Unsupervised Change Detection in High Spatial Resolution Optical Imagery Based on Modified Hopfield Neural Network

Keming Chen¹, Chunlei Huo¹, Zhixin Zhou¹², Hanqing Lu¹

¹ National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China, 100190;
² Beijing Institute of Remote Sensing, Beijing, China, 100854
E-mail: {kmchen, chlhuo, luhq @nlpr.ia.ac.cn}, zhixin.zhou@ia.ac.cn

Abstract

This paper addresses the problem of unsupervised change detection in high spatial resolution optical remote sensing images based on Hopfield neural network (HNN). An optimization relaxation approach based on the analysis of a modified Hopfield neural network is proposed for solving the change detection problem. The modified Hopfield neural network is designed to characterize a texture in terms of spatial-contextual information included in the neighborhood of each pixel within each color plane and interaction between different color planes. The network topology is built on the difference image so that each pixel in the RGB color planes is represented as a node in the network which is connected to its neighborhood units both in its own plane and other two planes. Each node is represented by its state which characterizes the pixel changed or unchanged, and an energy function is derived to represent the overall status of the whole network. Change detection maps are obtained by iteratively updating the output status of the neurons until the network converges. The main contribution of this paper lies in the construction of a novel continuous Hopfield-type neural network on the RGB image for solving the unsupervised image change detection problem. Experiments results obtained on two sets of remote sensing imagery confirm the effectiveness of the proposed approach.

1. Introduction

Change detection is one of the most important applications of the remote sensing technology and it plays a more and more important role in a variety of fields [1] [2]. Usually, Change detection aims at discerning areas of change on two registered remote sensing images acquired in the same geographical area at two different times. Two main approaches, supervised and unsupervised, are used to detect the change. The former is based on supervised classification methods, which require the availability of a ground truth in order to derive a suitable training set for the learning process of the classifiers. The latter performs change detection without any additional information besides the raw images considered. The effective unsupervised change-detection method is fundamental in many applications in that the suitable ground-truth information is not always available.

With the advent of high resolution (HR) satellite imagery, not only the spectral feature but also other spatial information (color, texture, context) play their relevant roles. Most of the existing techniques described in literature model the spatial-contextual information included in the neighborhood of each pixel by using statistical models. Markov Random Fields (MRFs) is a commonly used generative model to incorporate contextual information included in the neighborhood of each pixel [7].

The Hopfield neural network (HNN) paradigm initially proposed by Hopfield [3] has been widely used for solving optimization problems. In [5], a context-sensitive technique for unsupervised change detection based on a discrete HNN is proposed. A difference image is obtained by the change vector analysis (CVA) technique and each spatial position in it is represented as a neuron in the HNN. The network topology is built on a difference image. Each neuron corresponds to a pixel of the difference image and is connected to all the neurons in the neighborhood. Each node is characterized by its state, which determines if a pixel has changed. In [6], an optimization relaxation approach based on the analog Hopfield neural network (HNN) for change detection problem has developed. An energy function is derived making use of both the contextual information and the self-data information in difference image. All those techniques have ignored the color information included in the image. In this paper, our HNN model is constructed on each plane of the RGB image to characterize both the spatial-contextual information included in the neighborhood of each pixel within the same color plane and the interaction between different color planes.

The major contribution of the proposed technique is that a modified HNN is designed to characterize the spatial-contextual information included in the neighborhood of each pixel within each color plane and the interaction between different color planes in an optical remote sensing imagery. The paper is organized as follows: Section II provides a brief review of the Hopfield...
neural network. Section III describes the proposed change detection technique. The data sets used in the experiments and the performance of the method are described in Section IV. Finally, conclusions and perspectives are drawn in Section V.

