An Extraction method for Water Body of Remote Sensing Image Based on Oscillatory Network

Min Li
1. College of Computer and Information, Hohai University, Nanjing, China
Email: lm_0711@hhu.edu.cn

Lizhong Xu1, 2 and Min Tang1, 2
2. Institute of Communication and Information System Engineering, Hohai University, Nanjing, China
Email: lzxu@hhu.edu.cn, hhice@126.com

Abstract—How to avoiding the disadvantage of the pixel-wise image process methods has been a open problem. An object-wise method was proposed to extract the water body of remote sensing image, which is locally excitatory globally inhibitory oscillator networks (LEGION). This oscillation network could bind an object based on the similar property. The oscillation network in this paper used the normalized difference water index (NDWI) to encode the binding of pixels of water body. Because of the good coherence of spectrum of water body on the remote sensing image, the pixels of the water surface have strong couple weight and can overcome the globally inhibition in the network. Additionally the special spectrum property also was considered, it could reduce the disturber of the other ground object which has the similar spectra property. The proposed method combined both pixel-wise and object-wise scale. Thus, the proposed extraction method was robust to noise, and could keep the detail information of water body. The outputs were according to the human perception mechanism and can be used in the higher image analysis and reorganization. The results of experiment validate the effective of the proposed method.

Index Terms—image classification, remote sensing image, neural network

I. INTRODUCTION

Remote sensing is the most effective means to catch a huge amount of valuable data about our earth surface which can keep the integrity of data. The remote sensing data can be used to analyze, detect and monitor earth surface globally. Especially for the water area, remote sensing data supplied for the macroscopically and nondestructively descriptions.

Recently powerful signal processing methods have been developed to explore the hidden information in the remote sensing image and to take full advantage of the spectrum data. There have been plentiful achievements for remote sensing technology on water quality retrievals, water body extraction, coastline changes detection, the area of water body change detection and so on. As well, these were important for the prediction and monitoring of flood disasters.

A. The recent extraction method of remote sensing water body

It had important significance for the hydrology research to estimate the area of water and extract the edge of water body. The mostly used methods for extraction of water body were threshold method and multi-bands method.

The threshold method was based on the ground object spectra characteristic of every single band, chose an appropriate threshold to extraction the water body. Multi-band method was to construct an index to enlarge the distance between water and other ground objects, which was based on the difference of the spectra change trend of all the ground objects through these bands. At present most water extraction method was based on the index method or multi-band method. For the water area extraction there were spectrum property method [1], normalized difference vegetation index (NDVI) [2], normalized difference water index (NDWI) [3-4] and so on.

These index were normalized in the range of [-1, 1], and zero was chosen as the threshold to distinguish water from vegetation and ground. However, because of the complexity of nature environment, different ground objects would have the same spectrum characters. Only one kind of index couldn’t extract water body effectively from any kind of circumstance, so there were many improved indices (MNDWI) [4,5] proposed to get better performance in a certain environment. Therefore one kind of index method could only extract water bodies under some conditions instead of everyone.

Generally speaking, either the threshold method or the multi-band method was pixel-wise method. This kind method had the advantages of rapid computation, simple theory and achieve easily, so that they become the main extraction method of the water body. But the application range of this kind of methods was limited.

Beside above methods, there were many other novel methods, such as decision tree, step iterative method [6-8] and so on.

B. The existed problems
There existed complex effects among different ground objects, which affected the spectrum of water area. Remote sensing image is different from the common scene image, they describe the ground object on a large scale, and the resolution of remote sensing image is lower. Every pixel element contains more than one kind of object and the complexity relationship among them. Beside those, because of the effect of adjacent objects and the complexity of imaging process, the spectrum value of objects is obtained from the effect of surrounding environment and noise. The surface information contains some overlap which is recorded by every pixel and its adjacent pixel on the remote sensing image, and this is the result of many facts (sensor feature and atmosphere effect) [9]. The pixel was named mixed pixel, which contains more than one kind of ground objects. These kinds of pixel were difficult points of the analysis process. It often involved the problem which kinds of ground object the mixed pixel belonged to. The TM image is taken for example. The resolution of TM image is 30 meters, so there isn’t an explicit pixel corresponding to the river, bridge or road which is with the width of lower than 30 meters. Because of that, the pixel-wise method can’t distinguish the attribution of objects only by one pixel.

