Road Detection from High Resolution Satellite Imagery Using Texture Parameters in Neural Network

M. Mokhtarzade
PhD Student, K.N Toosi University of Technology, Geodesy and Geomatics Faculty
(m_mokhtarzade@yahoo.com)

H. Ebadi
Assistant Professor, K.N Toosi University of Technology, Geodesy and Geomatics Faculty
(Ebadi@kntu.ac.ir)

M.J. Valadan Zoej
Associate Professor, K.N Toosi University of Technology, Geodesy and Geomatics Faculty
(ValadanZoej@kntu.ac.ir)

Abstract: In this paper, neural networks are applied on high resolution IKONOS images for road detection. It was tried to optimize neural network's functionality using a variety of texture parameters with different window sizes and gray level numbers. Both the source image and pre-classified image were used for texture parameter extraction. The obtained results were compared in terms of road and background detection accuracy. It was concluded that using texture parameters from the source image could improve road detection ability of the neural networks, while using the results of texture analysis of the pre-classified image develops the background detection accuracy.

Keywords: Road detection, texture parameters, co-occurrence matrix, neural networks, back propagation.

1 Introduction

High resolution commercial satellite lunches have made available imagery at resolutions close to that of aerial photographs, which can then be used in wide variety of applications such as preparing and updating maps. Extraction of the road networks was mostly carried out manually by the operators so far. However, considerable skill was necessary for such operations, efficiency was never very high, and this method is costly and time consuming. Road detection can be considered as the first step in road extraction process and is defined as the process of assigning a value to each pixel that can be used as a criterion to extinguish road and non-road pixels. Vigorous methods have been proposed for automatic and semi-automatic extraction of road networks from satellite images. Recently, these methods are more focused on high resolution satellite images due to their outstanding characteristics in mapping from space.

A comprehensive review on the proposed methods for road extraction could be found in [1] where these methods are categorized from different aspects and a broad reference list is presented. The idea of geometrical and topological analysis of high resolution binary images for automatic vectorization of segmented road networks was presented in [2]. In [3] a new fuzzy segmentation method is proposed for road detection in high resolution satellite images that needed only a few number of road samples. Recently, the idea of using contextual information for improving segmentation process of road regions have been tested by many researchers. As a good example of exploiting texture information in road extraction, the research presented in [4] could be mentioned. Also, in [5], the effectiveness of angular texture signature was evaluated to discriminate between parking lots and roads using high resolution satellite images. In the present research, road detection is performed on high-resolution pan-sharpened RGB Ikonos satellite images, using texture parameters in artificial neural network algorithms. At first, road detection has been performed using only spectral information. Then
different texture parameters from source image and pre-classified road raster map were produced. Extracted texture parameters were integrated with the spectral information to evaluate the impact of contributing texture parameters in road detection.

2 Road Detection Using Neural Network's Approaches

Neural Networks, inspired from biological neural systems. There are two main stages in the operation of an ANN classifier: learning and recalling. Learning (training) is the process of adapting or modifying connection weights between neurons so that the network can fulfill a specific task while recalling refers to the generalization ability of the trained network.

Back-propagation is the most common learning algorithm which is an iterative gradient algorithm designed to minimize the error function. The error function is shown in Equation (1).

\[
E = \frac{1}{2} \sum_{j=1}^{L} (d_j - o_j^M)^2
\]  

(1)

In above equation \(d_j\) and \(o_j\) represent the desired output and current response of the neuron “j” in the output layer respectively and “L” is the number of neurons in the output layer.

Road detection from satellite images can be considered as a classification process in which pixels are divided into road and background classes. For using neural networks in road detection, input layer consists of neurons which are the same number as input parameters and output layer is made up of just one neuron that shows the response of the neural network whether the input parameters can represent a road pixel or not.

An important factor in road detection using ANNs is to decide what type of information should be extracted from input image to be fed through the network as its input parameters. The discrimination ability of the network is highly affected by chosen input parameters. Apart from the spectral information, textural parameters and contextual information are usually used by human being in object recognition from images. This is an inspiring point that encourages the use of texture parameters in artificial neural networks for road detection purposes. In other words, contributing texture information in the neural network input parameters seems to be an improving idea for road detection from satellite images that is going to be inspected in this research.

Road detection from IKONOS and Quick-Bird images using ANNS was tested in [6] where a variety of input parameter sets of spectral information and different network structures were evaluated.

