Extraction of Hydrographic Regions from Remote Sensing Images Using an Oscillator Network with Weight Adaptation

Xiuwen Liu, Ke Chen, and DeLiang L. Wang

Technical Report OSU-CISRC-4/99-TR12
Department of Computer and Information Science
The Ohio State University, Columbus, OH 43210-1277

Available at anonymous ftp site **ftp.cis.ohio-state.edu**
Directory **pub/tech-report/1999**
File in gzipped postscript format **TR12.ps.gz**

April 20, 1999

Authors are with the Department of Computer and Information Science, and K. Chen and D. L. Wang are also with the Center for Cognitive Science, at The Ohio State University, Columbus, OH 43210, USA (e-mail: {liux, dwang}@cis.ohio-state.edu). K. Chen's current address is National Laboratory of Machine Perception and Center for Information Science, Peking University, Beijing 100871, China (e-mail: chen@cis.pku.edu.cn).

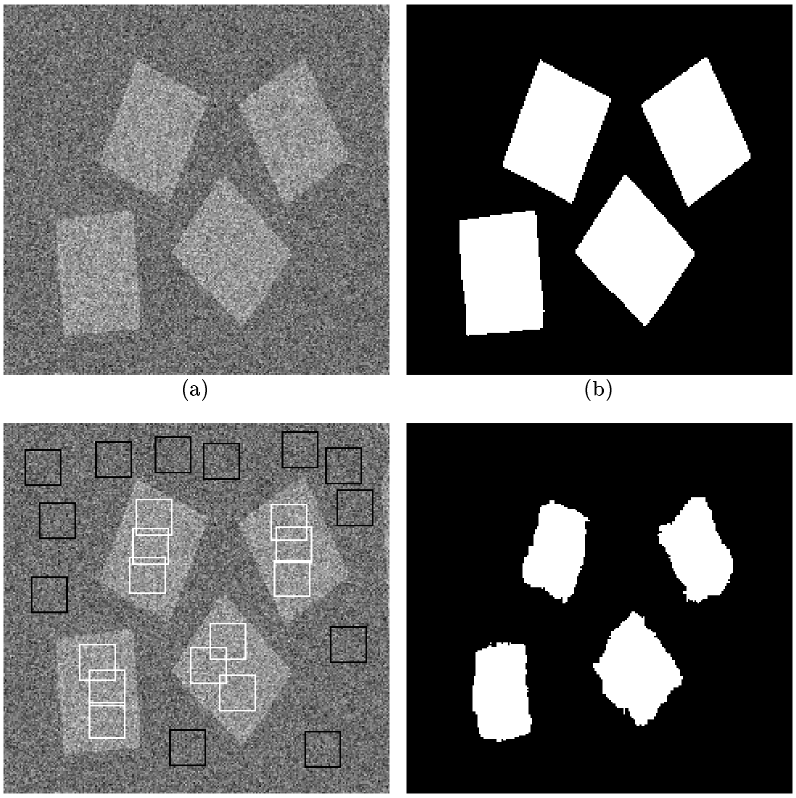
Abstract

We study the extraction of objects with accurate boundaries from remote sensing images. We use a locally excitatory globally inhibitory oscillator network (LEGION) as a representational framework to combine the advantages of classification-based methods and locally coupled networks. A multi-layer perceptron is used to select seed points within regions to be extracted. The boundaries of the extracted regions are accurately located through a topology preserving LEGION network. A novel weight-adaptation method, which preserves significant boundaries between regions and smoothes details due to variations and noise, is employed to increase the robustness of the system. Together, these provide a generic framework for feature extraction. A functional system has been developed and applied to hydrographic region extraction from Digital Orthophoto Quarter-Quadrangle (DOQQ) images. Experimental results show that the extracted regions are comparable with existing topographic maps and can be used for map revision.

Index Terms — **LEGION, Weight adaptation, Multilayer perceptron, Automated feature extraction, Hydrographic region extraction, Classification, Map revision.**

I. Introduction

With the availability of remotely sensed high-resolution imagery and advances in computing technologies, cost-effective and efficient ways to generate accurate geographic information are possible. Because geographic information is implicitly encoded in images, the critical step is how to extract geographic information from images and make it explicit. While humans can efficiently and robustly extract desired features from images, image understanding is regarded as one of the most difficult problems in machine intelligence. For remote sensing applications, classification is one of the most commonly used techniques to extract quantitative information from images. Because multilayer perceptrons can potentially approximate complex decision functions, they have been widely used for solving many practical problems and specifically in remote sensing applications [8] [25] [1] [23]. The major advantages of neural network approaches are that no prior information is required and parameters can be obtained automatically through training [20]. A major disadvantage is that neural networks classify each pixel individually and do not incorporate *contextual information*, i.e., the relationship among neighboring pixels, resulting in relatively poor performance when the variations within a class are large. To illustrate the problem, Fig. 1(a) shows a noisy synthetic image and Fig. 1(b) shows the ground truth image. A three-layer perceptron is trained using a standard back-propagation algorithm [11] with 12 positive and 12 negative examples as shown in Fig. 1(c). Figure 1(d) shows the classification result. While the central regions are classified correctly, the boundaries are not located accurately, resulting in a large classification error as shown in Table I.



(a) (b)

(c) (d)

Fig. 1. Classification result of a noisy synthetic image using a three-layer perceptron. (a) The input image with size of 230×240. (b) The ground truth image. (c) Positive and negative training samples. Positive examples are shown as white and negative ones as black. (d) Classification result from a three-layer perceptron.

Classification can be posed as a statistical mapping from observations to specified classes. In general, the design of a statistical pattern classification algorithm requires some prior statistical information. With assumptions of specific distribution and error functions, many statistical classifiers have been proposed [5] [16] and have been widely used in remote sensing applications [2] [16] [9]. The statistical formulation provides a unified way to incorporate prior knowledge and contextual information. Contextual classifiers explicitly incorporate contextual information in classification [6] [12], resulting in significant performance improvement. Markov Random Fields, as a special case in specifying contextual information, have been successful in image restoration, modeling, and segmentation [7]. Essentially by specifying joint distributions through Gibbs' distribution, the central computation task is an optimization problem for non-convex energy functions with high dimensionality. Because ideal optimization is computationally prohibitive, in practice only a local optimum can be obtained. With approximate but more efficient optimization algorithms, Markov Random Fields have been applied widely in remote sensing applications [17] [22]. As we can see from the development of classification algorithms, one of the critical issues in improving classification performance is how to incorporate contextual information.

We attempt to develop a framework for automated feature extraction that can derive accurate geographic information for map revision and be applied to very large images, such as DOQQ images. In this chapter, we pose the automated feature extraction problem as a binding problem. Pixels that belong to desired regions should be bound together to form desired objects. We use a LEGION network [24] [28] [29], which is a generic framework for binding and image segmentation. As shown analytically, LEGION networks can achieve both synchronization within an oscillator group representing a region and desynchronization among different oscillator groups rapidly. This offers a theoretical advantage over pure local networks such as Markov Random Fields [7], where efficient convergence has not been established in general. To improve the performance, we incorporate contextual information through a weight adaptation method proposed by Chen et al [4]. As multiple scales in edge detection algorithms [13] can be viewed as a way of incorporating contextual information by relating detected structures at different scales, weight adaptation can be viewed as a concrete multiple scale approach with two scales. Instead of applying the same operators at different scales, in weight adaptation, statistical information from a larger scale is derived and mainly used to govern a locally coupled adaptation process, resulting in accurate boundary localization and robustness to noise. We assume that features to be extracted are specified through examples and use a multilayer perceptron trained through back-propagation [11] for seed selection. To reduce the number of necessary training samples and increase the generalization of the classification method, instead of classifying the entire image, we use a multilayer perceptron to identify seed points only. LEGION is then used to provide a framework for integration. We have developed a functioning system using the proposed method for hydrographic feature extraction from DOQQ images and have obtained satisfactory results.

