

Journal of
Applied Remote Sensing

**Object-oriented change detection
approach for high-resolution remote
sensing images based on multiscale
fusion**

Chao Wang
Mengxi Xu
Xin Wang
Shengnan Zheng
Zhenli Ma

Object-oriented change detection approach for high-resolution remote sensing images based on multiscale fusion

Chao Wang,^a Mengxi Xu,^b Xin Wang,^{a,c} Shengnan Zheng,^b and Zhenli Ma^a

^aHohai University, College of Computer and Information Engineering, Nanjing 211100, China

hhu_wch@163.com

^bNanjing Institute of Technology, School of Computer Engineering, Nanjing 211167, China

^cHohai University, Engineering Research Center of Sensing and Computing, Nanjing 211100, China

Abstract. Aiming at the difficulties in change detection caused by the complexity of high-resolution remote sensing images that exist in varied ecological environments and artificial objects, in order to overcome the limitations in traditional pixel-oriented change detection methods and improve the detection precision, an innovative object-oriented change detection approach based on multiscale fusion is proposed. This approach introduced the classical color texture segmentation algorithm J-segmentation (JSEG) to change detection and achieved the multiscale feature extraction and comparison of objects based on the sequence of J-images produced in JSEG. By comprehensively using the geometry, spectrum, and texture features of objects, and proposing two different multiscale fusing strategies, respectively, based on Dempster/Shafer evidence theory and weighted data fusion, the algorithm further improves the divisibility between changed and unchanged areas, thereby establishing an integrated framework of object-oriented change detection based on multiscale fusion. Experiments were performed on high-resolution airborne and SPOT 5 remote sensing images. Compared with different object-oriented and pixel-oriented detection methods, results of the experiments verified the validity and reliability of the proposed approach. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JRS.7.073696](https://doi.org/10.1117/1.JRS.7.073696)]

Keywords: high-resolution remote sensing image; object-oriented; change detection; J-segmentation; multiscale fusion.

Paper 13174 received May 18, 2013; revised manuscript received Aug. 12, 2013; accepted for publication Aug. 22, 2013; published online Sep. 24, 2013.

1 Introduction

As one of the most popular research topics in current application of remote sensing, change detection for multitemporal remote sensing images is essentially a process of determining the information of geophysical changes using remote sensing images of the same area at different temporals.¹ The fields of application include a city's dynamic development and geographical information databases update, etc. As a major application field, urban change detection has played an important role in city planning and management. For moderate- and low-resolution remote sensing images, various effective change detection methods have been proposed by scholars, and most of these methods can achieve reliable results by comparing each pixel in images.²⁻⁸

In recent years, meter and submeter high-resolution remote sensing images represented by SPOT5, Quick-Bird, IKONOS, etc., have been widely applied.⁹ Improvement in spatial resolution not only provides more spectrum, texture, and geometrical information, but also brings about new challenges. First, the phenomenon of "the same object with different spectrums" is much more serious, and the phenomenon of "the same spectrum with different objects" still exists, so that it is difficult to differentiate changed areas from unchanged areas.¹⁰ Second, urban landscapes include various ecological environments and complex artificial objects.

Consequently, it is hard for traditional pixel-oriented change detection methods to incorporate the concept of “object,” and the traditional method has poor robustness on the pseudochange caused by the slight spectrum difference inside the “object.” In addition, pixel-oriented change detection methods have high requirements on registration accuracy, radiometric correction, and viewpoint changes. Finally, topographic shadow, clouds covering, etc., can also cause difficulty in change detection. Therefore, there is great difficulty in directly applying the traditional pixel-oriented change detection methods to high-resolution remote sensing change detection.¹¹

Compared with traditional pixel-oriented method, object-oriented change detection (OOCD) method chooses geographic object as basic unit for change detection and provides a new solution to the mentioned difficulties. OOCD method extracts the object’s features based on its natural shape and size, thus improving the category divisibility of different geographic objects and facilitating the deep analyzing of change information inside objects.^{12,13} Scholars have proposed some effective OOCD methods;^{14–18} e.g., Miller et al. proposed a method to detect blobs changes between gray-scale images, that is, first using connectivity analysis to obtain objects and then finding the matching object of each object in another image to make a comparison.¹⁴ Lefebvre et al. further validated the application of geometry (i.e., size, shape, and location) and content (i.e., texture) information in OOCD algorithm.¹⁵

Currently, there are several major challenges/issues of OOCD methods for high-resolution remote sensing images. First, meaningful image-objects should be completely extracted by typical segmentation to represent geographic objects in OOCD. However, currently no specific algorithm can be claimed to be adaptable for all OOCD algorithms. And in most OOCD algorithms, a great deal of spectrum or texture information generated during image segmentation is merely used to extract the objects and is still not fully exploited, especially for object-based features extraction.¹⁶ Second, since singly using the spectrum feature of an image to describe the change information in objects has to face the high requirement of image registration precision, and further the detection result is vulnerable to the effect of noise, the extra features, especially the texture features, are applied more and more in change detection. Therefore, a suitable combination of multiple object-based spectrum and texture features can effectively improve the accuracy and reliability of the algorithm.^{17–19} Finally, the results of change detection are related with the scale; that is, a single scale is insufficient to capture all the characteristics of objects within different sizes, shapes, etc. Based on the human visual system and expert knowledge, the combination of multiscale analysis tools and OOCD could more deeply analyze the changes between each object in different temporal and produce more reliable results than single-scale analysis.^{20–22} Therefore, designing an effective fusion strategy becomes another critical issue.

Based on the above analysis, this paper proposes a new OOCD approach for high-resolution remote sensing images based on multiscale fusion. Currently, J-segmentation (JSEG) algorithm²³ is one of the most popular methods for color image segmentation. The proposed approach uses the JSEG algorithm to extract the image-objects, perform multiscale feature extraction and object comparison on the sequence of J-images that are generated in the segmentation process. Then, two fusion strategies are presented to construct an integrated change detection framework and derive the final detection results. Experiment shows that both strategies can produce satisfying results and have their respective advantages in false and miss detection. At last, the detection results classify object areas under different change intensities.

This paper consists of four sections. The basic principles and specific implementations of the approach will be introduced in the next section. Section 3 makes an analysis and a comparison of the experiment results, and the last section provides the conclusion.

2 Method

In order to effectively extract, describe, and compare geographic objects from high-resolution remote sensing images, the method proposed in this paper mainly includes three components: object extraction, object analysis and comparison, multiscale fusion.

