International Journal of Approximate Reasoning
50 (2009) 37–50

An unsupervised context-sensitive change detection technique based on modified self-organizing feature map neural network

Susmita Ghosh [a], Swarnajyoti Patra [a], Ashish Ghosh [b,*]

[a] *Department of Computer Science and Engineering, Jadavpur University, Kolkata 700 032, India*
[b] *Machine Intelligence Unit and Center for Soft Computing Research, Indian Statistical Institute, 203 B.T. Road, Kolkata 700 108, India*

Accepted 23 January 2008
Available online 29 February 2008

Abstract

In this paper, we propose an unsupervised context-sensitive technique for change-detection in multitemporal remote sensing images. Here a modified self-organizing feature map neural network is used. Each spatial position of the input image corresponds to a neuron in the output layer and the number of neurons in the input layer is equal to the number of features of the input patterns. The network is updated depending on some threshold value and when the network converges, status of output neurons depict a change-detection map. To select a suitable threshold of the network, a correlation based and an energy based criteria are suggested. The proposed change-detection technique is unsupervised and distribution free. Experimental results, carried out on two multispectral and multitemporal remote sensing images, confirm the effectiveness of the proposed approach.

Keywords: Remote sensing; Change-detection; Multitemporal images; Self-organizing feature map; Thresholding

1. Introduction

In remote sensing applications, change-detection is the process of identifying differences in the state of an object or phenomenon by analyzing a pair of images acquired on the same geographical area at different times [1]. Such a problem plays an important role in many different domains, like studies on land-use/land-cover dynamics [2], monitoring shifting cultivations [3], burned area assessment [4], analysis of deforestation processes [5], identification of vegetation changes [6], monitoring of urban growth [7], etc. Since all these applications usually require an analysis of large areas, development of completely automatic change-detection techniques became of high relevance in order to reduce the effort required by manual image analysis.

In the literature [2–17] several supervised and unsupervised techniques for detecting changes in remote sensing images have been proposed. The supervised methods require the availability of a "ground truth" from which a training set, containing information about the spectral signatures of the changes that occurred in the considered area between the two dates, is generated. The statistics of the classes can be more easily estimated, given the a priori information. Moreover, it is also possible to estimate the kind of changes that occurred. In contrast, unsupervised approaches perform change-detection without using any additional information, besides the raw images considered. The difficulty with collecting ground truth information regularly in time makes it mandatory to develop unsupervised change-detection methods to support the analysis of temporal sequences of remote sensing images. Change-detection problem can be defined as an unsupervised classification problem where a "changed" class and an "unchanged" class have to be distinguished, given the input images.

Most widely used unsupervised change-detection techniques are based on a three-step procedure [1]: (i) pre-processing; (ii) pixel-by-pixel comparison of two multitemporal images; and (iii) image analysis. The aim of the pre-processing step is to make the considered images as comparable as possible with respect to operations like co-registration, radiometric and geometric corrections and noise reduction. The comparison step aims at producing a further image called "difference image", where differences between the two considered acquisitions are highlighted. Different mathematical operators can be adopted to perform image comparison. Once image comparison is performed, the change-detection/image analysis process can be carried out adopting either context-insensitive or context-sensitive procedures. The most widely used context-insensitive analysis techniques are based on histogram thresholding [8,11,17]. Thresholding procedures do not take into account the spatial correlation between neighboring pixels in the decision process. To overcome this limitation, different context-sensitive change-detection procedures based on Markov random fields (MRF) have been proposed in the literature [8,10,12,18]. These approaches require the selection of a proper model for the statistical distributions of changed and unchanged pixels. The EM algorithm [19] has been employed for estimation of these distributions under different assumptions for class distributions, e.g. Gaussian [8], generalized Gaussian [17] and mixture of Gaussian [12].

In order to overcome the limitations imposed by the need of selecting a statistical model for changed and unchanged class distributions, in this article we propose an unsupervised and context-sensitive change-detection technique which is distribution free. The presented technique automatically detects the changes in the difference image using a modified self-organizing feature map (SOFM) neural network [20]. The network has two layers, input and output. The number of neurons in the input layer is equal to the dimension of the input patterns. The input patterns are generated considering each pixel in the difference image along with its neighboring pixels, in order to take into account the spatial contextual information from the neighborhood. In the output layer, each neuron corresponds to a pixel in the difference image. The network is updated (until convergence) to generate a change-detection map for a particular assumed threshold. Different change-detection maps are generated by varying the threshold. To select an appropriate threshold automatically a correlation based and an energy based criteria are proposed. The major advantages of the proposed technique are (i) it is distribution free (both the proposed threshold selection strategies of the network are not based on any specific parametric model for the distributions of classes); (ii) unlike the techniques proposed in [8,12], it does not require manual setting of any input parameter, so it is completely unsupervised.

In order to assess the effectiveness of the proposed technique, we considered two multitemporal data sets corresponding to the geographical areas in Mexico and the Island of Sardinia, Italy, and compared the results produced by the proposed technique with those obtained by one of the popular methods published in the literature [8].

The article is organized into seven sections. Section 2 addresses the generation of the difference image. Section 3 provides a brief description of Kohonen's self-organizing feature map neural network. Section 4 describes the proposed change-detection technique. In Section 5, the data sets used for the experiments are described. Experimental results are reported in Section 6. Finally, in Section 7, conclusions are drawn.

2. Generation of difference image

In the literature, concerning unsupervised change-detection, several techniques are used to generate the "difference image". The difference image is computed in such a way that pixels associated with land-cover

changes present gray values significantly different from those associated with unchanged areas. Difference operator is most popularly used to generate the difference image. This operator can be applied to: (i) a single spectral band (univariate image differencing) [1,21,22]; (ii) multiple spectral bands (change vector analysis) [1,8]; (iii) vegetation indices (vegetation index differencing) [1,23] or to other linear (e.g. tasselled cap transformation [21]) or nonlinear combinations of spectral bands. Among these, the most popular change vector analysis (CVA) technique is used here to generate the difference image. This technique exploits a simple vector subtraction operator to compare two multispectral images, under analysis, pixel-by-pixel. In some cases, depending on the specific type of changes to be identified, the comparison is made on a subset of the spectral channels. The difference image is computed as the magnitude of spectral change vectors obtained for each pair of corresponding pixels.

Let us consider two co-registered and radiometrically corrected γ-spectral band images X_1 and X_2, of size $p \times q$, acquired over the same area at different times T_1 and T_2, and let $D = \{l_{mn}, 1 \leqslant m \leqslant p, 1 \leqslant n \leqslant q\}$ be the difference image obtained by applying the CVA technique to X_1 and X_2. Then

$$l_{mn} = (\text{int})\sqrt{\sum_{\alpha=1}^{\gamma}\left(l_{mn}^{\alpha}(X_1) - l_{mn}^{\alpha}(X_2)\right)^2}.$$

Here $l_{mn}^{\alpha}(X_1)$ and $l_{mn}^{\alpha}(X_2)$ are the gray values of the pixel at the spatial position (m,n) in αth band of images X_1 and X_2, respectively.