There are mainly four disadvantages of the pixel-wise methods: Firstly this kind of method is depended on the resolution, the luminance condition and the content of the image; secondly the water bodies extracted by the pixel-wise methods are lack of integrity; thirdly it is easy to be disturbed by noise and other ground object; fourthly the result of these methods can’t supply the available input to the higher process.

Differing from the pixel-wise method, the object-wise method is based on the macroscopic analysis of the image. These kinds of method analyze the image from object level, and constructs feature functions based on the properties of image objects. These functions are independent of resolution and lamination. Thus, with different spatial scale, the description will remain unchanged by feature functions. Additionally they are robust to noise and disturb. The results of these methods consist with the manner of perception and understanding of human brain. So the object oriented methods have become the research key point of remote sensing image extraction methods. But the feature functions are difficult to construct, and this is the common disadvantage of the object level methods.

2000, Baatz M. et al [10] had proposed an local mutual best fitting area increasing strategy which was an object-oriented optimum segmentation method of the remote sensing image. This method had been used in a remote sensing classification software eCongnition. And 2007 Gamanya R. R. et al [11] improved the method of Baatz. M. and proposed an auto-segmentation method. Both of these methods cut the object body from object-level.

Locally excitation and globally inhibition oscillation network (LEGION) is a network which use the couple weight to binding one kind of object tightly by oscillating the corresponding cells synchronized. The cells corresponding to different objects will oscillate asynchronous with the weaker couple weight. At the same time there is a global inhibition to control the largest coherent area to win the competition. This inhibition is increasing as the coherent areas increasing, and reset when there are one object win out. This network is a kind of cooperation and competition neural network. This network provides a novel and effective framework for image segmentation and figure-ground segmentation.

This paper would construct the couple weight based on NDWI instead of the luminance of pixels. For the water body of the remote sensing image has the characters of stronger coherence and larger areas, the couple weight of water body areas should be larger. Compared to other ground objects, water body could win in the competition firstly. At the same time the tight coupled weight made the water body to be robust.

In order to inhibit other ground objects which have the same spectrum characters of water, this paper also brings the special spectrum property into the oscillation network. This property is the special characters of water, but also can’t get the water body independently. Combining these two methods offset the shortages of each other.

From the above analysis, LEGION can extracted the water body on the object level, and the couple weight based on the NDWI is easy to achieve, and the spectrum property can distinguish the building and water, so the proposed method can obtain the whole body of water accurately without the complexity feature function.

II. THE INTRODUCTION OF LEGION

Inspired by biological visual system and brain, Wang and Terman et.al [12] proposed the locally exciting globally inhibition oscillation network (LEGION). This neural model is originated from correlation theory of Von Der Malsburg et al [13-15] proposed at 1994. They assert that the brain can bind the similar features of objects by firing activities the scattered cells from different cortex areas. Won Der Malsburg [13] considered this integrated perceive function of brain as the temporal correlation theory, the base of brain perceive the object is the neural signal temporal structure. By the research of biological cortex oscillation, Eckhorn [16] and Gray [17] point out that oscillation correlation is the special form of temporal correlation at 1989. Oscillators with the same phase will oscillation correlation is the special form of temporal correlation at 1989. Oscillators with the same phase will oscillate synchronized whenever if they are close to each other. And the oscillation with different phase will oscillate asynchronous. So the network can separate different areas and connate same feature areas.

Recently Deliang Wang et al [18,19] had expanded the LEGION function and proposed this network to audio signal analysis, object extraction and scene analysis. They also applied this network to object extraction of moving image with spatial-temporal correlation. The LEGION segments image into a set of coherent patterns and this is the fundamental aspect of perception, and automatic target recognition and image processing.
The model of LEGION is described as follows:

\[
\begin{align*}
x'_{i,j} &= 3x_{i,j} - \frac{3}{2}y_{i,j} + 2 - y_{i,j} + I_{i,j} + S_{i,j} + \rho \\
y'_{i,j} &= \varepsilon (1 + \tanh(\frac{\alpha}{\rho}) - y_{i,j})
\end{align*}
\]