2.1 Object Extraction

The purpose of object extraction is to extract the areas belonging to the same geographic objects through segmentation. JSEG algorithm proposed by Deng and Manjunath is a multiscale color texture segmentation method that shows a strong detection capability for homogeneity of regional color texture features and has been successfully applied in remote sensing image segmentation.^{24,25}

During the process of JSEG, a sequence of multiscale J-images is generated. J-image reflects color distribution of the original image, which means it is in essence a gradient image with scale features. Therefore, for the J-images with the same scale from different multitemporal images, a similar description of a certain object from segmentation results based on gray values actually reflects the overall similarity of this object's spectrum, texture, and scale features in different temporal images. In this manner, the limitations mentioned above in just using the spectrum feature in the original image can be effectively overcome. On the other hand, it means that there is no need to recalculate the multiscale images for following multiscale change detection. Compared with famous commercial software like eCognition, JSEG algorithm not only can implement the precise image segmentation, but also can be used with J-images for further object-based features extraction and comparison. In addition, JSEG algorithm can effectively improve the proposed change detection framework with better transparency and robustness. For these reasons, objects in this paper will be extracted using JSEG algorithm, which includes two steps: color quantization and space segmentation.

The color quantization applied the method proposed by Deng et al.²⁶ First, the color space of the image will be converted to LUV color space. Then peer group filtering is used to perform image smoothing and denoising. Finally, the quantized image is obtained by applying the classic Hard C-means algorithm.

As for the space segmentation phase, a local homogeneity index J value is calculated based on the quantized image, thereby generating J-images sequence. The detailed process of the subsequent segmentation in JSEG can be found in Refs. 23 and 26.

In particular, J value is defined as follows: Let each pixel's location $Z(x, y)$ in the quantized image be the value of pixel z , and $Z(x, y) \in Z$. Z is the set of all pixels inside the specific-sized window with pixel z as the center. Figures 1 and 2 are shown with z as the center, and their sizes are 9×9 pixels and 18×18 pixels, respectively. In order to maintain the consistency in each direction, corners in each window are removed.

J value can be calculated according to the following formula:

$$J = (S_T - S_W) / S_W, \quad (1)$$

where S_T is the population variance of all pixels in Z , and S_W is the sum of all pixel variances in the same gray level. Using the same window size to calculate the J value of each pixel, which is to be its pixel value as well, can produce the J-image at a single scale. Therefore, image sequence

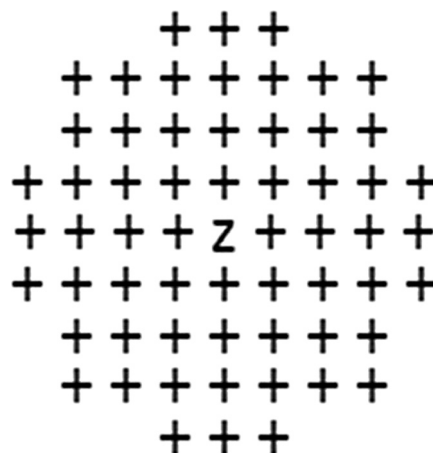


Fig. 1 Window of 9×9 pixels at z .

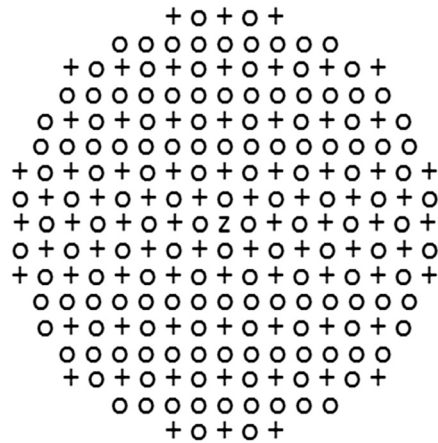


Fig. 2 Window of 18 × 18 pixels at z.

of multiscale J-images can be obtained by adjusting the size of windows. In this paper, the smallest scale is defined as the J-image calculated from minimum window size.

2.2 Object Analysis and Comparison

In view of the above-mentioned characteristics of J-image, we separately analyze and compare each object in multiscale J-images based on the segmentation results. At this point, it is critical to select an appropriate similarity measurement to describe the similarity of a certain object in different temporal. Common measurements include various “distances” such as Euclidean distance and Mahalanobis distance, histogram matching, covariance, etc. Structural similarity (SSIM)²⁷ first proposed by Wang et al. takes the mean value, variance, and covariance of vectors into account, and therefore, it can well express the similarity between vectors. SSIM between vector x and vector y is defined in Eq. (2).

$$S(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma, \tag{2}$$

where

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_2}, \tag{3}$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2\sigma_y^2 + C_2}, \tag{4}$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}, \tag{5}$$

In Eqs. (2) to (5), μ_x , μ_y , σ_x , σ_y , σ_x^2 , and σ_y^2 refer to mean value, standard deviation, and variance of x and y , respectively. σ_{xy} refers to a covariance between x and y . α , β , and γ are the weights of three vectors, and C_1 , C_2 , and C_3 are constants added to the formulas in order to prevent instability when denominator approximates to zero.

When $\alpha = \beta = \gamma = 1$, $C_3 = C_2/2$, Eq. (2) can be simplified as

$$S(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{6}$$

The larger $S(x, y)$ is, the smaller the change in object between multitemporal images and the higher the similarity. In addition, according to the definition thereof, SSIM has the following

characteristics: (1) it is bounded with $S(x, y) \in [0, 1]$; (2) it is symmetrical: $S(x, y) = S(y, x)$; (3) it has a unique maximum value, when and only when $x = y$, $S(x, y) = 1$. Normally, a similarity measurement satisfying the above three criteria is considered to describe vectors' similarity better.

Compared with SSIM, those various "distances" do not satisfy the characteristic of "bounded." The histogram matching is not symmetric, and the covariance does not meet the criterion "unique maximum value." Consequently, this paper selects SSIM to describe the similarity of each object between multitemporal images. For J-image at a certain scale, SSIM for all objects in segmentation results is calculated to obtain the change detection results at a single scale.

2.3 Multiscale Fusion

Considering the dependence of the objects and the changes on scale, and in order to improve change detection precision, two multiscale fusion strategies are presented in the proposed approach.

Fusion strategy 1 is based on Dempster/Shafer (D-S) evidence theory,²⁸ which analyzes the whole system through multisource information, thereby making the right decision. D-S evidence theory is an effective tool to solve uncertain reasoning problems; the basic concept of D-S evidence theory is explained below.