3. Kohonen's model of self-organizing feature map

Kohonen's self-organizing feature map (SOFM) network [24,25] consists of an input and an output layer. Each neuron in the output layer is connected to all the neurons in the input layer, i.e. the y-dimensional input signal $\vec{U} = [u_1, u_2, \ldots, u_y]$ can be passed to all the output neurons. Let the synaptic weight vector of an output neuron j be denoted by $\vec{W}_j = [w_{j1}, w_{j2}, \ldots, w_{jy}]^{\mathrm{T}}, j = 1, 2, \ldots M$, where M is the total number of neurons in the output layer and w_{jk} is the weight of the jth unit for the kth component of the input. If the synaptic weight vector $\vec{W}_i$ of output neuron i is best matched with input vector $\vec{U}$, then $\sum_{k=1}^{y} w_{ik} \cdot u_k$ will be maximum among $\sum_{k=1}^{y} w_{jk} \cdot u_k, \forall j$. The neuron i is then called the *wining neuron* for the input vector $\vec{U}$. The *wining neuron* is located at the center of a topological neighborhood of cooperating units. Let $h_i(itr)$ denote the topological neighborhood of *wining neuron* i at epoch number itr. There are several ways to define a topological neighborhood [26] such as Gaussian, rectangular, etc. The size of the topological neighborhood shrinks with increase in itr. Fig. 1 shows how the size of the topological neighborhood of the *wining neuron* i decreases over itr (in case of rectangular topological neighborhood). For the ith unit and all its neighbors (within a specified radius defined by $h_i(itr)$) the following weight updating rule is applied

$$\vec{W}_i(itr+1) = \vec{W}_i(itr) + \eta(itr)h_i(itr)(\vec{U} - \vec{W}_i(itr)). \tag{1}$$

Here, $\eta(0 < \eta < 1)$ is the learning rate that determines how rapidly the system adjusts over itr. The learning rate parameter η also decreases as itr increases. The term $h_i(\cdot)$ ensures updating the synaptic weights of the

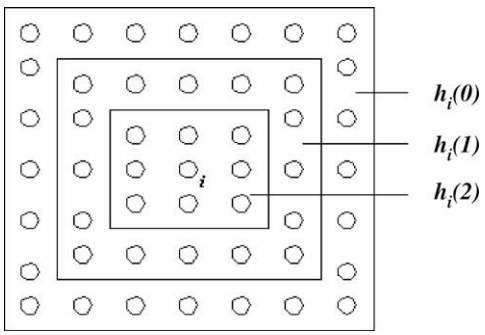

Fig. 1. Rectangular topological neighborhood of neuron i over itr.

neurons inside the topological neighborhood of *wining neuron i* only. The above updating procedure moves the weight vector towards the input vector. So repeated presentation of training patterns tries to make the synaptic weight vectors tend to follow the distribution of the input vectors.

4. Proposed change detection technique using modified SOFM

In this section, we propose an unsupervised context-sensitive technique for change-detection in multitemporal remote sensing images based on modified SOFM neural network. The basic idea exploited in the proposed approach is inspired by the principle used in [20] for object background classification.

4.1. Description of the network architecture

Let $D = \{l_{mn}, 1 \leqslant m \leqslant p, 1 \leqslant n \leqslant q\}$ be the difference image obtained by applying the CVA technique on X_1 and X_2. In order to use a SOFM network for solving the change-detection problems by exploiting both image radiometric properties and spatial-contextual information, we assign a neuron in the output layer of the network corresponding to each spatial position (m,n) in D. The spatial correlation between neighboring pixels is modeled by generating the input patterns corresponding to each pixel in the difference image D, considering its spatial neighborhood systems N of order d. For a given spatial position $(m,n), N^d(m,n)$ will be as follows: $N^d(m,n) = \{(m,n) + (i,j), (i,j) \in N^d\}$. In greater detail, if the input patterns take first order neighborhood information (N^1), then the input pattern $\vec{U}_{mn}$ corresponding to the spatial position (m,n) is generated by considering the grey value of the pixel at position (m,n) and its four nearest neighbors, i.e. $N^1 = \{(\pm 1, 0), (0, \pm 1)\}$. Similarly, with the second order neighborhood (N^2), a pixel at position (m,n) considers its eight nearest neighboring pixels, i.e. $N^2 = \{(\pm 1, 0), (0, \pm 1), (1, \pm 1), (-1, \pm 1)\}$. Fig. 2 depicts a second order neighborhood (N^2) of a pixel at position (m,n). In the present study we used a SOFM network where the number of neurons in the output layer is equal to the number of pixels in the difference image D and the number of neurons in the input layer is equal to the dimension of the input patterns.

Let $\vec{U}_{mn}$ and $\vec{W}_{mn}$, respectively, be the y-dimensional input and weight vectors corresponding to the neuron (m,n) located at mth row and nth column of the output layer, i.e. $\vec{U}_{mn} = [u_{mn,1}, u_{mn,2}, \ldots, u_{mn,y}]$ and $\vec{W}_{mn} = [w_{mn,1}, w_{mn,2}, \ldots, w_{mn,y}]^{\mathrm{T}}$. Note that in the feature mapping algorithm, the maximum lengths of the input and weight vectors are fixed. Let the maximum length for each component of the input vector be unity. To keep the value of each component of the input less than or equal to 1, let us apply a mapping function f, where $f : [c_{\min}, c_{\max}] \rightarrow [0, 1]$.

Here $c_{\min}$ and $c_{\max}$ are the global minimum and maximum component (feature) values present in the input vectors. The initial component of the weight vectors are chosen randomly in $[0,1]$.