(1)

where \(I_{i,j}\) is the stimulate input of pixel \((i,j)\) and \(y_{i,j}\) are two oscillators on different temporal scale. \(\rho\) is the variance of Gaussian noise. \(S_{i,j}\) was the couple equation:

\[
S_{i,j} = \sum_{k,j \in N(i,j)} W_{ij,kl} H(x_{k,l}) + W_p H(p_{i,j} - \theta_p) - W_z H(z - \theta_z)
\]

(2)

Where \(W_{ij,kl}\) is the couple weight of pixel \((i,j)\) and \((k,l)\). \(H\) is the step function. \(W_p, W_z\) are the weight coefficient correspond to \(P_{i,j}\), \(z\) is the globally inhibition. \(P_{i,j}\) is the later potential of pixel \((i,j)\), \(\theta_p, \theta_z\) are the threshold of \(P_{i,j}\) and \(z\).

\[
P'_{i,j} = (1 - P_{i,j}) H \left( \sum_{k,j \in N_p(i,j)} H(x_{k,l}) - \theta_p \right) - \varepsilon P_{i,j}
\]

(3)

And the inhibition equation is:

\[
z' = H \left( \sum_{k,l} H(x_{k,l}) - 1 \right) - z
\]

(4)

Equation (1), (3) and (4) are differential equations. The physical significance of these equations had been analyzed particularly and comprehensively by Terman and Wang in [12]. In order to reduce the computation of differential equations, Terman and Wang [18] had proposed a simplified LEGION method. The steps of the simplified LEGION were following:

First of all, compute couple weight of every pixel with their neighbors; then sum the weight in the neighborhood \(N(i,j)\), which is centered on the pixel \((i,j)\). If the weighted sum was larger than the threshold \(\theta_p\), the pixel \((i,j)\) will oscillate. So the pixel which has the largest weight summation will become the leader and excite firstly. Other pixels around the leader will compete to be the followers of the leader, which satisfied the second largest weight summation. Then the oscillating areas are enlarged and those pixels are grouped. There are two group rules: one of which chose the pixel with the largest weight sum and the other of which chose the pixel that has the largest weight with leader of the neighborhood \(N_g\).

According to the second rule, the coherent area will expand along the strongest couple weight. As the coherent area enlarged, the globally inhibition will increase. Only the largest area would overcome the inhibition to excite, other areas would be inhibited to oscillation. So the largest and most coherent area can win the competition. If there is one area winning, the inhibition reset to the original state, and the rest coherent areas will restart to compete.

The globally inhibition controlled different object body areas to excite, according to the magnitude of areas.

The couple weight coefficient \(W_{ij,kl}\) was the important factor, it decided what feature will be chosen to cut the image, or group the object body in the image. There were many kinds of feature used such as color, contrast and luminance of the stay images; motion orientation, speed of the moving images. All of these features could be brought into the LEGION and guide the network to extract different objects, on the other word with different \(W_{ij,kl}\), image would be extracted according to different features. Take the luminance as the simplest example, \(W_{ij,kl}\) was defined as:

\[
W_{ij,kl} = \frac{I(i,j) + I(k,l)}{|I(i,j) - I(k,l)|}
\]

(5)

Where \(I(i,j)\) is the pixel of the neighborhood \(N_g\) which was centered on the \((i,j)\). The \(W_{ij,kl}\) measured the similarity of \(I(i,j)\) and \(I(k,l)\). When they had the same luminance the \(W_{ij,kl}\) will be large. So the pixels with the coherent luminance would make the corresponded oscillators \(x(k,l)\) and \(x(i,j)\) excite synchronized, and the smaller \(W_{ij,kl}\) would make \(x(k,l)\) and \(x(i,j)\) lose the connection and oscillate asynchroneously.

### III. THE LEGION BASE ON NDWI

#### A. The model of NDWI

As the analysis about the water spectrum is not only referring to the water body-self, but also to the environment where water body located on. The NDWI is the index which is based on the spectrum of every ground object.