This chapter is organized as follows. Section II provides an overview of LEGION networks. Section III introduces weight adaptation. Section IV describes a multilayer perceptron for automated seed selection. Section V provides experimental results using DOQQ images. Section VI concludes the chapter with further discussions.

II. Grouping based on Oscillatory Correlation

Through theoretical investigation of brain functions, von der Malsburg [27] hypothesized that temporal correlation provides a generic framework for feature binding. In temporal correlation, an object is represented by the correlation of firing activities of cells that encode the object. Recently, this hypothesis has received direct support from neurobiological experiments [10]. Terman and Wang proposed *locally excitatory globally inhibitory oscillator networks* (LEGION) [24] [28] [29] based on oscillatory correlation as a computational framework for feature binding. In oscillatory correlation, which is a computationally effective form of temporal correlation, an object is represented by a group of synchronized oscillators and different objects are represented by groups that are desynchronized. It has been shown analytically that LEGION can rapidly achieve both synchronization in a locally coupled oscillator group and desynchronization among different oscillator groups [24] [29]. The LEGION network provides a computational framework for feature binding and image segmentation. LEGION has been successfully applied to segmenting gray level images [29], texture images [3], range images [14], and medical images [21].

A. Single Oscillator

As the building block of LEGION, a single oscillator is defined as a feedback loop between an excitatory unit x and an inhibitor unit y [24] [29]:

$$\frac{dx}{dt} = 3x - x^3 + 2 - y + I\,H(p - \theta) + S + \rho, \tag{1a}$$

$$\frac{dy}{dt} = \epsilon(\alpha(1 + tanh(x/\beta)) - y). \tag{1b}$$

Here H is the Heaviside step function, I represents external stimulation to the oscillator, and S represents overall coupling from other oscillators in the network. The *potential* associated with the oscillator, p, is introduced to distinguish fragment regions from major regions [29], and θ is a threshold. Basically, only oscillators within a large homogeneous region can develop a potential larger than threshold θ and thus receive effective external stimuli. These oscillators are referred to as *leaders*. Oscillators that belong to noisy regions cannot develop high potentials and are suppressed into a background region. See Wang and Terman [29] for more details regarding the potential term. The parameter ρ denotes a Gaussian noise term and is introduced to test the robustness of the system and to actively desynchronize different input patterns.

The parameter ϵ is chosen to be a small positive number, i.e., $0 < \epsilon \ll 1$. Under this condition, without any coupling S or noise ρ, (1) corresponds to a standard relaxation oscillator [26]. The x-nullcline of (1), $dx/dt = 0$, is a cubic curve; the y-nullcline, $dy/dt = 0$, is a sigmoid, whose steepness is controlled by the parameter β. For a positive external stimulus, i.e., $I > 0$, the $x-$ and $y-$nullcline intersect at the middle branch of the cubic, as shown in Fig. 2(a), and (1) has a stable periodic solution for all sufficiently small ϵ. The periodic solution alternates between the *active phase* and the *silent phase* (see Fig. 2(a)). Within these two phases, (1) exhibits near steady state behavior and the oscillator travels on a slow time scale. The parameter α in (1) controls the relative times that an oscillator spends in the two phases. In contrast, the transition between two phases occurs on a fast time scale. Figure 2(b) shows typical temporal activity of the excitatory unit x. The highly nonlinear behavior results from ϵ and leads to desirable properties for binding and segmentation. For $I < 0$, the two nullclines intersect on the left branch of the cubic, and (1) has a stable fixed point. In this case, the oscillator needs to receive excitatory coupling from other

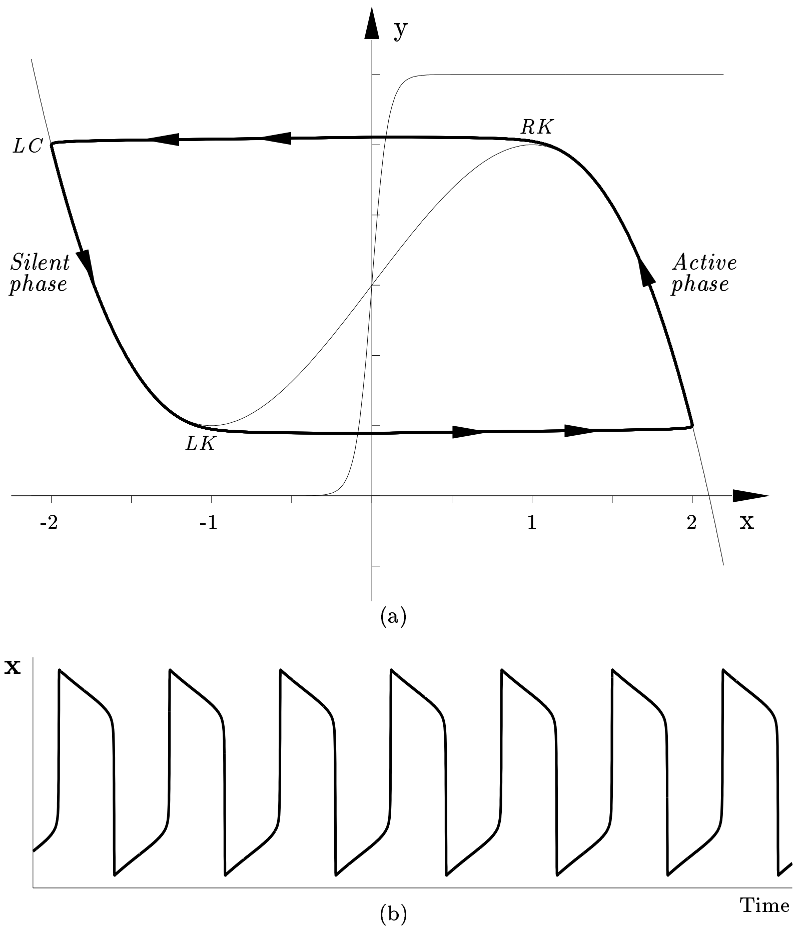


Fig. 2. Temporal properties of a single oscillator. (a) A typical stable limit cycle for a single relaxation oscillator. Thin solid lines show the x and y nullclines, corresponding to $dx/dt = 0$ and $dy/dt = 0$. Thick solid line represents the stable limit cycle. The double arrow indicates much faster motion than the single arrow. The double arrow branches and single arrow ones correspond to fast and slow time scales respectively of an relaxation oscillator. (b) Temporal activity of excitatory unit x, which exhibits two dramatically different time scale. The parameter values used are $I = 1.0$, $\epsilon = 0.02$, $\alpha = 3.0$, and $\beta = 0.1$.

oscillators in order to oscillate.

