weighted Dempster–Shafer fusion theory [22], and majority voting fusion [29,56,57] can greatly save time and efficiency. As reported in many references, the effectiveness of providing accurate results is different for different types of CD approaches, and the ensemble idea is considered as a key solution to reach a high CD accuracy. Du et al. [5] discussed that the change detection effects of different fusion strategies, i.e., the feature-level fusion, the decision-level fusion, and the improved CD results could be achieved when compared to the CD results of a single approach. Much more effort was made in this direction to find an improved fusion strategy [57–60]. In high-resolution remote sensing images, the spectral characteristics of ground objects can reflect the rich information on the categories and attributes of objects, while the spatial features can help identify buildings and roads. They complement each other and jointly reveal the rich information on land cover contained in the HR remote sensing images [61]. Furthermore, using other change information to optimize and supplement the detection results based on image features is also a new idea for improving the accuracy of change detection. The use of multi-information and decision fusion strategies is verified to be helpful in obtaining accurate change detection results. The object-oriented method can overcome the uncertainty of the ground targets and further improve the accuracy of change detection [38,62].

Inspired by such research, in this paper we propose an object-oriented change detection algorithm to make a comprehensive application of various forms of information, to convert from a single detection method to a multi-method fusion, and to convert from the pixel level to the object level by decision fusion. Three main characteristics can be found in the proposed algorithm. First, the co-saliency change map of bi-temporal remote sensing images not only considers the contrast cues, spatial cues, and correlation cues, but also supplements the insufficiency of image features. Second, unlike other traditional methods that apply only a single feature, spectral–spatial–saliency change information is utilized comprehensively to overcome the shortcomings of a single factor. Third, in the proposed approach, the combination of feature-level and decision-level fusion is used. The most important contribution of the suggested framework lies in constructing a new object-based configuration based on spectral–spatial–saliency change information and fusion using the fuzzy integral decision theory, which plays a key role in the transition from pixel-level detection to object-level recognition. It should be noted that the initial pixel-based change results and object-based segmentations can be organically fused according to the fuzzy integral strategy, which can determine the change probability of land objects regardless of interference factors to achieve reliable detection results.

The remainder of this paper is organized as follows: Section 2 presents the proposed change detection approach. Section 3 shows the experimental datasets and configuration.

The experimental results are described in Section 4. A detailed discussion is addressed in Section 5 and the conclusion is drawn in Section 6.

2. Methodology

It is a classical strategy in the field of information fusion to integrate multiple forms of information and multiple methods to reach improved results. Inspired by this idea, we tried to assemble multiple pieces of change information into object-oriented change detection, and the flowchart of the proposed method is shown in Figure 1. The proposed approach consists of the following steps:

(1) To overcome the limitations of a single extraction method, spectral feature change is generated by three independent algorithms (IRMAD, ISFA, and PCA) as well as the majority voting fusion strategy.

(2) Considering the scarcity of only employing image features, the cluster-based co-saliency method is used to acquire the saliency change information of two temporal remote sensing images.

(3) The spatial feature sets of bi-temporal remote sensing images are constructed by using a histogram of oriented gradient, multi-scale grey-level co-occurrence matrix texture and rolling guidance filter, and then the spatial change information is obtained through optimal feature selection and adaptive threshold segmentation.

(4) Multi-scale segmentation is performed on the superimposed first principal component image, and the optimal segmentation result is obtained by the scale parameter determination strategy.

(5) Initial pixel-level change information and the segmentation results are combined by fuzzy integral decision fusion to obtain the final land cover change results.
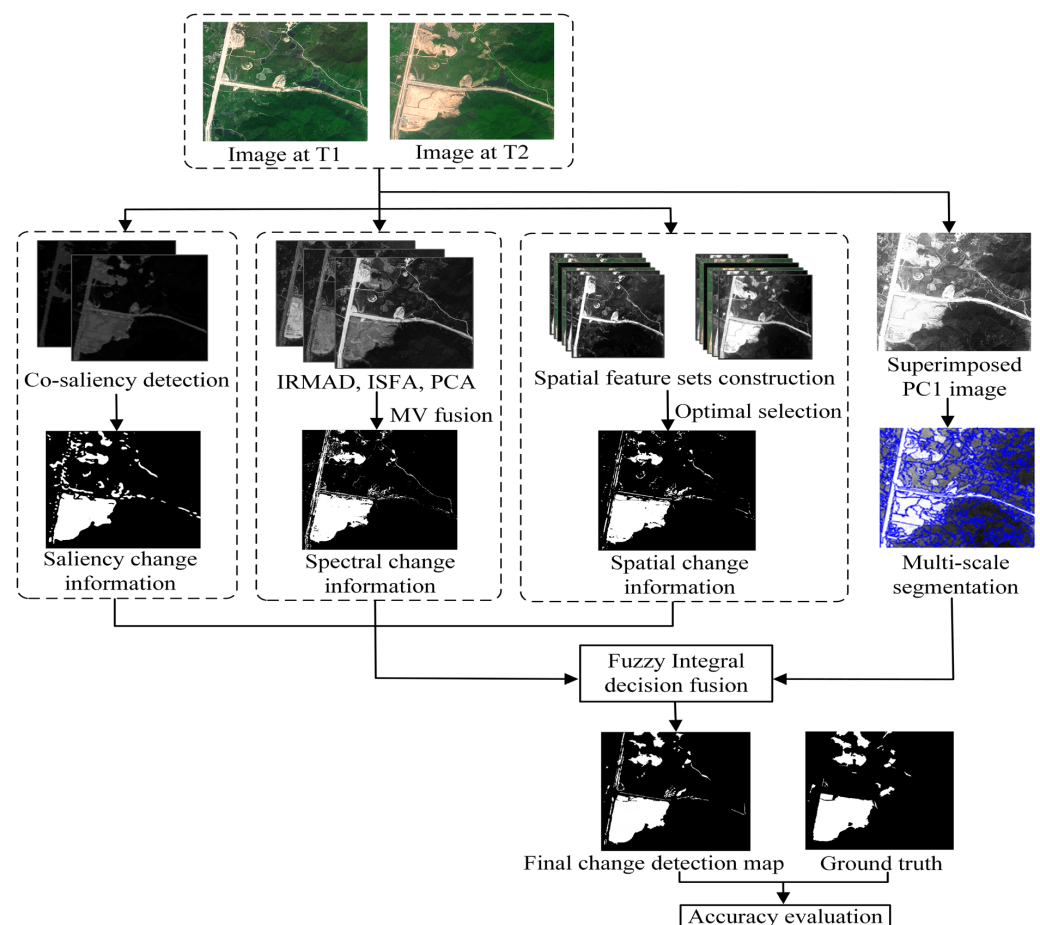

Figure 1. Framework of the proposed change detection approach.

2.1. Change Information Generation

The change information generation is composed of the first three phases previously exposed, and are described in detail below.

2.1.1. Spectral Change Information

The rich spectral characteristics of ground objects in remote sensing images provide a good reference for change detection. However, due to the false phenomenon caused by illumination conditions or radiation differences, the accuracy of a single spectral extraction method is limited. Therefore, the proposed algorithm integrates the initial results of three spectral change detection methods (IRMAD, ISFA, and PCA) to obtain accurate and comprehensive spectral change information after majority voting decision analysis.

IRMAD [22] is a typical algorithm based on spectral transformation. In the change detection problem, a random variable related to the correlation between the two-phase images is introduced in the basic function, and then the chi-square distribution function is used to iteratively reweight the pixels. Pixels with changed spectral characteristics will receive a smaller weight [21], and the new weight will be used for the next iteration until convergence. The spectral difference of IRMAD is calculated by the chi-square distance:

$$X_{IRMAD} = \sum_{K=1}^{N} \left(\frac{U_k - V_k}{\sigma_k} \right)^2 \tag{1}$$

$$U_k = a_k^T X_1 \tag{2}$$

$$V_k = b_k^T X_2 \tag{3}$$

where, σ_k is the standard deviation of the k-th band; a_k^T and b_k^T are the transformation vectors calculated by canonical correlation analysis [21]; k is the number of bands.

ISFA is similar to IRMAD. Compared with changed pixels, invariant pixels keep the spectral characteristics unchanged or changed weakly in bi-temporal images. When the spectral invariant components are extracted, the images will be converted into a new feature space. The ISFA method applies the idea of iterative weighting to assign larger weights to unchanged pixels in the iterative process, so that invariant pixels become more and more important in the calculation, and improve the separability of changed and unchanged pixels in the feature difference [20]. The chi-square distance is also utilized to calculate the spectral difference map:

$$X_{ISFA} = \sum_{j=1}^{N} \left(\frac{SFA_j}{\phi_j} \right)^2 \tag{4}$$

$$SFA_j = w_j^T X_1 - w_j^T X_2 \tag{5}$$

where ϕ_j is the variance of the j-th SFA characteristic band and j is the number of bands. w_j represents the transformation vector in ISFA, which satisfies the constraint optimization conditions of zero mean, unit variance, and de-correlation [54].

PCA method performs precisely in preventing both zero-mean Gaussian noise and speckle noise. It is based on the difference image $X_d = |X_2 - X_1|$ between bi-temporal images acquired at the same geographical area but at a different time. Then, the difference image X_d is segmented into 5×5 non-overlapping blocks so that the PCA algorithm can be applied to extract eigenvectors. The feature vectors for each pixel at a spatial location will be projected into a feature vector space [26]. Therefore, spectral difference images based on the PCA method can be calculated using the following formula:

$$X_{PCA} = e^T X_d - m \tag{6}$$

where e^T is the eigenvector of the covariance matrix, m is the mean vector.

After obtaining three spectral change magnitude images (CMI), the adaptive threshold determination algorithm is utilized to generate three spectral change maps:

$$X' = \begin{cases} 255, if\ CMI \geq (\mu + T \cdot \sigma) \\ \\ 0,\ otherwise \end{cases} \tag{7}$$

where μ and σ are the mean and standard deviation of CMI, and T is the threshold parameter to be established. When the discriminant criteria are met, the pixel is determined as the changed pixel and assigned with a value of 255, otherwise it will be assigned with a value of 0.

On the basis of ensuring the detection accuracy and effectively reducing the noise interference, the majority voting fusion can be adopted to remove the discrete points in the initial results and obtain more accurate spectral change results. For a certain pixel (i, j) in the same position of the three initial detection images, if the value of two or three of the three initial results $X'_{IRMAD}(i,j)$, $X'_{ISFA}(i,j)$, and $X'_{PCA}(i,j)$ is 255, this pixel is considered to be the changed pixel after fusion. Otherwise, it is considered an unchanged pixel. Based on this decision fusion strategy, the final spectral change result is acquired.

