

# Fuzzy Clustering With a Modified MRF Energy Function for Change Detection in Synthetic Aperture Radar Images

Maoguo Gong, *Member, IEEE*, Linzhi Su, Meng Jia, and Weisheng Chen, *Member, IEEE*

**Abstract**—In this paper, we put forward a novel approach for change detection in synthetic aperture radar (SAR) images. The approach classifies changed and unchanged regions by fuzzy c-means (FCM) clustering with a novel Markov random field (MRF) energy function. In order to reduce the effect of speckle noise, a novel form of the MRF energy function with an additional term is established to modify the membership of each pixel. In addition, the degree of modification is determined by the relationship of the neighborhood pixels. The specific form of the additional term is contingent upon different situations, and it is established ultimately by utilizing the least-square method. There are two aspects to our contributions. First, in order to reduce the effect of speckle noise, the proposed approach focuses on modifying the membership instead of modifying the objective function. It is computationally simple in all the steps involved. Its objective function can just return to the original form of FCM, which leads to its consuming less time than that of some obviously recently improved FCM algorithms. Second, the proposed approach modifies the membership of each pixel according to a novel form of the MRF energy function through which the neighbors of each pixel, as well as their relationship, are concerned. Theoretical analysis and experimental results on real SAR datasets show that the proposed approach can detect the real changes as well as mitigate the effect of speckle noises. Theoretical analysis and experiments also demonstrate its low time complexity.

**Index Terms**—Fuzzy clustering, image change detection, Markov random field (MRF), synthetic aperture radar (SAR).

## I. INTRODUCTION

IMAGE change detection, which means detecting regions of change in images of the same scene taken at different times, is of widespread interest due to a large number of applications in diverse disciplines, such as remote sensing [1]–[9], medical diagnosis [10], [11], and video surveillance [12], [13]. Especially, when a natural catastrophe strikes, an effective and efficient change detection task appears critical when lives and properties

are at stake. The images that are generated by synthetic aperture radars (SARs) are of great use due to their independence of atmospheric and sunlight conditions; therefore, they have become valuable and indispensable sources of information in change detection.

Generally speaking, change detection in SAR images is the process of the analysis of two coregistered SAR images that are acquired over the same geographical area at different times. Such analysis is unsupervised when it aims to discriminate between two opposite classes (which represent unchanged and changed areas) with no prior knowledge about the scene (i.e., no ground truth is available to model the classes) [3].

As is mentioned in [1], the procedure of change detection in SAR images can be divided into three steps: 1) image preprocessing; 2) generation of a difference image (DI) from multitemporal images; and 3) analysis of the DI. Geometric correction and registration are usually implemented in the first step to align two images in the same coordinate frame. Here, we emphasize on the second and third steps.

It is important to cope with speckle noise. Were there no or little noise pollution in the two SAR images, it would be fairly easy for the change detection process. Unfortunately, however, the way of originating SAR images is so special that they are usually corrupted by speckle noise, and its existence makes it rather difficult to discern the two classes. Therefore, utilizing a relatively primary approach will not be capable of undertaking the analysis so well. To overcome the defect inherently characterizing in SAR images, in many literatures, researchers have tried to utilize different kinds of algorithms to reduce the corruption of the speckle noise. In the DI-generation step, the log-ratio operator is often used because of its robustness and nonsensitivity to speckle noise [14]. In addition, the DI-analysis step in fact can be looked on as the process of image segmentation, and two conventional methods, the threshold method and the clustering method, have been widely used. In the threshold method, some essential models are usually established to search for the best threshold to divide the DI into two classes, and in the clustering method, we do not need to establish a model; therefore, it seems to be more convenient and feasible. One of the most popular clustering methods for image segmentation is the fuzzy c-means (FCM) algorithm, which can retain more image information than hard clustering in some cases.

In the standard FCM algorithm, a function that is related to the membership and dissimilarity is minimized in each iteration process, and the function is what is usually referred to as the objective function. Being able to retain more information from the

Manuscript received August 15, 2012; revised November 30, 2012; accepted January 31, 2013. Date of publication February 26, 2013; date of current version January 30, 2014. This work was supported by the National Natural Science Foundation of China under Grant 61273317, the National Top Youth Talents Support Program of China, and the Fundamental Research Fund for the Central Universities under Grant K50510020001 and Grant K5051202053.

The authors are with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, China (e-mail: gong@ieee.org; 412481579@qq.com; jiameng112@163.com; wshchen@126.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TFUZZ.2013.2249072

original image, FCM has robust characteristics for ambiguity. However, the standard FCM algorithm is very sensitive to noise since it considers no information about spatial context. In recent years, many researchers have incorporated the local spatial and local gray-level information into the original FCM algorithm to compensate for this defect of FCM [14]–[22]. Some novel approaches were developed, and the utilization of the spatial context was embodied basically in the modification of the objective function. In [15], Cai *et al.* proposed the fast generalized FCM algorithm (FGFCM) for image segmentation which incorporates the spatial information, the intensity of the local pixel neighborhood, and the number of gray levels in an image. FGFCM can significantly reduce the execution time by clustering on gray-level histogram rather than on pixels; meanwhile, it is less sensitive to noise to some extent because of the introduction of local spatial information. In [17], Krindis and Chatzis proposed a robust fuzzy local information C-means clustering algorithm (FLICM) for image segmentation. The characteristic of FLICM is the use of a fuzzy local similarity measure which is aimed to guarantee noise insensitiveness and image detail preservation. In particular, a novel fuzzy factor is introduced into its objective function to enhance the clustering performance. Recently, in [14], we proposed an improved FLICM to classify changed and unchanged regions of the change detection problem. The reformulated FLICM (RFLICM) incorporates the information about spatial context by adding a new fuzzy factor into its objective function for the purpose of enhancing the changed information and reducing the effect of speckle noise. It was emphasized on the modification of the objective function. In [18], another way of utilizing the spatial context was proposed for a spatiotemporal fuzzy-control system. In [18], the spatially constrained fuzzy c-means (SCFCM) method is presented by both adding some terms in the objective function and modifying the way to compute the clustering centers. The SCFCM method achieved some excellent results in its intended realm. Ju and Liu [19] applied FCM to human-hand motion recognition by introducing a new dissimilarity function. In [20], Coletta *et al.* offered some augmentation of FCM algorithms used to cluster distributed data by arriving at some constructive ways of determining essential parameters of the algorithms and forming a set of systematically structured guidelines. The objective function in [20] was also exhibited in a modified appearance by adding some terms with a few parameters involved.

A Markov random field (MRF) serves as an opportune tool to introduce information about the mutual influences among image pixels in a powerful and formal way. In [16], Chatzis and Varvarigou proposed a novel fuzzy objective function regularized by Kullback–Leibler divergence information. Their algorithm was facilitated by the application of a mean-field-like approximation of the MRF prior. In [21], Markov spatial constraint field and the fuzzy segmentation information resulting from FCM are fused. In [22], FCM with the MRF was applied in wavelet domain for image segmentation. Its label field of image was characterized by the MRF. The modified objective function with locally spatial constraint was introduced by the initial label of different scale wavelet coefficients.

Following the work mentioned above, in this study, we propose an SAR image change detection approach based on the FCM algorithm by adding the MRF with a novel form of energy function. The proposed approach (known as MRFFCM for short) does not improve FCM by modifying the objective function as in the aforementioned literature. Instead, it focuses on the modification of the membership to reduce the effect of speckle noise. It is of computational simplicity in all the steps involved, and its objective function can just return to the original form of FCM which leads to it being less time consuming than that of some obviously recently improved FCM algorithms. It modifies the membership of each pixel by introducing the information provided by the spatial context, i.e., the neighbors of the central pixel, as well as their interrelationship, are concerned in the process of using the MRF. What is more, theoretical analysis and experimental results show that the new approach not only does well in reducing the effect of speckle noise but has low time complexity as well.

The rest of this paper is organized as follows: Section II describes our motivation and the main procedure of the proposed approach. Section III describes the proposed novel energy function in the MRF. The datasets used are presented in Section IV, which also contains a description of the experiment setup as well as the evaluation criteria. Experimental results on real multitemporal SAR images are described in Section V. Finally, conclusions are drawn in Section VI.

## II. MOTIVATION AND PROCEDURE

As is mentioned in Section I, the process of change detection includes the generation of the DI and the analysis of it. Let us consider two coregistered SAR images:  $I_1 = \{I_1(h, l), 1 \leq h \leq A, 1 \leq l \leq B\}$ , and  $I_2 = \{I_2(h, l), 1 \leq h \leq A, 1 \leq l \leq B\}$ . Both of them are of size  $A \times B$  and are acquired by an SAR sensor over the same geographical area at two different times  $t_1$  and  $t_2$ . Then, the conventional log-ratio operator is applied to generate a DI, which is denoted by  $I_X = \{I_X(h, l), 1 \leq h \leq A, 1 \leq l \leq B\}$ . The next task is to analyze the DI.