4.2. Learning of the weights

$\vec{U}_{mn}$ and $\vec{W}_{mn}$ be the input and weight vectors of the output neuron (m,n), respectively. Now their dot product $x(m,n)$ is

(m-1,n-1)	(m-1,n)	(m-1,n+1)
(m,n-1)	(m,n)	(m,n+1)
(m+1,n-1)	(m+1,n)	(m+1,n+1)

Fig. 2. N^2 neighborhood of pixel (m,n).

$$x(m,n) = \vec{U}_{mn} \cdot \vec{W}_{mn} = \sum_{k=1}^{y} u_{mn,k} \cdot w_{mn,k}. \tag{2}$$

Only those neurons for which $x(m,n) \geqslant t$ (here t is assumed to be a pre-defined threshold) are allowed to modify (update) their weights along with their neighbors (specified by $h_{mn}(\cdot)$). Consideration of a set of neighbors enables one to grow the region by including those which might have been dropped out because of the initial randomness of weights (i.e. if the assigned weights are such that $x(m,n) \geqslant t$ for a few (m,n), then the weights of these neurons and their neighboring neurons will also be updated and subsequently categorized into changed region, if they originally belonged to changed region). The threshold t is varied from 0 to 1. To keep the value of dot product $x(m,n)$ for all (m,n) in $[0,1]$, we normalize each components of the weight vector $\vec{W}_{mn}$ of neuron (m,n) such that $\sum_{k=1}^{y} w_{mn,k} = 1$ (here all components of the input and weight vectors are nonnegative). The kth component of the weight vector $\vec{W}_{mn}$ is normalized as follows:

$$w_{mn,k} = \frac{w_{mn,k}}{\sum_{k=1}^{y} w_{mn,k}}. \tag{3}$$

The weight updating procedure is performed using Eq. (1). The value of learning rate parameter $\eta(itr)$ and the size of the topological neighborhood $h_{mn}(itr)$ decreases over itr. To check the convergence, total output $O(itr)$ for each epoch number itr is computed as follows:

$$O(itr) = \sum_{x(m,n) \geqslant t} x(m,n). \tag{4}$$

The updating of the weights continue until $|O(itr) - O(itr-1)| < \delta$, where δ is a preassigned small positive quantity. After the network is converged, the pixel at spatial position (m,n) in D is assigned to *changed region* if $x(m,n) \geqslant t$, else to *unchanged region*. The network converges for any value of t (proof is available in [20]). Table 1 described this learning technique algorithmically.

Note that, unlike the conventional SOFM, in the present network instead of giving the same input to all output neurons and finding out the *wining neuron*, here different input is given to different output neurons and weight updating is performed based on the considered threshold t.

4.3. Proposed threshold selection techniques

As stated in the previously mentioned algorithm, the updating of weights depends on the threshold value t. Initially for a particular threshold t, the network is updated till convergence. After convergence, if $x(m,n) \geqslant t$, then make the output value V_{mn} of neuron (m,n) 1, else make it -1. Thus in the output layer the neurons are divided into two groups G_u (represents *unchanged* regions) and G_c (represents *changed* regions) and a change-detection map corresponding to threshold t is generated. By varying t (threshold values are varied by an amount $1/L$, where L is the maximum gray value of D), different change-detection maps are generated. To select a threshold t_1 near the minimum error threshold t_0 (corresponding change-detection results provide minimum error), we propose two different criteria which are described below.

Table 1
Learning algorithm of the presented network

Step 1: Initialize $\vec{W}_{mn}$ randomly (in $[0,1]$) for each output neuron (m,n)
Step 2: Set $itr = 0$. Initialize $\eta(itr)$ and $h_{mn}(itr)$
Step 3: Set $O(itr) = 0$
Step 4: Select an input vector $\vec{U}_{mn}$ and corresponding normalize weight vector $\vec{W}_{mn}$
Step 5: Compute their dot product $x(m,n)$ using Eq. (2)
Step 6: If $x(m,n) \geqslant t$ then update the weight vector $\vec{W}_{mn}$ along with the weight vectors of the neighboring neurons of the (m,n)th neuron using Eq. (1) and set $O(itr) = O(itr) + x(m,n)$
Step 7: Repeat Steps 4–6 for all input patterns, completing one epoch
Step 8: If $(|\, O(itr) - O(itr-1) \,| < \delta)$ then goto Step 11
Step 9: $itr = itr + 1$
Step 10: Decrease the value of $\eta(itr)$ and $h_{mn}(itr)$ and goto Step 3
Step 11: Stop

4.3.1. Correlation maximization criterion

In this case, the correlation coefficient [27] between the input sequence (difference image) and the output sequence (change-detection map generated assuming threshold t) is maximized to select the near optimal threshold t_1. As the threshold t is varied, the change-detection map also varies, thereby the value of the correlation coefficient is changed. When the threshold corresponds to the boundary between changed and unchanged pixels, the correlation coefficient becomes maximum. Let Y and Z be two random variables that correspond to the input and output sequences, respectively. The correlation coefficient between Y and Z for threshold t (denoted as $R_{Y,Z}(t)$) is defined as

$$R_{Y,Z}(t) = \frac{\mathrm{cov}(Y,Z)}{\sigma_Y \cdot \sigma_Z}.$$

Here $\mathrm{cov}(Y,Z)$ is the covariance of Y and Z; σ_Y and σ_Z are the standard deviations of Y and Z, respectively. As l_{mn} and V_{mn} are the values of the difference image and output image (assuming threshold t) at the spatial position (m,n), respectively. The above formula may be rewritten as

$$R_{Y,Z}(t) = \frac{\sum_{m,n} l_{mn} \cdot V_{mn} - \frac{1}{p\times q}\sum_{m,n} l_{mn} \cdot \sum_{m,n} V_{mn}}{\sqrt{\sum_{m,n} l_{mn}^2 - \frac{1}{p\times q}\left(\sum_{m,n} l_{mn}\right)^2}\sqrt{\sum_{m,n} V_{mn}^2 - \frac{1}{p\times q}\left(\sum_{m,n} V_{mn}\right)^2}}.$$

We compute the correlation coefficient assuming different threshold and t_1 is assumed to be a near optimal threshold, if the correlation coefficient $R_{Y,Z}$ is maximum at t_1. Considering this automatically derived threshold t_1, we update the weights of the network and when the network converges the output layer implicitly generates a change-detection map.

4.3.2. Energy based criterion

In this case, a near optimal threshold is found out by analyzing the overall status of the change-detection maps obtained with different thresholds. To describe the overall status of the change-detection map generated by any threshold t, we used the following expression of energy [28]:

$$E(t) = -\sum_{m=1}^{p}\sum_{n=1}^{q}\left(\sum_{(i,j)\varepsilon N^2(mn)} V_{mn} \cdot V_{ij}\right) - \sum_{m=1}^{p}\sum_{n=1}^{q} V_{mn}^2. \tag{5}$$