B. LEGION for Image Segmentation

A LEGION network for image segmentation consists of locally coupled relaxation oscillators and a global inhibitor. A typical architecture of a two-dimensional LEGION network is shown in Fig. 3, which is used throughout this chapter. In the two-dimensional case, an oscillator is referred by (i, j), where i and j indicate the row and column of the oscillator. In LEGION, the

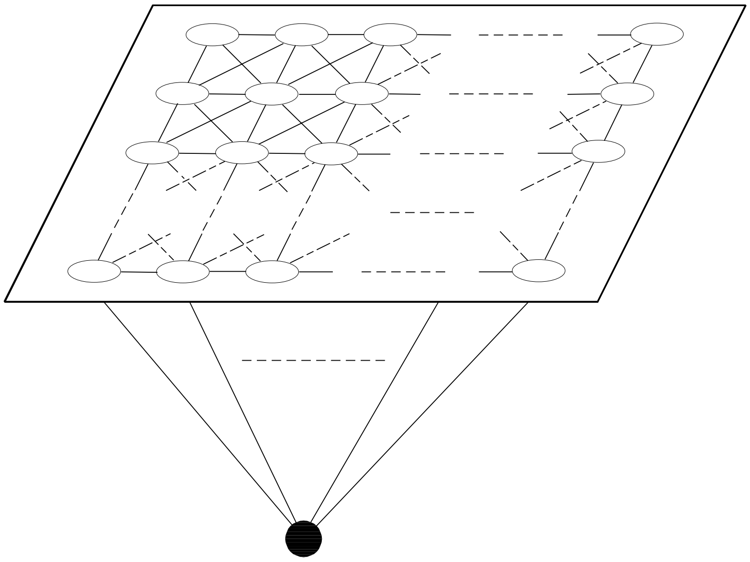


Fig. 3. Architecture of a two-dimensional LEGION network for image segmentation. An oscillator is represented by an open circle, and the global inhibitor is represented by the filled circle.

coupling term S_{ij} for oscillator (i,j) in (1) is given [4]:

$$S_{ij} = \frac{\sum_{(k,l)\in N(i,j)} H(x_{kl})/(1+|W_{ij;kl}|)}{\log\left(\sum_{(k,l)\in N(i,j)} H(x_{kl})+1\right)} - W_z H(z-\theta_z), \tag{2}$$

where the first term is total excitatory coupling that oscillator (i,j) receives from the oscillators in a local neighborhood $N(i,j)$, and $W_{ij;kl}$ is the dynamic connection from oscillator (k,l) to (i,j). Here $W_{ij;kl}$ encodes dissimilarity to simplify the equations for weight adaptation, hence the reciprocal[1]. Note that, unlike in [29], the first term implements a logarithmic grouping rule [4], which generates better segmentation results in general. In the second term, θ_z is a threshold, and the global inhibitor z is defined as

$$\frac{dz}{dt} = \phi(\sigma_\infty - z), \tag{3}$$

where $\sigma_\infty = 1$ if $x_{ij} \geq \theta_z$ for at least one oscillator (i,j) in the entire network, and $\sigma_\infty = 0$ otherwise. The parameter ϕ determines the rate at which the global inhibitor reacts to such stimulation. If the activity of at least one oscillator is above threshold θ_z, the global inhibitor approaches to 1, and in turn it sends inhibition to the entire network. W_z in (2) is the weight of inhibition from the global inhibitor. This leads to desynchronization among different oscillator groups [24] [29].

The computation of LEGION can be summarized as follows. An oscillator triggers the global inhibitor when it jumps to the active phase. As a result, the global inhibitor is activated and exerts

[1]This interpretation is different from a previous one used in [24] [29]. After algebraic manipulations, the equations can be re-written in terms of the previous, more conventional interpretation.

inhibition on all oscillators in the network. On the other hand, an oscillator in the active phase propagates its activation to its local neighbors through the local coupling as shown in the first term of (2), and the excitation will propagate until all the oscillators representing the same object are activated. Thus the underlying dynamics of LEGION rapidly achieves both synchronization by local cooperation through excitatory coupling and desynchronization by global competition via the global inhibitor. For rigorous analysis, see Terman and Wang [24] and Wang and Terman [29]

III. Weight Adaptation

Given LEGION dynamics, to extract desired features, we need to form local connections $W_{ij;kl}$ in (2) based on the input image. As illustrated in Fig. 3, effective couplings in a neighborhood are used. Without introducing assumptions about the desired features, $W_{ij;kl}$ in general can be formed based on the intensity values at the corresponding pixels (i,j) and (k,l) in the input image. However, due to variations and noise in real images, individual pixel values are not reliable, and the resulting connections would be noisy and give undesirable results. Figure 4(a) shows a one-dimensional signal which is a row from the image shown in Fig. 1(a). As shown in Fig. 4(b), the connections formed based on input intensity values are noisy, which lead to undesired region boundaries.

To overcome this problem, we use a weight adaptation method for noise removal and feature preservation [4]. For each oscillator in the network, two kinds of connections, namely, *fixed* and *dynamic* connections, are introduced. For oscillator (i,j), the fixed connectivity specifies a group of neighboring oscillators which affect the oscillator, and the associated neighborhood is called *lateral* neighborhood $N_l(i,j)$. On the other hand, the dynamic connectivity encodes the transient relationship between two oscillators in a local neighborhood during weight adaption, and the associated neighborhood is called *local* neighborhood $N(i,j)$. To achieve accurate boundary localization, in this chapter, $N(i,j)$ is defined as the eight nearest neighborhood of (i,j), as depicted in Fig. 3. Fixed connection weights are established based on the input image, while dynamic connection weights adapt themselves for noise removal and feature preservation, resulting interactions between two scales. Intuitively, dynamic weights between two oscillators should be adapted so that the absolute dynamic weight becomes small if the corresponding pixels are in a homogeneous region, while the weight should remain relatively large if the corresponding pixels cross a boundary between different homogeneous regions. Based on the observation that most of the discontinuities in the lateral neighborhood $N_l(i,j)$ correspond to significant features, such discontinuities should remain unchanged and be used to control the speed of weight adaptation for preserving features. Such discontinuities in the lateral neighborhood are called *lateral discontinuities* [4]. Furthermore, because proximity is a major grouping principle [18], we use another measure that reflects *local discontinuities* sensitive to the changes of local attributes among local oscillators. The lateral