2.1.2. Co-Saliency Change Information

If the ground objects in the bi-temporal remote sensing images do not change, the co-saliency map of two images can be considered to be the same [63]. This algorithm fully considers the contrast information, spatial information, and correlation information. The advantage of using the co-saliency detection method is that it can extract the most likely changed areas and make up for the inadequacy of change detection based on image features. The specific steps of the cluster-based co-saliency method are as follows:

1. The bi-temporal images X_1 and X_2 are divided into k clusters by the K-means method.
2. The contrast cues $\varphi^c(k)$ and the spatial cues $\varphi^s(k)$ of each cluster are calculated as follows:

$$\varphi^c(k) = \sum_{i=1, i \neq k}^{K} \left(\frac{n^i}{N} \|\mu^k - \mu^i\|_2 \right) \tag{8}$$

$$\varphi^s(k) = \frac{1}{n^k} \sum_{j=1}^{2} \sum_{i=1}^{N_j} \left\{ \Upsilon \left(\|z_i^j - o^j\|^2 \Big| 0, \sigma^2 \right) \cdot \delta \left[b \left(p_i^j \right) - C^k \right] \right\} \tag{9}$$

where n^* is the number of pixels in the cluster C^*; N represents the total number of pixels; $\|\cdot\|_2$ is utilized to calculate the feature space; μ^* means the cluster centers of C^*; N_j is the image lattice of image X_j; $\Upsilon(\cdot)$ can be used to calculate the Euclidean distance between the pixel z_i^j and the image center o^j; σ^2 is the normalized radius of the input image; $\delta(\cdot)$ indicates the Kronecker delta function; p_i^j is the pixel i in the input image X_j; $b(\cdot)$ refers to the clustering index.

3. The following formula is used to fuse the contrast cue and the spatial cue:

$$P\left(C^k\right) = \prod \varphi_i(k) \tag{10}$$

4. The co-saliency map of two temporal images can be obtained through the following formula:

$$S_j = \sum_{k=1}^{K} P\left(x\Big|C^k\right) P\left(C^k\right) \tag{11}$$

After co-saliency detection, the direct difference method is employed for S_1 and S_2 to generate the change magnitude image, and the final co-saliency change information is obtained by the aforementioned adaptive threshold determination algorithm.

2.1.3. Spatial Change Information

For high-resolution remote sensing images, different factors, such as solar altitude angles, sensors, and imaging time will lead to the nonlinear characteristics of the surface object's spectral mixture, resulting in errors in the change detection results. Therefore, it is impossible to obtain comprehensive detection results by only considering the spectral information of land objects. In addition, for buildings, structural and textural features play a key role in determining whether they changed or not. To make use of all the spatial information contained in the high-resolution remote sensing images, multiple spatial features (HOG, GLCM, and RGF) are exploited to improve the accuracy of change detection.

The HOG spatial feature is often used to extract the structural contour information from remote sensing images. Its basic principle is to describe the contour features of the local targets in the image to be measured completely using the gradient or edge directional density distribution [30]. First, the amplitude and direction of each pixel gradient in the input image are calculated to obtain texture and shape information and reduce the interference of illumination. Second, the gradient histograms of all pixels are put through cumulative projection and normalization processing to form a HOG feature vector. Therefore, the structural features of T1 and T2 remote sensing images are extracted, respectively, according to:

$$G_x(x,y) = f(x+1,y) - f(x-1,y) \tag{12}$$

$$G_y(x,y) = f(x,y+1) - f(x,y-1) \tag{13}$$

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \tag{14}$$

$$\alpha(x,y) = arctan\big(G_x(x,y)/G_y(x,y)\big) \tag{15}$$

where, $G_x(x,y)$ and $G_y(x,y)$ are the horizontal and vertical gradient values of the input image at pixels (x,y), respectively. $G(x,y)$ represents the gradient magnitude, and $\alpha(x,y)$ is the direction of the gradient.

To meet the requirements of different image change detection tasks, texture features are extracted by exploiting six grey-level co-occurrence matrix (GLCM) [27] statistics, namely variance, entropy, contrast, correlation, second moment, and dissimilarity. To comprehensively select the best texture features, this article extracted the texture information of T1 and T2 remote sensing images with pixel window sizes of 3×3, 5×5, and 7×7, respectively, and then obtained 18 GLCM texture features, respectively. In addition, another spatial feature extraction method, namely the rolling guide filter (RGF) [29], is used effectively to obtain the details, texture, and structure information of bi-temporal remote sensing images. In the end, 22 spatial features are acquired.

After extracting spatial features, the multi-feature sets of bi-temporal high-resolution images are constructed by direct superposition combination, that is, the feature vectors of a 22×1 dimension are combined. After the normalization of the feature sets, principal component transformation is performed, respectively. It is mapped to linearly independent variables in k dimensions, that is, the k-dimensional variables obtained after transformation are irrelevant when projected onto an orthogonal basis. In this paper, k is set to 3, and the first three principal components retain the most abundant spatial feature information from the original images. After optimal feature selection, the difference feature map is constructed, and the adaptive threshold algorithm above is employed for binary classification in order to obtain the spatial change result.

2.2. Multi-Scale Segmentation

The aim of image segmentation is to divide the image into different regions with special semantics according to a certain similarity criterion in order to separate the target of interest in the complex background. The multi-scale segmentation method not only obtains the segmentation results rapidly and effectively, but also maintains the object boundary

by considering the spatial structure information of high-resolution remote sensing images. The bi-temporal images are first processed by superposition, and then the first principal component (PC1) image of the superimposed image is segmented by the fractal evolution net approach (FENA) [29,54]. For multi-scale segmentation, in object-based remote sensing information extraction, the optimal segmentation scale of each object is relative. The ideal result of image segmentation is that the object obtained by segmentation exhibits good homogeneity inside and good heterogeneity between adjacent objects. Three parameters (shape, compactness, and scale) are covered in multi-scale segmentation to control the quality of the objects. In particular, the idea of optimal scale determination is mainly to control the shape and compactness in order to find the optimal scale. When the segmentation scale matches the real objects, the heterogeneity between different objects will reach the maximum, resulting in the maximum value of the local variance measurement [64].

2.3. Decision Fusion Using Fuzzy Integral

Considering that any single change detection operator has limitations, multiple pieces of change information can be integrated on the basis of ensuring detection accuracy and effectively reducing the interference of noise. On the other hand, the advantages of different change information detection methods are used comprehensively to improve the capacity to detect different objects. The fuzzy integral decision theory is expressed through a fuzzy measure [5], in the form of expert decisions and evaluations of the performance of different detection results.

In the change detection process, $Z = \{Z_1, \cdots, Z_n\}$ constitutes the set of initial change results. $h_k(Z_n)$ represents the classification results of Z_n in class $K \in \{0,1\}$, and $g_k(Z_n)$ is the performances of Z_n in class $K \in \{0,1\}$. Where, 0 represents the unchanged class and 1 represents the changed class. A set function $g : 2^Z \to [0,1]$ represents a fuzzy measure if it satisfies several properties [5,65]:

(1) $g(\varnothing) = 0$;
(2) $g(Z) = 1$;
(3) $g(Z_m) \leq g(Z_n)$ if $Z_m \subset Z_n$.

The fuzzy measure represents the degree of interaction between two elements according to a λ parameter.

$$g(Z_m \cup Z_n) = g(Z_m) + g(Z_n) + \lambda \cdot g(Z_m) \cdot g(Z_n) \tag{16}$$

For the object O_i, if $h_k^i(Z_1) \geq \cdots \geq h_k^i(Z_n) \geq 0$ is satisfied, the set of initial change results are then rearranged, and the fuzzy measures of the new order $A_m = \{Z_1, \cdots, Z_m\}$ are constructed according to:

$$g_k^i(A_1) = g_k^i(Z_1) \tag{17}$$

$$g_k^i(A_m) = g_k^i(A_{m-1} \cup Z_m) = g_k^i(A_{m-1}) + g_k^i(Z_m) + \lambda \cdot g_k^i(A_{m-1}) \cdot g_k^i(Z_m) \tag{18}$$

where, λ means the unique root of an $n-1$ degree equation that verifies $\lambda \in [-1, +\infty]$ and $\lambda \neq 0$. It is confirmed by solving the following formula:

$$\lambda + 1 = \prod_{m=1}^{n} \left(1 + \lambda \cdot g_k^i(Z_m) \right) \tag{19}$$

For the object O_i in each class $K \in \{0,1\}$, the fuzzy integral is calculated as:

$$FI_K^i = \sum_{m=1}^{n} \text{Max} \left[\text{Min} \left(h_k^i(Z_m), g_k^i(A_m) \right) \right] \tag{20}$$

After obtaining the fuzzy integral of the object O_i with respect to the changed and unchanged class, the object change probability can be obtained by the formula:

$$results = \begin{cases} 255, & FI_1^i > FI_0^i \\ 0, & FI_1^i \leq FI_0^i \end{cases} \tag{21}$$

Therefore, based on the above FI decision fusion process, the change probability of each object in the image can be determined to obtain the final object-oriented change detection results. The advantage of different initial pixel-level change information is integrated under the restriction of the segmentation objects to acquire complete land cover change.

2.4. Assessment of the Change Detection Processes

In terms of accuracy evaluation, five indices related to experimental results and ground truths were adopted, including the missed detection rate (MR), false alarm rate (FAR), overall accuracy (OA), kappa coefficient and F1 score. Table 1 shows the common quantitative evaluation index: TP, FP, FN, and TN [23,32,46].

Table 1. Quantitative evaluation index of change detection.

True Detection	Change	Unchange
Change	TP	FP
Unchange	FN	TN
Total	N_c	N_u

Specifically, the MR refers to the probability that actually changed pixels are detected as invariant pixels, which is defined as:

$$MR = FN/N_c \tag{22}$$

where FN is the number of unchanged pixels in the change detection result that were classified as the changed class in the ground truth image, and N_c is the total number of changed pixels counted in the ground truth image.

The FAR is the probability that actually unchanged pixels are detected as changed pixels, which is indicated as

$$FAR = FP/N_u \tag{23}$$

where FP is the number of changed pixels in the change detection result that were classified as the unchanged class in the ground truth image, and N_u is the total number of unchanged pixels counted in the ground truth image.

The OA represents the probability of the correctly detected part, which is calculated as:

$$OA = (TP + TN)/(N_c + N_u) \tag{24}$$

where TP is the number of changed pixels both in the change detection result and the ground truth image, and TN is opposite.