### A. Motivation of Using FCM Modified by the Markov Random Field

The purpose of the DI analysis is to distinguish the changed class from the unchanged one, and the process actually belongs to the domain of image segmentation. In the realm of image segmentation, FCM is especially well-known and popular [14]–[26].

As mentioned in Section I, the MRF provides a basis to model information about the mutual influences among image pixels. A paramount important issue of the MRF is the energy function that directly characterizes the way to utilize spatial context. Considering the severe speckle noise in SAR images, the relationship among pixels is of more complexity than in some other kinds of images. Such complexity appears as two aspects: first, in the homogeneous region in the DI, outliers tend to disturb (even sabotage) the normal use of the energy function, and it is not easy to stem such corruption; second, in the heterogeneous region in the DI, an obscure boundary will emerge between two

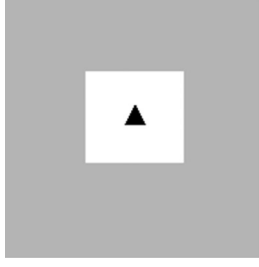


Fig. 1. Example to illustrate how the MRF deals with spatial context.

classes instead of an exact one. Therefore, enlightened by [16], [21], and [22], in order to reduce the effect of speckle noise, we propose a novel form of the energy function of the MRF to modify the membership of the FCM algorithm instead of modifying the objective function. Furthermore, the MRF can be utilized to consider the spatial context and, thus, to enhance the traditional FCM algorithm without engendering much time complexity. This can be demonstrated by a comparison of the time complexity. It shows that in the process of computing the objective functions in [14] or [17], the time complexity is  $o((2m + 4)N)$  due to the computation of the fuzzy factor  $G_i$ , and the whole time complexity is  $o((2m + 7)N)$ , the updating process included. Here,  $N$  and  $m$  denote the number of pixels in an image and in a neighborhood, respectively. However, the proposed MRFFCM is of computational simplicity in all the steps involved. The objective function can just return to its original form, and no complicated summation equation appears in the entire steps. Therefore, in the process of computing the objective function here, the time complexity appears as  $o(4N)$ . The disparity above demonstrates that MRFFCM will be less time-consuming than the improved FCM in [14]. Experimental results in Section V also demonstrate the low time complexity of MRFFCM.

### B. Brief Introduction to the Markov Random Field

In general, an image  $I = \{I(h, l), 1 \leq h \leq A, 1 \leq l \leq B\}$  itself can be viewed as a field, and each pixel of image is an element. If and only if some property of each element is only related to the neighborhood ones and is of no relationship to the other ones in the field, we call the random field  $p(x)$  an MRF. To emphasize the concept, let  $\partial$  be a neighborhood system lying on the field, and then, the previously considered random field  $p(x)$  is an MRF with respect to the introduced neighborhood system  $\partial_j$  if [27]

$$p(x_j | x_{I-\{j\}}) = p(x_j | x_{\partial_j}) \quad \forall j \in I \quad (1)$$

where  $j = (h, l)$  denotes the position of a certain element in the field, and  $x_{I-\{j\}}$  means the property of the whole elements in the field except the pixel  $j$ . Fig. 1 is used to illustrate the concept. In the figure, the black triangle represents the central element  $j$ , the white area means the neighborhood of  $j$  ( $\partial_j$ ), and the gray area is the rest part of the field. What (1) demonstrates is that some properties of the black triangle are not independent but rely on the field in which it locates. However, it is assumed that they depend only on the white area instead of on the other part of the field. Thus, the field can be viewed as an MRF.

The Hammersley–Clifford theorem [28] has proved that a given random field is an MRF if and only if its (joint) probability distribution  $p(x)$  is a Gibbs distribution, i.e., the joint distribution of an MRF is given by [29]

$$p(x|\beta) \triangleq Z(\beta)^{-1} \cdot \exp[-E(x|\beta)] \quad (2)$$

where  $Z(\beta)$  is a normalizing constant (the partition function) given by

$$Z(\beta) = \sum_{x \in \Omega} \exp[-E(x|\beta)]. \quad (3)$$

$E(x|\beta)$  is the energy function comprised of some clique potentials, and  $\beta$  is an adjusting parameter. The significance of the theorem is that it converts the abstract MRF expression into a computable expression.

From the theory mentioned above, the key to apply the MRF is to establish an appropriate energy function in a certain scenario. Some universal models, such as the Ising model, Potts model, MLL model, and quad-tree model, are widely applied to image processing. In the change detection task, just let  $\Omega = \{u, c\}$ , which stands for the label set of unchanged class and the changed class. In addition, we use the mean-field-like approximation serving as a simulation of the MRF, whose validity has been proved in [16].

### C. Main Procedure of Markov Random Field Fuzzy c-Means

The proposed MRFFCM improves FCM by modifying the membership of each pixel according to the MRF-based spatial context. The spatial context is contained in the pivotal energy function through the use of the neighborhood system. The information provided by the neighborhood pixels serves as the spatial context, and from Fig. 1, it is considered of relevance to the central pixel in the FCM progress. The energy function is established, and then, the Gibbs expression in (2) is computed to generate the so-called pointwise prior probability before updating the membership, and thus, the aim to deal with spatial context is reached by adding the MRF into FCM.

The main procedure of MRFFCM is as follows.

- 1) In the first iteration ( $k = 1$ ), derive the mean  $\mu_i^1$  and the standard deviation  $\sigma_i^1$  of the two classes through the K&I method introduced in [3]. In addition, the initial membership matrix  $\{u_{ij}^1\}$  is generated by utilizing the original FCM algorithm unmodified ( $i = u, c$ ). Then, by means of hard division (the threshold of which is 0.5), generate the same-kind-number matrix  $\{n_{i \in \partial_j}^1\}$ , and each element of the matrix denotes the number of the neighborhood pixels belonging to  $i$ .
- 2) In the  $k$ th iteration, establish the energy matrix  $\{E_{ij}^k\}$ . This step is the key step to utilize the spatial context. In Section III, we will introduce a novel energy function with an additional term.
- 3) Using Gibbs expression, compute the pointwise prior probabilities of the MRF, and get the pointwise prior

probability matrix  $\{\pi_{ij}^k\}$

$$\pi_{ij}^k = \frac{\exp(-E_{ij}^k)}{\exp(-E_{uj}^k) + \exp(-E_{cj}^k)}. \quad (4)$$

- 4) Compute the conditional probability  $\{p_i^k\}$  that is given by (5), shown below, and then, generate the distance matrix  $\{d_{ij}^k\}$  that is given by (6), shown below:

$$p_i^k(y_j | \mu_i^k, \sigma_i^k) = \frac{1}{\sigma_i^k \sqrt{2\pi}} \exp \left[ -\frac{(y_j - \mu_i^k)^2}{2(\sigma_i^k)^2} \right] \quad (5)$$

$$d_{ij}^k = -\ln [p_i^k(y_j | \mu_i^k, \sigma_i^k)]. \quad (6)$$

- 5) Compute the objective function  $J_{ij}^k$  that is given by (7), shown below, where  $I_X$  denotes the DI generated by the log-ratio operator. In case of convergence (which is presented in (8), shown below, where  $\delta$  is the convergence threshold), exit and output  $\{u_{ij}^k\}$ ; otherwise, go to 6. It is worth noting that in [16], another objective function with an artificial parameter is used. We conceive that function, but find it hard to search for an appropriate value of the parameter. Therefore, the original FCM objective function is used

$$J_{ij}^k = \sum_{i=u,c} \sum_{j \in I_X} (u_{ij}^k)^2 (d_{ij}^k)^2 \quad (7)$$

$$|J_{ij}^k - J_{ij}^{k-1}| \leq \delta. \quad (8)$$

- 6) Compute the new membership that is given by (9), generating the new membership matrix  $\{u_{ij}^{k+1}\}$ , which is to be used in the next iteration process

$$u_{ij}^{k+1} = \frac{\pi_{ij}^k \exp(-d_{ij}^k)}{\pi_{uj}^k \exp(-d_{uj}^k) + \pi_{cj}^k \exp(-d_{cj}^k)}. \quad (9)$$

- 7) Update the mean and the standard deviation as  $\mu_i^{k+1}$  and  $\sigma_i^{k+1}$ , respectively, as is given in (10), shown below.  $k := k + 1$ . Then, return to 2:

$$\begin{cases} \mu_i^{k+1} = \frac{\sum_{j \in I_X} (u_{ij}^k y_j)}{\sum_{j \in I_X} (u_{ij}^k)} \\ \sigma_i^{k+1} = \sqrt{\frac{\sum_{j \in I_X} [u_{ij}^k (y_j - \mu_i^{k+1})^2]}{\sum_{j \in I_X} (u_{ij}^k)}} \end{cases} \quad (10)$$

### III. DESCRIPTION OF THE IMPROVED ENERGY FUNCTION

From what has been described above, it is crucial to establish a proper energy function by which the pointwise prior probabilities are of high accuracy. It serves as an operator that utilizes the information provided by the neighborhood pixels. As is mentioned above, a simple form of energy function usually means an inadequate use of the information, and thus, the result will be affected by the survived noise after the log-ratio operator. Only by discovering some further relationship, we can make adequate use of the information, and stem the impact of noise. It is worth

noting that the utilization entails that we enhance the benefit of the information and hold back the defect of that. In practice, several kinds of information are used to restrain the drawbacks of each other to the greatest extent.