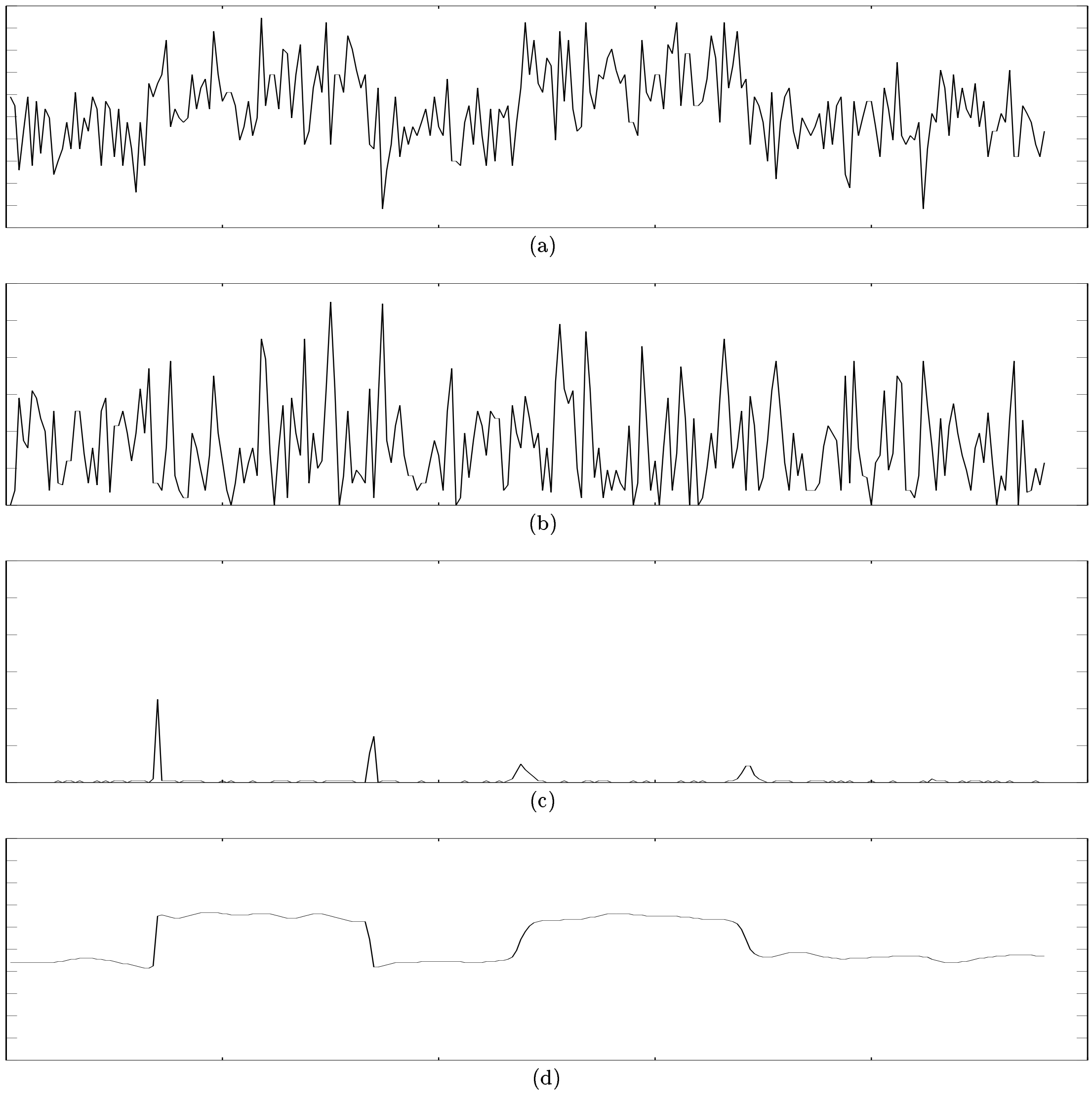


Fig. 4. Lateral connection evolution through weight adaptation illustrated using the 170th row from the image shown in Fig. 1(a). (a) The original signal. (b) Initial connection weights. (c) Connection weights after 40 iterations. (d) Corresponding smoothed signal.

neighborhood provides a more reliable statistical context, by which the weight adaptation algorithm is governed. The local neighborhood utilizes the statistical context and local geometrical constraints to adaptively change the local connections. These two discontinuity measures are jointly incorporated in weight adaptation.

Mathematically, the weight adaption method is formulated as follows. For oscillator (i, j), the weight of its fixed connection from oscillator (k, l) in $N_l(i, j)$, $T_{ij;kl}$, is defined as the difference between the external stimuli received by (i, j) and (k, l), i.e.

$$T_{ij;kl} = I_{kl} - I_{ij}. \tag{4}$$

Here I_{ij} and I_{kl} are the intensities of pixels (i,j) and (k,l) respectively. For oscillator (i,j), the fixed connections exist only in $N_l(i,j)$ and $T_{ij;kl} = -T_{kl;ij}$. On the other hand, to achieve accurate boundary localization, a dynamic connection weight from oscillator (k,l) to oscillator (i,j), $W_{ij;kl}$, is defined only within a local neighborhood $N(i,j)$ and initialized to the corresponding fixed weight, i.e., $W_{ij;kl}^{(0)} = T_{ij;kl}$. Dynamic weight $|W_{ij;kl}^{(t)}|$ encodes the dissimilarity between oscillators (i,j) and (k,l) at time t.

First, the variance of all the fixed weights associated with an oscillator is used to measure its lateral discontinuities. For oscillator (i,j), the mean of its fixed weights on $N_l(i,j)$, μ_{ij}, is calculated using

$$\mu_{ij} = \frac{\sum_{(k,l)\in N_l(i,j)} T_{ij;kl}}{|N_l(i,j)|}. \tag{5}$$

Accordingly, we compute the variance of its fixed weights, σ_{ij}^2, by

$$\begin{aligned} \sigma_{ij}^2 &= \frac{\sum_{(k,l)\in N_l(i,j)} \left(T_{ij;kl} - \mu_{ij}\right)^2}{|N_l(i,j)|} \\ &= \frac{\sum_{(k,l)\in N_l(i,j)} T_{ij;kl}^2}{|N_l(i,j)|} - \left(\frac{\sum_{(k,l)\in N_l(i,j)} T_{ij;kl}}{|N_l(i,j)|}\right)^2. \end{aligned} \tag{6}$$

The variance, σ_{ij}^2, is normalized through

$$\tilde{\sigma}_{ij}^2 = \frac{\sigma_{ij}^2 - \sigma_{min}^2}{\sigma_{max}^2 - \sigma_{min}^2}. \tag{7}$$

Here σ_{max}^2 is the maximal variance across the entire image and σ_{min}^2 the minimal. Intuitively, $\tilde{\sigma}_{ij}^2$ encodes the lateral discontinuities for oscillator (i,j). Oscillators corresponding to significant features tend to have large $\tilde{\sigma}_{ij}^2$ and vice versa. Based on this observation, the local discontinuity of an oscillator with a high lateral discontinuity should be preserved; the local attributes of an oscillator with a low lateral discontinuity should adapt towards homogeneity.

To preserve accurate region boundaries during weight adaptation, local discontinuities are detected along four orientations, namely vertical (V), horizontal (H), diagonal (D), and counter-diagonal (C), respectively. Accordingly, we define four detectors for oscillator (i,j) as

$$D_{H_{ij}} = |W_{ij;i-1,j} - W_{ij;i+1,j}|, \tag{8a}$$

$$D_{V_{ij}} = |W_{ij;i,j-1} - W_{ij;i,j+1}|, \tag{8b}$$

$$D_{C_{ij}} = |W_{ij;i-1,j-1} - W_{ij;i+1,j+1}|, \tag{8c}$$

$$D_{D_{ij}} = |W_{ij;i-1,j+1} - W_{ij;i+1,j-1}|. \tag{8d}$$

If there is an edge through (i,j) in one of these four orientations, the corresponding detector will respond strongly. Based on the responses from the four detectors, a measure of local discontinuity for oscillator (i,j) is defined as

$$D_{ij} = \frac{D_{H_{ij}} + D_{V_{ij}} + D_{C_{ij}} + D_{D_{ij}}}{4}. \tag{9}$$

Here D_{ij} is sensitive to local discontinuity along all the orientations.

To integrate the lateral and local attributes of oscillator (i,j) and realize noise removal and feature preservation, $V_{ij}^{(t)}$ is introduced based on $\tilde{\sigma}_{ij}^2$ and D_{ij}:

$$V_{ij}^{(t)} = \frac{\sum_{(k,l)\in N(i,j)} \exp\left[-\left(\kappa\Phi(\tilde{\sigma}_{kl}^2, \theta_\sigma) + D_{kl}^{(t)}/s\right)\right] W_{ij;kl}^{(t)}}{\sum_{(k,l)\in N(i,j)} \exp\left[-\left(\kappa\Phi(\tilde{\sigma}_{kl}^2, \theta_\sigma) + D_{kl}^{(t)}/s\right)\right]}, \tag{10}$$