U is defined as a recognition framework. Define basic probability assignment formula (BPAF) as a function $m: 2^U \rightarrow [0, 1]$ in 2^U , and m satisfying

$$m(\emptyset) = 0 \sum_{A \subseteq U} m(A) = 1, \quad (7)$$

where A satisfying $m(A) > 0$ is called a focal element, $m(A)$ represents a trust measurement of evidences on A . As for $\forall A \subseteq U$, the Dempster's combinational rule is defined as follows:

$$m = m_1 \oplus m_2 \oplus \dots \oplus m_n(A) = \frac{1}{K} \sum_{\cap A_i = A} \prod_{1 \leq d \leq n} m_d(A_d), \quad (8)$$

where K is the normalization constant, which reflects the extent of conflicts between evidences, and can be defined as follows:

$$K = \sum_{\cap A_i \neq \emptyset} \prod_{1 \leq d \leq n} m_d(A_d). \quad (9)$$

In fusion strategy 1, D-S theory framework is defined as $U: \{JL, MX, N\}$, where JL stands for dramatically changed objects, MX refers to obvious changed objects, and N means unchanged objects. Thus, nonempty subsets of 2^U include $\{JL\}$, $\{MX\}$, $\{N\}$, and $\{JL, MX, N\}$. For each object R_i ($i = 1, 2, 3, \dots, P$, with P being the total number of objects in the segmentation results), define S_{ik} as the SSIM of R_i between multitemporal J-images at the same scale k , and the corresponding BPAF is established through the following formula:

$$m_{ik}(\{JL\}) = (1 - S_{ik}) \times T \times \alpha_k, \quad (10)$$

$$m_{ik}(\{MX\}) = (1 - S_{ik}) \times (1 - T) \times \alpha_k, \quad (11)$$

$$m_{ik}(\{N\}) = S_{ik} \times \alpha_k, \quad (12)$$

$$m_{ik}(\{JL, MX, N\}) = 1 - \alpha_k, \quad (13)$$

where threshold T determines the change intensity of pixels that belong to the dramatically changed objects and $\alpha_k \in (0, 1)$ ($k = 1, 2, 3, \dots, M$, with M being the total number of scales

in the segmentation), represents the credibility of each scale in decision. As shown in Figs. 1 and 2, the small scale is suited to be used to detect the detail changes of objects, while the detection in large scale can effectively reduce the interference from noise and isolated points. Consequently, the values of parameters in the approach need to be set manually by experience or actual requirements in special applications.

Based on the established BPAF Eqs. (10) to (13), the decision rule for fusion strategy 1 can be explained as follows:

Step 1: For each R_i , calculate $m_i(\{JL\})$, $m_i(\{MX\})$, $m_i(\{N\})$, and $m_{ik}(\{JL, MX, N\})$ by S_{ik} from different scales according to Eq. (8). $m = m_1 \oplus m_2 \oplus \dots \oplus m_M$.

Step 2: If $m_i(\{JL\}) > 0.8$ or $m_i(\{MX\}) > 0.2$, and $m_i(\{JL\}) > 0.6$, then R_i is an object with dramatic change.

Step 3: If $m_i(\{MX\}) > 0.4$ or $m_i(\{N\}) < 0.7$, then R_i is an object with an obvious change.

Step 4: Otherwise, R_i is unchanged.

Step 5: Repeat steps 1 to 4 until all objects in the segmentation results are gone through.

In order to further confirm that, compared with single-scale detection, multiscale fusion strategy can effectively improve detection precision and yield more reliable results, fusion strategy 2 uses weighted data fusion. Define $\alpha_l \in (0, 1)$ ($l = 1, 2, 3, \dots, M$) as the weight value for detection results at each scale. Decision rule for fusion strategy 2 can be explained as follows:

Step 1: For each R_i , S_{ik} is combined with rule $S_i = \alpha_1 \times S_{i1} + \alpha_2 \times S_{i2} + \dots + \alpha_M \times S_{iM}$.

Step 2: If $S_i \in [0.85, 1]$, then R_i is unchanged.

Step 3: If $S_i \in [0.3, 0.85)$, R_i shows obvious changes.

Step 4: Otherwise, R_i is considered to be dramatically changed.

Step 5: Repeat steps 1 to 4 until all objects in the segmentation results are gone through.

2.4 Specific Implementation of Approach

As presented above, the specific implementation process of the proposed approach is illustrated in Fig. 3.

As shown in Fig. 3, the two temporal remote sensing images first need to be radiometrically corrected and geometrically registered. Then, JSEG algorithm is used to extract objects. It should be noted that the temporal image with less noise or shadows in multitemporal images will be chosen to be segmented in the proposed approach. In order to extract the same geographic objects, we directly map these boundaries of segmentation results to all J-images from different temporal images based on the registration results. On the other hand, we can also separately segment each temporal image and directly map all the segmented boundaries to J-images

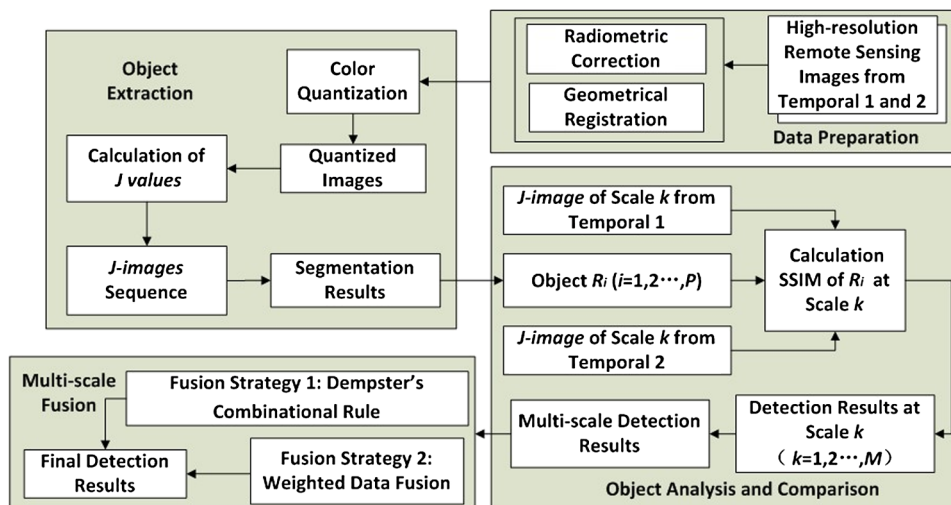


Fig. 3 Flow chart of approach.

from different temporal images based on registration results. No matter in which segmentation ways the J-image sequence should be calculated by the same set of window sizes from multi-temporal images, and both segmentation ways are allowed in the proposed change detection framework.

In the phase of object analysis and comparison, we can find the corresponding region for each object R_i in every J-image from different temporal images based on the segmentation and registration results. Based on this, the SSIM for each object R_i at single scale is calculated according to Eq. (6). Finally, the detection results are fused from multiscales according to the fusion strategies proposed in Sec. 2.3 and the entire detection process is accomplished.

3 Experiment Results and Analysis

For the purpose of comprehensively analyzing the performance of the proposed approach, this paper not only compares the method with traditional pixel-oriented and OOC algorithm, but also analyzes the effects that the change of scale and fusion strategy have on the detection results. In addition, in order to further test the validity and reliability of the approach on remote sensing images from different sensors, two different types of datasets are selected for this experiment.

For pixel-oriented change detection, we choose the classic change vector analysis (CVA) method and the improved CVA-expectation-maximization (CVA-EM) algorithm²⁹ proposed by Bruzzone et al. for comparison. CVA-EM algorithm uses the difference image generated by CVA method and introduces EM algorithm to estimate the relevant parameters of Gaussian model, which obviously yields a higher detection precision. Experiments were performed on both datasets, with the branch number of Gaussian mixture model defined as $p = 2$. The initial value for EM algorithm was set the same way as in Ref. 29.