In this section, a novel form of energy function with an additional term will be put forward. We omit the superscript  $k$  that denotes the number of iteration if no confusion is generated.

#### A. Appearance of the Energy Function in the Markov Random Field Fuzzy c-Means Algorithm

As stated in Section II, in order to make full use of the relationship among neighbors, we need to find a robust energy function to undertake the analysis of a change detection task for SAR images. The aim here is to make full use of the information provided by the neighborhood pixels.

We select the most commonly used second-order (the adjacent eight neighbors) neighborhood system in the following discussion. From (9), we get the conclusion that the pointwise prior probability is of positive correlation to the new membership, which indicates that we are able to modify the new membership by modifying the pointwise prior probability indirectly.

Since we view the mean-like field as an approximation of the MRF, considering the structure of (4), the simplest and the most convenient approach is to get the mean of the membership of all of the neighbors

$$E_{ij} = -\ln(mu_{ij}) \quad (11)$$

where

$$mu_{ij} = \text{mean}_{m \in \partial_j} \{u_{im}\}. \quad (12)$$

Obviously, such an operator utilizes the membership of the neighbors and reduces the effect of noise surviving after the generation of the DI. Nevertheless, there still exist two main drawbacks. First, in the homogeneous region, the result tends to be influenced by some outliers (i.e., noise pixels). Second, in the heterogeneous region, the accuracy of the edge will likely to decline because of the uncertainty of neighbors. Such drawbacks emerge because of the inappropriate and inadequate utilization of the information provided by neighbors. From (11) and (12), such information merely contains the memberships without a further consideration. In order to overcome such defects, we change (12) into

$$mu_{ij} = \text{mean}_{m \in \partial_j} \{u_{im} | x_j, x_m = i\}. \quad (13)$$

It can be seen from (13) that only the memberships of the pixels that belong to the same class as the central one are utilized. Hence, by adding the information of classification, the error engendered by the other pixels will be reduced. Accordingly, it holds back the adverse factor to some extent. However, still in the homogeneous region, the result will be affected by some erratic pixel (which can be also called outliers) whose membership is close to the threshold of hard division (i.e., 0.5). To get a better and more accurate result, an additional term  $A_{ij}$ , which is anticipated to adjust the mean membership automatically, is



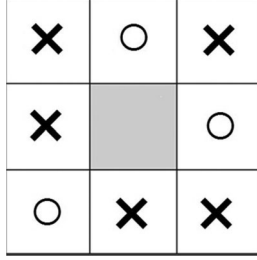


Fig. 2. Certain neighborhood in the  $\{mu_{u,j}\}$  matrix where the gray part is the central pixel; “x” and “o” denote the two classes after the hard division process. In the neighborhood, the number of the pixels labeled by “x” is 5, while the number of the pixels labeled by “o” is 3. If the central pixel is labeled by “x,” then we calculate the mean of the pixels labeled by “x” as  $mu_{u,j}$ , and then,  $mu_{c,j} = 1 - mu_{u,j}$ . If the central pixel is labeled by “o,” then we calculate the mean of the pixels labeled by “o” as  $mu_{c,j}$ , and then,  $mu_{u,j} = 1 - mu_{c,j}$ .

added to (11):

$$E_{ij} = -\ln(mu_{ij}) + A_{ij}. \quad (14)$$

$A_{ij}$  is to decide whether the neighborhood system belongs to a homogeneous region or a heterogeneous region according to the numbers. It serves as a modifier to modify the membership better. When  $n_{i \in \partial j} < 4$ , it can adjust the membership to 0.5 in different degrees; when  $n_{i \in \partial j} > 4$ , it can adjust the membership to 0 or 1 in different degrees; when  $n_{i \in \partial j} = 4$ , there will be no adjustment, and it is high time that we preserved the primordial membership. Except the last situation, the degree is determined according to the different value of  $n_{i \in \partial j}$ . This way, further information (the same class number) is discovered and utilized.

In practice, we need two mean matrix  $\{mu_{u,j}\}$  and  $\{mu_{c,j}\}$  to compute the energy, but at the same time, we notice that  $n_{u \in \partial j}$  is always equivalent to  $n_{c \in \partial j}$ , and thus, the function of  $A_{ij}$  will dwindle (even futile). Therefore, again we convert (13) into another form in which the mean membership representing the other class can be calculated directly. This way, (13) is altered as

$$\begin{cases} mu_{ij} = \text{mean}_{m \in \partial_j} \{u_{im} | x_m = u\} \\ mu_{pj} = 1 - mu_{ij} \end{cases}, \quad \text{if } x_j = i \quad (15)$$

where  $p \in \Omega - \{i\}$ .

After the conversion, the following equation can be obtained:

$$n_{u \in \partial j} + n_{c \in \partial j} = 8. \quad (16)$$

It is worth noting that (16) only applies to the pixels locating in the central part of the image. When the pixel is located in the four edges of the image, the sum number is 5, while when the pixel is located in the four corners of the image, the sum number is 3. In practice, we change the sum numbers according to where the pixel locates.

Fig. 2 shows an example to illustrate the above.

### B. Establishment of the Additional Term With the Least-Square Method

Here, a multiplicative model is established to comprise  $A_{ij}$ , as is shown in

$$A_{ij} = \beta_j \cdot t_{qj} \cdot n_{ij} \quad (17)$$

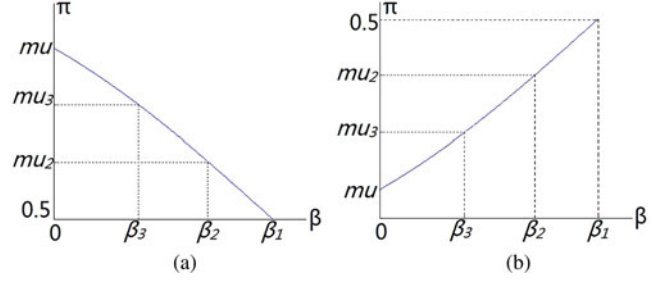


Fig. 3. Schematic curves of the scenarios where  $n = 1, 2$ , and  $3$ . (a) Situation where  $mu > 0.5$ . (b) Situation where  $mu < 0.5$ .

where

$$q = x_j. \quad (18)$$

$t_{qj}$  adjusts the sign of term that is only concerned with the class that the central pixel belongs to and the location of it. It is defined as

$$t_{qj} = \text{sgn}(u_{qj} - 0.5). \quad (19)$$

$n_{ij}$  is short for  $n_{i \in \partial j}$ .  $\beta_j$  is a parameter that is not concerned with  $i$ , as the two classes are to be adjusted to the same degree. However,  $\beta_j$  is not an artificial parameter, and it is established by the relationship between  $\beta_j$  and  $n_{ij}$  and  $mu_{ij}$ . Note that here, we only consider the pixels in the central part of the image, and the analysis of the pixels located in the edges and corners is very similar to the ones in the central.

Upon substituting (14)–(17) into (4), we obtain

$$\pi_{ij} = \frac{1}{1 + \frac{1 - mu_{ij}}{mu_{ij}} \cdot \exp[-t_{qj} \cdot (8 - 2 \cdot n_{ij}) \cdot \beta_j]}. \quad (20)$$

Evidently, we can view the pointwise prior probability  $\pi_{ij}$  as a function of the independent variable  $\beta_j$  with two parameters:  $n_{ij}$  and  $mu_{ij}$ . As when  $n_{ij} = 0$ , the central pixel can be viewed as an outlier; in the following discussion, we only consider the other eight values (1, 2, 3, 4, 5, 6, 7, and 8).

After a further study of the function  $\pi_{ij}(\beta_j | n_{ij}, mu_{ij})$ , we are able to draw the curve in eight different scenarios each of which contains two situations that are shown in the following figures. We will divide these scenarios into three groups. Our task is to get an ideal value of  $\beta_j$  in the light of  $n_{ij}$ . In the following discussion, we omit the subscripts  $i$  and  $j$  to avoid confusion.

1) *Scenarios Where  $n = 1, 2$ , and  $3$* : In these scenarios, the schematic curves are presented as two different appearances according to two different situations, as shown in Fig. 3.