where s $(s > 0)$ is a parameter that determines to what extent local discontinuities should be preserved, and κ $(\kappa > 0)$ is to control to what extent features should be preserved in terms of lateral discontinuities during weight adaptation. The function $\Phi(\nu, \theta)$ is a rectification function, defined as $\Phi(\nu,\theta) = \nu$ if $\nu \geq \theta$ and $\Phi(\nu, \theta) = 0$ otherwise. To deal with variations due to noisy details, θ_σ $(0 \leq \theta_\sigma \leq 1)$ is introduced to alleviate the influence of noise in the estimation of lateral discontinuities. The degree of lateral discontinuities in an image gives a measure of the significance of the corresponding discontinuities. In (10), if the lateral discontinuities of all the oscillators in $N(i,j)$ are similar, their local discontinuities should play a dominant role in updating $V_{ij}^{(t)}$. In this case, the contribution of dynamic weight $W_{ij;kl}^{(t)}$ in updating $V_{ij}^{(t)}$ is determined by the corresponding local discontinuity $D_{kl}^{(t)}$ of oscillator (k,l) in $N(i,j)$. If $D_{kl}^{(t)}$ is small, $W_{ij;kl}^{(t)}$ has a large contribution, and vice versa. The local attributes of oscillator (i,j) are changed through updating $V_{ij}^{(t)}$ so that the dissimilarity between (i,j) and its neighboring oscillators is reduced with respect to s. Noise removal along with feature preservation is achieved through the reduction of dissimilarity in terms of local discontinuities. On the other hand, when neighboring oscillators of (i,j) have different lateral discontinuities, both lateral and local discontinuities must be employed in determining the contribution from $W_{ij;kl}^{(t)}$ to $V_{ij}^{(t)}$. In this case, when the overall discontinuities associated with oscillator (k,l) are relatively small, $W_{ij;kl}^{(t)}$ makes a large contribution. Lateral and local discontinuities jointly provide a robust way to realize feature preservation and noise removal for those oscillators associated with large lateral discontinuities, i.e. $\tilde{\sigma}_{kl}^2 \geq \theta_\sigma$. The local attributes of oscillator (i,j) are adapted towards reduction of the dissimilarity between (i,j) and the neighboring oscillators with relatively small overall discontinuities. The dissimilarity between (i,j) and those with relatively large overall discontinuities tends to remain relatively large.

Based on (10), We define weight adaptation for $W_{ij;kl}$

$$W_{ij;kl}^{(t+1)} = W_{ij;kl}^{(t)} + \left[\exp\left(-\kappa\Phi(\tilde{\sigma}_{kl}^2(R), \theta_\sigma)\right) V_{kl}^{(t)} - \exp\left(-\kappa\Phi(\tilde{\sigma}_{ij}^2(R), \theta_\sigma)\right) V_{ij}^{(t)}\right]. \tag{11}$$

Here, $W_{ij;kl}^{(t+1)}$ is updated based on the local attributes associated with (i,j) and (k,l). The lateral discontinuity further determines gain control during weight adaptation.

Figure 4(c) and (d) show the adapted connection weights and the smoothed signal after 40 iterations. From Fig. 4(c), one can see that pixels belonging to the same region are strongly coupled, while couplings between pixels belonging to different regions are weak (As pointed out earlier, in this chapter the coupling is the reciprocal of the connection weight). This greatly improves LEGION segmentation performance when noise is large.

This method can be viewed as a simplified and efficient way of utilizing and integrating information from multiple scales. However, it is different from general multiple scale approaches [13]. Instead of applying the same operators on different scales, in lateral neighborhood, statistical information is derived and mainly used to guide the local weight adaptation; in local windows, geometrical constraints are enforced through local coupling, preserving significant boundaries precisely. The weight adaptation scheme is closely related to nonlinear smoothing algorithms [15]. It preserves significant discontinuities while adaptively smoothing variations caused by noise. Compared to existing nonlinear smoothing methods, the weight adaptation method offers several distinctive advantages [4]. First, it is insensitive to termination conditions while many existing nonlinear methods critically depend on the number of iterations. Second, it is computationally fast. Third, by integrating information from different scales, this method can generate better results. Quantitative comparisons with other methods, including various smoothing algorithms, are provided in [4].

IV. Automated Seed Selection

Both LEGION networks and weight adaptation methods are generic approaches, where no assumptions about the features being extracted are made. To extract desired features from remote sensing images, we need to specify relevant features. One way to do this is to use certain parametric forms based on assumptions about the statistical properties of features and noise. However, for map revision and other remote sensing applications, images may be acquired under different conditions and even through different sources, such as DOQQ images. These make it very difficult to model the features using parametric forms. A more feasible way is to specify the desired features through positive and negative examples. There are essentially two broad categories of approaches for solving the problem. Statistical based approaches need knowledge of statistical distribution [2] [20], which may not be available or may be difficult to obtain due to different data sources and acquisition conditions. Artificial neural networks provide an alternate approach. Because they can approximate arbitrary complex functions, artificial neural networks are suitable for describing features being extracted in remote sensing applications, where many factors are

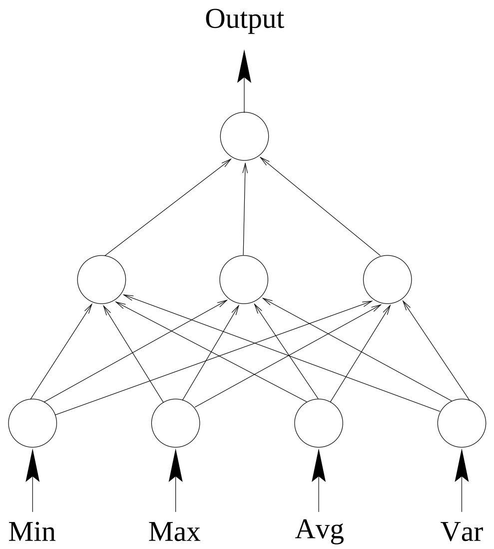


Fig. 5. Architecture and local features for the seed selection neural network.

involved in generating the features in input images. The major advantages of using neural networks include that no prior knowledge of statistical distribution is needed and parameters can be obtained automatically through training [2] [20]. Artificial neural networks, especially multilayer perceptrons, have been widely used in remote sensing applications [8] [25] [1] [23].

To apply a multilayer perceptron, a number of design choices must be made. These choices may greatly affect the convergence of the network and learning results. For our task, we use a three-layer perceptron, with four input units, three hidden units, and one output unit, as shown in Fig. 5. It is trained using a standard back-propagation algorithm [11]. If we present the pixels in the training examples directly to the network, we observe that many training examples are necessary to achieve good results. Due to the potential conflicting conditions, the network often does not converge. To achieve rotational invariance and reduce the necessary number of training samples, instead of presenting the training windows directly to the network we extract several local attributes from training windows as input to the network. More specifically, we use the average value, minimum value, maximum value, and variance from training samples. These values are normalized to improve training and classification results.

To further reduce the number of necessary training samples and deal with variations within features being extracted, we use a three-layer perceptron for seed selection only. In other words, instead of classifying all the pixels directly, we train the network to find pixels that are within a large region to be extracted. The accurate boundaries are derived using an oscillator network with weight adaptation. As demonstrated in the next section, a small of training samples are sufficient for very large images with significant variations. Our training methodology offers a distinctive advantage in reducing necessary training examples. In contrast, many experiments often divide the entire data into training and test sets with approximately the same size to achieve good performance [22], ending up using many more training samples.

V. Experimental Results

We have developed a fully functioning system with a user-friendly graphical interface using the proposed method. As shown in Fig. 3, we construct a two-dimensional LEGION network, where each pixel in the input image corresponds to an oscillator. Seed points are extracted using a three-layer perceptron, the output of which is used to determine the leaders in the LEGION network. Oscillators in a major seed region develop high potentials and thus are oscillating. In this chapter, a seed region is considered to be a major one if its size is larger than a threshold θ_p. The dynamic connection $W_{ij;kl}$ between oscillator (i,j) and (k,l) is established based on the weight adaptation method presented in Section III. We have developed a computationally faster LEGION algorithm for grouping, compared to the one given [29]. The result of region extraction corresponds to all the oscillating oscillators.