TM image supplied seven bands spectrum of ground object. Every band corresponds to a certain rang of light wave, water has different reflection characteristics on seven bands. So the spectrum value of water and other ground object will change on seven bands. Among these bands, water have the highest reflect rate on the 5-th band, so the spectrum value of water on the band is the largest, and on the near infrared band, water have the highest absorbency on this band and the spectrum value of water near to zeros.

NDWI is a utility method to extraction water body of remote sensing image. But different environment around water will bring different effects on the water body. For example when the water is located in the urban, the NDWI must consider more on the spectrum difference of building and water. The spectrum of water and building will effect on each other on the edge of the water body; when the image is about rural and open country, the extraction of water will mainly concentrate on the effect of vegetation or soil. The NDWI was defined as follow:
Green represent the green band, NIR represent the near infrared band. Because of the water absorb rate of green light is high, but lower on the near infrared band. Opposite to water, building have higher absorb rate on near infrared band, and lower on the green band. So the NDW will be larger on the water position, and lower on the other ground objects. Based on the NDWI, the distance between water and other ground objects was enlarged. Zero is chosen as the threshold to get the water body which is larger than zero.

But the spectrum of water often was affected by the pollution or the magnitude of sandiness and biased from the ideal value, then the coherence on the water surface will be reduced and the NDWI fluctuated above zero. Beside these, the other ground objects have the same spectrum value also disturb the extraction of water body exactly.

Fig.1 shows the spectrum change trend of water area, vegetation area and building area around the river. The NDWI considered the different spectrum change trend between vegetation area and water area. So it was based on the fourth band TM4 and the second band TM2. Like the analysis in paper [5], we can see that the spectrum of building on the forth band (TM4) and the second band (TM2) had the similarity change trend with the water. They were both decreasing on these two bands. And the paper[5] proposed a modified NDWI to construct MNDWI. So only the choice of a threshold of NDWI couldn’t separate the building objects from the river body, and the extracted river body had coarse edge. Although the NDWI of water and building were both larger than zero, they must have different properties of NDWI.

NDWI could show the spectrum change trend of building area and water area, but it didn’t show the difference of the mean spectrum of building area and water area. This property should be considered, and it could be used to separate these two bodies. So 15 points were sampled on the water body and building body separately, and these points of NDWI were plotted in Fig.2. Let us see the Fig.2. Although both of them fluctuated around their mean value because of the pollution and the different sample areas, the mean NDWI of water and building were different apparently. The NDWI of water was around 0.5, and the building was around 0.25. The variance within the water body and the building body were smaller, and the variance between water and building was larger.

Because of the different mean of spectrum value, it implied that the spectrum value of every ground object was aggregated to a value. The NDWI could only describe the different trend of change, but not the different convergence value. Bring the variance between categories will improve the performance of NDWI.

B. The couple weight based on NDWI

The LEGION based on the biology cortex of brain can make the cells oscillate synchronized on which have the same feature on different areas, and asynchronous oscillate on which have different features. So the beginning of oscillation mainly depends on the relationship of pixels. The object on the image which had the largest area and had the higher coherence will win the competition. Relative to the WTA network [20] (winner take all), the LEGION can constructed the weight represent the relationship of scene pixels and use the relationship to bind the object area tightly. This kind of understanding of object was according to the analysis and understanding process of human brain.

From the above analysis, LEGION extracted object on the object level, so it was robust to the noise and disturbing. Then we could conclude that: LEGION can reduce the small variance which was caused by noise, but remain the large variance which was caused by the feature variance of different objects. Construct one couple weight based on NDWI to describe the relationship of pixels between the different objects or within one object. The couple coefficient reflected the aggregation on the feature level of objects.

$$W_{NDWI}(i,j;k,l) = \frac{N(i,j)+N(k,l)}{N(i,j)+N(k,l)+\varepsilon}$$

Where $N(i,j)$ was the NDWI value of pixel $I(i,j)$.

![Figure.1 The spectrum change trend of water area, vegetation area and building area around the river](image)
Pixel $I(k, l)$ was a random pixel in neighborhood $N_h$ which was centered on pixel $I(i,j)$. The value of $W_{NDWI}(i,j;k,l)$ reflected the similarity of NDWI between pixel $I(i,j)$ and pixel $I(k,l)$. The larger $W_{NDWI}(i,j;k,l)$ represented these two pixels were from the same object.