The energy value defined in Eq. (5) has two parts. In terms of images, the first part can be seen as the impact of the gray values of the neighboring pixels, whereas the second part can be attributed to the gray value of the pixel under consideration. As in the present context, the value of $V_{mn} = \pm 1, \forall(m,n)$, the above expression takes the minimum value when the generated change-detection map totally belongs to either *unchanged* regions ($V_{mn} = -1, \forall(m,n)$) or *changed* regions ($V_{mn} = +1, \forall(m,n)$). It takes the maximum value when the number of *unchanged and changed* regions are very high. In the proposed procedure, we first compute the energy (at convergence) of each change-detection map generated by different thresholds and plot these status values with the corresponding threshold value (see Fig. 3). By analyzing the behavior of this graph, one can see that initially the value of $E(t)$ increases with t (as the number of *changed and unchanged* regions increases. After a certain threshold value, $E(t)$ decreases (as some *unchanged* regions are merged together). After that the energy does not change significantly with change in threshold, i.e. *changed and unchanged* regions are not significantly altered. We expect that this stable behavior of the energy function $E(t)$ is reached around optimal initialization threshold t_0 (see Fig. 3). If the threshold value increases more, the energy changes slowly and reaches a minimum when the whole output image belongs to the *unchanged* class. By observing this general behavior, we propose a heuristic technique that first generates the smallest convex curve $E_1(\cdot)$ containing the energy curve $E(\cdot)$ using the concavity analysis algorithm [29] and then exploiting these two curves, a threshold t_1, close to the optimal threshold t_0, is automatically detected. Considering this automatically derived threshold t_1 we update the weights of the network and at the convergence of the network the output layer implicitly generates a change-detection map. The corresponding algorithm is described in Table 2.

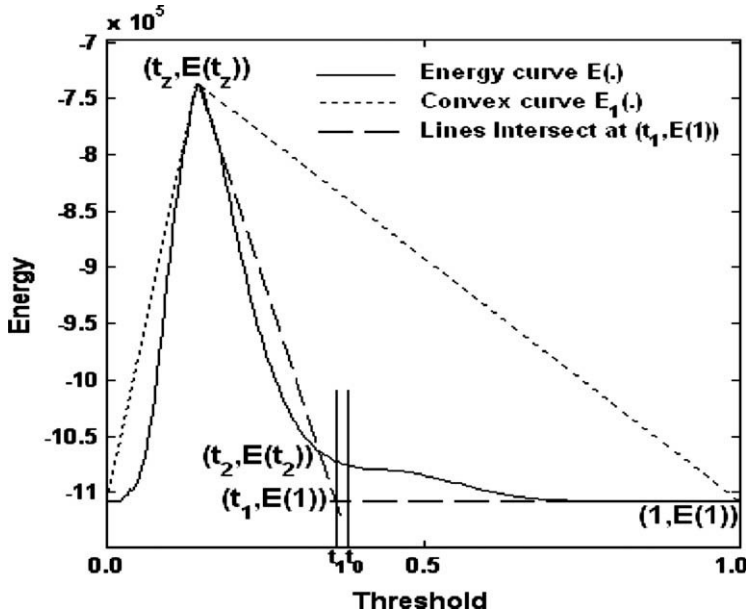

Fig. 3. Behavior of energy value with threshold. Threshold t_1 is detected by the proposed approach (t_0 is the optimal).

Table 2
Algorithm for automatically detecting the near optimal threshold t_1

Phase 1: Generate the smallest convex curve $E_1(\cdot)$ containing the energy curve $E(\cdot)$
Step 1: Initialize $k = 0.0$
Step 2: While $k \neq 1.0$
Step 3: For $i = k$ to 1.0 step $\frac{1}{L}$ (L is the maximum gray value of D)
Step 4: Compute the gradient (slope) of the line passing through the points $(k, E(k))$ and $(i, E(i))$
Step 5: Find out a point $(j, E(j)), k < j \leqslant 1.0$, such that the slope of the line passing through the points $(k, E(k))$ and $(j, E(j))$ is maximum
Step 6: Join the two points $(k, E(k))$ and $(j, E(j))$ by a straight line (this line is a part of the convex curve $E_1(\cdot)$)
Step 7: Reset $k = j$
Step 8: End while
Phase 2: Derive the initialization threshold value t_1
Step 9: Find the point $(t_z, E(t_z))$ on the energy curve where energy value is maximum (see Fig. 3)
Step 10: Select t_2, so that $\{E_1(t_2) - E(t_2)\} = \max_i\{E_1(i) - E(i)\}, t_z \leqslant i \leqslant 1.0$
Step 11: Select the threshold t_1, at the intersection between the straight line connecting $(t_z, E(t_z))$ and $(t_2, E(t_2))$ and the straight line parallel to the abscissa and passing through the minimum energy value $E(1)$ (see Fig. 3)
Step 12: Stop

5. Description of the data sets

In order to carry out an experimental analysis aimed at assessing the effectiveness of the proposed approach, we considered two multitemporal data sets corresponding to geographical areas of Mexico and Island of Sardinia, Italy. A detailed description of each data set is given below.

5.1. Data set related to Mexico area

The first data set used in the experiment is made up of two multispectral images acquired by the Landsat Enhanced Thematic Mapper Plus (ETM+) sensor of the Landsat-7 satellite in an area of Mexico on 18th April 2000 and 20th May 2002. From the entire available Landsat scene, a section of 512 × 512 pixels has been selected as test site. Between the two aforementioned acquisition dates a fire destroyed a large portion of

the vegetation in the considered region. Fig. 4a and b shows channel 4 of the 2000 and 2002 images, respectively. In order to be able to make a quantitative evaluation of the effectiveness of the proposed approach, a reference map was manually defined (see Fig. 4d) according to a detailed visual analysis of both the available multitemporal images and the difference image (see Fig. 4c). Different color composites of the above mentioned images were used to highlight all the portions of the changed area in the best possible way. This procedure resulted in a reference map containing 25,599 changed and 236,545 unchanged pixels. Experiments were carried out to produce, in an automatic way, a change-detection map as similar as possible to reference map that represents the best result obtainable with a time consuming procedure.

Analysis of the behavior of the histograms' of multitemporal images did not reveal any significant difference due to light and atmospheric conditions at the acquisition dates. Therefore, no radiometric correction algorithm was applied. The 2002 image was registered on the 2000 one using 12 ground control points. The procedure led to a residual average misregistration error on ground control points of about 0.3 pixels.

5.2. Data set related to Sardinia Island, Italy

The second data set used in the experiment is composed of two multispectral images acquired by the Landsat Thematic Mapper (TM) sensor of the Landsat-5 satellite in September 1995 and July 1996. The test site is a section of 412×300 pixels of a scene including lake Mulargia on the Island of Sardinia (Italy). Between the two aforementioned acquisition dates the water level in the lake increased (see the lower central part of the image). Fig. 5a and b shows channel 4 of the 1995 and 1996 images. As done for the Mexico data set, in this case also a reference map was manually defined (see Fig. 5d) according to a detailed visual analysis of both the available multitemporal images and the difference image (see Fig. 5c). At the end, 7480 changed and 116,120 unchanged pixels were identified. As histograms did not show any significant difference, no radiometric correction algorithms were applied to the multitemporal images. The images were co-registered with 12 ground control points resulting in an average residual misregistration error of about 0.2 pixels on the ground control points.