The kappa coefficient is utilized to reflect the consistency of the classification results and the reference change map, which is defined as:

$$kappa = (OA - P_e)/(1 - P_e) \tag{25}$$

$$P_e = ((TP + FP)N_c + (FN + TN)N_u)/(N_C + N_u)^2 \tag{26}$$

The F1 score is an indicator that comprehensively considers the detection rate p and the recall rate r. It is calculated as:

$$F1\ score = 2pr/(p+r) \tag{27}$$

$$p = TP/(TP+FP) \tag{28}$$

$$r = TP/(TP+FN) \tag{29}$$

3. Experimental Datasets and Experimental Configuration
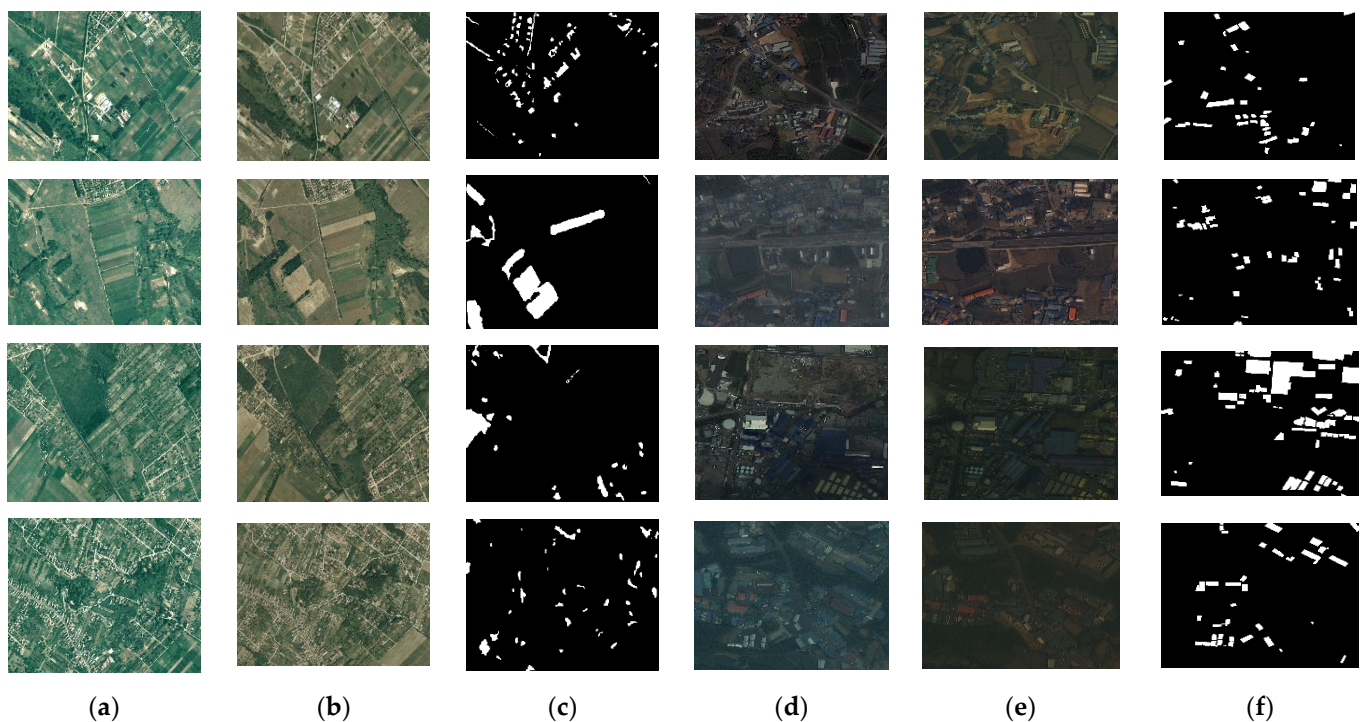
3.1. Experimental Datasets

Three datasets with different spatial resolutions, different sensors, and different ground objects were selected for change detection experiments to verify the comprehensive detection performance of the proposed algorithm. As shown in Figure 2, each dataset includes two bi-temporal remote sensing images and a reference change map. The reference change map was manually delineated by specialists according to visual interpretation, while the changed regions were marked in white. All datasets were pre-processed with image registration and radiometric correction.

DS1 (a) (b) (c)

DS2 (d) (e) (f)

DS3 (g) (h) (i)
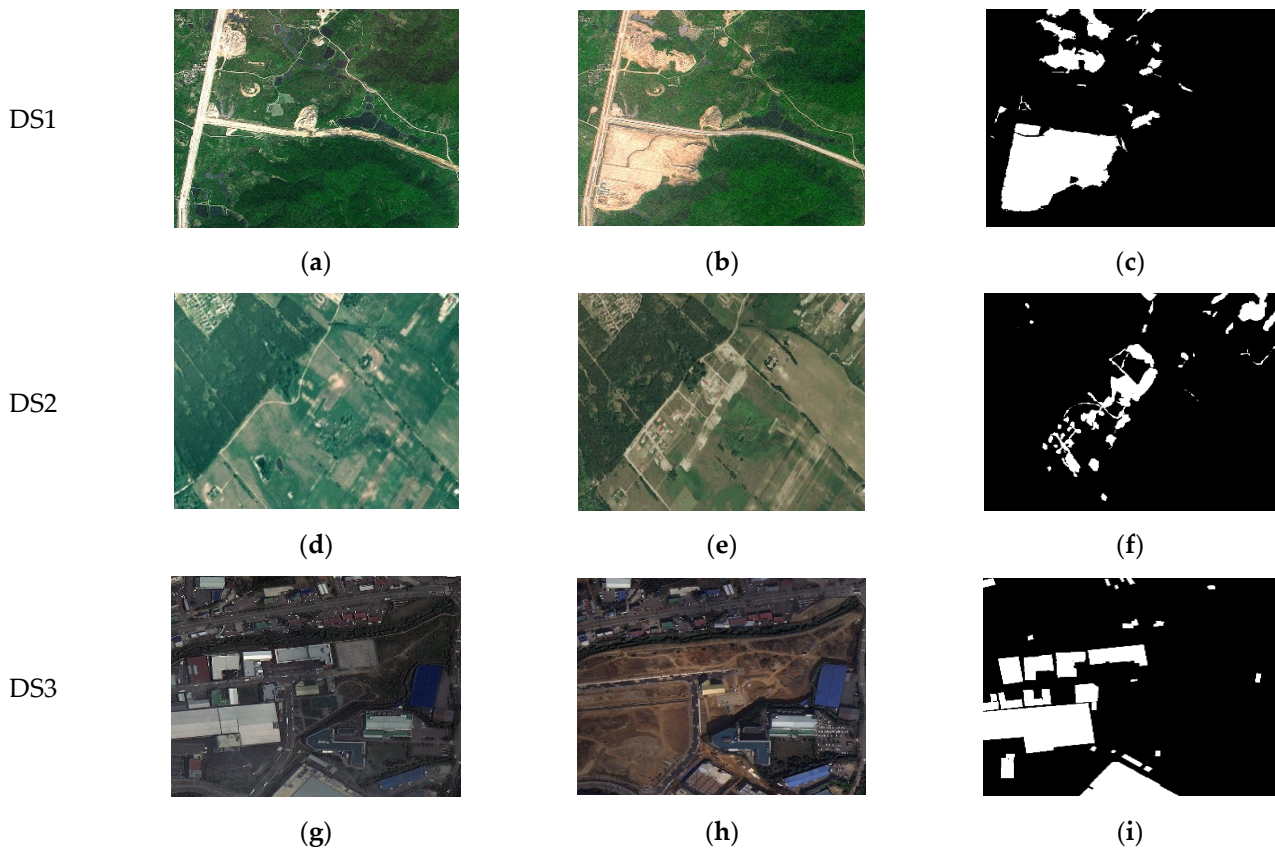
Figure 2. Experimental datasets and ground truth images for change detection: (**a**) SPOT images acquired in 2006 with a spatial resolution of 2.5 m; (**b**) SPOT images acquired in 2007 with a spatial resolution of 2.5 m; (**c**) reference change map of DS1. (**d**) aerial images acquired in 2000 with a spatial resolution of 1.5 m; (**e**) aerial images acquired in 2005 with a spatial resolution of 1.5 m; (**f**) reference change map of DS2. (**g**) side-view satellite images acquired in 2017 with a spatial resolution of 0.5~0.8 m; (**h**) side-view satellite images acquired in 2020 with a spatial resolution of 0.5~0.8 m; and (**i**) reference change map of DS3.

The first dataset (DS1) is formed by images captured by the SPOT satellite in 2006 and 2007, with a size of 877 × 738 pixels and a spatial resolution of 2.5 m. The main change

categories are bare land and vegetation, with strong representation in the wild. The second dataset (DS2) is the SZADA dataset [66], which represents a section (952 × 640 pixels) of optical aerial images with a spatial resolution of 1.5 m, taken in 2000 and 2005, respectively. The change information involved different land-use types, such as buildings, bare land, roads, and vegetation. The third dataset (DS3) is the side-looking dataset [67], which was derived from side-view satellite remote sensing images of rural areas captured at different off-nadir angles in 2017 and 2020. The image size is 1024 × 1024 pixels, and the spatial resolution is 0.5~0.8m. The change information consists of large areas with newly constructed and demolished buildings. For a clearer explanation, some indicative images of the public SZADA dataset and side-looking dataset are shown in Figure 3.

(a)   (b)   (c)   (d)   (e)   (f)

Figure 3. Indicative images from (**a**–**c**) the SZADA dataset and (**d**–**f**) the side-looking dataset: (**a**,**d**) image at time t1; (**b**,**e**) image at time t2; and (**c**,**f**) reference change map.

3.2. Experimental Configuration

Three experiments were designed to evaluate the effectiveness of the proposed algorithm. The first experiment is used to verify the performance of multi-method fusion by comparing the proposed algorithm with the other three methods. The second experiment aims to test the superiority of the object-oriented method compared to five pixel-level change detection methods, including IRMAD, ISFA, PCA, CVA-SIFCM [63], and DI-Kmeans [13]. Finally, to demonstrate the superiority of fusing multi-feature and multi-information for change detection, different object-level spectral–spatial feature combination methods were compared in the third experiment, including: spectral feature alone, spectral and texture fusion, spectral and structure fusion, and the proposed method of fusing multiple pieces of change information.

In the three experiments, the relevant parameters were set as follows: In the IRMAD and ISFA iteration, the maximum number of iterations was 100 and the convergence threshold was 10^{-5}. When extracting HOG features, considering the actual situation and applicability, it is more appropriate to set the pixel numbers of each cell unit as 10, 15, and 13, respectively, in DS1, DS2, and DS3. As for the adaptive threshold algorithm, the parameter T was set to 2, 1.5, and 2.5, respectively, in the three datasets. For multi-scale segmentation, the shape and compactness were fixed at 0.3 and 0.5. Furthermore, the

optimal scale determination strategy proved that desirable results were obtained when the segmentation scales were 87, 73, and 139, respectively, in three datasets.