As for object-oriented method, this paper uses the multiscale object-specific approach (MOSA)³⁰ proposed by Hall et al. for comparison. MOSA extracts objects using multiscale marker-controlled watershed segmentation. It then calculates the difference image by adaptive threshold and obtains the final change results, which can effectively identify the change information related to scale. Hall believes that for MOSA method, the finest scale produces the best detection results. Therefore, this paper only evaluates the detection precision of MOSA at this particular scale.

3.1 Analysis of Experiment Results on Dataset 1

Image #1 and image #2 have been selected as dataset 1 to perform the experiments, as shown in Figs. 4(a) and 4(b). Images #1 and #2 are the airborne remote sensing digital ortho-photo map images acquired in March 2009 and Feb 2012, respectively, at the location of Jiangning campus of Hohai University, Nanjing city, Jiangsu province, China. Dataset 1 is at a spatial resolution of 0.5 m, and the size of image is 512×512 pixels.

Images from datasets 1 and 2 (see Fig. 5) were acquired in early spring (February to March) and late spring (June to July), respectively, which means that vegetation types are similar and

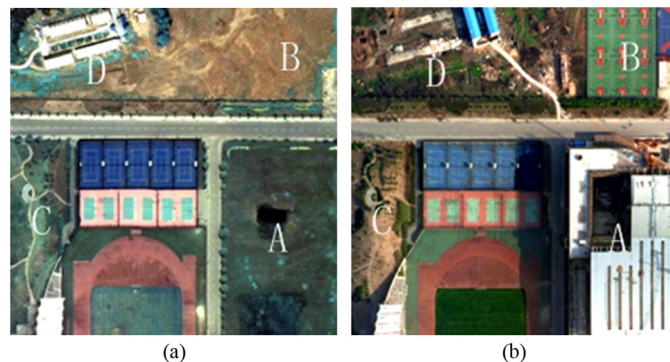


Fig. 4 Dataset 1 Airborne remote sensing images. (a) #1 Jiangning campus of Hohai University in 2009. (b) #2 Jiangning campus of Hohai University in 2012.

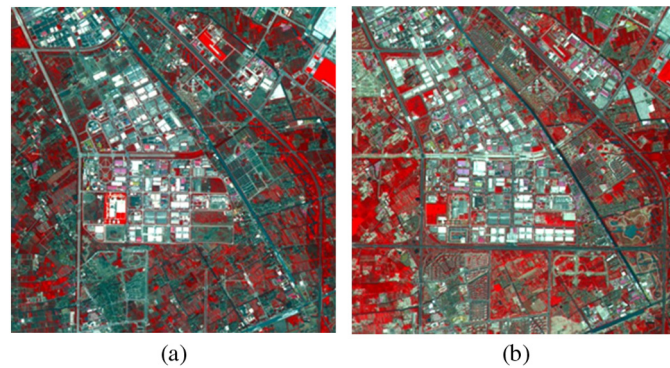


Fig. 5 SPOT 5 pan-sharpened multispectral images. (a) Image #3 Shanghai, 2004. (b) Image #4 Shanghai, 2008.

therefore helpful for change detection. The matching precision for these two datasets is maintained within 0.5 pixel after the radiation correction and the geometric accuracy correction. Comparison between these two datasets, as shown in Figs. 4 and 5, indicates several aspects of the complexity and typicality of the scenes in these images: they all include typical changes, i.e., obvious changes of complex artificial objects in large areas and small changes as in tiny plants etc.; images in both datasets contain various geometric objects like vegetation, lakes, roads and buildings, etc. In addition, affected by illumination changes, there are large areas of shadow in image #2 in dataset 1; image #1 was therefore segmented.

The set of window sizes for J value was set as 20×20 , 10×10 , and 5×5 pixels, so $M = 3$. Figures 6(a) and 6(b) show the maximal scale J-image by 20×20 pixels window, namely scale 1.

The extracted boundaries and an object R_i are shown in Fig. 7. Figure 8 shows the corresponding region of R_i at scale 2 (window size for J value is 10×10 pixels) in image #2.

In Eq. (6), let $C_1 = 0.2$ and $C_2 = 0.8$. In fusion strategy 1, let threshold $T = 0.3$, $\alpha_1 = 0.7$, $\alpha_2 = 0.8$, and $\alpha_3 = 0.9$. In order to fairly compare the two strategies, the weight value α_l in fusion strategy 2 were set as same as α_k in strategy 1.

The final change detection results of two fusion strategies in dataset 1 are shown in Figs. 9(a), 9(b), and 9(c). In the figures below, areas with different colors refer to the objects belonging to dramatically changed areas, obviously changed areas, and unchanged areas, respectively.

Figures 10(a), 10(b), and 10(c) present change detection results using MOSA, CVA, and CVA-EM algorithms, respectively.

For the convenience of visual analysis, as shown in dataset 1, the locations of typically changed ground objects for Jiangning campus of Hohai University during 2009 to 2012 are marked with letters A to D. Changed items include buildings, basketball court, vegetation, and other irregular artificial objects. Location A is the newly built gymnasium of the university.

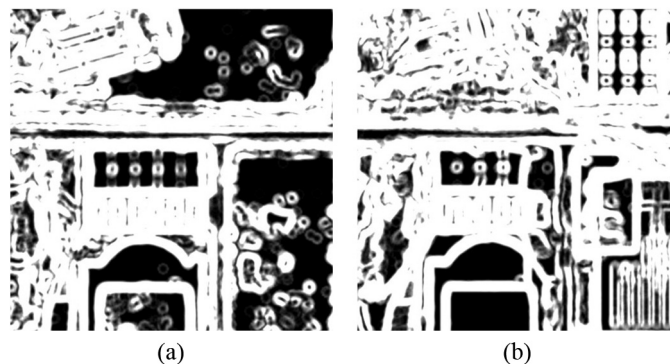


Fig. 6 J-image from multitemporal images. (a) Image #1 J-image at scale 1. (b) Image #2 J-image at scale 1.