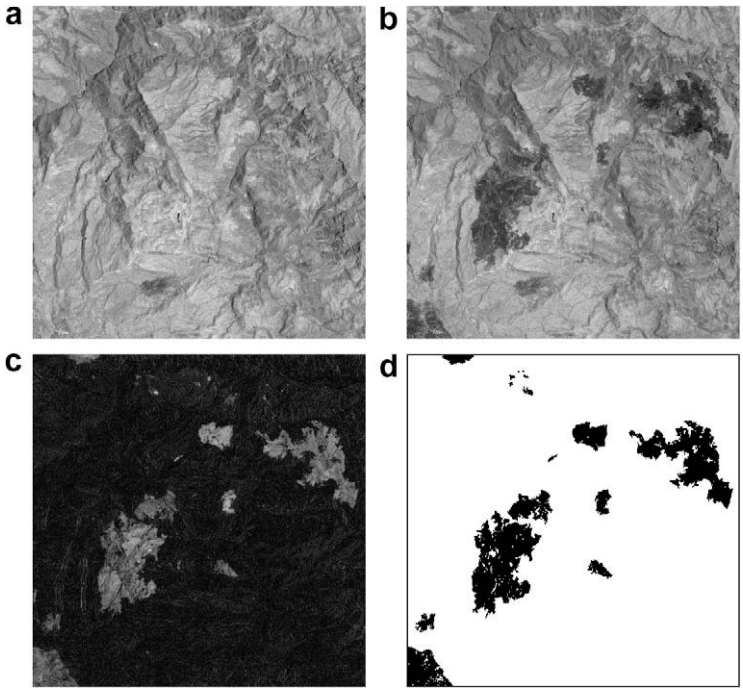

Fig. 4. Image of Mexico. (a) Band 4 of the Landsat ETM+ image acquired in April 2000; (b) band 4 of the Landsat ETM+ image acquired in May 2002; (c) corresponding difference image generated by CVA technique; and (d) reference map of the changed area.

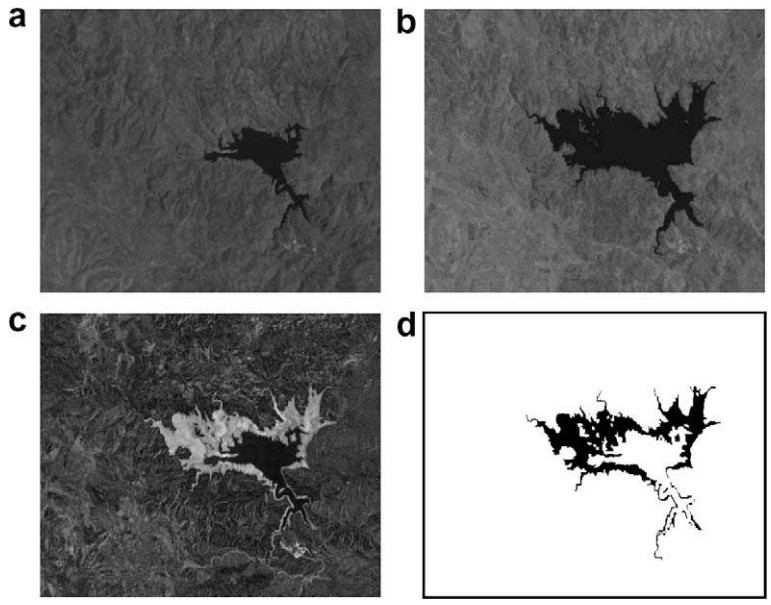

Fig. 5. Image of Sardinia Island, Italy. (a) Band 4 of the Landsat TM image acquired in September 1995; (b) band 4 of the Landsat TM image acquired in July 1996; (c) difference image generated by CVA technique using bands 1, 2, 4, & 5; and (d) reference map of the changed area.

6. Experimental results

In this study, the input vectors $\vec{U}_{mn}$ (for each pixel $(m,n) \in D$) have nine components considering the gray value of the pixel (m,n) and the gray values of its eight neighboring pixels (here N^2 neighbor is considered). To map the value of each component of the input vectors in $[0,1]$, the following formula is used:

$$\frac{u_{mn,k} - c_{\min}}{c_{\max} - c_{\min}},$$

where $u_{mn,k}$ is the kth component value of input vector $\vec{U}_{mn}$ and $c_{\max}, c_{\min}$ are the maximum and minimum available gray values of the difference image D. The weight vectors also have nine components. Initial weights are assigned randomly in $[0,1]$. The learning rate parameter η is chosen as $\eta(itr) = \frac{1}{1+itr}$, i.e. the value of η at the itrth epoch is taken as $\frac{1}{1+itr}$. This ensures $0 < \eta \leqslant 1$ and it decreases with itr. Initial size of the topological neighborhood $h_{mn}(itr), \forall(m,n)$ was taken as a 11×11 rectangular window, and gradually reduced to 3×3 after 5 epochs and kept constant for the remaining epochs (until converges). The value of the convergence parameter δ is considered as 0.01.

The first experiment aims at assessing the validity of the proposed threshold selection criteria (as described in Section 4.3). To this end, the optimal threshold t_0 is chosen by a trial-and-error procedure where the change-detection maps (at convergence of the network) are generated by varying threshold t and computing the change-detection error corresponding to each threshold with the help of the reference map (please note that the reference map is not available in real situation). The threshold t_0 corresponds to the minimum change-detection error. The change-detection results obtained considering the threshold detected by the proposed criteria were compared with the change-detection results produced by assuming optimal threshold t_0.

In order to establish the effectiveness of the proposed technique, the second experiment compares the results (change-detection maps) provided by our method with a context-insensitive manual trial-and-error thresholding (MTET) technique [17] and a context-sensitive technique presented in [8] based on the combined use of the EM algorithm and Markov random fields (MRF) (we refer to it as EM + MRF technique). The MTET technique generates a minimum error change-detection map under the hypothesis of spatial independence among pixels by finding a minimum error decision threshold for the difference image. The minimum error decision threshold is obtained by computing change-detection errors (with the help of the reference map) for all values

of the decision threshold. Note that, this minimum error context-insensitive threshold is different from the context-sensitive optimal threshold t_0 as obtained in the first experiment. Comparisons were carried out in terms of both overall change-detection error and number of false alarms (i.e. unchanged pixels identified as changed ones) and missed alarms (i.e. changed pixels categorized as unchanged ones).

6.1. Results on Mexico data

In order to determine the most effective spectral bands for detecting the burned area in the considered data set we performed some trials. On the basis of the results of these trials, we found that band 4 is very effective to locate the burned area. Hence we generated the difference image by considering only spectral band 4.