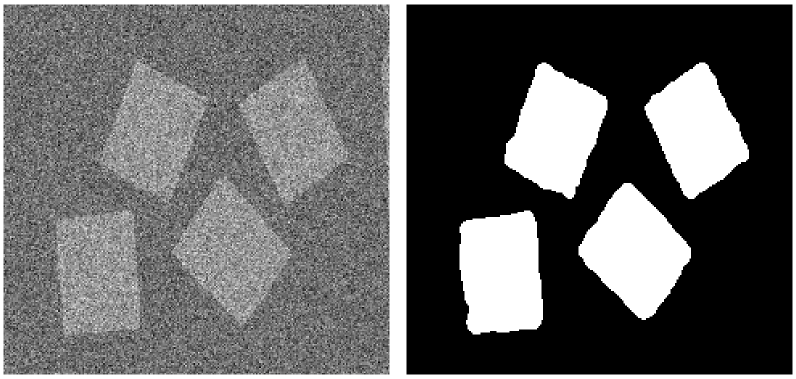
A. Parameter Selection

Our method consists of three relatively independent systems. For LEGION networks, except W_z and θ_p, all other parameters are system parameters and application independent. For weight adaptation, three parameters are involved. In order to achieve optimal results, these parameters are application dependent and subject to tuning. However, as demonstrated in [4], these parameters can be fixed for images of a similar type. In this chapter, they are fixed for all the experiments: $s = 10.0$, $\kappa = 60.0$, and $\theta_\sigma = 0.02$. The lateral neighborhood $N_l(i,j)$ consists of 7×7 oscillators. $W_{ij;kl}$ is updated for 40 iterations using (11).

The three-layer perceptron for seed selection is trained using a standard back-propagation with momentum. The learning rate is set to 0.025 and the momentum parameter is set to 0.9. The activation function used is a sigmoid function, the steepness parameter of which is 0.5. The training is terminated when the error from all the samples is less than 0.05.

B. Synthetic Image

To demonstrate the capability of the proposed method, we have applied the system to a noisy synthetic image shown in Fig. 1(a). A three-layer perceptron is trained using 12 positive and 12 negative samples, as shown in Fig. 1(c). Each sample is a window with 23×23 pixels from the original image. Local attributes are first calculated from the training samples and then fed to a three-layer perceptron as input vectors. The classification result is shown in Fig. 1(d). While the region boundaries are not localized precisely, seed points within central regions are found



(a) (b)

Fig. 6. Segmentation result using the proposed method for a synthetic image. (a) A synthetic image as shown in Fig. 1(a). (b) The segmentation result from the proposed method. Here $W_z = 0.25$ and $\theta_p = 100$.

TABLE I
COMPARISON OF ERROR RATES USING NEURAL NETWORK CLASSIFICATION AND THE PROPOSED METHOD.

Dataset	False nontarget rate		False target rate	
	Classification only	Proposed method	Classification only	Proposed method
Synthetic image	45.07 %	4.89 %	0.00 %	0.23 %
Washington East	51.95 %	11.70 %	1.97 %	0.75 %
Damascus	32.56 %	8.42 %	1.79 %	0.18 %

correctly. Leaders in the LEGION network are then determined based on the classification result from the three-layer perceptron. With weight adaptation, the regions are extracted with quite precise boundaries, as shown in Fig. 6(b). As shown in Table I, the classification error rate is reduced significantly. The total error rate of the proposed method is 5.12% while the total error rate is 45.07% when multilayer perceptron classification is applied only. The obtained result is also comparable with the best result obtained by an edge-based approach [19] with carefully tuned parameters.

C. Hydrographic Region Extraction from DOQQ Images

We have applied the proposed method to extract hydrographic regions from several DOQQ images. High resolution DOQQ images are readily available from many commercial sources and a national coverage for the United States will be available in the near future. Map revision and other remote sensing applications using DOQQ will have significant values since geographical structures undergo constant alterations due to seasonal and other natural and man-made changes. Because it requires very large memory if we process the entire DOQQ image once, DOQQ images are partitioned into tiles with user specified sizes. This partitioning may cause a boundary problem where some tiles may contain only small parts of hydrographic regions and produce no seeds by themselves using our seed selection criterion. This problem is resolved by running the system for several additional iterations. At each iteration, the input image is re-partitioned so that tiles sufficiently overlap with those from the previous partition and seeds are generated using the extracted hydrographic regions so far.

The first DOQQ image, shown in Fig. 7, is from the Washington East, D.C.-Maryland area and the size is 6204×7676 pixels. A three-layer perceptron is trained using 19 positive and 28 negative representational examples where each sample is a 31×31 window. The trained network is then applied to classify the entire DOQQ image, and seed points from classification are shown in Fig. 8. While the pixels in central and major river regions are correctly classified, the river boundaries are rough. Also there are pixels that are misclassified as hydrographic seed points even though they do not belong to any hydrographic regions. As shown in Table I, the false target rate is 1.97%. Here the false target rate is the ratio of the number of non-hydrographic pixels which are misclassified to the total number of true hydrographic pixels. Similarly, the false nontarget rate is the ratio of the number of hydrographic pixels that are misclassified to the total number of true hydrographic pixels. The ground truth, as shown in Fig. 10 is generated by manual seed selection based on a 1:24,000 topographic map from the United States Geological Survey (see Fig. 12(c) for example) and DOQQ image. Figure 11(b) is generated from Fig. 11(a) using the same procedure. Leaders in LEGION are determined with $\theta_p = 4000$. Because noisy seed points cannot develop high potentials, no hydrographic regions are extracted around those pixels. We apply a LEGION network with weight adaptation where there are leaders in the tile being processed. Figure 9 shows the result from our system. As shown in Table I, both the false target and false nontarget rates are reduced dramatically. The false target rate is reduced to 0.75%. Also the hydrographic region boundaries are localized much more accurately, and thus the false nontarget rate is reduced. Mainly because the Kenil Worth Aquatic Gardens, magnified in Fig. 11(a), is not extracted, the false nontarget rate of the proposed method still stands at 11.70%. Because the aquatic area is statistically very similar to soil land, it is not possible to classify it correctly solely using the perceptron.

If we assume that seed points are correctly detected in the area, the proposed method can correctly extract the aquatic region with high accuracy. The result, shown in Fig. 11(b), is generated by manually selecting seed points in the area. This would reduce the error rate of the proposed method even further.

More importantly, major hydrographic regions are extracted with accurate boundaries and cartographic features, such as bridges and islands, are preserved well. These are critically important

Fig. 7. A DOQQ image with size of 6204 × 7676 pixels of the Washington East, D.C.-Maryland area.

Fig. 8. Seed pixels obtained by applying a trained three-layer perceptron to the DOQQ image shown in Fig. 7. Seed pixels are marked as white and superimposed on the original image. The network is trained using 19 positive and 28 negative samples, where each sample is a 31×31 window.

Fig. 9. Extracted hydrographic regions from the DOQQ image shown in Fig. 7. Hydrographic regions are marked as white and superimposed on the original image to show the accuracy of the extracted result. Here $W_z = 0.15$ and $\theta_p = 4000$.

Fig. 10. A grouth truth generated by manually placing seeds based on a 1:24,000 USGS topographic map and DOQQ image. The result was manually edited.

(a) Input image. (b) Extracted regions.