4. Results

4.1. First Experiment Results

As can be seen in Figure 4, the change maps acquired by co-saliency, spectral feature, and spatial features had a good detection effect on DS1, DS2, and DS3, but there was still several obvious "salt and pepper" noises and false detections, especially for buildings. It is worth noting that areas most likely to change were extracted by co-saliency detection, which continues to compensate for the shortcomings of spectral-spatial change information. The spectral change results illustrate that the strategy integrating three independent methods through the MV decision reduced noise interference and retained the real spectral variation areas. The spatial change maps had a better effect on edge detection of changed areas, but were worse at anti-noise performance. It can be found that the proposed method had significant advantages in terms of visual effects. The object-oriented results after fuzzy integral fusion not only effectively employed the advantages of different feature extraction methods, but also decreased false alarms and missed detections. We observe that the proposed approach had the potential power to detect correctly all the kinds of changes in three datasets.
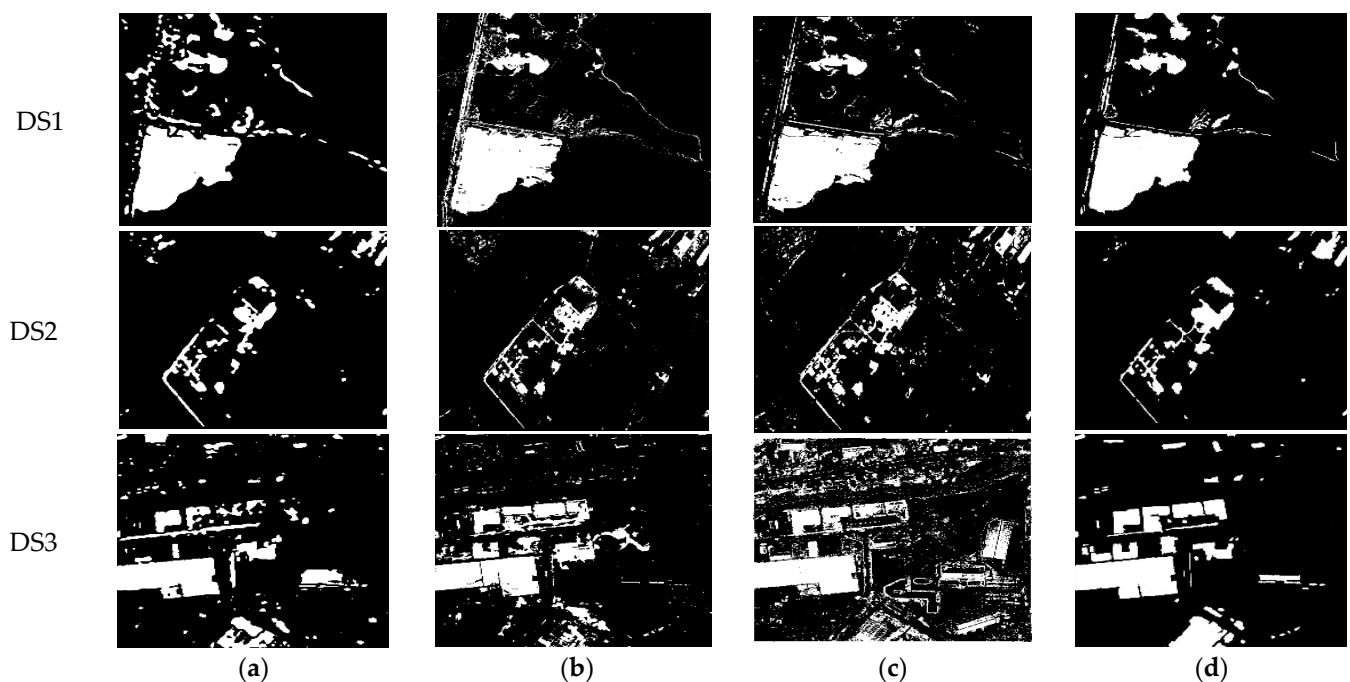
Figure 4. Change detection results in the first experiment: (**a**) co-saliency change; (**b**) spectral change; (**c**) spatial change; and (**d**) proposed method.

Table 2 shows the quantitative evaluation of the results achieved previously. In particular, higher accuracy can be reached by the proposed approach, with 2.424% to 13.902% increase in overall accuracy, 0.074 to 0.362 increase in the kappa coefficient and the F1 score, while 0.012 to 0.203 decrease in missed detection rate and false alarm rate. This result not only proves the high performance of the proposed method, but also certifies the effectiveness of the Fuzzy Integral decision fusion theory in the transition from pixel-level detection to object-level recognition. It is not difficult to observe that almost all the changed areas could be obtained, the whole structure was more complete, and the speckle noise was significantly reduced. The comprehensive application of combining multiple change information could realize the complementary advantages among features, which was helpful in obtaining more robust change results and the highest detection accuracy.

Therefore, it can be seen that the overall accuracy of the proposed framework was greater than 95% and the kappa coefficient, as well as the F1 score were higher than 0.78 in the three datasets.
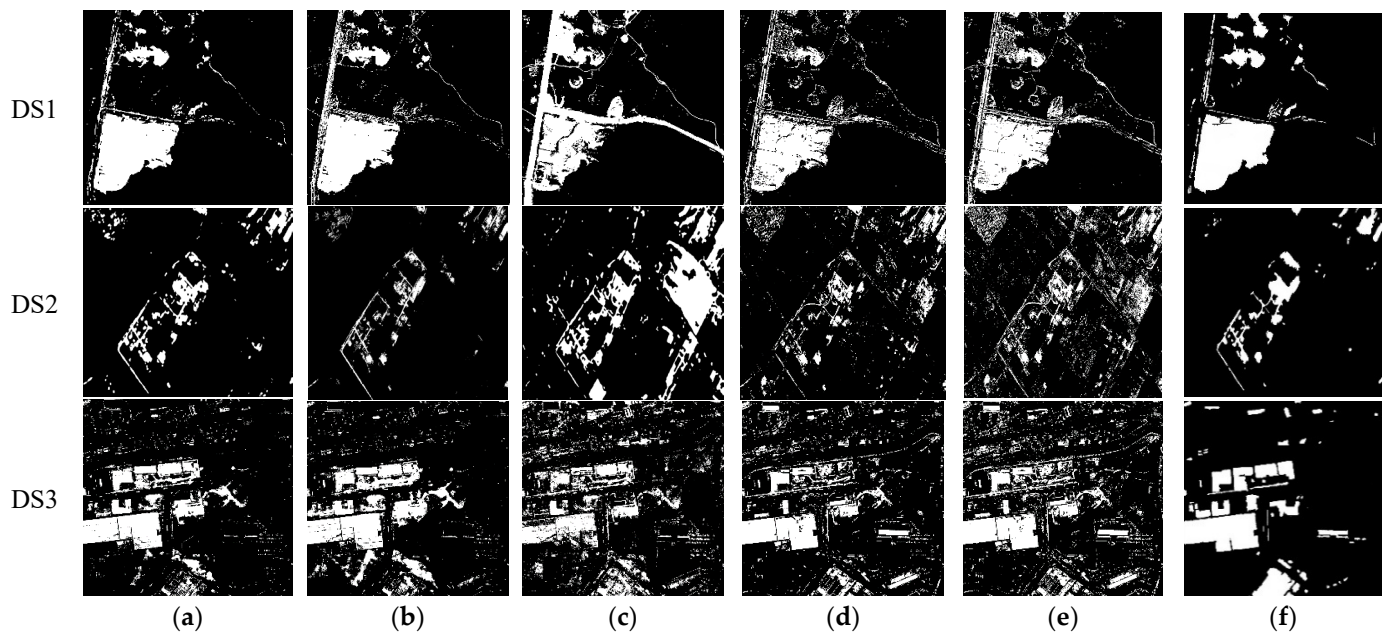
Table 2. Quantitative evaluation of change detection results in the first experiment.

Dataset	Method	Quantitative Evaluation Index				
		MR	FAR	OA (%)	Kappa	F1 Score
DS1	Co-saliency change	0.190	0.042	93.629	0.754	0.792
	Spectral change	0.203	0.040	93.531	0.748	0.787
	Spatial change	0.208	0.023	94.892	0.793	0.823
	Proposed method	**0.129**	**0.011**	**97.316**	**0.881**	**0.897**
DS2	Co-saliency change	0.337	0.024	95.775	0.622	0.644
	Spectral change	0.340	0.024	95.729	0.618	0.641
	Spatial change	0.288	0.035	95.054	0.598	0.624
	Proposed method	**0.261**	**0.005**	**98.479**	**0.781**	**0.789**
DS3	Co-saliency change	0.441	0.061	87.667	0.524	0.596
	Spectral change	0.396	0.057	88.764	0.570	0.637
	Spatial change	0.327	0.158	81.366	0.429	0.541
	Proposed method	**0.238**	**0.016**	**95.268**	**0.791**	**0.818**

4.2. Second Experiment Results

Figure 5 represents the change detection results for the five pixel-level methods and the proposed method. As can be seen from the detection results, the IRMAD method had a good detection effect, but is still affected by noise. The performance of the ISFA algorithm was similar to the IRMAD achievement, due to the iterative weighting strategy. However, the ISFA method was revealed to have a high false alarm rate and poor internal integrity of the changed objects. The experimental results of the three datasets proved that the PCA method not only detected relatively complete changed areas, but also produced a large range of false alarms, which seriously affected the accuracy of the change detection. The reason is that after the principal component transformation, the false detection phenomenon caused by illumination conditions or radiation differences cannot be eliminated. Furthermore, the detection effect of CVA-SIFCM and DI-Kmeans was poor; there was a large number of false changed areas and speckle noises, which seriously affected the change detection accuracy.

It can be seen that the edge contour of the changed objects in the detection results of the five pixelwise methods was not clear enough, and there were many holes and gaps inside. Especially for the third dataset, due to the changes of a large number of buildings involved in the images, pixel-level methods had a poor detection effect. Compared with the above methods, the proposed object-oriented approach reflected the real changes better, reduced noise interference, and obtained accurate changed objects with clear edges. The reason is that the proposed method can integrate the advantages of all initial change information and transform change detection from the pixel level to the object level in order to obtain comprehensive change results from high-resolution remote sensing images. The results of the accuracy evaluation (Table 3) correspond to the visual effects. It can be observed that the proposed approach achieved the highest overall accuracy, the kappa coefficient and the F1 score, the lowest missed detection rate, and the false alarm rate, which showed that the object-level fusion strategy was helpful in identifying changed and unchanged land cover areas.

Figure 5. Change detection results in the second experiment: (**a**) IRMAD; (**b**) ISFA; (**c**) PCA; (**d**) CVA-SIFCM; (**e**) DI-Kmeans; and (**f**) proposed method.

Table 3. Quantitative evaluation of change detection results in the second experiment.