Here, several parameters are set in Fig. 3, as in

$$\begin{cases} mu_n = 0.5 + \frac{(mu - 0.5) \cdot (n - 1)}{3}, & mu > 0.5 \\ mu_n = mu + \frac{(0.5 - mu) \cdot (4 - n)}{3}, & mu < 0.5 \end{cases} \quad (21)$$

where  $n = 1, 2$ , and  $3$ . The setting of the parameters actually presents a linear modification perspective in which we divide the closed interval  $[0.5, mu]$  or  $[mu, 0.5]$  lying on the  $\pi$ -axis into three tantamount sections by setting  $mu_n$  through (21). After

that, we calculate  $\beta_n$  in the light of the inverse function  $\beta = \beta^{-1}(\pi)$ .

Here,  $mu_n$  serves as a modification of the mean probability according to different values of  $n$  (here,  $n$  is of value 1, 2, or 3), and the modified result should be close to 0.5 in different degrees. When  $n = 1$ , we can suspect that the central pixel is polluted by noise; therefore, we modify the probability value to  $mu_1$  (whose value is 0.5), which both manifests our suspicion and preserves the possibility that it belongs to the class it appears to belong to (by comparing  $mu$  with 0.5 directly). As for the situation where  $n = 2$  or 3, we are not so fully convinced whether the central pixel is polluted by noise or lies in a heterogeneous region as in the scenario where  $n = 1$ ; therefore, instead of modifying the probability that much, we preserve the mean probability to two different degrees, which may serve to an eclectic result. It is worth noting that for simplicity, we draw the three groups of value together in Fig. 3 that contains two merely schematic curves, and it does not purport that we obtain the three groups of values in a fixed curve. As is demonstrated in (20), when  $n$  is of different values, the curve is also different.

We can divide the closed interval  $[0, 1]$  into say 30 equivalent segments, and at intervals of  $1/30$ , we get a value serving as a possible value of  $mu$ . Then, for every  $n$  valuing 1, 2, and 3, 21 functions are obtained, and 21 values of  $\beta_n$  can be selected in the way mentioned above. Therefore, all together we get a table of the size  $31 \times 3$  with 63 elements in it. For each  $n$  with a different value, we are able to find the relationship between  $\beta_n$  and  $mu$  with the well-known least-square method (LSM). Thus, for each  $mu$  ranging from 0 to 1 and each  $n$  ranging from 1 to 3, a binary function can be established:

$$\beta = \beta(mu, n), 0 \leq mu \leq 1, n = 1, 2, 3. \quad (22)$$

It is worth noting that the interval  $[0, 1]$  can also be divided into ten segments (at intervals of 0.1), or 50 segments (at intervals of 0.02), etc. Let  $R$  be the number of the equivalent segments in the following experiments, and the size of the table is  $(R + 1) \times 3$ . Besides, the practice utilization of the LSM is of diversity, such as linear fitting, polynomial fitting, exponential fitting, etc. A further study suggests that it is more accurate to use the LSM at intervals of  $[0, 0.5]$  and  $[0.5, 1]$  than in the entire interval  $[0, 1]$ . This is because of the different appearance of the curve in two situations.

In practice, the process of the LSM is undertaken before the change detection task. Having obtained the binary function established by the LSM, we are able to compute the energy to each pixel.

2) *Scenario Where  $n = 4$* : From (20), we can get the knowledge that when  $n = 4$ , the function degenerates to a constant function, and Fig. 4 itself also shows this.

In this scenario, the parameter  $\beta$  loses the effect to adjust, and it is also the scenario we have anticipated. If  $n = 4$ , we deem that the central pixel locates in a heterogeneous region, and it is better to preserve the primordial mean probability. Since the value of pointwise probability has nothing to do with the value of  $\beta$ , it can be set as any value.

3) *Scenarios Where  $n = 5, 6, 7$ , and 8*: Similar to the scenarios, where  $n = 1, 2$ , and 3, first, we use the schematic curves

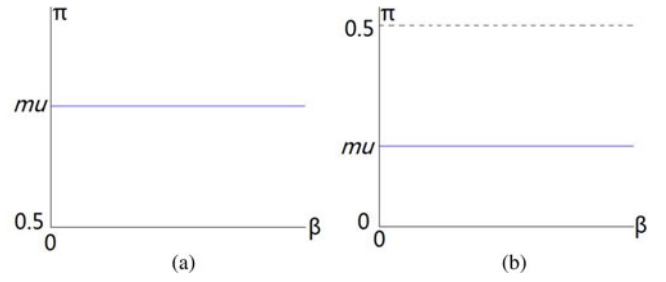


Fig. 4. Schematic curves of the scenarios where  $n = 4$ . (a) Situation where  $mu > 0.5$ . (b) Situation where  $mu < 0.5$ . Both of them are constant functions.

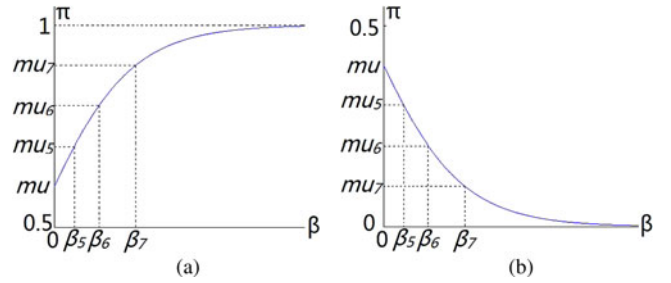


Fig. 5. Schematic curves of the scenarios where  $n = 5, 6, 7$ , and 8. (a) Situation where  $mu > 0.5$ . (b) Situation where  $mu < 0.5$ .

below to demonstrate the scenarios in two situations that are shown in Fig. 5.

Then, we obtain  $mu_n$  through

$$\begin{cases} mu_n = mu + \frac{(1 - mu) \cdot (n - 4)}{4}, & mu > 0.5 \\ mu_n = \frac{mu \cdot (8 - n)}{4}, & mu < 0.5 \end{cases} \quad (23)$$

where  $n = 5, 6, 7$ , and 8. In the same way above, we use the inverse function to get  $\beta_n$ . It is worth noting that the value of  $mu_8$  is 1 and 0 when  $mu > 0.5$  and  $mu < 0.5$ , respectively, and actually,  $\beta_8$  is the positive infinity; therefore, in practice, we only need to satisfy  $\beta_8 \gg \beta_7$ .

When  $n = 5, 6$ , or 7, the central pixel tends to locate in a homogeneous region, and the probability should be modified into the two poles (0 and 1) in different degrees according to different values of  $n$ , which has been shown in Fig. 4. In addition, similarly, we get another table of the size  $(R + 1) \times 3$ . Again, by using LSM, we obtain

$$\beta = \beta(mu, n), 0 \leq mu \leq 1, n = 5, 6, 7. \quad (24)$$

### C. Final Form of the Novel Energy Function

To summarize, after the establishment of the nonartificial parameter  $\beta_j$ , the additional term  $A_{ij}$  in the energy function is also established. Then, the energy function is

$$E_{ij} = -\ln(mu_{ij}) + \beta_j(mu_{ij}, n_{ij}) \cdot t_{qj} \cdot n_{ij} \quad (25)$$

where the expression of  $mu_{ij}$  has been given by (15). It is worth noting that the data to calculate  $\beta_j$  with different  $mu_{ij}$  and  $n_{ij}$  through the LSM are stored before the task and that whenever

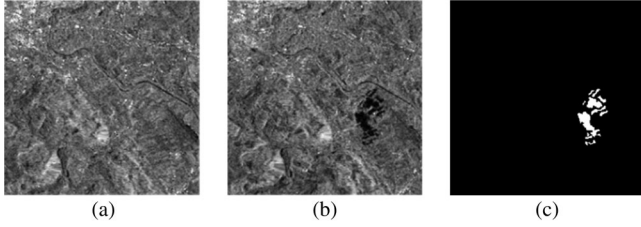


Fig. 6. Bern dataset. (a) Image acquired in April 1999. (b) Image acquired in May 1999. (c) Ground-truth image.

there is a requirement to calculate  $\beta_j$ , all we need to do is to access the memory and get the data generated in advance.

A qualitative analysis of time complexity in the entire process is made here. (25) can be calculated in only one loop; therefore, the time complexity of (25) is  $o(mN)$ . After adding the complexity concerning the objective function (7) and the updating formula (10), the entire complexity is  $o((m+7)N)$ . In addition, obviously, in the second-order neighborhood system,  $m = 8$ .

#### IV. DATASETS DESCRIPTION AND EXPERIMENTAL SETTINGS

In order to assess the effectiveness of the proposed approach, in this section, three datasets with different characteristics are considered in the experiments. We will also introduce the comparison experiments as well as some evaluation criteria of the experiment results.