2. A Review on the HNN

A Hopfield neural network consists of a set of neurons. The output of each neuron is feedback to each of the other neurons in the network so that the response of such a network is dynamic. An energy function of the network is defined by using the network architecture, i.e., the number of neurons, their output functions, threshold values, connection between neurons, and the strength of the connections [4]. This means that after applying a new input, the output is calculated and fed back to modify the input. The output is then recalculated, and the process is repeated. Successive iterations produce smaller and smaller output changes, until eventually the outputs become constant, or achieve an acceptable stability. Therefore, the computational complexity is inherently reduced by the interactions among all the units. The total input \( U_i \) to node \( i \) is given by

\[
U_i = \sum_{j \neq i} W_{ij} V_j + I_i
\]

where the weight \( W_{ij} \) representing the interconnection strength between the node \( i \) and node \( j \) in the network are assumed to be symmetric, and \( n \) is the total number of nodes in the neighborhood. The output \( V_i \) of neuron \( i \) is defined as

\[
V_i = g(U_i)
\]

where \( g(\cdot) \) is an activation function.

There are two kinds of Hopfield networks, continuous and discrete, which differ on the output values a neuron can take. In [3], it has shown that continuous networks perform better since they have the ability to smooth the surface of the energy function and to prevent the system from being stuck in minor local minima.

For continuous Hopfield networks, the dynamic energy of a node is defined by

\[
\frac{dU_i}{dt} = -\frac{U_i}{R_i} + \sum_{j \neq i} W_{ij} V_j + I_i
\]

where \( R_i \) is a time constant.

3. The Proposed Change Detection Technique Based on Modified HNN

3.1. Description of the Network Architecture

Figure 1. First-order topological network. Each neuron in the network is connected neighbors both within its own color plane and in other two color planes. Neurons are represented by circles, and lines represent connections between neurons.

Let us consider two georeferenced and coregistered optical images \( X_1 \) and \( X_2 \) acquired over the same geographical area but at two different time \( t_1 \) and \( t_2 \) respectively. \( D = \{ x_s \mid s \in S \} \) is the difference image, where \( S \) is a set of sites contained within an image and \( x_s \) is corresponding to the color feature. Our aim is to generate a change map that represents changes that occurred on the ground between the acquisition dates.

Under the Continuous HNN, we consider the output image as a network of nodes where each node is associated to a pixel location in the difference image \( D \). Each node is characterized by its state value, ranging in \([-1, 1]\). The spatial correlation between neighboring pixels is modeled by defining the spatial neighborhood systems \( N \) of order \( d \), for a given spatial position \((m, n)\) as \( N = \{ (m, n) + (u, v), (u, v) \in N \} \). The neuron in position \((m, n)\) is connected to its neighboring units included in \( N \) within its own color plane and between different color planes. Fig. 1 depicts a first-order \((N^1)\) topological network. Let \( W_{ij} \) be the connection strength between the \((m, n)\)th and \((u, v)\)th neurons in the neighborhood \( N \). Hence, the presented architecture can be seen as a modified version of the Hopfield network in which the connection strength to all neurons outside the neighborhood \( (N^d) \) is zero. It can be seen that the output of a neuron depends not only on the elements in its own plane but also on the ones in the other two planes, which is more reasonable for RGB imagery.

In this way, the network architecture is intrinsically able to model the spatial-contextual information of each pixel. Therefore, the network state is characterized by the states of the nodes. After the iterative energy optimization process, the nodes have achieved its final state value when the network stability is reached.

3.2. Network Initialization

The network initialization is carried out by a simple thresholding method. Three thresholds \( \theta \) are produced by a simple thresholding algorithm on each RGB plane respectively. Given a pixel \((m, n)\) in one plane of the difference image, its associated node \((m, n)\) in the network is initialized as follows:
\[ V_i = \begin{cases} \frac{(D(m,n) - th_i)}{(L_{max} - th_i)}, & D(m,n) \geq th_i \\ \frac{(D(m,n) - th_i)}{th_i}, & D(m,n) < th_i \end{cases} \] (4)

Where \( D(m,n) \) is the gray value of pixel \((m,n)\) in \(i\) plane of the difference image, and \(x_{th}\) is the threshold value. \(i \in \{R, G, B\}\), denoting the each plane in difference image.