One of the advantages of neural networks for road detections, as compared to conventional statistical methods, is that they are distribution-free operators and so more capable of dealing with complicated high resolution satellite images. In addition, neural networks are highly capable of dealing with multi-source data, because they do not require explicit modeling of the data from different sources. Therefore, there is no need to treat them independently as in the case of many statistical methods. In other words, they do not suffer the problem in statistical multi-source analysis of specifying how much influence each data source should have in classification process. This characteristic makes ANNs a good choice for integrating different data sources for road detection purposes.

3 Co-occurrence Matrix Texture Analysis

Texture analysis is a common method of generating region descriptors in digital images. The texture of an image region is determined by the way the gray levels are distributed over the pixels in that region. In performing texture analysis, gray value relationships between the current pixel and its neighbor pixels are evaluated and one or some certain texture measures are calculated.

Gray level co-occurrence matrix method is the most common statistical method of texture analysis that takes into account not only the distribution of gray levels, but also the position of pixels with respect to each other. This task is performed by considering the pixels in pair in a predefined neighborhood of the current pixel. The relative positions of pixels could be introduced by the relative distance measured in pixel numbers (d=1 for neighboring pixels) and an orientation parameter \(\phi\) which is commonly quantized in four directions: horizontal, diagonal, vertical and anti-diagonal \((0^\circ, 45^\circ, 90^\circ, 135^\circ)\).

Co-occurrence matrix has the following form in Equation.
\[ A = [a_{i,j}] ; \quad a_{i,j} = P(d, \phi, I_i, I_j) ; \quad i, j = 0,1,2, ..., N_{g-1} \quad (2) \]

Where \( P(d, \phi, I_i, I_j) \) shows the number of times that pixel pairs with intensity values of \((I_i, I_j)\) are located at the relative position defined by \(d\) and \(\phi\) divided by the whole number of possible pairs. The term \(N_g\) is the number of gray levels in the interest neighborhood \(\{I(m,n) \in [0,1,2, ..., N_g]\}\).

Apparently, the size of co-occurrence matrix and therefore the required computation effort is directly related to the number of gray levels in the input image and gray level reduction is a common practice in texture analysis.

Having defined the co-occurrence matrix in a predefined neighborhood and based on a distinct relationship defined by \((d)\) and \((\phi)\), texture parameters could be computed for the interest pixel. The most common texture parameters, computed at each pixel, are shown as Equation 3, 4, 5 and 6:

\[
\text{Energy} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} a_{i,j}^2 \quad (3)
\]

\[
\text{Contrast} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - j)^2 a_{i,j} \quad (4)
\]

\[
\text{Entropy} = -\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} a_{i,j} \log(a_{i,j}) \quad (5)
\]

\[
\text{Homogeneity} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{a_{i,j}}{1 + |i-j|} \quad (6)
\]

Each parameter is usually represented as an output image with the same size as input image where the gray level of each pixel is the local texture criterion at the corresponding pixel in the source image.

Different texture parameters show different aspects of textural behavior in a defined neighborhood of a given pixel. Apparently, each texture parameter is affected by the chosen window size, the distance and orientation parameters and the level of gray reduction introduced in texture analysis. Furthermore, using different views of the input image (e.g. the source or pre-classified image) can lead to diverse results in texture parameter evaluation. Therefore, it is an important issue to make the right decision about the mentioned factors in texture analysis, which is going to be dealt with in this research.

### 4 Implementation and Practical Results

In this research, a 500*550 pan-sharpened IKONOS image from Kish Island in Iran was chosen as input data. Figure 1 shows the selected image and its manually produced reference map which is used for accuracy assessment.

![Figure 1: a) Pan-Sharpened IKONOS image from Kish Island in Iran. b) Manually produced reference map.](image)

1000 road and 1000 background pixels were selected from the input image to be used as training data in the learning stage of neural networks. Rather the same numbers of road and background pixels were extracted manually in order to be used as test data in evaluating the performance of the neural networks. It was tried to choose training and test data so that all the objects available in the source image have some representative pixels.

In section 4.1, the result of implementing simple neural network, that just use spectral information as its input parameters, is presented and accuracy assessment parameters are introduced. The following sections describe different cases of participating texture parameters in neural network classifiers and the obtained results are presented.