To assess the validity of the automatically derived threshold t_1, in the first experiment a comparison was carried out between the optimal threshold t_0 and the thresholds t_1 detected by the proposed correlation based criterion and energy based criterion. Fig. 6a and b shows the variation of correlation coefficient and energy value, respectively with threshold. From Fig. 6, it is seen that the correlation criterion automatically selects a threshold t_1 which is on the right side of t_0 and the energy based criterion selects a threshold t_1 which is on the left side of t_0. But both the thresholds are near to the optimal one. Table 3 shows the change-detection results produced by the proposed technique assuming optimal threshold t_0 and automatically detected thresholds t_1. As the correlation based criterion selects higher threshold value (with $t_1 = 0.232$), it generates higher missed alarms and lower false alarms as compared to energy based one (with $t_1 = 0.183$). Since correlation based criterion is able to detect a threshold t_1 which is closer to the optimal threshold t_0 (with $t_0 = 0.216$), the overall error produced by this criterion (3217 pixels) is close to the optimal one (2979 pixels).

As mentioned, in the second experiment, the change-detection results produced by the proposed context-based approaches are compared with those obtained by the context-insensitive MTET and context-sensitive EM + MRF (see [8]) techniques. The results are put in Table 4. It is seen from the table that the overall error obtained by the proposed method (using two threshold selection criteria) is much smaller than the overall error incurred by MTET technique. Concerning the error typology, the proposed technique based on correlation criterion and energy criterion resulted in 3217 and 3512 pixels as overall error, respectively, whereas the MTET involves 4591 pixels. For a better understanding of the behavior of the different methods the change-detection maps produced by them are depicted in Fig. 7. A visual comparison points out that the proposed approach exploits the spatio-context information for reducing the noise present in the maps. Table 4 also presents the best change-detection result obtained by the context-sensitive EM + MRF technique, when the parameter β of MRF [8] was set to 1.5 (this value was defined manually and corresponds to the minimum possible error). The overall error obtained by the proposed technique based on correlation criterion is similar to the existing EM + MRF technique; whereas the technique based on energy criterion produces slightly poor

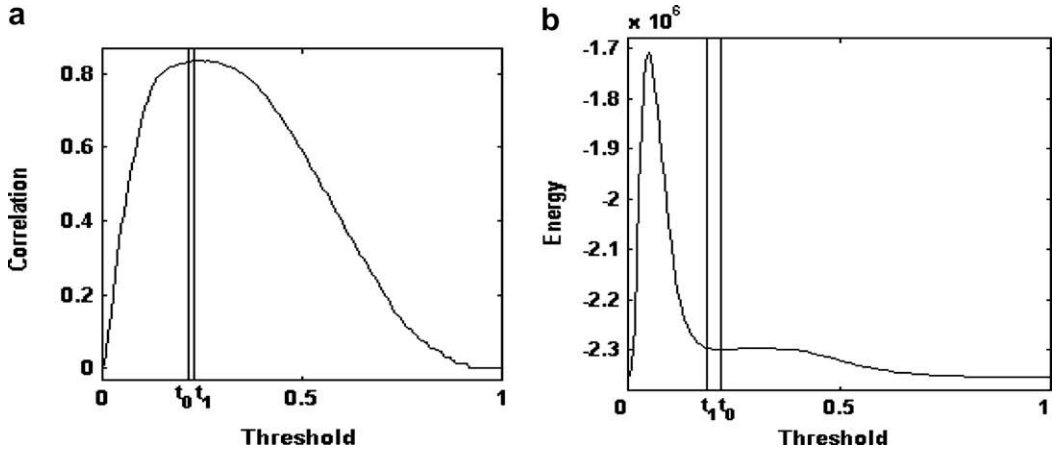

Fig. 6. Variation of (a) correlation coefficient and (b) energy value with threshold. Detected threshold t_1 is close to the optimal threshold t_0 (band 4, Mexico data).

Table 3
Change-detection results obtained by the proposed technique considering the optimal threshold t_0 and the automatically detected thresholds t_1 using correlation and energy based criteria (band 4, Mexico data set)

Techniques used	Detected thresholds	Missed alarms	False alarms	Overall error
Optimal	0.216 (t_0)	1406	1573	2979
Correlation	0.232 (t_1)	2039	1178	3217
Energy	0.183 (t_1)	583	2929	3512

Table 4
Overall error, false alarms, and missed alarms resulting from the proposed context-sensitive technique using correlation and energy criteria

Techniques used	Missed alarms	False alarms	Overall error
Correlation	2039	1178	3217
Energy	583	2929	3512
MTET	2404	2187	4591
EM + MRF ($\beta = 1.5$)	946	2257	3203

The table also gives the errors associated with the context-insensitive MTET and context-sensitive EM + MRF techniques (band 4, Mexico data).

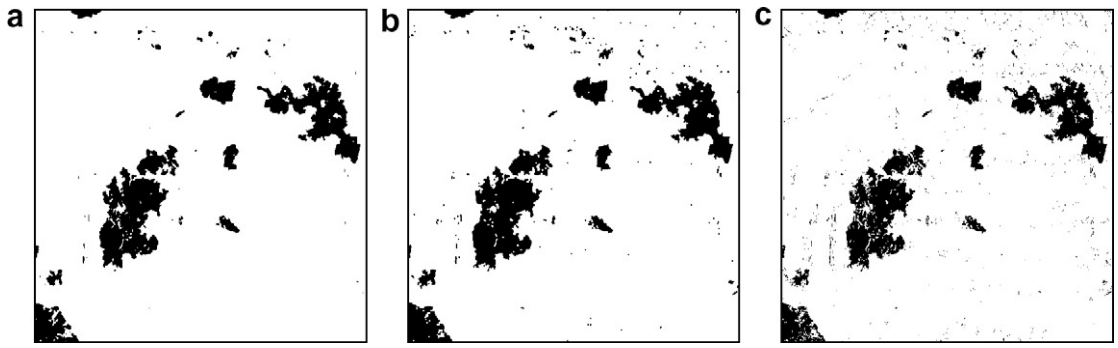

Fig. 7. Change-detection map obtained for Mexico data by (a) the proposed technique based on correlation criterion; (b) the EM + MRF technique; and (c) the context-insensitive MTET technique.

results as compared to the existing EM + MRF technique. In particular, the correlation based technique produced 3217 overall error (with 2039 missed alarms and 1178 false alarms) and the energy based technique generated 3512 overall error (with 583 missed alarms and 2929 false alarms), whereas the EM + MRF produced 3203 overall error (with 946 missed alarms and 2257 false alarms). Note that the proposed technique does not require the optimization of the parameter β, thus resulting in a more practical tool for applications.

6.2. Results on Sardinia Island data

We applied the CVA technique on spectral bands 1, 2, 4, and 5 of the two multispectral images to generate the difference image, as preliminary experiments show that the above channels contain useful information of the changes in water level.