Dataset	Method	Quantitative Evaluation Index				
		MR	FAR	OA (%)	Kappa	F1 Score
DS1	IRMAD	0.181	0.028	94.850	0.796	0.826
	ISFA	0.196	0.051	92.700	0.723	0.767
	PCA	0.349	0.120	84.546	0.466	0.557
	CVA-SIFCM	0.286	0.054	91.149	0.654	0.707
	DI-Kmeans	0.241	0.069	90.528	0.649	0.705
	Proposed method	**0.129**	**0.011**	**97.316**	**0.881**	**0.897**
DS2	IRMAD	0.328	0.031	95.212	0.593	0.619
	ISFA	0.465	0.027	94.778	0.514	0.542
	PCA	0.290	0.147	84.439	0.382	0.345
	CVA-SIFCM	0.464	0.056	92.023	0.396	0.437
	DI-Kmeans	0.414	0.078	90.230	0.361	0.409
	Proposed method	**0.261**	**0.005**	**98.479**	**0.781**	**0.789**
DS3	IRMAD	0.398	0.064	88.127	0.552	0.623
	ISFA	0.383	0.089	86.290	0.512	0.594
	PCA	0.434	0.103	84.334	0.447	0.541
	CVA-SIFCM	0.460	0.092	84.756	0.444	0.536
	DI-Kmeans	0.433	0.122	82.741	0.413	0.517
	Proposed method	**0.238**	**0.016**	**95.268**	**0.791**	**0.818**

4.3. Third Experiment Results

In terms of object-level methods, based on the optimization of spectral features, the influence of different combinations of feature factors on the precision of change detection was analyzed. Details can be found in Figure 6 and Table 4, where the different feature combination methods with different effects on the performance of change detection are illustrated. However, the best accuracy was obtained by the proposed approach. Furthermore, the accuracy of change detection was improved by combining texture or structure feature alone, while the results were significantly improved by combining simultaneously spatial features and co-saliency detection results. The main reason is that the spatial features help to distinguish the buildings from other impermeable surfaces and prevent false

alarm detection. Co-saliency detection was helpful in optimizing the feature extraction results and increasing the reliability of the final results. The method utilizing the spectral feature alone had the most obvious salt and pepper noise, resulting in the lowest detection accuracy. Combined with the spectral difference and GLCM texture methods, the spectral information extracted by the difference method was highly mixed with false changes, resulting in missed and error detections. However, employing texture features can promote the detection of buildings and roads, so the overall accuracy, kappa coefficient, and F1 score of change detection results in DS3 were significantly improved. After combining spectral and structure features, the changed objects with clear boundaries can be detected, but several false detection regions should not be neglected.
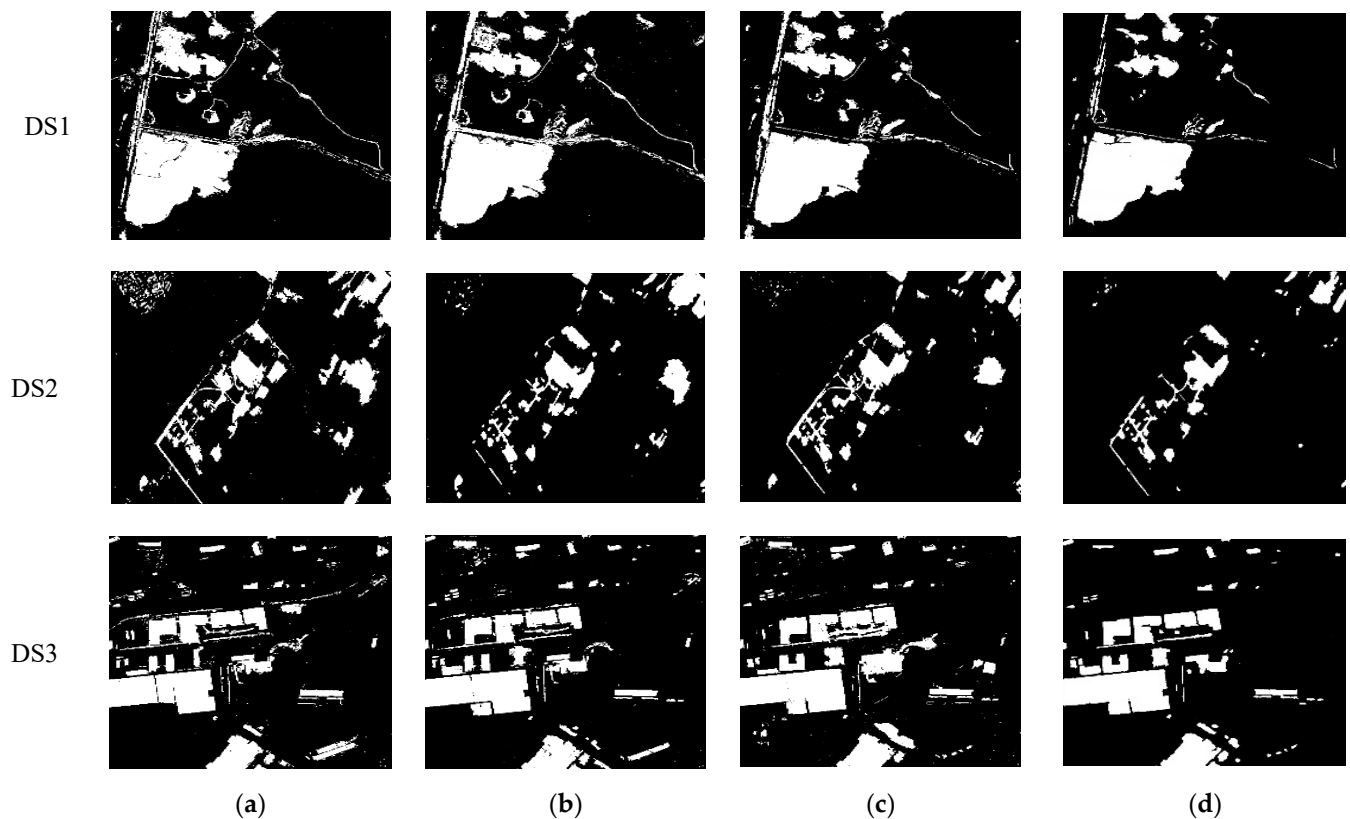
Figure 6. Change detection results in the third experiment: (**a**) spectra; (**b**) spectra + texture; (**c**) spectra + structure; and (**d**) proposed method.

Compared with methods that applied the spectral feature alone, the spatial feature coupled with the spectral feature acquired the clear boundaries and complete changes of the land cover objects. From the indicators of accuracy evaluation, it can be seen that the OA, Kappa coefficient, and FAR were improved significantly, i.e., with a 0.077% to 5.525% increase in overall accuracy, 0.001 to 0.237 increase in kappa coefficient, and a 0.009 to 0.050 decrease in false alarm rate. In addition, visually compared with detection results based on spectra and texture, as well as on spectra and structure, the use of co-saliency detection and fuzzy integral fusion can effectively avoid more noise interference and false alarms. To be specific, higher accuracy can be obtained by the proposed approach, with a 2.287 % to 4.774% increase in overall accuracy, 0.077 to 0.244 increase in kappa coefficient, and F1 score, as well as a 0.026 to 0.059 decrease in false alarm rate.

Table 4. Quantitative evaluation of change detection results in the third experiment.

Dataset	Method	Quantitative Evaluation Index				
		MR	FAR	OA (%)	Kappa	F1 Score
DS1	Spectra	0.128	0.065	92.564	0.733	0.778
	Spectra + texture	**0.110**	0.070	92.641	0.734	0.779
	Spectra + structure	0.112	0.048	94.203	0.786	0.820
	Proposed method	0.129	**0.011**	**97.316**	**0.881**	**0.897**
DS2	Spectra	0.292	0.081	90.667	0.423	0.467
	Spectra + texture	0.280	0.050	93.705	0.537	0.569
	Spectra + structure	**0.175**	0.031	96.192	0.660	0.680
	Proposed method	0.261	**0.005**	**98.479**	**0.781**	**0.789**
DS3	Spectra	0.323	0.058	89.865	0.625	0.685
	Spectra + texture	0.274	0.049	91.463	0.684	0.735
	Spectra + structure	0.259	0.066	90.299	0.655	0.714
	Proposed method	**0.238**	**0.016**	**95.268**	**0.791**	**0.818**

5. Discussion

The results of the accuracy evaluation show that the overall accuracy of the proposed method was above 95% and the kappa coefficient and the F1 score were the highest in the three datasets. Furthermore, the accuracy evaluation results were also consistent with the visual interpretation analysis. Through the change detection experiments of three datasets of different sensors and different resolutions, it is observed that the proposed algorithm can effectively integrate the advantages of multiple features, and retain more accurate land cover change information. To accelerate the application and robustness of the proposed framework in practical problems, this section discusses the major achievements of this research.

First, scholars noticed and applied the idea of employing multiple pieces of information to improve the accuracy of change detection in the past ten years. Three fusion levels, i.e., data-level, feature-level, and decision-level, were discussed in some literature. However, there was no deterministic strategy to find the most appropriate method to implement the change detection process. Du et al. [5] found that the fusion of feature levels and decision-level fusion led to an increase in overall accuracy, which is consistent with our research. In addition, feature-level fusion can effectively reduce omission errors, and decision-level fusion is good at restraining commission errors [65]. Different fusion strategies are still necessary to find the appropriate algorithm for the detection of heterogeneous situations, such as building extraction during the urban expansion process, and to improve the accuracy of the change detection results.

In the first experiment, there were differences in the change detection results of different types of regions, such as wild, rural, and urban. It is observed that the change detection results in rural areas are closest to the reference change map, with the highest overall accuracy and the lowest false detection rate. In particular, with a 1.163% to 3.211% increase in overall accuracy and a 0.006 to 0.011 decrease in false alarm rate. False detections occurred on country trails in wild areas due to the difference in illumination conditions, resulting in an increased false alarm rate. However, other changes in wild regions can be detected and the boundary of ground objects was relatively complete, resulting in the missed detection rate being at its lowest, and the kappa coefficient, as well as the F1 score, being higher than 0.88. In urban areas, false detection and missed detection existed at the same time, and changes in several small buildings were ignored. Specifically, lower accuracy can be reached in urban area detection results, with a 2.048% to 3.211% decrease in overall accuracy and a 0.005 to 0.109 increase in missed detection rate and false alarm rate. In terms of high-resolution remote sensing images, the spectral characteristics of ground objects can reflect the rich information of land cover, the texture features can reflect the relationship between neighborhood pixels, and the attributes of structural

features to identify buildings and roads. Therefore, the spatial feature and the spectral feature complement each other and jointly reveal the information on land cover from remote sensing images. In the third experiment, compared to the raw spectral feature, the addition of spatial information, such as texture and structure features, can eliminate the phenomenon of salt and pepper noise and obtain accurate changed land objects, resulting in an improvement in OA, the kappa coefficient, and the F1 score. Furthermore, to overcome the insufficiency of spectral–spatial feature extraction methods, the co-saliency detection algorithm that considers contrast cues, spatial cues, and correlation cues, plays an important role in optimizing the feature extraction results and improving the accuracy of the final detection results. The comprehensive use of multi-feature and multiple pieces of change information showed extraordinary advantages in the application of three remote sensing images with different resolutions and different changes.