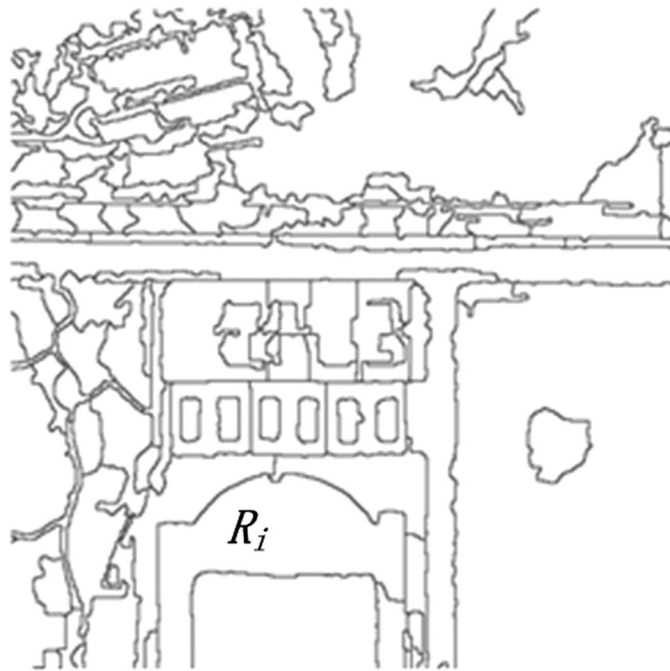


Fig. 7 Segmentation results for image #1.

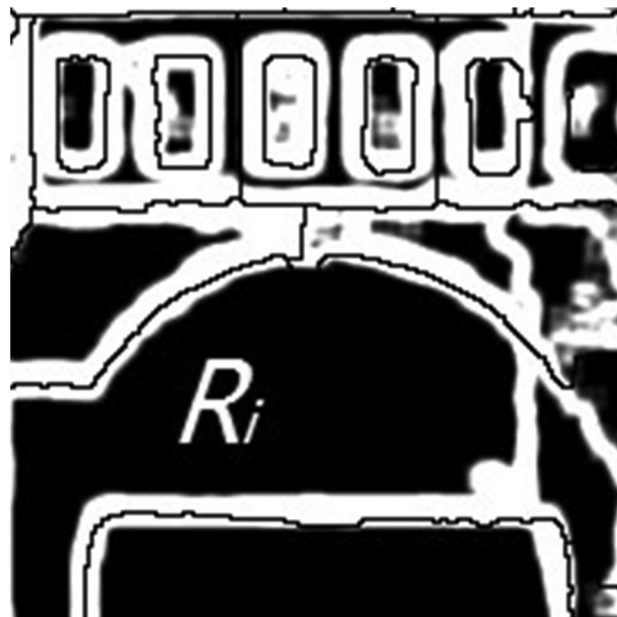


Fig. 8 Corresponding region of R_i at scale 2 in image #2.

Location B is the new basketball court adjoined to the new handball field. Location C is the degraded lawn, and D is the temporary house.

Visual observation and comparison among Figs. 4, 9, and 10 reveal that the following: (1) Both CVA and CVA-EM algorithms mainly miss the basketball court and handball field at location B. MOSA method performs poorly in detecting changes in complex structures like location D. (2) Both fusion strategies can effectively detect change information at the four marked locations. Detection results under two strategies show that the detection results on regular anthropogenic objects like A and B are substantially the same, while the difference

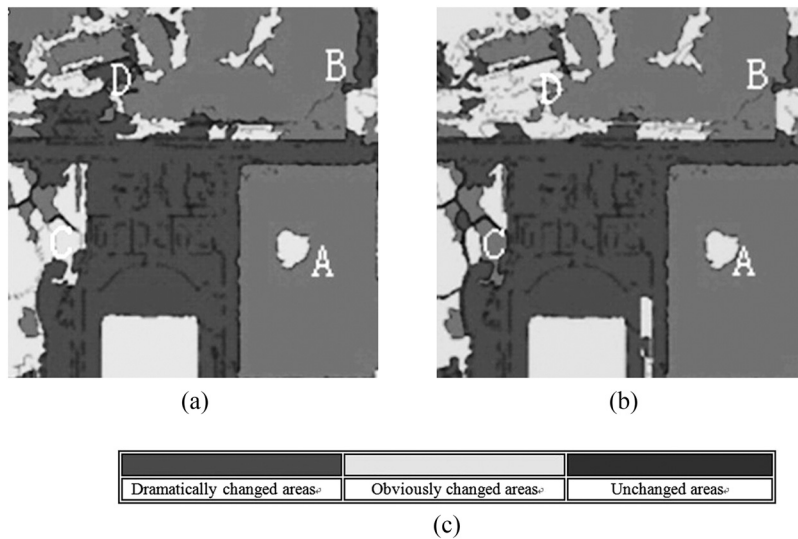


Fig. 9 Detection results of dataset 1 with proposed approach. (a) Fusion strategy 1. (b) Fusion strategy 2. (c) Change intensities with different colors.

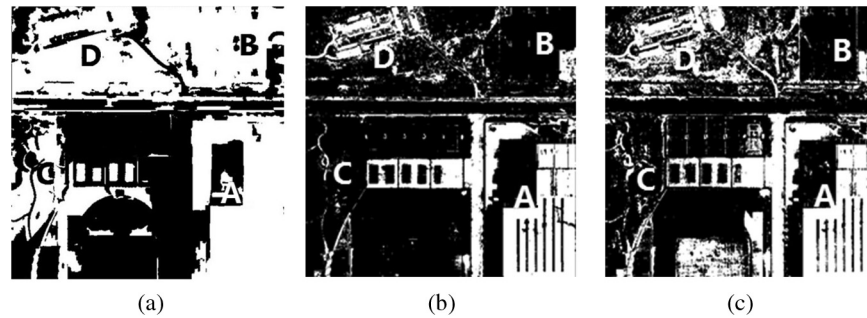


Fig. 10 Detection results of different algorithms. (a) MOSA. (b) CVA. (c) CVA-EM.

is in areas of complex background mixing various objects, such as location D. They also have different determination on the intensity levels of changes in some areas such as location C. Also, fusion strategy 2 detects more changed areas in the whole scene. (3) Large blocks of shadow in image #2 result in significant amount of false alarms with CVA and CVA-EM method. However, object-oriented MOSA and algorithm proposed in this paper can effectively reduce the interference from shadows, like road areas on the right side of location A.

In order to further quantitatively analyze the performance of different detection methods, on the basis of field visits and visual observation of detection results, a sample dataset of 7523 changed pixels and 8861 unchanged pixels is selected as the real sample data. Overall accuracy, false alarm rate, miss detection rate, and Kappa index are calculated to evaluate the performance of each method with results listed in Table 1.

Based on the above table, the following can be observed: (1) The O OCD approach proposed in this paper is obviously better than MOSA and the other two pixel-oriented detection methods, and is consistent with results of visual analysis. The overall accuracy and Kappa indexes for the two fusion strategies are 87.3%, 0.7212 and 86.8%, 0.7074, respectively, and the false alarm rates are considerably lower than the two pixel-oriented algorithms. Even though fusion strategy 1 has a slightly higher miss detection rate than MOSA algorithm, its false alarm rate is even lower and overall accuracy is higher. (2) Strategy 1 applies decision fusion based on D-S evidence theory, and yields the best performance in the experiments even though it has a slightly higher miss detection rate than strategy 2. (3) Strategy 2 adopts weighted data fusion on detection results at different scales, and its false alarm rate is a little higher than that of CVA-EM method, but the miss detection rate is the lowest in the experiments.