In the first experiment, the results obtained by the proposed approach with the automatically detected threshold t_1 using correlation criterion and energy criterion were compared with the results produced by considering the optimal threshold t_0. Fig. 8a and b shows the variation of correlation coefficient and energy value versus threshold. From an analysis of Fig. 8 one can deduce that the proposed technique, based on both the correlation and energy criteria, select the thresholds which are on the left side of the optimal threshold t_0; but both the thresholds are near to the optimal one. Table 5 shows the change-detection results produced by the proposed technique for the optimal threshold t_0 and automatically detected thresholds t_1. As the threshold (with $t_1 = 0.356$) detected by the energy based criterion is higher than the threshold (with $t_1 = 0.337$) detected

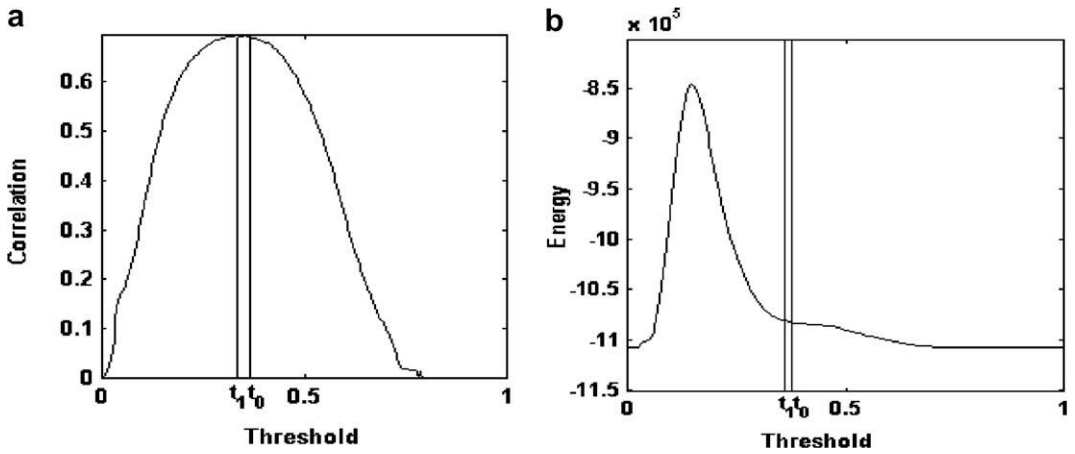

Fig. 8. Variation of the (a) correlation coefficient and (b) energy value with threshold. Detected threshold t_1 is close to the optimal threshold t_0 (Sardinia Island data).

Table 5
Change-detection results obtained by the proposed technique considering the optimal threshold t_0 and the automatically detected threshold t_1 using correlation and energy based criteria (Sardinia Island data)

Techniques used	Detected thresholds	Missed alarms	False alarms	Overall error
Optimal	0.368 (t_0)	1090	558	1648
Correlation	0.337 (t_1)	721	1164	1885
Energy	0.356 (t_1)	935	729	1664

by correlation based criterion, the energy based technique generates higher missed alarms and lower false alarms. As the energy based criterion detects a threshold t_1 which is more close to the optimal threshold (with $t_0 = 0.368$), the produced overall error (1664 pixels) is more close to the optimal one (1648 pixels).

Concerning the second experiment, the change-detection results produced by the proposed approach was compared with the change-detection results obtained by the context-insensitive MTET procedure and context-sensitive EM + MRF technique (see [8]). Table 6 shows the change-detection results obtained applying different techniques. By analyzing those results, one can deduce that the overall change-detection error obtained by the proposed technique based on both the criteria are less than the overall error incurred by the MTET technique. For example, the overall error obtained by the proposed technique based on correlation criterion and energy criterion are 1885 pixels and 1664 pixels, respectively; whereas the overall error yielded by the MTET technique is 1890 pixels. Fig. 9 shows the change-detection maps produced by the different techniques. A visual comparison points out that the proposed approach exploits the spatio-contextual information for reducing the noise present in the maps. Table 6 also presents the best change-detection result obtained by the context-sensitive EM + MRF technique, when the parameter β of MRF [8] was set to 2.2. Although the

Table 6
Overall error, false alarms, and missed alarms resulting from the proposed context-sensitive technique using correlation and energy criteria

Techniques used	Missed alarms	False alarms	Overall error
Correlation	721	1164	1885
Energy	935	729	1664
MTET	1015	875	1890
EM + MRF ($\beta = 2.2$)	592	1108	1700

The table also gives the errors associated with the context-insensitive MTET and context-sensitive EM + MRF techniques (Sardinia Island data).

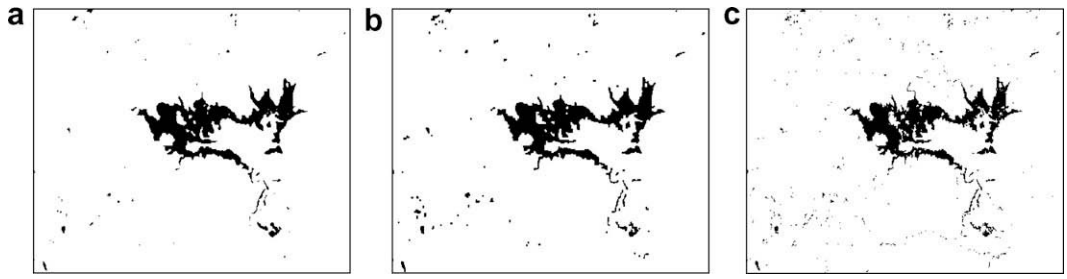

Fig. 9. Change-detection map obtained for the data set related to the Island of Sardinia, Italy by (a) the proposed technique based on energy criterion; (b) the EM + MRF technique; and (c) the context-insensitive MTET technique.

overall error obtained by the proposed technique based on correlation criterion is slightly higher than the existing EM + MRF technique, the proposed technique based on energy criterion generates lower overall error. In particular, the correlation based technique produced 1885 overall error (i.e. 721 missed alarms and 1164 false alarms) and the energy based technique generated 1664 overall error (i.e. 935 missed alarms and 729 false alarms), whereas the EM + MRF produced 1700 overall error (i.e. 592 missed alarms and 1108 false alarms).

7. Discussion and conclusion

In this article, an unsupervised and context-sensitive technique for change-detection in multitemporal remote sensing images has been proposed. The technique discriminates the *changed and unchanged* regions in the difference image by using a modified SOFM network implemented according to a specific architecture. Here the number of input neurons is equal to the dimension of the input patterns and the number of output neurons is equal to the number of pixels in the difference image, i.e. one neuron is assigned to each pixel of the difference image.