##### A. Datasets

The first dataset is called the Bern dataset, which presents a section ( $301 \times 301$ ) of two SAR images acquired by the European Remote Sensing 2 satellite SAR sensor over an area near the city of Bern, Switzerland, in April and May 1999. Between the two dates, River Aare flooded parts of the cities of Thun and Bern and the airport of Bern entirely. Hence, the Aare valley between Bern and Thun was selected as a test site to detect flooded areas. The images and the available ground truth which is obtained by integrating prior information with photointerpretation are shown in Fig. 6.

The second dataset, the Ottawa dataset, is a section ( $290 \times 350$ ) of two SAR images over the city of Ottawa acquired by RADARSAT SAR sensor. They were provided by the Defence Research and Development Canada, Ottawa. This dataset contains two images acquired in May and August 1997 and presents the areas where they were once afflicted with floods. The images and the available ground truth which is created by integrating prior information with photo interpretation are shown in Fig. 7.

The third dataset used in the experiments is the Yellow River dataset that consists of two SAR images acquired by Radarsat-2 at the region of the Yellow River Estuary in China in June 2008 and June 2009. The original size of these two SAR images acquired by Radarsat-2 is  $7666 \times 7692$  as shown in Fig. 8(a) and (b). They are too huge to detail the information in such small pages, and it is of great difficulty to collect the integrating prior information in so large area; therefore, we select one representative area of size  $257 \times 289$  [see Fig. 8(c) and (d)] to compare the change detection results that are obtained by different ap-

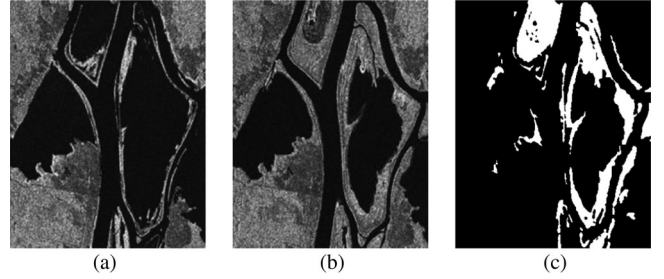


Fig. 7. Ottawa dataset. (a) Image acquired in May 1997. (b) Image acquired in August 1997. (c) Ground-truth image.

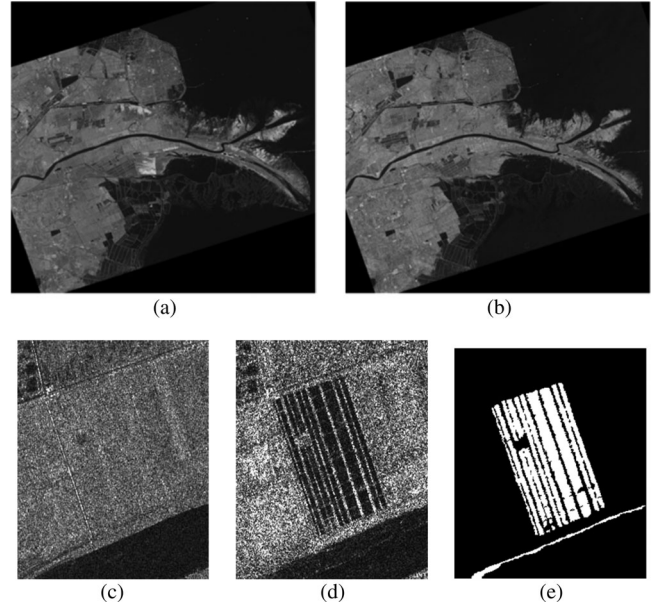


Fig. 8. Yellow River dataset. (a) Original image acquired in 2008. (b) Original image acquired in 2009. (c) Selected area in (a). (d) Selected area in (b). (e) Ground-truth image of the selected area.

proaches. The available ground truth of the selected area can be created through the integrating prior information without much difficulty, and it is shown in Fig. 8(e). It can be seen that the changed area appears as a newly reclaimed farmland and the borderline of the river. It is worth noting that the two images are a single-look image and a four-look image, which means that the influence of speckle noise on the image acquired in 2009 is much greater than that of the one acquired in 2008. Such huge disparity of speckle noise level between the two images may aggravate the difficulties met in the process of change detection.

##### B. Compared Algorithms and Parameter Test

In order to assess the effectiveness of the proposed change detection approach (written as MRFFCM), a parameter test and six compared experiments will be carried out.

In Section III, we divide the closed interval  $[0, 1]$  into  $R$  segments, and how the value of the parameter  $R$  influences the final result will be tested before the compared algorithms. This test also provides an appropriate value for the experiments to be carried out.



The first compared algorithm is the common FCM algorithm (serving as a fundamental one). This experiment is carried out to prove whether adding the information provided by the neighborhood pixels will yield better results.

The second and third compared algorithms are the FLICM and RFLICM algorithms that were introduced in [16] and [14], respectively. The two algorithms are characterized by their objective functions with no artificial parameter involved but in a complicated form. The two experiments are designed to demonstrate the low time complexity of MRFFCM.

The fourth, fifth, and sixth compared algorithms are about the energy function. In the second compared algorithm (written as MRFEM), we utilize (11) and (12) as the energy function to get the mean of the membership of the entire neighbors without considering whether the central pixel locates in a homogeneous region or a heterogeneous region. In the third compared algorithm (written as MRFSM), (11) and (13) are utilized, entailing that we consider the neighbors which belong to the same class as the central pixel. In the fourth compared algorithm (written as MRFN), the energy function below mentioned in [16] is considered:

$$E_{ij} = -r \sum_{m \in \partial_j} \delta(x_j - x_m) = -r \cdot n_{ij} \quad (26)$$

where  $r$  is a changing parameter in the iteration process and yields

$$r^{k+1} = \underset{r}{\operatorname{argmax}} \sum_{i=u,c} \sum_{j \in I} [u_{ij}^k \cdot \pi_{ij}^k(r)]. \quad (27)$$

These experiments are to demonstrate the superiority of the energy function in MRFFCM to those in the other algorithms.

### C. Evaluation Criteria

The results of the experiments are shown in two ways: One is to exhibit the final binary maps, and the other is to provide some values of criteria to evaluate the final map.

In the map of the available reference ground truth, we first count the number of the entire pixels  $N$ . Then, according to the map, we count the actual number of pixels belonging to the unchanged class and changed class (written as  $Nu$  and  $Nc$ , respectively). On the other hand, we compare the reference map with the one generated from a certain approach pixel by pixel. This way, we get the number of the pixels belonging to the unchanged class but falsely classified as changed class and of the pixels belonging to the changed class but falsely classified as unchanged class, and they are written as  $FP$  and  $FN$  for short, respectively. Another two number  $TP$  and  $TN$  are calculated as follows:

$$\begin{cases} TP = Nc - FN \\ TN = Nu - FP. \end{cases} \quad (28)$$

$TP$  and  $TN$  represent the number of changed pixels and unchanged pixels that are correctly detected, respectively.

To evaluate the result further, the number is not enough; therefore, we define the percentage correct classification ( $PCC$ )

as [30]

$$PCC = \frac{TP + TN}{N} \quad (29)$$

and it expresses the correct rate of the result roughly. As a matter of fact, due to the large value of  $N$ , in common cases, the values of  $PCC$  by different approaches are very similar, and it is difficult to observe the specific discrepancy through  $PCC$ . Therefore, overall error ( $OE$ ) is applied and is defined as

$$OE = FP + FN. \quad (30)$$

Moreover, as is mentioned in Section II, the analysis of the DI can be regarded as an image segmentation process; therefore, it is proper to use Kappa coefficient ( $KC$ ) as an overall evaluation criterion.  $KC$  is used to evaluate the effect of the result in the domain of image segmentation (here is the segmentation of the DI). The higher the value of  $KC$ , the better the segmentation result.  $KC$  is calculated as (30):

$$KC = \frac{PCC - PRE}{1 - PRE} \quad (31)$$

where

$$PRE = \frac{(TP + FP) \cdot Nc + (FN + TN) \cdot Nu}{N^2}. \quad (32)$$

The value of  $KC$  usually ranges from 0 to 1. It can be seen from (28), (29), (31), and (32) that  $KC$  depends on the dependent values of  $TP$  and  $TN$ , whereas  $PCC$  only depends on the sum value of  $TP$  and  $TN$ ; therefore,  $KC$  is a more cogent coefficient than  $PCC$  here because more detailed information of the classification is involved. In the following experiments, it can be found that the values of  $PCC$  or  $OE$  by some algorithms are very similar, but those of  $KC$  are of large discrepancy.

In addition, the time  $T$  elapsed in the whole process is also an important criterion. It is listed to compare the time complexity of different approaches. Here, the unit of time is the second.