### 3.3. Energy Functions Formation

Considering the dynamic energy of a single node defined in (3), the energy function \(E\) of the continuous model is given as follow [5]:

\[ E = \sum_{k \in \{R, G, B\}} \left( -\sum_{j \in N_i} W_{ij} V_j - \sum_{j \in N_i} I_i V_j + \beta \sum_{j \in N_i} \int g^{-1}(V) dv \right) \] (5)

The energy function has three parts: the first part models the local field, whereas the second and third parts correspond to the threshold value \(I\) of each neuron in the network and the input bias, respectively. In terms of the change detection for optical images in this paper, we apply the following three principles [6]: 1) proximity, changed/unchanged pixels that lie close in space tend to group; 2) similarity, changed/unchanged pixels with similar spectral values usually are labeled with the same label; 3) connectedness, changed/unchanged pixels that lie inside the same connected region are generally grouped into the same cluster.

Based on the above principles, the energy function can be composed of three parts: the first part models the local field, whereas the second and third parts correspond to the threshold value \(I\) of each neuron in the network and the input bias, respectively. In terms of the change detection for optical images in this paper, we apply the following three principles [6]: 1) proximity, changed/unchanged pixels that lie close in space tend to group; 2) similarity, changed/unchanged pixels with similar spectral values usually are labeled with the same label; 3) connectedness, changed/unchanged pixels that lie inside the same connected region are generally grouped into the same cluster.

The data consistency between the nodes \(i\) and node \(j\) is mapped into the coefficient \(W_{ij}\) as follows:

\[ W_{ij} = \begin{cases} \frac{1 - f(i,j)}{dist(i,j)}, & j \in N_i \\ 0, & j \notin N_i \end{cases} \] (6)

Where \(f(i,j)\) measures the difference in the color feature of pixels corresponding to nodes \(i\) and node \(j\), gives the spatial distance between node \(i\) and node \(j\). According to [6], the data information is mapped through the a posteriori probabilities that given a pixel value in the difference image \(D\), which represents the statistics of the gray level in the difference image.

\[ f(i,j) = p(H_k^i|\chi^j_i) - p(H_k^i|\chi^j_{ij}) \] (7)

Where \(p(H_k^i|\chi^j_i)\) is the color feature posterior probability of label \(H_k\) conditioned on \(\chi^j_i\), \(H_k \in \{H_0,H_1\}\) denote the label of the change map. In this paper, \(p(H_k|\chi^j_i)\) is modeled in terms of the Gaussian distribution, and is estimated through the EM algorithm.

\[ g(U_i) = \tanh(U_i / \beta) \] (8)

Where \(g\) is decreased as the iteration number increases. A more detail about \(\beta\) can be found in [6].

At each iteration \(itr\), the external input bias \(I_{mn}(itr)\) of the neuron at position \((m,n)\) updates its value by taking the previous iteration output value \(V_{mn}(itr-1)\). As the network begins to update its state, the energy value is gradually reduced until the minimum is reached.

### 4. Experiments

#### 4.1. Design of experiments

In order to assess the performances of the proposed approach, we consider two different data sets related to Indonesia and Lebanon. All change maps obtained for each data set were generated by using the continuous model. The whole change detection procedure is summarized as following six steps:

First, the difference images were obtained from the original images which were preprocessed by weighted smooth with a window size of \(7 \times 7\). Second, each node in the HNN was associated to a pixel location in the RGB plane of difference image \(D\), and a topological network was constructed as shown in figure 1. Third, each component of the RGB vector in the difference image was segmented with a simple thresholding algorithm respectively. Each node \((m, n)\) in the network is initialized through (4). Fourth, the coefficients \(W_{ij}\) were computed according to (5). Here we consider the first- and second-order neighborhood in the plane the node \((m, n)\) located, while only consider a first-order neighborhood plus the corresponding node \((m, n)\) in the other two planes. Density function \(P(D(m,n)|H_k)\) and priori probability \(P(H_k)\) were estimated through the EM algorithm with the initialization obtained by the above mentioned thresholding method. The gain \(\beta\) followed the strategy reported in [6]. Fifth, for each node, the energy was computed according to (5). Finally, update the change map \(V\) until no node changing, or reaching the maximum number of iterations allowed. The final change map is obtained with the majority vote method by combining RGB three change maps.