In the presented technique two different criteria are proposed to select a near optimal threshold. The proposed technique shows the following advantages with respect to the context-sensitive change-detection method based on the EM + MRF [8]: (i) it is distribution free, i.e. it does not require any explicit assumption on the statistical model of the distributions of classes of *changed and unchanged* pixels and (ii) it is completely automatic, i.e. the proposed technique does not require the setting of any input parameters manually (the EM + MRF technique requires the definition of the regularization parameter β that tunes the effect of the spatio-contextual information). Although SOFM itself has some parameters (like h, η), but there is no need to set them manually. In terms of computational complexity, the time requirement of the proposed technique is comparable to that of [8]. It is also worth noting that for the considered kind of application it is not fundamental to produce results in real time.

Experimental results obtained on different real multitemporal data sets confirm the effectiveness of the proposed approach, which significantly outperforms the standard optimal manual context-insensitive technique and provides an overall change-detection error comparable to the best one achieved with the context-sensitive EM + MRF technique.

In order to take spatial contextual information of the difference image, in this work the gray value of the pixel and the gray values of its neighboring pixels are used as input. In future, further experiments can be done by taking wavelet based or gray level co-occurrence matrix based features as input.

Acknowledgements

The authors wish to thank the anonymous referees for their constructive criticism and valuable suggestions. Authors would also like to thank the Department of Science and Technology, Government of India and University of Trento, Italy, the sponsors of the ITPAR program and Prof. L. Bruzzone, the Italian collaborator of this project, for his comments on this article and providing the data.

References

[1] A. Singh, Digital change detection techniques using remotely sensed data, Int. J. Remote Sensing 10 (6) (1989) 989–1003.
[2] J. Cihlar, T.J. Pultz, A.L. Gray, Change detection with synthetic aperture radar, Int. J. Remote Sensing 13 (3) (1992) 401–414.
[3] L. Bruzzone, S.B. Serpico, An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images, IEEE Trans. Geosci. Remote Sensing 35 (4) (1997) 858–867.
[4] L. Bruzzone, D.F. Prieto, An adaptive parcel-based technique for unsupervised change detection, Int. J. Remote Sensing 21 (4) (2000) 817–822.
[5] T. Hame, I. Heiler, J.S. Miguel-Ayanz, An unsupervised change detection and recognition system for forestry, Int. J. Remote Sensing 19 (6) (1998) 1079–1099.
[6] P.S. Chavez Jr., D.J. MacKinnon, Automatic detection of vegetation changes in the southwestern United States using remotely sensed images, Photogram. Eng. Remote Sensing 60 (5) (1994) 1285–1294.
[7] K.R. Merril, L. Jiajun, A comparison of four algorithms for change detection in an urban environment, Remote Sensing Environ. 63 (2) (1998) 95–100.
[8] L. Bruzzone, D.F. Prieto, Automatic analysis of the difference image for unsupervised change detection, IEEE Trans. Geosci. Remote Sensing 38 (3) (2000) 1171–1182.
[9] S. Gopal, C. Woodcock, Remote sensing of forest change using artificial neural networks, IEEE Trans. Geosci. Remote Sensing 34 (2) (1996) 398–404.
[10] T. Kasetkasem, P.K. Varshney, An image change detection algorithm based on Markov random field models, IEEE Trans. Geosci. Remote Sensing 40 (8) (2002) 1815–1823.
[11] F. Melgani, G. Moser, S.B. Serpico, Unsupervised change-detection methods for remote-sensing data, Opt. Eng. 41 (2002) 3288–3297.
[12] L. Bruzzone, D.F. Prieto, An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images, IEEE Trans. Image Process. 11 (4) (2002) 452–466.
[13] M.J. Canty, Image Analysis, Classification and Change Detection in Remote Sensing, CRC Press, Taylor & Francis, 2006.
[14] S. Ghosh, L. Bruzzone, S. Patra, F. Bovolo, A. Ghosh, A context-sensitive technique for unsupervised change detection based on Hopfield-type neural networks, IEEE Trans. Geosci. Remote Sensing 45 (3) (2007) 778–789.
[15] S. Patra, S. Ghosh, A. Ghosh, Unsupervised change detection in remote-sensing images using modified self-organizing feature map neural network, in: International Conference on Computing: Theory and Applications (ICCTA-2007), Kolkata, India, IEEE Computer Society Press, 2007, pp. 716–720.
[16] S. Ghosh, S. Patra, A. Ghosh, A neural approach to unsupervised change detection of remote-sensing images, in: B. Prasad, S.R.M. Prasanna (Eds.), Speech, Audio, Image and Biomedical Signal Processing using Neural Networks, Springer-Verleg, Berlin, 2008, pp. 239–264.
[17] Y. Bazi, L. Bruzzone, F. Melgani, An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images, IEEE Trans. Geosci. Remote Sensing 43 (4) (2005) 874–887.
[18] R. Wiemker, An iterative spectral-spatial Bayesian labeling approach for unsupervised robust change detection on remotely sensed multispectral imagery, in: Proceedings of the CAIP, 1997, pp. 263–270.
[19] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J. Royal Stat. Soc. 39 (1) (1977) 1–38.
[20] A. Ghosh, S.K. Pal, Neural network, self-organization and object extraction, Pattern Recogn. Lett. 13 (1992) 387–397.
[21] T. Fung, An assessment of TM imagery for land-cover change detection, IEEE Trans. Geosci. Remote Sensing 28 (4) (1990) 681–684.
[22] D.M. Muchoney, B.N. Haack, Change detection for monitoring forest defoliation, Photogram. Eng. Remote Sensing 60 (1994) 1243–1251.
[23] J.R.G. Townshend, C.O. Justice, Spatial variability of images and the monitoring of changes in the normalized difference vegetation index, Int. J. Remote Sensing 16 (12) (1995) 2187–2195.
[24] T. Kohonen, Self-organized formation of topologically correct feature maps, Biol. Cybernet. 43 (1982) 59–69.
[25] T. Kohonen, Self-Organizing Maps, second ed., Springer-Verlag, Berlin, 1997.
[26] Z.-P. Lo, Y. Yu, B. Bavarian, Analysis of the convergence properties of topology preserving neural networks, IEEE Trans. Neural Networks 4 (1993) 207–220.
[27] S.M. Ross, Introduction to Probability and Statistics for Engineers and Scientists, Wiley, New York, 1987.
[28] A. Ghosh, N.R. Pal, S.K. Pal, Object background classification using Hopfield type neural network, Int. J. Pattern Recogn. Artif. Intell. 6 (5) (1992) 989–1008.
[29] A. Rosenfeld, P. De La Torre, Histogram concavity analysis as an aid in threshold selection, IEEE Trans. Syst. Man Cybernet. SMC-13 (3) (1983) 231–235.