The values of the six main criteria of evaluation ( $FN$ ,  $FP$ ,  $PCC$ ,  $OE$ ,  $KC$ , and  $T$ ) are organized in tabular form.

## V. EXPERIMENTAL RESULTS

### A. Test of the Parameter $R$

The first experiment is a test of  $R$ , which stands for the number of subintervals that the closed interval  $[0, 1]$  is divided into. It is selected as the test parameter because it is related to both the accuracy of the final result and the amount of storage. We want to know whether it is worthwhile to get a more accurate result by sacrificing the storage in the memory. In this experiment,  $R$  ranges from 10 to 100 and is of some discrete values. The three datasets are experimented on, and the deliberate and cogent criterion  $KC$  is employed here. The results are shown in Fig. 9.

It is quite clear from the three curves that when the value of  $R$  is less than 20, the value of  $KC$  varies conspicuously, whereas when it is larger than 30, the value nearly stays constant in both of the datasets. At the same time, it is noticed that the higher the value of  $R$ , the more the storage space occupied in the memory. Therefore, it is not necessary that  $R$  be selected as so high a value. Hence, upon a comprehensive analysis, it is appropriate



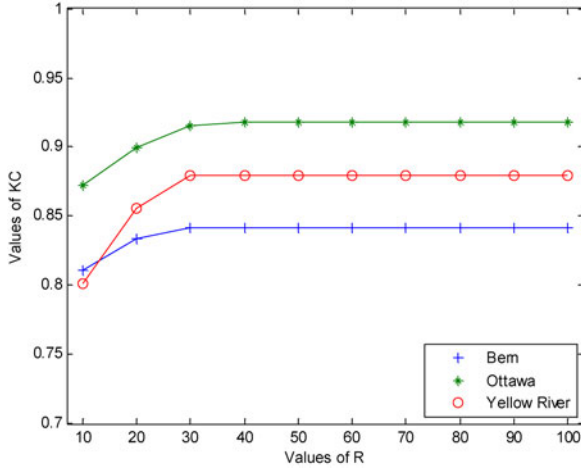


Fig. 9. Testing curve of the parameter  $R$  on the three datasets.

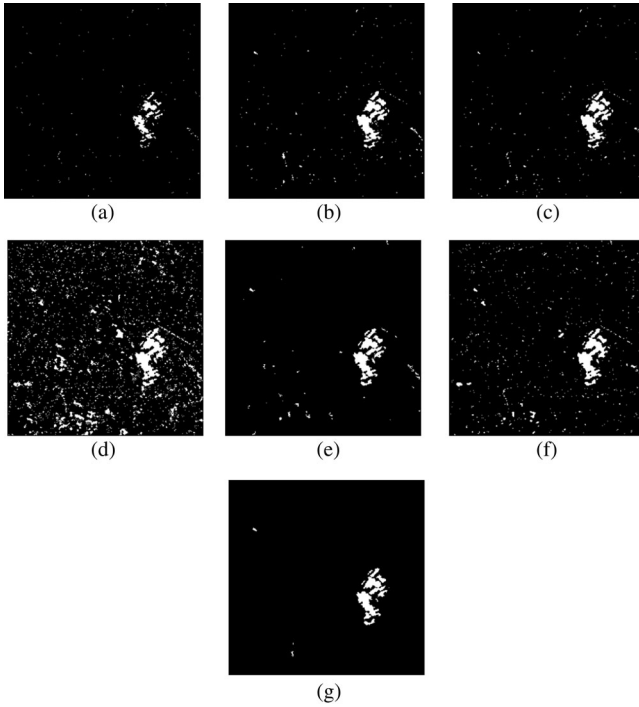


Fig. 10. Final maps of the Bern dataset generated by (a) FCM; (b) FLICM; (c) RFLICM; (d) MRFEM; (e) MRFSM; (f) MRFN; and (g) proposed MRFFCM.

to select  $R$  as a moderate value, say 30. The value of  $R$  in the experiments is 30.

### B. Results on the Bern Dataset

The results of the main experiment and the six comparison experiments are exhibited in two ways: The final maps in figure form and the criteria in tabular form. Fig. 10 shows the final maps of the experiments on the Bern dataset, and Table I lists the values for evaluation. In the table, the results of the novel approach are written in bold.

Fig. 10 and Table I show that for the original FCM, some vital changed regions are not detected (appearing as a high

TABLE I  
VALUES OF THE EVALUATION CRITERIA OF THE BERN DATASET

	$FP$	$FN$	$OE$	$PCC$	$KC$	$T/s$
FCM	190	349	539	0.9941	0.7464	10.6
FLICM	724	84	808	0.9911	0.8045	137.6
RFLICM	723	61	784	0.9913	0.8132	139.4
MRFEM	6390	26	6416	0.9292	0.2436	63.4
MRFSM	651	45	696	0.9923	0.7576	66.1
MRFN	1756	36	1792	0.9802	0.5471	54.9
<b>MRFFCM</b>	<b>364</b>	<b>47</b>	<b>411</b>	<b>0.9955</b>	<b>0.8413</b>	<b>73.9</b>

value of  $FN$ ). It proves that the use of the information provided by the neighbors positively affects the results in this scenario. The two methods, FLICM and RFLICM in which the objective function is modified, do generate better results than FCM due to their conspicuous high values of  $KC$ . The results are even better than those by MRFEM, MRFSM, and MRFN. However, the result of the proposed approach outperforms all the compared algorithms in both  $PCC$  and  $KC$ . In Fig. 10(d), many white noise spots that do not actually exist emerge on the black background, which is due to MRFEM in which all the memberships of the neighbors are counted. Therefore, the new membership should be considerably influenced by the adjacent points polluted by noise. Without using the pixels which do not belong to the same class as the central one, MRFSM and MRFN appear to be much better than MRFEM. These two experiments distinguish the useful information from the useless (even adverse) information to some extent. With an additional term, MRFFCM appears more effective as is seen from both the figure and the table. In Table I,  $FN$  and  $FP$  of MRFFCM do not exhibit the best, whereas  $PCC$  and  $KC$ , both of which serve as an overall evaluation, are the best. It is worth noting that although the values of  $PCC$  in FCM and MRFFCM are very similar, those of  $FN$  and  $OE$  differ remarkably. Besides, the values of  $KC$ , which serves as a more cogent criterion than  $PCC$ , are of large discrimination.

In addition, the computational times of FLICM and RFLICM are nearly twice as long as that of the methods in which the MRF are involved. These results demonstrate that applying the MRF to FCM is less time-consuming and verify our qualitative analysis of time complexity. The results indicate that the novel approach, indeed, is capable of improving detecting accuracy effectively and do not engender a high time complexity.

### C. Results on the Ottawa Dataset

The results of the experiments on the Ottawa dataset are shown and listed in Fig. 11 and Table II.

For the Ottawa dataset, Fig. 11 and Table II show that the proposed MRFFCM also obtains the best values of  $FP$ ,  $FN$ ,  $OE$ ,  $PCC$ , and  $KC$ , for it adds an additional term which can adjust the probability automatically in the light of  $n_{i \in \partial_j}$ . This number may be even more indispensable from the fact that the value of  $KC$  of MRFN is larger than that of MRFSM. In addition, the fact that the value of  $KC$  of MRFFCM is larger than that of MRFN illustrates that modifying probability by the number of pixels is also of effectiveness. To sum up, the proposed approach will fit

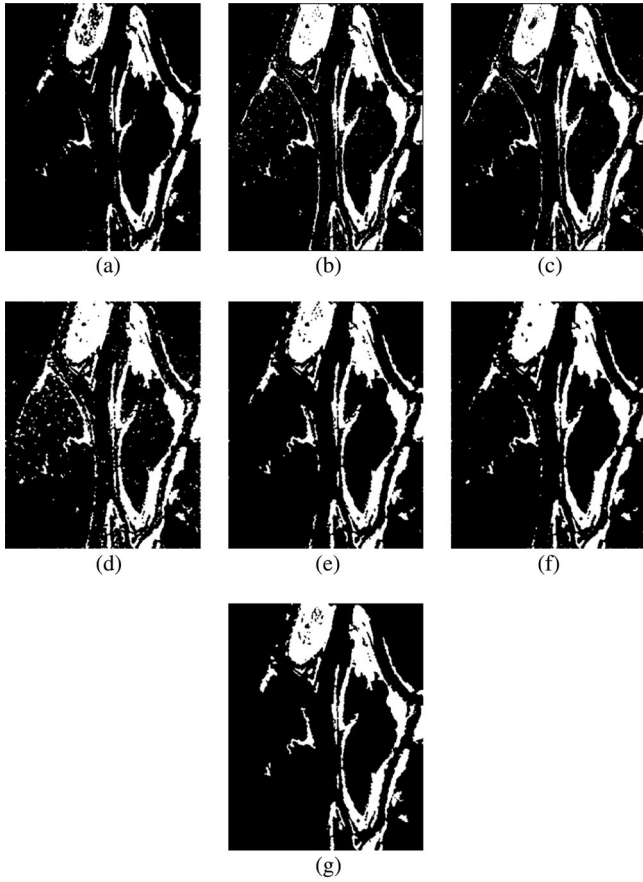


Fig. 11. Final maps of the Ottawa dataset generated by (a) FCM; (b) FLICM; (c) RFLICM; (d) MRFEM; (e) MRFSM; (f) MRFN; and (g) proposed MRFFCM.