Table 1 Detection accuracy of different methods for dataset 1.

Methods/parameters	Overall accuracy/%	False alarm rate/%	Miss detection rate/%	Kappa index
Fusion strategy 1	87.3	11.12	17.21	0.7212
Fusion strategy 2	86.8	12.95	15.96	0.7074
MOSA	84.64	14.35	16.49	0.6878
CVA	81.2	13.29	30.5	0.6531
CVA-EM	83.5	12.59	23.32	0.6796

3.2 Analysis of Experiment Results on Dataset 2

Dataset 2 uses SPOT 5 pan-sharpened multispectral images #3 and #4 with a spatial resolution of 5 m and size of 1024×1024 pixels as shown in Figs. 5(a) and 5(b). They are fused with four wave bands in SPOT 5 including panchromatic band, red band, green band, and near-infrared band. Images #3 and #4 were acquired in June 2004 and July 2008, respectively, in Shanghai, China.

Compared with dataset 1, dataset 2 has lower space resolution and more complex background. Therefore, smaller windows were used for object extraction in the experiment with the proposed approach: 9×9 , 7×7 , and 5×5 pixels. Set $C_1 = 0.2$ and $C_2 = 0.8$; the threshold value was set as $T = 0.4$, with $\alpha_1 = 0.8, \alpha_2 = 0.9, \alpha_3 = 0.95$. Detection results are shown as Figs. 11(a), 11(b), and 11(c).

Detection results under MOSA, CVA, and CVA-EM algorithms are shown in Figs. 12(a), 12(b), and 12(c).

With reference to the previous experiments, a dataset containing 7523 changed pixels and 8861 unchanged pixels in the image are selected to be real change results. Accuracy parameters are calculated for different methods as shown in Table 2.

Table 2 summarizes the results in the following aspects: (1) The performance of different algorithms on dataset 2 are basically the same as the conclusion obtained from dataset 1; thus, it further validates the effectiveness and reliability of the proposed method. Obviously, compared with traditional pixel-oriented change detection methods, the method proposed in this paper can significantly improve detection precision for high-resolution remote sensing images. In addition, compared with traditional object-oriented method MOSA, the detection

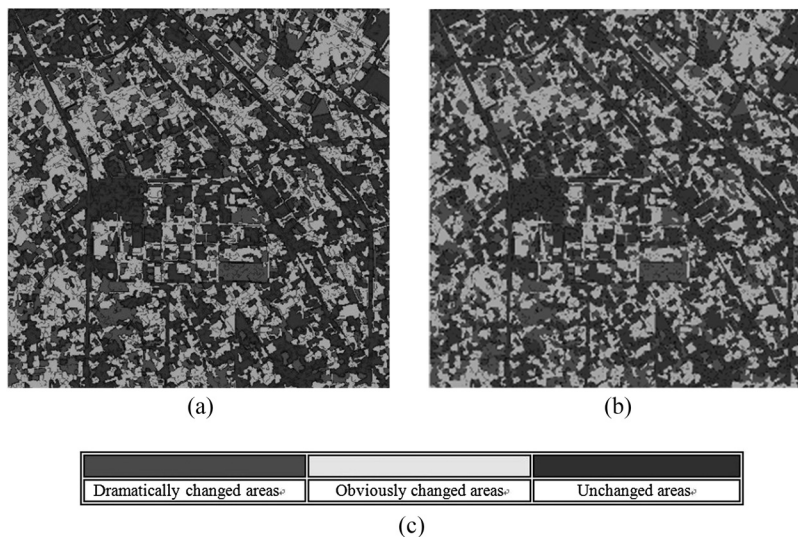


Fig. 11 Detection results of dataset 2 with proposed approach. (a) Fusion strategy 1. (b) Fusion strategy 2. (c) Change intensities with different colors.

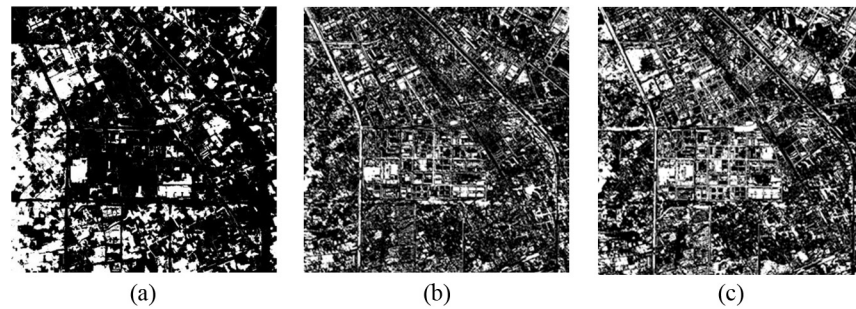


Fig. 12 Experiment results of different algorithms. (a) MOSA. (b) CVA. (c) CVA-EM.

Table 2 Detection accuracy of different methods for dataset 2.

Methods/parameters	Overall accuracy%	False alarm rate%	Miss detection rate%	Kappa index
Fusion strategy 1	85.2	13.75	16.18	0.7058
Fusion strategy 2	85.1	14.83	15.42	0.6996
MOSA	83.7	15.02	16.05	0.6857
CVA	80.6	16.39	29.32	0.6531
CVA-EM	81.2	14.59	24.57	0.6647

algorithm proposed in this paper yields better accuracy parameters except a slightly higher miss detection rate in fusion strategy 1. (2) The overall detection accuracy of each algorithm in dataset 2 is lower than that in dataset 1, which is mainly driven by low spatial resolution of images in dataset 2. The reduction in resolution leads to the increase of the proportion of mixed pixels that contain multiple objects in the scene. (3) Results of the two datasets indicate that fusion strategy 1 can effectively control the false alarm rate, while strategy 2 can effectively reduce the miss detection rate.

3.3 Scale Dependence and Fusion Strategy Analysis

In order to analyze the dependence of change on scale and the effects of the two fusion strategies on detection results, further comparisons are performed in two aspects: accuracy parameters of detection results and area proportion of regions with different change intensity.

With reference to the previous two experiments, detection results at each scale J -image and the acquired accuracy parameters are illustrated in Figs. 13(a), 13(b), 13(c), and 13(d). In these figures, the dotted curve represents dataset 1 and the full curve represents dataset 2.

The following conclusion can be drawn based on the comparison between detection accuracy parameters in Fig. 13 at different scales and under different fusion strategies: change detection results at each single scale differ obviously, and the corresponding detection precisions are lower than those under fusion strategies. Therefore, applying multiscale fusion to single-scale detection results can effectively improve detection precision and reliability of algorithm. Comparison between Table 1, Table 2, and Fig. 13 indicates that the overall accuracy under single-scale object-oriented method in this paper is still obviously better than that under CVA and CVA-EM algorithm.