For the water spectrum has the greatest coherence, the largest couple weight will appear on the water surface. The equation (2) could be rewritten as follow:

$$S_{i,j} = \sum_{k,l\in N(i,j)} W_{NDWI}(i,j;k,l) H(N(i,j)) H((I_3(k,l)+I_2(k,l))\cdot\theta_p)+W_p H(p_i-j-0.5)-W_z H(z-0.5)$$

(8)

This equation had three terms. Where the first term was the relationship of NDWI between every two adjacent pixels; the second term gave the position with the greatest coherence, which satisfied $P_{i,j} > \theta_p$; and the third term was the globally inhibition which made the largest coherent area to excite.

C. LEGION based on NDWI

From the analysis of NDWI in section two, the aggregate property of NDWI could be used to separate the water body from other ground objects by LEGION. Although the LEGION network can inhibit the most of disturb and noise, but the larger variance of building made some building body still remained, where the value of NDWI was close to the water. So the special spectrum property was introduced in to the proposed extraction method which was found by Zhou. The special spectrum property of water is:

$$(34) \quad B_3 \text{green} > B_4 \text{SW} \& B_2 \text{red} > B_1 \text{IR}$$

(9)

This special spectrum property could enhance the accuracy of water body extraction method and eliminate the disturbing of building body and vegetation body which had the closer NDWI. So the equation (8) could be rewritten as follow:

$$S_{i,j} = \sum_{k,l\in N(i,j)} W_{NDWI}(i,j;k,l) H(N(i,j)) H((I_3(k,l)+I_2(k,l))\cdot\theta_p)+W_p H(p_i-j-0.5)-W_z H(z-0.5)$$

(10)

Where $\theta_p=0.5$ and $\theta_z=0.5$ ; $I_3(k,l)$ represent the spectrum luminance of pixel $I(k,l)$ of the first band. $I_4(k,l)$ represented the same pixel spectrum value of the forth band, so does $I_2(k,l)$ and $I_3(k,l)$ . $H(x)$ was the step function, and $H(N(i,j))$ compared the NDWI value of pixel $I(i,j)$ with zero, if it was larger than zero, then $H(N(i,j))$ was set to one, otherwise $H(N(i,j))$ was set to zero. Adding $H(N(i,j))$ onto the second term was to make the object to be remained which was with the NDWI larger than zero. And other equations of network still remained unchanged.

IV. EXPERIMENTS AND ANALYSIS

A. Experiment Data

In this paper, we chose the Landsat TM image date to complete the experiments. This remote image contains the area of Tai lake and the part of Chang Jiang River which was cut on the north of the lake.

1) River part

This cut part contains Yangzhou city and other small towns. The size of image was 800°810 pixels.

Fig. 3 showed the water body extraction results based on NDWI method and the spectrum property. We can see that the results were both according to the conclusion of paper [5]. The building bodies were remained because the NDWI of building was closer to the water. So the NDWI couldn’t separate water body from building...
bodies, and the river edge was coarse because there were buildings around the river.

Fig. 4 were the two enlarged parts of original image, they corresponds to the red block and yellow block separately. In order to look clearly, this paper used the enlarged part of fifth band (TM5) to show the detail. Fig. 4(a) was the enlarged yellow block. There had a bridge cross the river in Fig. 4(a). And the Fig. 4(b) was the enlarged red block, where contains the transition part between river and land. Because there didn’t have the clear edge, so the results of the transition part were ambiguity.

Fig.5 shows the results of the LEGION method. The parameter was $W_z = 30; \theta_p = 1100$, and the iterative time was 80. Compared the output of Fig. 4(b) with Fig. 5, we could point out that where was utmost correspondent to water and where was utmost correspondent to land. At the same time the bridge was extracted clearly in these two results.

Fig.5 could extract the water body with vivid edge on the transition part and the integrity bridge body. Although the result of LEGION can mainly extract the river body area in Fig. 5(a), but still remain some part of other objects. It was seen that the result of Fig. 5(a) contained a little of land which was with the higher coherency on the top right corner. Be different from Fig. 5(a), Fig. 5(b) considered the spectrum properties of water and got the water body accurately without other ground objects. At the same time the edge and coherency of Fig. 5(b) was better than Fig. 5(a).

