FEATURE RECOGNITION FROM POTENTIAL FIELDS
USING NEURAL NETWORKS

Yi Guo
A thesis submitted to the Faculty and Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science, Geophysics.

Golden, Colorado
Date 02/17/1992

Signed: Yi Guo

Approved: Dr. Richard O. Hansen
Thesis Advisor

Golden, Colorado
Date 20 February 1992

Philip R. Romig, Head
Department of Geophysics
ABSTRACT

Structural analysis and lithologic mapping are important parts in the interpretation of potential fields, but are labor-intensive and time consuming due to the complication of potential fields and the ambiguity of the potential field inverse problem. The successful manual analysis and interpretation of potential fields depends on professional experience and knowledge of the specific geographic area, which may result in large differences between the final interpretations prepared by different people.

It is difficult to extract features from potential fields using statistical classification methods, due to the complication of potential fields and ambiguity of the inversion problem. Neural network models show great potential in pattern recognition. One of the advantages of neural networks over the statistical classification methods is that neural networks are distribution-free, so no prior knowledge about the statistical distributions of classes is needed. The neural network method also determines automatically how much weight each pattern or class should have in the recognition process.

Based on a number of experiments using synthetic data and real data, a technique was developed, which can be used to
automatically extract features (lineaments and lithologic classifications) from magnetic fields using back-propagation neural networks. The results indicate that if we can choose the proper parameters for the neural networks and construct appropriate model patterns, we can obtain reasonable maps of the structural lineaments and a magnetic field classification that is useful for structural analysis and geophysical mapping.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>x</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 1 MAGNETIC FIELDS</td>
<td>5</td>
</tr>
<tr>
<td>1.1 Magnetism of rocks</td>
<td>5</td>
</tr>
<tr>
<td>1.2 General characteristics of magnetic fields of rocks</td>
<td>7</td>
</tr>
<tr>
<td>1.3 General characteristics of geologic structure from magnetic fields</td>
<td>9</td>
</tr>
<tr>
<td>CHAPTER 2 NEURAL NETWORKS</td>
<td>11</td>
</tr>
<tr>
<td>2.1 General introduction to neural networks</td>
<td>11</td>
</tr>
<tr>
<td>2.2 Back propagation neural networks (BPNN)</td>
<td>13</td>
</tr>
<tr>
<td>2.3 Experiments on the effects of BPNN parameters</td>
<td>18</td>
</tr>
<tr>
<td>2.3.1 Number of nodes in the hidden layer</td>
<td>19</td>
</tr>
<tr>
<td>2.3.2 Slope of the sigmoid activity function</td>
<td>24</td>
</tr>
<tr>
<td>2.3.3 Learning rate coefficient</td>
<td>27</td>
</tr>
<tr>
<td>2.3.4 Other considerations</td>
<td>29</td>
</tr>
<tr>
<td>CHAPTER 3 LINEAMENT RECOGNITION</td>
<td>31</td>
</tr>
<tr>
<td>3.1 Experiments on the possibilities for lineament recognition</td>
<td>31</td>
</tr>
<tr>
<td>3.2 Design and construction of model patterns</td>
<td>34</td>
</tr>
<tr>
<td>3.3 Normalization for lineament recognition</td>
<td>37</td>
</tr>
<tr>
<td>3.4 Principles of lineament recognition</td>
<td>39</td>
</tr>
<tr>
<td>3.5 Test of lineament recognition using real data</td>
<td>43</td>
</tr>
<tr>
<td>CHAPTER 4 LITHOLOGIC CLASSIFICATION</td>
<td>49</td>
</tr>
<tr>
<td>4.1 The possibilities and principles of lithologic classification</td>
<td>49</td>
</tr>
<tr>
<td>4.2 Parameters of BPNN and normalization</td>
<td>52</td>
</tr>
</tbody>
</table>
4.3 Tests on synthetic data ............................... 53
4.4 Tests of lithologic classification
using real data ....................................... 57

CHAPTER 5 FEATURE RECOGNITION FROM THE AEROMAGNETIC
FIELD OF THE PUEBLO QUADRANGLE, COLORADO ...... 62

5.1 Geology ................................................. 62
5.2 Lineament recognition ............................... 66
5.3 Lithologic classification ............................... 69