The accuracy of the results are evaluated in terms of: 1) error rate (PE); 2) false alarm rate (PF); 3) missed alarm rate (PM).

#### 4.2. Experimental Results on Indonesia Data Set

The Indonesia data set relate to the Indian Ocean tsunami in 2004. It contains a pair of Ikonos 2m resolution optical images which were acquired before and after the tsunami on January 10, 2003 and December 29, 2004 respectively, over Aceh, Sumatra, Indonesia. After
co-registration, each image had the size of 256×256 pixels. Fig. 2(a) and (b) shows the image taken before and after the tsunami, respectively. This area in the images was mainly covered by green vegetables, and the remarkable changes were that original covered vegetation submerged and seacoast eroded by flood. To make a quantitative evaluation of the effectiveness of the proposed approach, the reference map was obtained by Manual Trial-and-Error thresholding Procedure (MTEP), showing in Figure 2(d). Figure 2(c) illustrates the change map was produced by our proposed technique. Compared with the reference map, the change map obtained by our approach had an obvious false alarm (red circle in figure 2(a)), the reasons of which might be that there was a shadow on the image obtained before the tsunami. Considering the error typology, the number of missed alarms PM was 236, and the number of false alarms PF was 194. The missed alarms mainly located along the interface between the eroded earth and the sea (blue circle in figure 2(b)) where the sea water was blended with the mud.

![image](image1)

Figure 2. Indonesia data set. (a) Ikonos image acquired on January 10, 2003, (b) Ikonos image December 29, 2004, (c) change map obtained by our proposed technique, (d) reference map of the change area by using MTEP technique.

4.3. Experimental Results on Lebanon Data Set

For Lebanon data set, a pair of Quickbird 0.4m resolution color images. The images were taken over downtown area in Beirut, on November, 2003 and February, 2005 respectively, and the image size was 285 × 285 pixels. Between the two acquisition dates a building was erected in the square. The corresponding images are shown in Fig. 3(a) and (b). As done for the Indonesia data set, in this case, a reference map was manually defined [see figure 3(d)]. In this experiment, the changes caused by the small cars were ignored, because those changes were not the main arm of our detection. Compared with the reference map, the false alarms PF were 299 pixels and the missed alarms PM were 238 pixels. Error rate is 537 pixels. The false alarms (red circle in figure 3(a) and 3(b)) could be explained by the influence of illumination and shadow. The missed alarms mainly lay on the edge of the new erected building.

![image](image2)

Figure 3. Lebanon data set. (a) Quickbird image acquired on November, 2003, (b) Quickbird image February, 2005, (c) change map obtained by our proposed technique, (d) reference map of the change area by using MTEP technique.

5. Conclusions and Discussions

In this paper, we present a novel unsupervised change for high spatial resolution optical imagery based on modified Hopfield neural network. The technique models the spatial-contextual information between neighboring pixels located both in its own color plane and other two color planes by using a modified Hopfield neural network constructed on each pixel in the difference image, where the connection coefficients are defined considering their spatial distance from each other. An energy function is also formed by integrating the data and contextual information together. The state of the network is initialized by a simple thresholding method and three initial change maps are produced. The final change map is obtained with the majority vote method by combining those three change maps. Experiments results obtained on different data set confirm the effectiveness of the proposed approach, which also indicate the HNN approach is a suitable strategy to embed both spatial-contextual and self-data information. It's also proven that the proposed approach has exploited the abundant
correlative information provided by the RGB image which makes it more robust again noise.

The main contribution of this paper lies in the construction of a continuous Hopfield-type neural network on the RGB image to characterize the spatial-contextual information included in the neighborhood of each pixel within each color plane and the interaction between different color planes for solving the unsupervised image change detection problem. The main drawback is that a local minimum of energy function is often reached and the final change map depends much on the initialization.

Future research may be related to: 1) a more sophisticated technique for a proper initialization of the network, which helps the energy function converge more rapid and reach a more approximate minimum; 2) a more accurate model can be developed to compute the interconnection strength between different nodes; 3) other color space may be considered and more color information can be exploited.

ACKNOWLEDGMENTS

This work was supported by the 863 Program of the People’s Republic of China.

REFERENCES