Fig. 11. Hydrographic region extraction result for an aquatic garden area with manually placed seed pixels. Due that no reliable seed region is detected, this aquatic region, which is very similar to soil regions, is not extracted from the DOQQ image as shown in Fig. 9. Extracted regions are marked as white and superimposed on the original image.

for deriving accurate spatial information. The major river, Anacostia River, is extracted correctly. Several roads crossing the river are preserved. To demonstrate the effectiveness of our method in preserving important cartographic features, Fig. 12(a) shows a magnified area around the Kingman Lake. Figure 12(b) shows the classification result, Fig. 12(c) shows the corresponding part of the USGS 1:24,000 topographic map, and Fig. 12(d) shows our result. Within this image, intensity values and local attributes change considerably as shown in Fig. 12(a). The boundaries of small islands are localized accurately even though they are covered by forests. In weight adaptation, the information from the lateral and local windows is jointly used when variances in a local neighborhood are large, resulting in robust feature preservation and noise removal. Similarly, the forests along the river banks are preserved well. A bridge connects the lake and the river is also preserved. As shown in Fig. 12(a), the bridge is spatially small and it would be very difficult for nonlinear smoothing algorithms to preserve this cartographic feature.

By comparing Fig. 12(c) and (a), one can see that hydrographic regions have changed from the map. Note, for example, the lower part of the left branch. This geographical change illustrates our previous point about the constant nature of such changes and the need for frequent map revision. With precise region boundaries from our method, we believe that our system is suited for frequent map revision purposes. While the major features are the same, the lake has shrunk in size, and such shrinkage is captured by our algorithm (see Fig. 12(d)). This suggests that our method can be used for monitoring changes of hydrographic features.

We have also applied our system to the DOQQ image of a rural area around Damascus, Pennsylvania-New York, with 5802×7560 pixels, shown in Fig. 13. Because this dataset os dramatically different from the Washington East dataset, the multilayer perceptron is retrained using 25 positive and 17 negative training examples. Figure 14 shows the result from our method. Table I gives the quantitative results from classification using the three-layer perceptron only and the proposed method compared against a ground truth. Again, the ground truth, shown in Fig. 15 is generated based on a 1:24,000 topographic map of the area. Generally speaking, our extraction results for the Damascus image are comparable with those for the Washington East image (In fact, a little better as revealed in Table I). While the major river, Delare River, is extracted out with branches with great accuracy, and small island features are preserved. However, a small river, the East Branch Callicoon Creek, located at upper right in the DOQQ image, is missing. When carefully checking the DOQQ image, it is clear that the river is similar to its surrounding areas, suggesting that the creek is not prominent.

To summarize, the results generated for both DOQQ images are comparable with the hydrographic regions shown in the United States Geological Survey 1:24,000 topographic maps. In certain cases, our results reflect better the current state of the geographical areas.

By using a multilayer perceptron, characteristics of hydrographic regions are captured with only a small of number of training samples, even though there are considerably variations within the regions. This greatly alleviates the problem of parameter selection. For the DOQQ images shown here, only two parameters, namely W_z and θ_p, need to be changed. This offers a distinctive advantage over many existing methods, where extensive parameter tuning is needed.

VI. Discussions

In this chapter, we present a novel computational framework for extracting geographic features in general from remote sensing images. We demonstrate the feasibility of the method using hydrographic region extraction that is important for remote sensing applications. By using LEGION as a segmentation framework, we combine advantages of different methods. Using a multilayer perceptron, parameters can be selected much more easily and the system can be adapted easily for other types of features. Because multilayer perceptrons do not incorporate geometrical constraints among neighboring pixels, we use a trained perceptron only for seed selection, which determines the leaders in a LEGION network. We have used a weight adaptation method to adaptively change local weights based on the statistical context provided by a large window. This preserves major region boundaries while smoothing out details due to variations and noise. It also preserves cartographic features such as bridges and islands, which are important spatial information. Because geometric constraints are incorporated in LEGION and weight adaptation, we achieve accurate region boundaries. As shown by the numerical comparison, the proposed method significantly improves the classification error from the multilayer perceptron.

Compared with existing classification approaches in remote sensing applications, our method offers several advantages. By using the multilayer perceptron for seed selection only, our method greatly improves generalization of the classification method and reduces the number of necessary training samples. As shown in the experimental results, the network is trained using only about 50 training samples and is successfully applied to classify images with $6,000 \times 8,000$ pixels. It would be very difficult, if not impossible, to train a network to achieve results comparable to ours due to inconsistency among different samples. Through weight adaptation, contextual information is incorporated more efficiently and effectively. More importantly, our method extracts boundaries that are comparable with features shown in topographic maps and thus can be used for map revision purposes. To our knowledge, this is the only system that is demonstrated for map revision using DOQQ images.

Practically, we demonstrate through a prototype system that hydrographic features can be extracted highly automatically from DOQQ images. We have applied the system to several DOQQ images and have obtained good results. With advances in computing technology, very affordable systems can be built for map revision and other feature extraction tasks. Compared with traditional map-making methods based on aerial photogrammetry, our method is computationally efficient. We believe that this kind of technology would be very useful for improving map revision and other remote sensing applications. Further more, because remotely sensed images can be captured more readily with high resolutions, efficient methods like the one proposed here are necessary for generating up-to-date and accurate geographic information.

There are a number of improvements that can be made for our prototype system. In the current version, features are extracted based on a single input image. While good results have been obtained, they can be further improved by using data from multiple sources. Feature extraction from multiple data sources has been studied extensively [2] [22] [1]. Our method can potentially be applied to feature extraction from multiple data sources. One possible way is to extend weight adaptation to vector data by viewing the data from different sources in a vector form. A similar extension has been done for nonlinear smoothing techniques [30]. Another constraint that is not utilized is the relationships among different features; for example, when a road crosses a river, there should be a bridge. In general, the knowledge concerning different features can provide contextual information at a higher level than that currently incorporated in our system. With these improvements, a complete feature extraction system could be feasible.