As proven in the previous sections, the proposed method achieved the best change detection accuracies for high-resolution remote sensing images. However, the detection performance of remote sensing images with different resolutions should also be considered in this paper. The experimental results of the SPOT images (DS1) show that when relatively complete changed areas were detected, some false detection areas were also generated, which affected the accuracy of the change detection. Aerial images with a spatial resolution of 1.5 m (DS2) presented less noise than the others and had the best visual interpretation effect of change detection. However, when the resolution of the images was further improved, due to the influence of shadows and spectral differences, there were several omissions and errors in the detection results of DS3. Compared to the reference change map, the unchanged buildings were incorrectly detected and the internal compactness of the buildings was not high. On the contrary, the change information in DS1 and DS2 can be obtained correctly. It can be seen from the quantitative evaluation that the overall accuracy of DS2 was the highest, and the false detection rate was the lowest.

The scale of multi-scale segmentation has an important influence on the result of object-level recognition. Taking a small area in experimental dataset 1 as an example, three segmentation scales (58, 87, and 124) were used to perform a comparative analysis on the influence of different scale parameters on the recognition of changed objects, as shown in Figure 7.

False alarms were detected for small segmentation scales, that is, non-changed ground objects were wrongly identified as changed objects, as shown in the green box in Figure 7. The reason is that a segmentation scale that is too small leads to fragmentation of ground object segmentation. However, due to the high false alarm rate of pixel-level change detection in high-resolution remote sensing images and the high proportion of changed pixels in the segmented objects, the use of segmented objects to screen changed land cover objects is ineffective. When comparing the performance of different methods, a segmentation scale that is too large can easily lead to a missed detection, that is, the changed ground objects were not correctly identified, as shown in the red box in Figure 7. The main reason is that the large segmentation scale leads to an over-segmentation of the surface objects, so those with a relatively small area and similar spectral characteristics to the neighboring ground objects are merged into the adjacent objects, reducing the pixel proportion of the sub-target level changed objects. The proposed approach takes advantage of the optimal scale estimation strategy to select the appropriate segmentation parameter, which provides a good basis for the final results of object-level change detection.
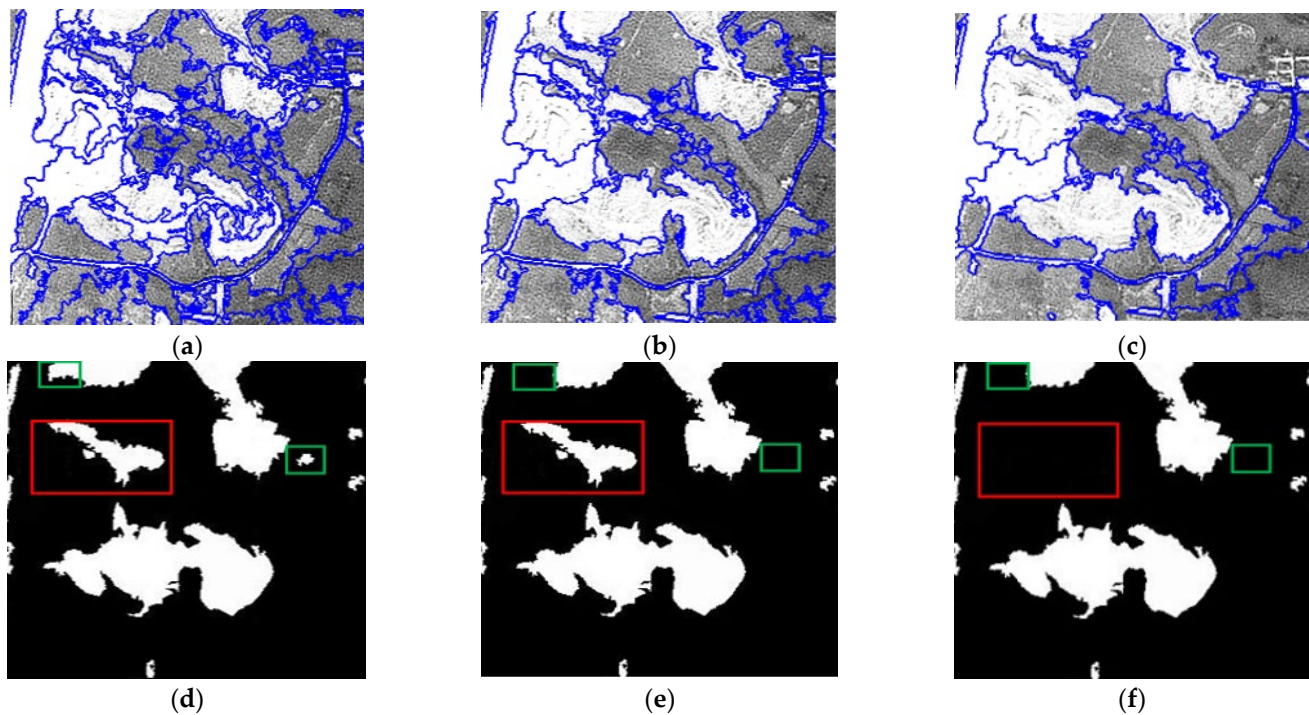
Figure 7. Comparison of recognition results at different segmentation scales: (**a**) image segmentation when scale = 58; (**b**) image segmentation when scale = 87; (**c**) image segmentation when scale = 124; (**d**) changed objects recognition when scale = 58; (**e**) changed objects recognition when scale = 87; and (**f**) changed objects recognition when scale = 124. Green box: false detection; Red box: missed detection.

Regarding the change detection post-processing, the proposed fusion procedure provides a new idea; that is, multi-scale segmentation of the first principal component superimposed image is selected to combine the initial pixel-level change information and generate the final object-oriented change detection map. To verify the advantage of the proposed fusion procedure, morphology post-processing was carried out on the images fused with multiple pieces of change information in the second experiment and the images only using spectral features in the third experiment. In the morphological processing, the opening and closing convolution kernels are set to 7×7 pixels and 5×5 pixels.

In general, post-processing of change detection eliminates or reduces the interference of "noise detection" by utilizing relevant knowledge of mathematical morphology. From the basic erosion and dilate tool of mathematical morphology, Figure 8 illustrates that this type of post-processing method will destroy the boundary of the actual ground object while removing the noise. To optimize the effect of pixel-based change detection results, the fuzzy integral strategy, under the restriction of multi-scale segmentation, was then applied. As shown in Figure 8, based on the advantages of multiple methods of extracting initial information, the proposed decision fusion procedure can remove the interference of "salt and pepper noise" and maintain the internal and boundary integrity of the actual ground objects. Furthermore, under the condition of morphological post-processing, the effect of multi-change information fusion was better than that of spectral features alone, which also proved the advantages of spectral–spatial–saliency change information fusion.
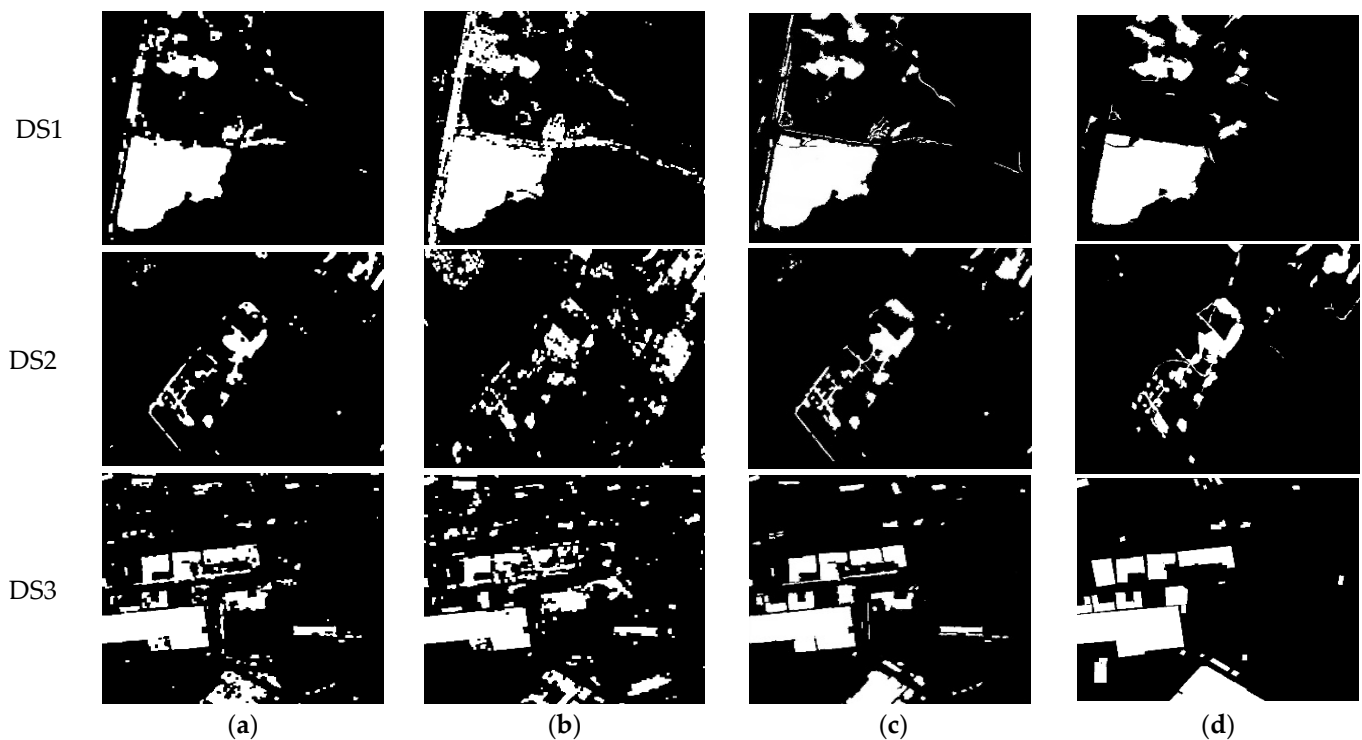
Figure 8. Comparison of post-processing methods: (**a**) multi-feature + morphology operation; (**b**) spectra + morphology operation; (**c**) proposed method; and (**d**) reference change map.