TABLE II  
VALUES OF THE EVALUATION CRITERIA OF THE OTTAWA DATASET

	<i>FP</i>	<i>FN</i>	<i>OE</i>	<i>PCC</i>	<i>KC</i>	<i>T/s</i>
FCM	422	2319	2741	0.9730	0.8935	12.3
FLICM	2608	369	2977	0.9707	0.9052	156.7
RFLICM	2381	469	2850	0.9719	0.9075	161.3
MRFEM	5397	298	5695	0.9439	0.8133	77.0
MRFSM	2855	487	3342	0.9671	0.8833	77.9
MRFN	2642	414	3056	0.9699	0.8929	63.5
<b>MRFFCM</b>	<b>1636</b>	<b>712</b>	<b>2348</b>	<b>0.9769</b>	<b>0.9151</b>	<b>83.8</b>

the two situations where the changed area appears centralized (the Bern dataset) and scattered (the Ottawa dataset), and it also verifies its relatively far-ranging applicability.

#### D. Results on the Yellow River Dataset

Unlike the Bern or Ottawa dataset, the influence of speckle noise on the image acquired in 2009 is much greater than the one acquired in 2008 since the two images considered are single-look image and four-look image. It represents a more complicated situation to assess the effectiveness of the proposed approach. Moreover, there is no ground-truth image provided for the whole Yellow River Estuary dataset to quantitatively evaluate the effectiveness of the proposed change detection methods.

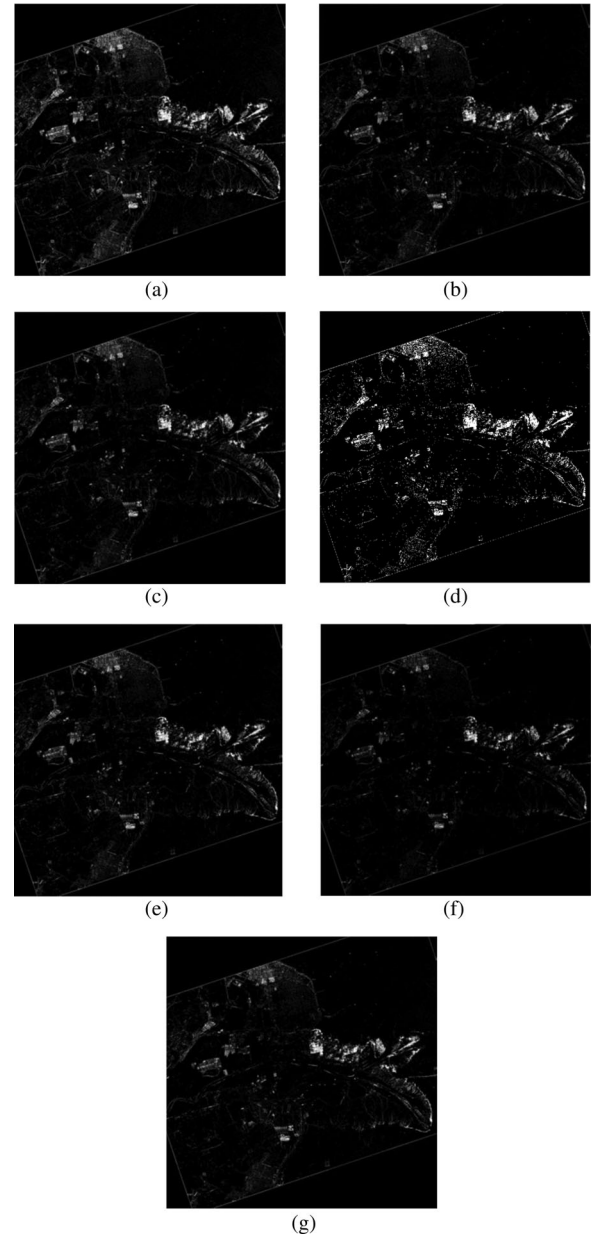


Fig. 12. Final maps of the whole Yellow River dataset generated by (a) FCM; (b) FLICM; (c) RFLICM; (d) MRFEM; (e) MRFSM; (f) MRFN; and (g) proposed MRFFCM.

Hence, only a visually analysis is available to analyze the change detection results. Fig. 12 illustrates the change detection results that are obtained by different approaches. Due to the large size of the original images ( $7666 \times 7692$ ), the detailed information in such small pages cannot be clearly exhibited, which is hardly able to verify the effectiveness of our proposed method. Therefore, the results of the selected area are given in Fig. 13 and Table III.

Fig. 12, Fig. 13, and Table III show that for the Yellow River dataset, many white spots appear in the final map generated by FCM, which is listed as a high value of *FP* in Table III. These white spots are eliminated to different degrees by modifying FCM. FLICM and RFLICM have reduced the value of *FP* to

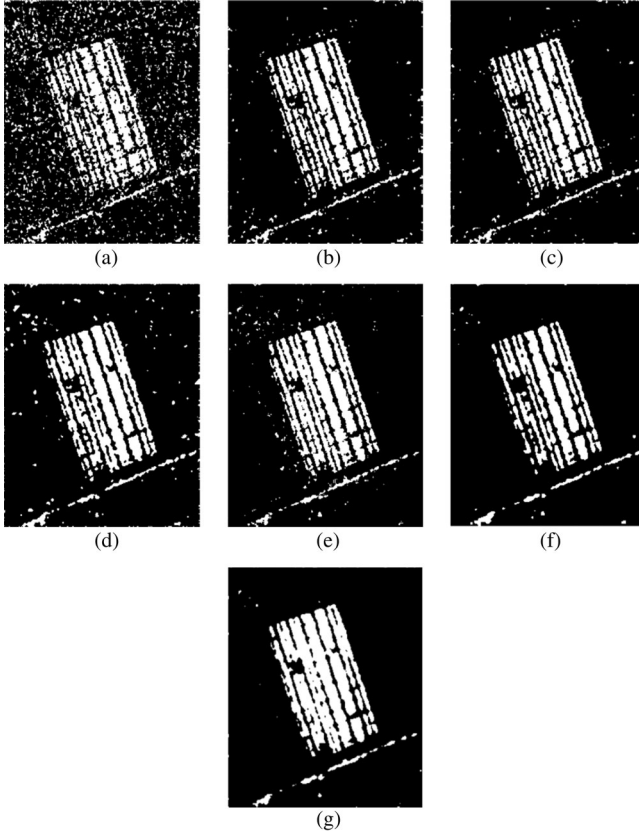


Fig. 13. Final maps of the selected area of the Yellow River dataset generated by (a) FCM; (b) FLICM; (c) RFLICM; (d) MRFEM; (e) MRFSM; (f) MRFN; and (g) proposed MRFFCM.

TABLE III  
VALUES OF THE EVALUATION CRITERIA OF THE SELECTED AREA OF THE  
YELLOW RIVER DATASET

	<i>FP</i>	<i>FN</i>	<i>OE</i>	<i>PCC</i>	<i>KC</i>	<i>T/s</i>
FCM	6914	3399	10313	0.8611	0.5747	8.39
FLICM	1415	2951	4366	0.9412	0.7695	110.4
RFLICM	1261	2947	4208	0.9433	0.7741	113.3
MRFEM	2602	2687	5289	0.9057	0.7391	50.9
MRFSM	2747	2645	5392	0.9105	0.7598	51.7
MRFN	1521	1651	3172	0.9463	0.8482	45.0
<b>MRFFCM</b>	<b>1260</b>	<b>1182</b>	<b>2442</b>	<b>0.9615</b>	<b>0.8791</b>	<b>59.9</b>

a large degree at the cost of a high time complexity (listed as great time consumption). The methods with the MRF are low in time complexity, and the proposed MRFFCM is the best, which verifies the effectiveness of the new algorithm.

## VI. CONCLUSION

In this paper, a novel change detection approach specifically toward the analysis of multitemporal SAR images has been presented. This approach is based on the universal utilized FCM algorithm and the MRF model. After generating the DI through the log-ratio operator, we add the MRF method in the procedure of FCM. The approach is of novelty in that the form of the energy function is altered by utilizing not only the memberships

but also the number of the same class of neighborhood pixels. Thus, we are able to decide whether the central pixel locates in a homogeneous region or a heterogeneous region. Great emphasis is put on the establishment of the additional term. By adding the term, the modified membership can be adjusted automatically according to both the membership and the number of the same class of neighborhood pixels. The well-known LSM is applied to get the relationship the membership and the adjusting parameter. The effectiveness is tested through four comparison experiments on three real SAR datasets.