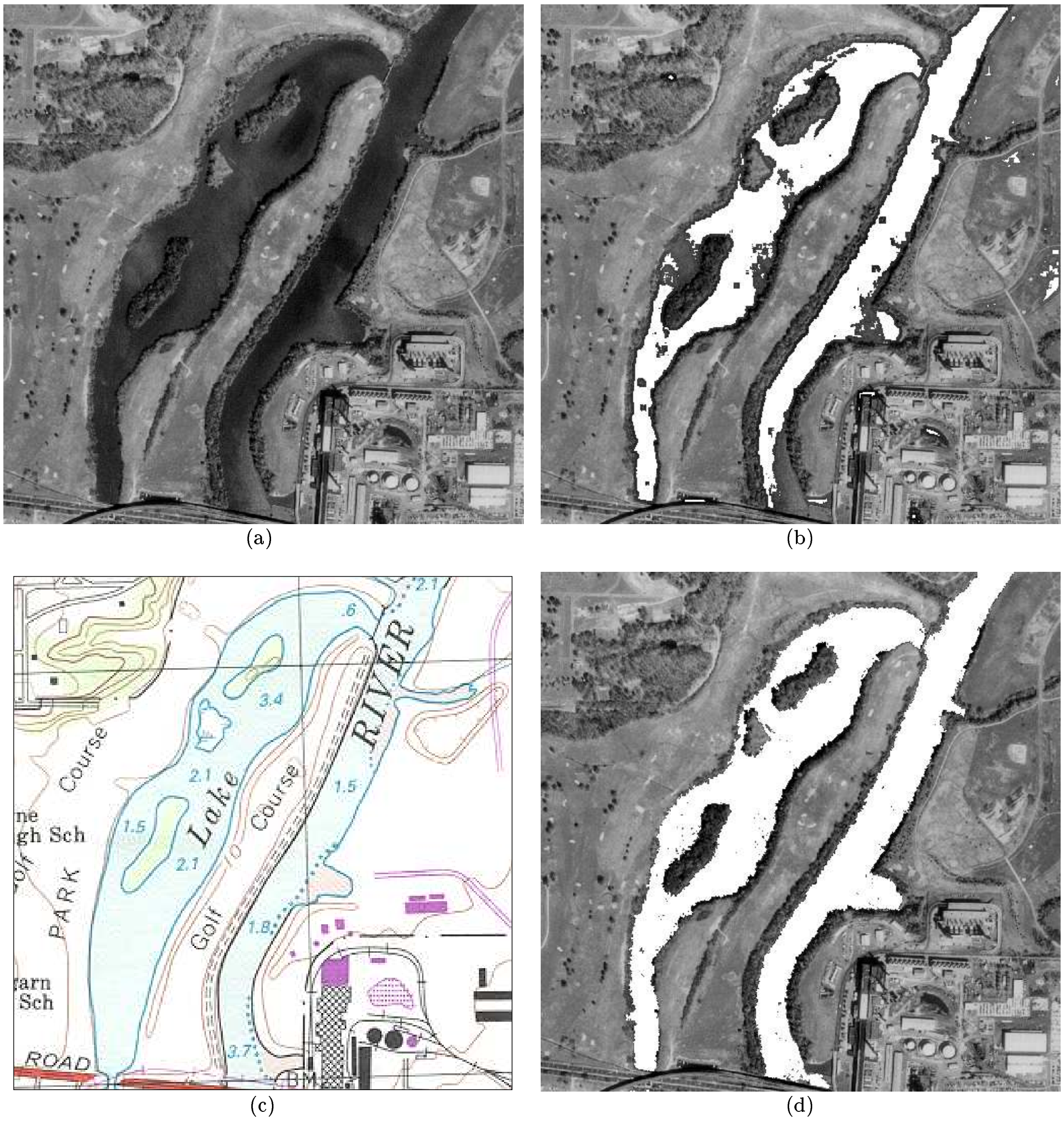


Fig. 12. Extraction result for an image patch from Fig. 7. (a) The input image. (b) The seed points from the neural network. (c) A topographic map of the area. Here the map is scanned from the chapter version and not wrapped with respect to the image. (d) Extracted result from the proposed method. Extracted regions are represented by white and superimposed on the original image.

Fig. 13. A DOQQ image with size of 5802×7560 pixels of Damascus, Pennsylvania-New York area.

Fig. 14. Extracted hydrographic regions from the DOQQ image shown in Fig. 13. The extracted regions are represented by white pixels and superimposed on the original image.

Fig. 15. A ground truth generated based on a 1:24,000 USGS topographic map and DOQQ image.

References

[1] J. A. Benediktsson and J. R. Sveinsson. "Feature extraction for multisource data classification with artificial neural networks". *International Journal on Remote Sensing*, 18(4):727–740, March 1997.

[2] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy. "Neural network approaches versus statistical methods in classification of multisource remote sensing data". *IEEE Transactions on Geoscience and Remote Sensing*, 28(4):540–552, July 1990.

[3] E. Cesmeli and D. L. Wang. "Texture segmentation using Gaussian Markov random fields and LEGION". In *Proceedings of the 1997 IEEE International Conference on Neural Networks*, pages 1529–1534, 1997.

[4] K. Chen, D. L. Wang, and X. Liu. Weight adaptation and oscillatory correlation for image segmentation. Technical Report OSU-CISRC-8/98-TR37, Department of Computer and Information Science, The Ohio State University, 1998.

[5] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis.* John Wiley and Sons, New York, 1973.

[6] K. S. Fu and T. S. Yu. *Statistical Pattern Classification using Contextual Information.* Research Studies Press, Chichester, England, 1980.

[7] S. Geman and D. Geman. "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

[8] S. Gopal and C. Woodcock. "Remote sensing of forest change using artificial neural networks". *IEEE Transactions on Geoscience and Remote Sensing*, 34(2):398–404, March 1996.

[9] B. Gorte and Alfred Stein. "Bayesian classification and class area estimation of satellite images using stratification". *IEEE Transactions on Geoscience and Remote Sensing*, 36(3):803–812, May 1998.

[10] C. M. Gray, P. Konig, A. K. Engel, and W. Singer. "Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties". *Nature*, 338:334–337, 1989.

[11] J. Hertz, A. Krogh, and R. G. Palmer. *Introduction to the Theory of Neural Computation.* Addison-Wesley, Reading, MA, 1991.

[12] Y. Jhung and P. H. Swain. "Bayesian contextual classification based on modified M-estimates and Markov Random Fields". *IEEE Transactions on Geoscience and Remote Sensing*, 34(1):67–75, January 1996.

[13] T. Lindeberg and B. M. ter Haar Romeny. Linear scale-space. In B. M. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, pages 1–41. Kluwer Academic Publishers, Dordrecht, Netherlands, 19994.

[14] X. Liu and D. L. Wang. "Range image segmentation using an oscillatory network". *IEEE Transactions on Neural Networks*, in press.

[15] P. Perona and J. Malik. "Scale space and edge detection using anisotropic diffusion". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:16–27, 1990.

[16] J. A. Richards. *Remote Sensing Digital Image Analysis.* Springer-Verlag, Berlin, 1993.

[17] E. Rignot and R. Chellappa. "Segmentation of polarimetric synthetic aperture radar data". *IEEE Transactions on Image Processing*, 1:281–300, July 1992.

[18] I. Rock and S. Palmer. "The legacy of Gestalt psychology". *Scientific American*, 263:84–90, 1990.

[19] S. Sarkar and K. L. Boyer. "On optimal infinite impulse response edge detection filters". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:1154–1171, 1991.

[20] J. Schurmann. *Pattern Classification: A Unified View of Statistical and Neural Approaches.* John Wiley and Sons, New York, 1996.

[21] N. Shareef, D. L. Wang, and R. Yagel. "Segmentation of Medical Images Using LEGION". *IEEE Transactions on Medical Imaging*, 18(1):74–91, January 1999.

[22] A. H. S. Solberg, T. Taxt, and A. K. Jain. "A Markov Random Field model for classification of multisource satellite imagery". *IEEE Transactions on Geoscience and Remote Sensing*, 34(1):100–113, January 1996.

[23] C. Sun, C. M. U. Neale, J. J. McDonnell, and H. D. Cheng. "Monitoring land-surface snow conditions from SSM/I data using an artificial neural network classifier". *IEEE Transactions on Geoscience and Remote Sensing*, 35(4):801–809, July 1997.

[24] D. Terman and D. L. Wang. "Global competition and local cooperation in a network of neural oscillators". *Physica D*, 81(1-2):148–176, 1995.

[25] D. Tsintikidis, J. L. Haferman, E. N. Anagnostou, W. F. Krajewski, and T. F. Smith. "A neural network approach to estimating rainfall from spaceborne microwave data". *IEEE Transactions on Geoscience and Remote Sensing*, 35(5):1079–1093, September 1997.

[26] B. van der Pol. "On 'relaxation oscillations'". *Philosophical Magazine*, 2(11):978–992, 1926.

[27] C. von der Malsburg. The Correlation Theory of Brain Function. Internal Report 81-2, Max-Planck-Institute for Biophysical Chemistry, 1981.

[28] D. L. Wang and D. Terman. "Locally excitatory globally inhibitory oscillator networks". *IEEE Transactions on Neural Networks*, 6(1):283–286, 1995.

[29] D. L. Wang and D. Terman. "Image segmentation based on oscillatory correlation ". *Neural Computation*, 9:805–836, 1997.

[30] R. Whitaker and G. Gerig. Vector-valued diffusion. In B. M. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, pages 93–134. Kluwer Academic Publishers, Dordrecht, Netherlands, 1994.