Table 3 (a) and (b) lists the proportion of areas of each change intensity level in the detection results of both fusion strategies.

As shown in the tables, dramatically changed areas are mostly overlapping (Figs. 9 and 11) and basically of the same size under both strategies for the same dataset (the proportions are 10.2 to 11.3% for dataset 1 and 16.1 to 18.7% for dataset 2). Thus, dramatically changed areas can be set as the areas where actual changes are most likely to occur and should therefore be the primary

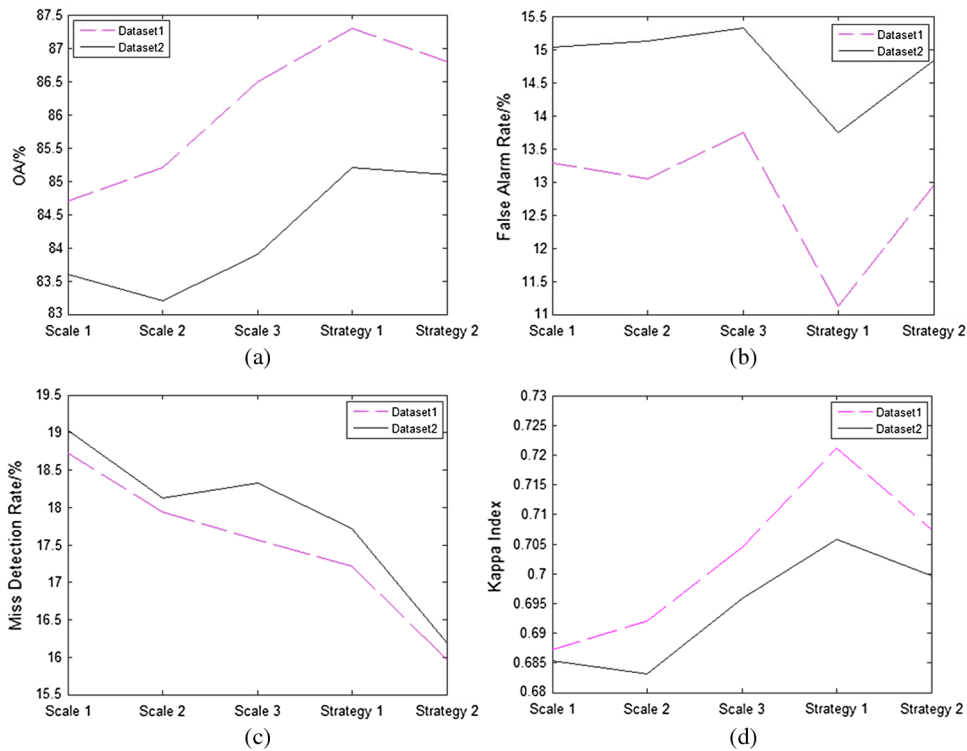


Fig. 13 Scale dependence and fusion strategy analysis. (a) Overall accuracy. (b) False alarm rate. (c) Miss detection rate. (d) Kappa index.

Table 3 Proportion of areas of different change intensity levels/%.

Fusion strategy/change intensity	Dramatic change/%	Obvious change/%	No change/%	Total/%
(a) Dataset 1				
Fusion 1	11.3	10.4	78.3	100
Fusion 2	10.2	16.6	73.2	100
(b) Dataset 2				
Fusion 1	16.1	20.8	63.1	100
Fusion 2	18.7	22.8	58.5	100

detection target in practice. Obviously changed areas can be set as the “hot areas” in next phase of field investigations.

4 Conclusions

This paper established an integrated OOCDF framework based on multiscale fusion and compared the detection performance of this framework. The following conclusions can be drawn:

1. The detection framework proposed in the paper is effective and reliable in urban change detection in high-resolution remote sensing images. The use of JSEG algorithm not only achieves the accurate extraction of objects in the scene, but also uses the multifeatures contained in the J-image sequence to perform change detection, and final results can be acquired by further applying two different fusion strategies. Experiment proves that this method overcomes the uncertainty of single-scale detection, thus producing detection

results that are closer to real changes. In addition, with *J*-image's multifeatures, calculation of SSIM between objects based on *J*-image is less susceptible to noises, and the interference from shadows in city scenes has been effectively reduced so that actual change location can be narrowed down and identified, thereby increasing detection precision.

2. Compared with the traditional pixel-oriented and object-oriented detection methods, the approach proposed in this paper has obviously higher precision. In the experiments conducted on two datasets, this algorithm performs better than two pixel-oriented detection algorithms even at single scale. Thus, it proves that pixel-oriented change detection algorithm can hardly satisfy the demands for high-resolution remote sensing images.
3. Both fusion strategies in the framework have their own advantages. Strategy 1 can effectively control the false alarm rate, while strategy 2 is better at reducing the missed detection rate. In practice, actual demands need to be taken into consideration in order to select the appropriate fusion strategy.
4. Dramatically changed areas detected by both strategies can serve as primary target areas for fieldwork; then, obviously changed areas can be examined as important prospecting areas. Division of change intensity can provide valuable reference information for fieldwork, thereby reducing workload and saving resources.

Hence, the future work will be focused on how to further improve the detection precision of the proposed framework and the application of multiscale analysis in OOD algorithms.

Acknowledgments

This study is supported by the National Natural Science Foundation of China (No. 61271386), the Industrialization Project of Universities in Jiangsu Province (No. JH10-9), and Youth Foundation of Nanjing Institute of Technology (No. QKJA201204).

References

1. H. A. Sing, "Digital change detection techniques using remotely sensed data," *Int. J. Remote Sens.* **10**(6), 989–1003 (1989), <http://dx.doi.org/10.1080/01431168908903939>.
2. M. Gabriele, A. Elena, and B. S. Sebastiano, "Multiscale unsupervised change detection on optical images by Markov random fields and wavelets," *IEEE Geosci. Remote Sens. Lett.* **8**(4), 725–729 (2011), <http://dx.doi.org/10.1109/LGRS.2010.2102333>.
3. F. Pacifici and F. D. Frate, "Automatic change detection in very high-resolution images with pulse-coupled neural networks," *IEEE Geosci. Remote Sens. Lett.* **7**(1), 58–62 (2010), <http://dx.doi.org/10.1109/LGRS.2009.2021780>.
4. L. Deren, "Remotely sensed images and GIS data fusion for automatic change detection," *Int. J. Image Data Fusion* **1**(1), 99–108 (2010).
5. M. D. Mura, J. A. Benediktsson, and F. Bovolo, "An unsupervised technique based on morphological filters for change detection in very high resolution images," *IEEE Geosci. Remote Sens. Lett.* **5**(3), 433–437 (2008), <http://dx.doi.org/10.1109/LGRS.2008.917726>.
6. F. Bovolo, L. Bruzzone, and M. Marconcini, "A novel approach to unsupervised change detection based on a semisupervised SVM and a similarity measure," *IEEE Trans. Geosci. Remote Sens.* **46**(7), 2070–2082 (2008), <http://dx.doi.org/10.1109/TGRS.2008.916643>.
7. S. Ghosh et al., "A context-sensitive technique for unsupervised change detection based on hopfield-type neural networks," *IEEE Trans. Geosci. Remote Sens.* **45**(3), 778–789 (2007), <http://dx.doi.org/10.1109/TGRS.2006.888861>.
8. F. Bovolo and L. Bruzzone, "A split-based approach to unsupervised change detection in large-size multitemporal images: application to tsunami-damage assessment," *IEEE Trans. Geosci. Remote Sens.* **45**(6), 658–670 (2007), <http://dx.doi.org/10.1109/TGRS.2007.895835>.
9. L.-Z. Xu et al., "Feature detector model for multi-spectral remote sensing images inspired by insect visual system," *Acta electronica sinica.* **39**(11), 2497–2501 (2011).