#### 4.1 Applying neural networks for road detection using only spectral information

A back propagation neural network with just one hidden layer and adaptive learning rate and momentum was self programmed in Delphi programming language to be used for road detection. At first, only spectral information was fed through the network and thus three neurons were designed in the input layer in charge of receiving R, G, and B bands. 10 neurons were put in the hidden layer and output layer consists of only one neuron that shows the network's response by a number between 0 and 1 as background and road pixel respectively. The network was trained 500 times using the training
set and considering the RMSE of test data, its convergence was confirmed.

The influence of network structure and different spectral information as network input parameters have been tested for road detection from high resolution satellite images in [6].

When the trained network is performed on entire pixels, a 2D matrix, which has the same size as input image is obtained that is, called output matrix. Multiplying the output matrix by 255 produces a gray level image that shows the network performance for road detection. Figure 3.a shows the obtained gray level image.

For accuracy assessment, a binary image was extracted from image-truth, assigning 1 to road and 0 to background pixels. Adding the multiplication of correspondent values in whole binary image and output matrix produce a value that when is divided to the road pixel number can be considered as road detection correctness coefficient (RCC). When the binary image is inversed, in a similar method, background detection correctness coefficient (BCC) can be obtained too. The third parameter is root mean square error (RMSE), assuming the difference between network’s responses and expected results (1 for road and 0 for background pixels) as error values. After putting a threshold on output matrix, overall accuracy can be calculated the same way as conventional classification methods. This parameter is actually the percentage of correctly classified pixels to all available pixels in entire image.

Accuracy assessment parameters of the obtained result of this section are 75.62 and 89.60 percent for RCC and BCC, 0.2287 for RMSE and 94.43 percent for overall accuracy.

4.2 Applying texture parameters of source image in neural networks

The source image was transformed into a gray level image with four gray values \{0,1,2,3\}. Four texture parameters introduced in section 3 including contrast, energy, entropy and homogeneity were computed for each pixel in a 3*3 window using co-occurrence matrix. Each texture parameter was computed in 4 directions (horizontal, diagonal, vertical and anti-diagonal) and the averaged gray level image was produced representing the relevant texture criterion at each pixel. In this way, each texture parameter could be regarded as a new information channel.

Therefore, each texture parameter can be put next to the spectral information in order to be sent to the neural network via input neurons. Figure 2 illustrates the way of combining spectral information and texture parameters in neural networks. Different combinations of spectral and texture parameters can be designed to be used as network input parameters.

Table 1 show that, participating texture information improves the network's road detection ability, while the decrease of BCC

<table>
<thead>
<tr>
<th></th>
<th>RCC</th>
<th>BCC</th>
<th>RMSE</th>
<th>O. A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Texture</td>
<td>75.62</td>
<td>89.60</td>
<td>0.2287</td>
<td>94.43</td>
</tr>
<tr>
<td>Contrast</td>
<td>77.99</td>
<td>86.23</td>
<td>0.293</td>
<td>91.65</td>
</tr>
<tr>
<td>Energy</td>
<td>80.82</td>
<td>82.97</td>
<td>0.325</td>
<td>82.97</td>
</tr>
<tr>
<td>Entropy</td>
<td>75.59</td>
<td>82.30</td>
<td>0.322</td>
<td>90.11</td>
</tr>
<tr>
<td>Homog.</td>
<td>77.60</td>
<td>82.33</td>
<td>0.313</td>
<td>90.47</td>
</tr>
<tr>
<td>All</td>
<td>78.78</td>
<td>82.97</td>
<td>0.312</td>
<td>91.45</td>
</tr>
</tbody>
</table>

Table 1. Accuracy assessment parameters of using source image texture parameters, texture window size 3*3, Ng=four.

Figure 2. Combining spectral and texture information as neural network's input parameters

At first, each texture image was added to the input parameter set individually. In this way, four neurons were designed into the neural network's input layer where the forth neuron was in charge of receiving texture parameter. All four texture images were used together making a seven neuron input layer. The network with 10 hidden neurons was trained by the prepared training set for 5000 times. Then, the trained network was implemented on the whole image for road detection and accuracy assessment parameters were evaluated. This procedure was repeated several times to make sure of network's stability.