This novel approach is basically built on the mathematical analysis such as the knowledge of elementary function and LSM. The new approach does not consider the use of any prior knowledge about the scene but consider only the use of the gray-level intensity and, therefore, is an unsupervised approach. In general, the main advantages of our change detection approach can be concluded as 1) superiority in reducing speckle noises (as the membership to each pixel is determined by two important factors comprehensively, and it has been validated by the experimental results), and 2) computational simplicity (the calculation of each step only includes the computation of some elementary functions, and theoretical analysis and experiments have demonstrated its low time complexity). In either scientific experiments or engineering practice, these advantages will play an indispensable role.

## ACKNOWLEDGMENT

The authors wish to thank the editors and anonymous reviewers for their valuable comments and helpful suggestions, which greatly improved the paper's quality.

## REFERENCES

- [1] L. Bruzzone and D. F. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multi-temporal remote-sensing images," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 452–466, Apr. 2002.
- [2] A. A. Nielsen, "The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 463–478, Feb. 2007.
- [3] Y. Bzai, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multi-temporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 874–887, Apr. 2005.
- [4] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2963–2972, Dec. 2005.
- [5] F. Bujor, E. Trounev, L. Valet, J. M. Nicolas, and J. P. Rudant, "Application of log-cumulants to the detection of spatiotemporal discontinuities in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2073–2084, Oct. 2004.
- [6] F. Chatelain, J.-Y. Tourneret, and J. Inglada, "Change detection in multisensor SAR images using bivariate gamma distributions," *IEEE Trans. Image Process.*, vol. 17, no. 3, pp. 249–258, Mar. 2008.
- [7] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432–1445, May 2007.
- [8] S. Marchesi, F. Bovolo, and L. Bruzzone, "A context-sensitive technique robust to registration noise for change detection in VHR multispectral images," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1877–1889, Jul. 2010.



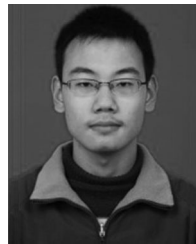
- [9] A. Robin, L. Moisan, and S. Hegarat-Masclé, "An a-contrario approach for subpixel change detection in satellite imagery," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 1977–1993, Nov. 2010.
- [10] M. Bosc, F. Heitz, J. P. Armspach, I. Namer, D. Gounot, and L. Rumbach, "Automatic change detection in multimodal serial MRI: Application to multiple sclerosis lesion evolution," *Neuroimage*, vol. 20, no. 2, pp. 643–656, Oct. 2003.
- [11] D. Rey, G. Subsol, H. Delingette, and N. Ayache, "Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis," *Med. Image Anal.*, vol. 6, no. 2, pp. 163–179, Jun. 2002.
- [12] D. M. Tsai and S. C. Lai, "Independent component analysis-based background subtraction for indoor surveillance," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 158–167, Jan. 2009.
- [13] S. S. Ho and H. Wechsler, "A martingale framework for detecting changes in data streams by testing exchangeability," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2113–2127, Dec. 2010.
- [14] M. Gong, Z. Zhou, and J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2141–2151, Apr. 2012.
- [15] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy C-means clustering algorithms incorporating local information for image segmentation," *Pattern Recogn.*, vol. 40, no. 3, pp. 825–838, Mar. 2007.
- [16] S. P. Chatzis and T. A. Varvarigou, "A fuzzy clustering approach toward hidden Markov random field models for enhanced spatially constrained image segmentation," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 5, pp. 1351–1361, Oct. 2008.
- [17] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1328–1337, May 2010.
- [18] X. Zhang, H. Li, and C. Qi, "Spatially constrained fuzzy-clustering-based sensor placement for spatiotemporal fuzzy-control system," *IEEE Trans. Fuzzy Syst.*, vol. 18, no. 5, pp. 946–957, Oct. 2010.
- [19] Z. Ju and H. Liu, "A unified fuzzy framework for human-hand motion recognition," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 5, pp. 901–913, Oct. 2011.
- [20] L. F. S. Coletta, L. Vendramin, E. R. Hruschka, R. J. G. B. Campello, and W. Pedrycz, "Collaborative fuzzy clustering algorithms: Some refinements and design guidelines," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 3, pp. 444–462, Jun. 2012.
- [21] S. Liu, X. Li, and Z. Li, "A new image segmentation algorithm based the fusion of Markov random field and fuzzy c-means clustering," in *Proc. IEEE Int. Symp. Commun. Inf. Technol.*, Oct. 2005, vol. 1, pp. 144–147.
- [22] X. Li and S. Bian, "Multiscale image segmentation using Markov random field and spatial fuzzy clustering in wavelet domain," in *Proc. Int. Workshop Intell. Syst. Appl.*, May 2009, pp. 1–6.
- [23] H. C. Huang, Y. Y. Chuang, and C. S. Chen, "Multiple kernel fuzzy clustering," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 1, pp. 120–134, Feb. 2012.
- [24] J. P. Mei and L. H. Chen, "A fuzzy approach for multitype relational data clustering," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 2, pp. 358–371, Apr. 2012.
- [25] E. Hullermeier, M. Rifqi, S. Henzgen, and R. Senge, "Comparing fuzzy partitions: A generalization of the rand index and related measures," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 3, pp. 546–556, Jun. 2012.
- [26] H. Cao, H.-W. Deng, and Y.-P. Wang, "Segmentation of M-FISH images for improved classification of chromosomes with an adaptive fuzzy c-means clustering algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 1, pp. 1–8, Feb. 2012.
- [27] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
- [28] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Roy. Statist. Soc. B*, vol. 36, pp. 192–236, 1974.
- [29] P. Clifford, "Markov random fields in statistics," in *Disorder in Physical Systems*, G. Grimmett and D. Welsh, Eds. Oxford, U.K.: Clarendon, 1990.
- [30] P. L. Rosin and E. Ioannidis, "Evaluation of global image thresholding for change detection," *Pattern Recogn. Lett.*, vol. 24, no. 14, pp. 2345–2356, 2003.



**Maoguo Gong** (M'07) received the B.S. degree in electronic engineering (with first-class Hons.) and the Ph.D. degree in electronic science and technology from Xidian University, Xi'an, China, in 2003 and 2009, respectively.

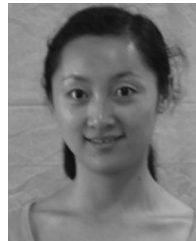
Since 2006, he has been a Teacher with Xidian University, where he is currently a Full Professor with the Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education. In 2008 and 2010, he was promoted as an Associate Professor and as a Full Professor, respectively, both with exceptive admission. His research interests include computational intelligence with applications. His current research interests include image segmentation and change detection.

Dr. Gong is a member of the IEEE Computational Intelligence Society, an Executive Committee Member of the Natural Computation Society of the Chinese Association for Artificial Intelligence, and a Senior Member of the Chinese Computer Federation. He received the New Century Excellent Talent in University of the Ministry of Education of China, the Eighth Young Scientist Award of Shaanxi, the New Scientific and Technological Star of Shaanxi Province, and the Science and Technology Award of Shaanxi Province (First Level, 2008 and 2010), etc.



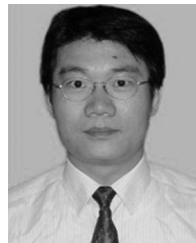
**Linzhi Su** received the B.S. degree in intelligence science and technology from Xidian University, Xi'an, China, in 2011, where he is currently working toward the Ph.D. degree.

His research interests include computational intelligence and image understanding.



**Meng Jia** received the B.S. degree in mechanical and electrical engineering and the M.S. degree in electronic engineering from Xidian University, Xi'an, China, in 2006 and 2009, respectively, where she is currently working toward the Ph.D. degree.

Her research interests include multitemporal remote sensing image analysis.



**Weisheng Chen** (M'09) received the B.Sc. degree from Qufu Normal University, Qufu, China, in 2000 and the M.Sc. and Ph.D. degrees from Xidian University, Xi'an, China, in 2004 and 2007, respectively.

From 2008 to 2009, he was a Visiting Scholar with the Automation School, Southeast University, Nanjing, China. Since 2009, he has been a Postdoctoral Fellow with the School of Electronic Engineering, Xidian University, where he is currently a Professor with the Department of Applied Mathematics and the Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education of China. He has authored or coauthored more than 60 journal and conference publications. His current research interests include neural networks, fuzzy systems, adaptive control, and learning control for uncertain nonlinear systems.