2) Tai lake part

This section used another part of Tai lake remote sensing image to validate the effectiveness of our proposed method. The experimental data was the cut of Tai lake area, because of the huge image date, we downsampled this image to one fourth of its original size. The parameters of Fig. 6(c) were $W_z = 25; \theta_p = 1100$ and the iterative time was 150 times, and Fig. 6(d) were $W_z = 12; \theta_p = 950$ and the iterative time was 200. Because of the additional constrain terms, the value of parameter $W_z$ and $\theta_p$ was smaller than Fig. 6(c).

The blue block in Fig. 6(b) marked the area of pond and enclosure aquacultures and some shallow water areas. The coherence of the blue block areas was lower, and there were regular texture on the aquaculture area. Although these areas also had the spectrum character of water, the results of proposed methods would effect by the coherence of these areas. The NDWI method of Fig. 6(b) could extract the blue block area with coarse edge by the spectrum value.

Fig.6(c) and Fig.6(d) made emphasis on the detail and texture of the lake surface. For example, the southeast part of Tai lake extracted by these proposed method was sparse, it represents that these areas was shallow and there were some artificial objects on the surface of the lake. Therefore, comparing Fig. 6(c)-(d) and Fig. 6(b), the former two methods could inhibit the building body mostly and keep the detail of water information. The red block was used to mark the bridge body in Fig. 6(c). Fig. 6(d) extracted the whole lake body and the accuracy was higher than Fig. 6(c), it needed more iterative times. Beside of these, because of the blank part on the left quarter, the Fig.6(c) extracted this part based on its coherence. But in the Fig. 6(d) this part wasn’t be extracted because it doesn’t satisfy the spectrum property. From this experiment we can see that the LGEION with spectrum property can extracted the water body more accurately. And the proposed oscillation couple network dependent on the coherent of object feature.

B. The convergence of the proposed Method

The areas of water body were enlarged as the LEGION iterated. This section talked about the convergence of LEGION method based on the river body. Fig. 7 showed the change of increase magnitude between two adjacent iteration times, red line represented the area
Figure 4  The enlarged parts of original image

(a) Enlarged yellow bridge  (b) Enlarged red block

Figure 5  The results of LEGION with NDWI

(a) The output based on LEGION (NDWI) method  (b) The result of LEGION with the spectrum property

(a) Original remote sensing data of Tai lake  (b) Result of NDWI
Figure 6  The results about Tai lake body. (a) was the original image about Tai lake area of the fifth band of remote sensing image; (b) was the extraction result based on NDWI method; (c) was the result based on the LEGION method and (d) was the result of LEGION and spectrum property method.

Figure 7  The change of areas of every two iterative times

increase of LEGION method and the blue line represented the area increase of LEGION method with spectrum property. From this figure, 40 times were an optimum iterative time and the areas of water body stay on a stable value. Fig.7 showed that the blue line converged quickly more than red line when the change below a certain value.

V. CONCLUSION

This paper presents a water body extraction method based on LEGION (NDWI). The experiment results indicated that the proposed method could extract the coherent water body with smooth and clear edge accurately. Without the analysis of water and its surround environment, the oscillation network could extract the water body successfully. Because this oscillation network was dependent on the feature function, it was important to construct an appropriate feature function which can bind the object tightly. Additionally the proposed extraction method needed expensive compute and a lot of iterative times. Because of the large magnitude of remote sensing data, the iterative computation would increase the burden of computation. So our future work should focus on developing a new fast enlarging mechanism which could extract the water body in a few iterative times.

ACKNOWLEDGMENT

This paper was supported by National Natural Science Foundation of China (No. 60774092 and 60901003).
REFERENCES


Min Li is a PhD student of Hohai University, and received the B.S degree in Communication Engineering from Hohai University, Nanjing, P. R. China in 2005 and 2007 respectively. Her main research interest is remote sensing images processing, bionic vision computation and analysis.

Min Tang received PhD degrees from Nanjing University of Science and Technology in 2006. Currently, she is a lecturer at the Hohai University. Her research interests include computer vision and computer graphics.