Table 1 shows the obtained accuracy assessment parameters in each case. The first row in this table shows the same parameters for the situation of using no texture parameters in road detection (section 4.1).
implies the contradictory result for background detection. The reason is the uniform pattern of roads in high-resolution satellite images, while background consisted of many different objects and thus it does not have a distinct texture behavior. Furthermore, it can be easily concluded that using all texture parameters produces better results than using them individually. This would be the common practice in the following sections to use all the texture parameters as a set. In order to evaluate the impact of texture window size and determine the optimum neighboring size in texture analysis for road detection purposes, all mentioned texture parameters were calculated in (5*5) and (7*7) window sizes and then the averaged texture images were produced. Table 2 presents the obtained results of contributing all texture parameters together, which are extracted from different window sizes.

Table 2. Checking the impact of texture window size

<table>
<thead>
<tr>
<th>W. S.</th>
<th>NG</th>
<th>RCC</th>
<th>BCC</th>
<th>RMSE</th>
<th>O. A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Texture</td>
<td>75.62</td>
<td>89.60</td>
<td>0.2287</td>
<td>94.43</td>
<td></td>
</tr>
<tr>
<td>3*3</td>
<td>4</td>
<td>78.78</td>
<td>82.97</td>
<td>0.3132</td>
<td>91.45</td>
</tr>
<tr>
<td>5*5</td>
<td>4</td>
<td>81.34</td>
<td>81.52</td>
<td>0.3409</td>
<td>89.19</td>
</tr>
<tr>
<td>7*7</td>
<td>4</td>
<td>76.44</td>
<td>80.55</td>
<td>0.3567</td>
<td>87.82</td>
</tr>
</tbody>
</table>

Table 2 shows that increasing the size of window size deteriorates BCC. The reason is that in larger window sizes, background texture irregularity becomes more. The comparison between RCC values for different window sizes shows that, as long as the whole window is located in the road part, window size increase causes the results to be improved.

When the window size becomes larger than the road width, edge effects cause the texture parameters to be weakened. In this case, (5*5) window size was chosen as the optimum window size. Figure 3 shows the obtained road raster map in this case and the result of no texture participation together. The comparison between two images in Figure 3 shows that using texture parameters of source image, only improves road detection ability of the network and in general deteriorates the results, especially for background detection. In the following sections, it is tried to use texture descriptors of classified image in order to improve the results.

### 4.3 Optimization of road detection using texture parameters of classified image

In this section, it is intended to implement the idea of Zhang [7] for optimization of road detection by combining neural network classification and texture filtering. Figure 4 shows the methodology of this method.

![Figure 4](image-url)
In this method, texture parameters are extracted from binarized result in section 4.1, which is the classified image, by neural networks using only spectral information (Figure 3a). Neighboring window size was chosen (3*3) and texture parameters were evaluated in all four directions (Horizontal, diagonal, vertical and anti-diagonal). Then, the averaged texture images were put in a clustering process considering two initial clusters. Based on training set, road clusters were determined and a segmented image was produced. The optimized road raster map was produced by introducing the primary classified image and segmented image in a logical AND operation.

Table 3 presents the computed accuracy assessment parameters in each case.

<table>
<thead>
<tr>
<th></th>
<th>RCC</th>
<th>BCC</th>
<th>RMSE</th>
<th>O. A</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Texture</td>
<td>75.62</td>
<td>89.60</td>
<td>0.2287</td>
<td>94.43</td>
</tr>
<tr>
<td>Optimized Road Raster Map</td>
<td>62.96</td>
<td>94.82</td>
<td>0.2380</td>
<td>94.26</td>
</tr>
</tbody>
</table>

The obtained results indicate that this method improves BCC to high extent but RCC has been decreased. This is mainly due to the edge effects in texture analysis of classified image that have different texture behavior. It causes road network in the optimized road raster map to be shrieked (because of the AND logical process) and thus RCC decreases.

5 Conclusions and suggestions
In this research, it was concluded that using texture parameters from the gray level source version of the input image makes the network's response stronger in road pixels but in general deteriorates the results especially in background detection. The optimum texture window size was determined to be equal to the road width. Texture parameters were extracted from binarized pre-classified road raster map and a road segmented image was produced by clustering the obtained texture images. The optimization of road raster map by road segmented image brought high improvements in background detection but roads were shrieked.

Considering the contradictory results of these methodologies, a good idea could be using texture parameters from classified image in neural network classification which is in the future plan of the authors.

References