The quantitative evaluation results in Table 5 were consistent with the visual interpretation in Figure 8. Specifically, the detection accuracies of the proposed method achieved the best accuracies in terms of MR, FAR, OA, kappa coefficient, and F1 score. It can be seen from Table 5 that compared to the morphology operation methods, there is a 1.933 % to 9.095% increase in overall accuracy and a 0.008 to 0.185 decrease in the false alarm rate and missed detection rate, further supporting the effectiveness and feasibility of the proposed post-processing framework. In summary, in the actual change detection application, pixel-based and object-based change detection processes can be organically combined according to different detection purposes. Therefore, the final change detection results not only correspond to the meaningful geographic entities but also effectively integrate the advantages of both strategies to obtain the best detection accuracy.

Table 5. Quantitative evaluation of post-processing methods.

Dataset	Method	Quantitative Evaluation Index				
		MR	FAR	OA (%)	Kappa	F1 Score
DS1	Multi-feature + morphology operation	0.170	0.034	94.558	0.788	0.820
	Spectra + morphology operation	0.170	0.084	90.349	0.663	0.720
	Proposed method	**0.129**	**0.011**	**97.316**	**0.881**	**0.897**
DS2	Multi-feature + morphology operation	0.379	0.013	96.546	0.657	0.675
	Spectra + morphology operation	0.367	0.076	90.664	0.423	0.439
	Proposed method	**0.261**	**0.005**	**98.479**	**0.781**	**0.789**
DS3	Multi-feature + morphology operation	0.340	0.049	90.324	0.633	0.690
	Spectra + morphology operation	0.423	0.083	86.173	0.494	0.576
	Proposed method	**0.238**	**0.016**	**95.268**	**0.791**	**0.818**

The time complexity of the proposed method was also investigated. Table 6 shows the processing time of different components of the proposed framework. It can be observed that the acquisition of spatial change information demanded more time, as the spatial

feature sets had to be constructed and the optimal features selected. Furthermore, the multi-scale segmentation processing time increased due to the determination of the optimal segmentation scale. DS3 required the longest processing time (Table 6), which may also be related to the larger image size (1024 × 1024 pixels).

Table 6. Running time of different components of the proposed method.

	Saliency Change	Spectral Change	Spatial Change	Multi-Scale Segmentation
DS1	12.1 s	37.2 s	336.0 s	67.2 s
DS2	13.0 s	28.2 s	298.1 s	64.2 s
DS3	13.5 s	55.4 s	463.1 s	76.9 s

The main contributions of the proposed framework are as follows: First of all, it should be noted that the comprehensive use of multiple pieces of change information can overcome the uncertainty of any single method. Unlike other traditional methods that use the raw spectral feature alone, spectral–spatial–saliency change information is employed comprehensively in this paper. The co-saliency detection can supplement the insufficiency of image features, and the advantages of different change maps are integrated to enhance the accuracy of change detection. Second, in the process of extracting spectral feature changes, the idea of integrating multiple spectral change detection methods (IRMAD, ISFA, and PCA) is adopted to overcome the limitation of a single operator and the influence of false alarms, such as salt and pepper noise, as well as obtain the optimal spectral difference information. Third, the combination strategies of both feature-level and decision-level fusion are utilized and verified in this article. It is worth noting that the fuzzy integral decision theory, which can determine the change probability of land objects by integrating the advantages of initial change results, improves the change detection accuracy.

6. Conclusions

Generally, a large amount of salt and pepper noise and the low accuracy of the detection of artificial objects frequently appear in methods based on a single spectral difference. In this paper, an object-level change detection approach was proposed that combines spectral–spatial–saliency change information and the fuzzy integral decision fusion algorithm. By combining three independent change results with the decision analysis strategy, real land cover change information was obtained. The proposed approach not only overcame the salt and pepper noise caused by illumination conditions or radiation differences, but also acquired the whole change object with a distinguishable boundary. The results of the three experiments showed that the proposed method could effectively obtain the changed objects. The overall accuracy of the proposed method was greater than 95%, the false alarm rate was lower than 0.016, and the kappa coefficient, as well as the F1 score, were higher than 0.78 in the three datasets. In addition, the detection accuracy of the proposed method improved significantly compared to other state-of-the-art methods.

The proposed method has the following findings: (1) The fusion of three spectral change detection results can overcome the influence of speckle noise and obtain optimal spectral difference information. (2) Spectral characteristics can reflect the rich land cover information, spatial features can display the domain and spatial relationship between pixels, while co-saliency detection considers contrast, spatial, and correlation information. The joint application of multiple pieces of change information can take advantage of the complementary features, which is useful for more robust change results and improves detection accuracy. (3) The fuzzy integral decision fusion strategy integrates the initial pixel-level results and determines the change probability of objects under the restriction of the multi-scale segmentation, which plays a key role in generating the final results.

However, regarding the change detection effect on buildings, the proposed framework was not satisfactory compared to other land cover objects. The limitation of the proposed

method is the selection of optimized parameters, which will have some influence on the final change detection results. In other respects, the time complexity of the proposed method should be taken into account since the method is composed of different components. Therefore, how to improve the applicability of the proposed algorithm for buildings change detection and effectively select optimal spatial features under appropriate running time for subsequent experiments remains a research topic that should be focused in the future.

Author Contributions: Conceptualization, C.G. and H.D.; methodology, C.G.; software, H.D.; validation, C.G. and H.D.; formal analysis, I.M.; investigation, Y.H.; resources, D.P.; writing—original draft preparation, C.G.; writing—review and editing, H.D. and I.M.; supervision, I.M. and Y.H.; project administration, H.D.; funding acquisition, Y.H. and D.P. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Publicly datasets were used in this study. The SZADA dataset can be found: http://web.eee.sztaki.hu/remotesensing/airchange_benchmark.html (accessed on 22 May 2022), and the side-looking dataset can be found: https://github.com/S2Looking/ (accessed on 22 May 2022).

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

References

1. Grogan, K.; Pflugmacher, D.; Hostert, P.; Kennedy, R.; Fensholt, R. Cross-border forest disturbance and the role of natural rubber in mainland Southeast Asia using annual Landsat time series. *Remote Sens. Environ.* **2015**, *169*, 438–453. [CrossRef]
2. Liu, J.; Li, P.; Wang, X. A new segmentation method for very high resolution imagery using spectral and morphological information. *ISPRS J. Photogramm. Remote Sens.* **2015**, *101*, 145–162. [CrossRef]
3. Pisek, J.; Rautiainen, M.; Nikopensius, M.; Raabe, K. Estimation of seasonal dynamics of understory NDVI in northern forests using MODIS BRDF data: Semi-empirical versus physically-based approach. *Remote Sens. Environ.* **2015**, *163*, 42–47. [CrossRef]
4. Wood, E.M.; Pidgeon, A.M.; Radeloff, V.C.; Keuler, N.S. Image texture as a remotely sensed measure of vegetation structure. *Remote Sens. Environ.* **2012**, *121*, 516–526. [CrossRef]
5. Du, P.; Liu, S.; Xia, J.; Zhao, Y. Information fusion techniques for change detection from multi-temporal remote sensing images. *Inf. Fusion* **2013**, *14*, 19–27. [CrossRef]
6. Fang, H.; Du, P.; Wang, X.; Lin, C.; Tang, P. Unsupervised Change Detection Based on Weighted Change Vector Analysis and Improved Markov Random Field for High Spatial Resolution Imagery. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]
7. Tarantino, C.; Adamo, M.; Lucas, R.; Blonda, P. Detection of changes in semi-natural grasslands by cross correlation analysis with WorldView-2 images and new Landsat 8 data. *Remote Sens. Environ.* **2016**, *175*, 65–72. [CrossRef]
8. Wu, T.; Luo, J.; Fang, J.; Ma, J.; Song, X. Unsupervised Object-Based Change Detection via a Weibull Mixture Model-Based Binarization for High-Resolution Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 63–67. [CrossRef]
9. Biao, W.; Seokkeun, C.; Younggi, B.; Soungki, L.; Jaewan, C. Object-Based Change Detection of Very High Resolution Satellite Imagery Using the Cross-Sharpening of Multitemporal Data. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1151–1155. [CrossRef]
10. Lv, Z.Y.; Shi, W.; Zhang, X.; Benediktsson, J.A. Landslide Inventory Mapping From Bitemporal High-Resolution Remote Sensing Images Using Change Detection and Multiscale Segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 1520–1532. [CrossRef]
11. Wang, H.; Qi, J.; Lei, Y.; Wu, J.; Li, B.; Jia, Y. A Refined Method of High-Resolution Remote Sensing Change Detection Based on Machine Learning for Newly Constructed Building Areas. *Remote Sens.* **2021**, *13*, 1507. [CrossRef]
12. Wang, J.; Yang, X.; Yang, X.; Jia, L.; Fang, S. Unsupervised change detection between SAR images based on hypergraphs. *ISPRS J. Photogramm. Remote Sens.* **2020**, *164*, 61–72. [CrossRef]
13. Xue, D.; Lei, T.; Jia, X.; Wang, X.; Chen, T.; Nandi, A.K. Unsupervised Change Detection Using Multiscale and Multiresolution Gaussian-Mixture-Model Guided by Saliency Enhancement. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 1796–1809. [CrossRef]
14. Chen, H.; Shi, Z. A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection. *Remote Sens.* **2020**, *12*, 1662. [CrossRef]
15. DeVries, B.; Verbesselt, J.; Kooistra, L.; Herold, M. Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series. *Remote Sens. Environ.* **2015**, *161*, 107–121. [CrossRef]
16. Lv, Z.; Liu, T.; Wan, Y.; Benediktsson, J.A.; Zhang, X. Post-Processing Approach for Refining Raw Land Cover Change Detection of Very High-Resolution Remote Sensing Images. *Remote Sens.* **2018**, *10*, 472. [CrossRef]
17. Mayes, M.T.; Mustard, J.F.; Melillo, J.M. Forest cover change in Miombo Woodlands: Modeling land cover of African dry tropical forests with linear spectral mixture analysis. *Remote Sens. Environ.* **2015**, *165*, 203–215. [CrossRef]

18. Saha, S.; Bovolo, F.; Bruzzone, L. Unsupervised Deep Change Vector Analysis for Multiple-Change Detection in VHR Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3677–3693. [CrossRef]

19. Solano-Correa, Y.T.; Bovolo, F.; Bruzzone, L. An Approach to Multiple Change Detection in VHR Optical Images Based on Iterative Clustering and Adaptive Thresholding. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1334–1338. [CrossRef]

20. Wu, C.; Du, B.; Zhang, L. Slow Feature Analysis for Change Detection in Multispectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2858–2874. [CrossRef]