10. L. Bruzzone and L. Carlin, "A multilevel context-based system for classification of very high spatial resolution images," *IEEE Trans. Geosci. Remote Sens.* **44**(9), 2587–2600 (2006), <http://dx.doi.org/10.1109/TGRS.2006.875360>.
11. P. Gong, X. Li, and B. Xu, "Research issues with high-resolution image interpretation theory and application methods," *J. Remote Sens.* **10**(1), 1–5 (2006).
12. L. Durieux, E. Lagabrielle, and A. Nelson, "A method for monitoring building construction in urban sprawl areas using object-oriented fusion of Spot 5 images and existing GIS data," *ISPRS J. Photogramm. Remote Sens.* **63**(4), 399–408 (2008), <http://dx.doi.org/10.1016/j.isprsjprs.2008.01.005>.
13. L.-Z. Xu, X.-F. Li, and S. X. Yang, "Intelligent information processing and system optimization," *Intell. Autom. Soft Comput.* **17**(7), 829–833 (2011), <http://dx.doi.org/10.1080/10798587.2011.10643191>.
14. O. Miller, A. Pikaz, and A. Averbuch, "Objects based change detection in a pair of gray-level images," *Pattern Recognit.* **38**(11), 1976–1992 (2005), <http://dx.doi.org/10.1016/j.patcog.2004.07.010>.
15. A. Lefebvre, T. Corpetti, and L. H. Moy, "Object-oriented approach and texture fusion for change detection in very high-resolution images," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium, Boston*, Vol. 6–11, pp. 663–666 (2008).
16. G. Chen et al., "Object-based change detection," *Int. J. Remote Sens.* **33**(14), 4434–4457 (2012), <http://dx.doi.org/10.1080/01431161.2011.648285>.
17. A. Shi, L. Xu, and F. Xu, "Multispectral and panchromatic image fusion based on improved bilateral filter," *J. Appl. Remote Sens.* **5**(1), 1–17 (2011), <http://dx.doi.org/10.1117/1.3616010>.
18. C. R. Neil and L. L. Grewe, "Change detection for linear features in aerial photographs using edge-finding," *IEEE Trans. Geosci. Remote Sens.* **39**(7), 1608–1612 (2001), <http://dx.doi.org/10.1109/36.934092>.
19. F. Huang et al., "Insect visual system inspired small target detection for multi-spectral remotely sensed images," *J. Commun.* **32**(9), 88–95 (2011).
20. O. Hall and J. Hay, "A multiscale object-specific approach to digital change detection," *Int. J. Appl. Earth Obs. Geoinf.* **4**(4), 311–327 (2003), [http://dx.doi.org/10.1016/S0303-2434\(03\)00010-2](http://dx.doi.org/10.1016/S0303-2434(03)00010-2).
21. G. Chen et al., "A multiscale geographic object-oriented image fusion (GEOBIA) to estimate lidar-measured forest canopy height using Quickbird image," *Int. J. Geogr. Inf. Sci.* **25**(6), 877–893 (2011), <http://dx.doi.org/10.1080/13658816.2010.496729>.
22. G. J. Hay et al., "An automated object-oriented approach for the multiscale image segmentation of forest scenes," *Int. J. Appl. Earth Obs. Geoinf.* **7**(4), 339–359 (2005), <http://dx.doi.org/10.1016/j.jag.2005.06.005>.
23. Y. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(8), 800–810 (2001), <http://dx.doi.org/10.1109/34.946985>.
24. Y. Song and J. Shan, "Building extraction from high-resolution color image based on edge flow driven active contour and JSEG," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.* **37**(28), 185–190 (2008).
25. T. Liu et al., "Region segmentation of multi-spectral remote sensing images using improved JSEG algorithm," *J. Remote Sens.* **13**(1), 30–34 (2009).
26. Y. Deng et al., "Peer group filtering and perceptual color image quantization," in *Proc. of the 1999 IEEE Int. Symp. on Circuits and Systems* Vol 4, pp. 21–24 (1999).
27. Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Process.: Image Commun.* **19**(2), 121–132 (2004), [http://dx.doi.org/10.1016/S0923-5965\(03\)00076-6](http://dx.doi.org/10.1016/S0923-5965(03)00076-6).
28. A. P. Dempster, "Upper and low probabilities induced by a multi-valued mapping," *Ann. Math. Stat.* **38**(2), 325–339 (1967), <http://dx.doi.org/10.1214/aoms/1177698950>.
29. L. Bruzzone and D. F. Prieto, "Automatic fusion of difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.* **38**(3), 1171–1182 (2000), <http://dx.doi.org/10.1109/36.843009>.

30. O. Hall and G. J. Hay, "A multiscale object-specific approach to digital change detection," *Int. J. Appl. Earth Obs. Geoinf.* **4**(4), 311–327 (2003), [http://dx.doi.org/10.1016/S0303-2434\(03\)00010-2](http://dx.doi.org/10.1016/S0303-2434(03)00010-2).



Chao Wang received his master's degrees in communication and signal processing from China University of Mining and Technology, China, in 2010. He is currently a Ph.D. candidate in computer science at Hohai University, China. His major research interests include remote sensing image processing and pattern recognition.



Mengxi Xu received her MSc degree from UNESCO-IHE in the Netherlands in 2007. Now she is a lecturer in the School of Computer Engineering at Nanjing Institute of Technology. Her current research interests include image processing, information processing system and its application.



Xin Wang is in the College of Computer and Information, Hohai University, China. She obtained her Ph.D. degree from Nanjing University of Science and Technology of China in 2010. Her research interests include image processing, pattern recognition, and computer vision.



Shengnan Zheng graduated from HOHAI University, Dept. of Electronic Information Engineering & Technology. Research direction is signal and information process (master's) in 2010, and orientation of research is digital image processing. She now is an assistant engineer of computer engineering at Nanjing Institute of Technology.



Zhenli Ma is an assistant professor and a tutor of master's students. Now she is working on teaching and researching projects about information and communication engineering.