21. Nielsen, A.A. The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. *IEEE Trans. Image Process.* **2007**, *16*, 463–478. [CrossRef] [PubMed]

22. Han, Y.; Javed, A.; Jung, S.; Liu, S. Object-Based Change Detection of Very High Resolution Images by Fusing Pixel-Based Change Detection Results Using Weighted Dempster–Shafer Theory. *Remote Sens.* **2020**, *12*, 983. [CrossRef]

23. Zhao, J.; Liu, S.; Wan, J.; Yasir, M.; Li, H. Change Detection Method of High Resolution Remote Sensing Image Based on D-S Evidence Theory Feature Fusion. *IEEE Access* **2021**, *9*, 4673–4687. [CrossRef]

24. Singh, A.; Singh, K.K. Unsupervised change detection in remote sensing images using fusion of spectral and statistical indices. *Egypt. J. Remote Sens. Space Sci.* **2018**, *21*, 345–351. [CrossRef]

25. Cai, L.; Shi, W.; Zhang, H.; Hao, M. Object-oriented change detection method based on adaptive multi-method combination for remote-sensing images. *Int. J. Remote Sens.* **2016**, *37*, 5457–5471. [CrossRef]

26. Leichtle, T.; Geiß, C.; Wurm, M.; Lakes, T.; Taubenböck, H. Unsupervised change detection in VHR remote sensing imagery—An object-based clustering approach in a dynamic urban environment. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *54*, 15–27. [CrossRef]

27. Xiao, P.; Zhang, X.; Wang, D.; Yuan, M.; Feng, X.; Kelly, M. Change detection of built-up land: A framework of combining pixel-based detection and object-based recognition. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 402–414. [CrossRef]

28. Ansari, R.A.; Buddhiraju, K.M.; Malhotra, R. Urban change detection analysis utilizing multiresolution texture features from polarimetric SAR images. *Remote Sens. Appl. Soc. Environ.* **2020**, *20*, 100418. [CrossRef]

29. Zheng, Z.; Cao, J.; Lv, Z.; Benediktsson, J.A. Spatial–Spectral Feature Fusion Coupled with Multi-Scale Segmentation Voting Decision for Detecting Land Cover Change with VHR Remote Sensing Images. *Remote Sens.* **2019**, *11*, 1903. [CrossRef]

30. Ye, Y.; Bruzzone, L.; Shan, J.; Bovolo, F.; Zhu, Q. Fast and Robust Matching for Multimodal Remote Sensing Image Registration. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9059–9070. [CrossRef]

31. Dalla Mura, M.; Benediktsson, J.A.; Waske, B.; Bruzzone, L. Morphological Attribute Profiles for the Analysis of Very High Resolution Images. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3747–3762. [CrossRef]

32. Xiao, P.; Yuan, M.; Zhang, X.; Feng, X.; Guo, Y. Cosegmentation for Object-Based Building Change Detection From High-Resolution Remotely Sensed Images. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 1587–1603. [CrossRef]

33. Zhang, X.; Xiao, P.; Feng, X.; Yuan, M. Separate segmentation of multi-temporal high-resolution remote sensing images for object-based change detection in urban area. *Remote Sens. Environ.* **2017**, *201*, 243–255. [CrossRef]

34. Zhang, Y.; Zhao, H. Land–Use and Land-Cover Change Detection Using Dynamic Time Warping–Based Time Series Clustering Method. *Can. J. Remote Sens.* **2020**, *46*, 67–83. [CrossRef]

35. Xing, H.; Zhu, L.; Chen, B.; Zhang, L.; Hou, D.; Fang, W. A novel change detection method using remotely sensed image time series value and shape based dynamic time warping. *Geocarto Int.* **2021**, 1–18. [CrossRef]

36. Afaq, Y.; Manocha, A. Analysis on change detection techniques for remote sensing applications: A review. *Ecol. Inform.* **2021**, *63*, 101310. [CrossRef]

37. Mandal, M.; Vipparthi, S.K. An Empirical Review of Deep Learning Frameworks for Change Detection: Model Design, Experimental Frameworks, Challenges and Research Needs. *IEEE Trans. Intell. Transp. Syst.* **2021**, 1–22. [CrossRef]

38. Shi, W.; Zhang, M.; Zhang, R.; Chen, S.; Zhan, Z. Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges. *Remote Sens.* **2020**, *12*, 1688. [CrossRef]

39. Khelifi, L.; Mignotte, M. Deep Learning for Change Detection in Remote Sensing Images: Comprehensive Review and Meta-Analysis. *IEEE Access* **2020**, *8*, 126385–126400. [CrossRef]

40. Chen, H.; Qi, Z.; Shi, Z. Remote Sensing Image Change Detection With Transformers. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–14. [CrossRef]

41. Pan, J.; Cui, W.; An, X.; Huang, X.; Zhang, H.; Zhang, S.; Zhang, R.; Li, X.; Cheng, W.; Hu, Y. MapsNet: Multi-level feature constraint and fusion network for change detection. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *108*, 102676. [CrossRef]

42. Pan, F.; Wu, Z.; Liu, Q.; Xu, Y.; Wei, Z. DCFF-Net: A Densely Connected Feature Fusion Network for Change Detection in High-Resolution Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 11974–11985. [CrossRef]

43. Zhang, C.; Yue, P.; Tapete, D.; Jiang, L.; Shangguan, B.; Huang, L.; Liu, G. A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2020**, *166*, 183–200. [CrossRef]

44. Ferraris, V.; Dobigeon, N.; Chabert, M. Robust fusion algorithms for unsupervised change detection between multi-band optical images—A comprehensive case study. *Inf. Fusion* **2020**, *64*, 293–317. [CrossRef]

45. Rokni, K.; Ahmad, A.; Solaimani, K.; Hazini, S. A new approach for surface water change detection: Integration of pixel level image fusion and image classification techniques. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 226–234. [CrossRef]

46. Xu, L.; Jing, W.; Song, H.; Chen, G. High-Resolution Remote Sensing Image Change Detection Combined With Pixel-Level and Object-Level. *IEEE Access* **2019**, *7*, 78909–78918. [CrossRef]

47. Zhang, C.; Li, G.; Cui, W. High-Resolution Remote Sensing Image Change Detection by Statistical-Object-Based Method. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 2440–2447. [CrossRef]

48. Hao, M.; Zhou, M.; Jin, J.; Shi, W. An Advanced Superpixel-Based Markov Random Field Model for Unsupervised Change Detection. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1401–1405. [CrossRef]

49. Jimenez-Sierra, D.A.; Quintero-Olaya, D.A.; Alvear-Munoz, J.C.; Benitez-Restrepo, H.D.; Florez-Ospina, J.F.; Chanussot, J. Graph Learning Based on Signal Smoothness Representation for Homogeneous and Heterogeneous Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–16. [CrossRef]

50. Shuai, W.; Jiang, F.; Zheng, H.; Li, J. MSGATN: A Superpixel-Based Multi-Scale Siamese Graph Attention Network for Change Detection in Remote Sensing Images. *Appl. Sci.* **2022**, *12*, 5158. [CrossRef]

51. Sun, Y.; Lei, L.; Guan, D.; Li, M.; Kuang, G. Sparse-Constrained Adaptive Structure Consistency-Based Unsupervised Image Regression for Heterogeneous Remote-Sensing Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–14. [CrossRef]

52. Zhao, R.; Peng, G.-H.; Yan, W.-d.; Pan, L.-L.; Wang, L.-Y. Change detection in SAR images based on superpixel segmentation and image regression. *Earth Sci. Inform.* **2020**, *14*, 69–79. [CrossRef]

53. Wu, J.; Li, B.; Ni, W.; Yan, W.; Zhang, H. Optimal Segmentation Scale Selection for Object-Based Change Detection in Remote Sensing Images Using Kullback–Leibler Divergence. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1124–1128. [CrossRef]

54. Luo, H.; Liu, C.; Wu, C.; Guo, X. Urban Change Detection Based on Dempster–Shafer Theory for Multitemporal Very High-Resolution Imagery. *Remote Sens.* **2018**, *10*, 980. [CrossRef]

55. Shao, P.; Yi, Y.; Liu, Z.; Dong, T.; Ren, D. Novel Multiscale Decision Fusion Approach to Unsupervised Change Detection for High-Resolution Images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]

56. Cui, G.; Lv, Z.; Li, G.; Atli Benediktsson, J.; Lu, Y. Refining Land Cover Classification Maps Based on Dual-Adaptive Majority Voting Strategy for Very High Resolution Remote Sensing Images. *Remote Sens.* **2018**, *10*, 1238. [CrossRef]

57. Du, P.; Liu, S.; Gamba, P.; Tan, K.; Xia, J. Fusion of Difference Images for Change Detection Over Urban Areas. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 1076–1086. [CrossRef]

58. Lal, A.M.; Margret Anouncia, S. Semi-supervised change detection approach combining sparse fusion and constrained k means for multi-temporal remote sensing images. *Egypt. J. Remote Sens. Space Sci.* **2015**, *18*, 279–288. [CrossRef]

59. Ye, S.; Chen, D.; Yu, J. A targeted change-detection procedure by combining change vector analysis and post-classification approach. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 115–124. [CrossRef]

60. Zou, B.; Li, H.; Zhang, L. Multilevel Information Fusion-Based Change Detection for Multiangle PolSAR Images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]

61. Liu, S.; Du, Q.; Tong, X.; Samat, A.; Bruzzone, L. Unsupervised Change Detection in Multispectral Remote Sensing Images via Spectral-Spatial Band Expansion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 3578–3587. [CrossRef]

62. Lv, Z.; Liu, T.; Benediktsson, J.A. Object-Oriented Key Point Vector Distance for Binary Land Cover Change Detection Using VHR Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 6524–6533. [CrossRef]

63. Huang, L.; Peng, Q.; Yu, X. Change Detection in Multitemporal High Spatial Resolution Remote-Sensing Images Based on Saliency Detection and Spatial Intuitionistic Fuzzy C-Means Clustering. *J. Spectrosc.* **2020**, *2020*, 2725186. [CrossRef]

64. Dragut, 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]

65. Nemmour, H.; Chibani, Y. Multiple support vector machines for land cover change detection: An application for mapping urban extensions. *ISPRS J. Photogramm. Remote Sens.* **2006**, *61*, 125–133. [CrossRef]

66. Benedek, C.; Sziranyi, T. Change Detection in Optical Aerial Images by a Multilayer Conditional Mixed Markov Model. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3416–3430. [CrossRef]

67. Shen, L.; Lu, Y.; Chen, H.; Wei, H.; Xie, D.; Yue, J.; Chen, R.; Lv, S.; Jiang, B. S2Looking: A Satellite Side-Looking Dataset for Building Change Detection. *Remote Sens.* **2021**, *13*, 5094. [CrossRef]