DISCUSSION & CONCLUSIONS ............................... 71

REFERENCES ............................................... 73
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. A taxonomy of six types of neural networks. Classical algorithms which are most similar to the neural net models are listed along the bottom (from Lippmann, 1987)</td>
<td>12</td>
</tr>
<tr>
<td>2.2. The architecture of a typical back propagation neural network with three layers</td>
<td>15</td>
</tr>
<tr>
<td>2.3. The principle of back propagation networks</td>
<td>17</td>
</tr>
<tr>
<td>2.4. Curves of the sigmoid activity function with different slopes C</td>
<td>26</td>
</tr>
<tr>
<td>3.1. 2-D linear anomaly model</td>
<td>32</td>
</tr>
<tr>
<td>3.2. Test results of lineament recognition using synthetic data</td>
<td>33</td>
</tr>
<tr>
<td>3.3. The model patterns for lineament recognition</td>
<td>36</td>
</tr>
<tr>
<td>3.4. The principle and equation of shaded relief normalization</td>
<td>38</td>
</tr>
<tr>
<td>3.5. Diagram of the principles of lineament recognition using BPNN</td>
<td>41</td>
</tr>
<tr>
<td>3.6. Contour map of the total magnetic field of Xinjiang, China. Contour interval 100 gammas</td>
<td>44</td>
</tr>
<tr>
<td>3.7. Lineament map of the total magnetic field using the window-by-window method, and shaded relief normalization (inclination: 30 degrees, azimuth: NS)</td>
<td>45</td>
</tr>
<tr>
<td>3.8. Lineament map of the total magnetic field using the window-by-window method, and analytic signal normalization</td>
<td>46</td>
</tr>
</tbody>
</table>
3.9. Lineament map of total magnetic field using the point-by-point method, and shaded relief normalization (inclination: 30 degrees, azimuth: NS) ........................................ 47

4.1. Diagram of the principles of lithologic classification ........................................ 51

4.2. Image of synthetic data for tests of lithologic classification ................................ 55

4.3. Image of the result of lithologic classification with synthetic data using BPNN ............. 56

4.4. Contour map of the aeromagnetic field from Xinjiang, China. Contour interval 100 gammas ...... 58

4.5. Classification map of the aeromagnetic field using the window-by-window method ............ 59

4.6. Classification map of the aeromagnetic field using the point-by-point method ............... 60

5.1. Contour map of the aeromagnetic field of the Pueblo Quadrangle, Colorado. Contour interval 200 gammas ........................................ 63

5.2. Geologic map of the Pueblo Quadrangle, Colorado ............................................. 64

5.3. Composite lineament map from the aeromagnetic field of the Pueblo Quadrangle, Colorado using BPNN (with shaded-relief normalization both northeast-southwest and north-south sun angles) .............................................. 67

5.4. Comparison between the geologic map and lineament map ...................................... 68

5.5. Lithologic classification map from the aeromagnetic fields of the Pueblo Quadrangle, Colorado using BPNN ................................. 70
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. Magnetic susceptibilities of various rock types (from Telford et al., 1974)</td>
<td>6</td>
</tr>
<tr>
<td>2.1. Result of a test of variation of the number of nodes in the hidden layer, where the number of input nodes is 49, the number of output nodes is one, and the slope of the sigmoid function is 1.0</td>
<td>22</td>
</tr>
<tr>
<td>2.2. Result of a test of variation of the number of nodes in the hidden layer, where the number of input nodes is 49, the numbers of output nodes is one, and the slope of the sigmoid function is 0.5</td>
<td>23</td>
</tr>
<tr>
<td>2.3. Result of a test of variation of the sigmoid function slope. The numbers of nodes in the three layers are 49, 15 and 1</td>
<td>25</td>
</tr>
<tr>
<td>2.4. The number of iterations with different ( \eta )</td>
<td>28</td>
</tr>
<tr>
<td>2.5. The number of iterations with different ( \eta ) using the improved method</td>
<td>30</td>
</tr>
<tr>
<td>4.1. Range and average values of susceptibility of four kinds of rocks (from Telford et al., 1974)</td>
<td>54</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

First, I would like to express my gratitude to Dr. Norman Harthill and Dr. Richard Hansen who accepted me as a research assistant at the Center for Potential Field Studies (CPFS) just after my arrival at the Colorado School of Mines; without the support of CPFS and all of the sponsors of CPFS, I could not have finished my thesis within one and a half years.

I specially thank my thesis advisor, Dr. Richard Hansen, for his great help in technique and valuable advice; my thesis committee member Dr. Norman Harthill, for his encouragement and assistance; and Dr. Ronald Knoshaug, for his guidance through his courses in knowledge engineering. My thanks also to Ms. Helen Leek's help with formatting and typing this dissertation.

Finally, I want to present the thesis to my parents in China.
INTRODUCTION

Of all the geophysical methods, magnetic and gravity methods are the oldest and have been widely used up to the present for exploration of mineral deposits and oil and gas. The main reason is that, first, potential fields can play an important role revealing geological information about deeper sources, especially in areas with overburden; and second, potential field exploration costs much less than other geophysical methods. Although gravity measurements are more costly than magnetic surveys, gravity is much less expensive than seismic exploration.

The analysis and interpretation of potential field data have until now been done by skilled geophysicists. It is hard work, and is also time consuming, due to the complication of potential fields and ambiguity of the inversion problem. The successful analysis and interpretation of potential fields depends on professional experience and knowledge of the specific area, which may result in large differences between the final interpretations prepared by different people. Geophysicists, especially those working on potential field exploration, have been looking for a technique which can be used to automatically extract useful geological information.

The conventional statistical pattern recognition methods such as Euclidean distance, Mahalanobis distance and the maximum likelihood algorithm (Fukunaga, 1972) have been
successfully applied to computerized information extraction from remotely sensed imagery to solve the problems of geological structure analysis and lithology classification. For example, the remote sensing data associated with forest maps and digital terrain can be used for forest classification (Franklin et al., 1986) and cloud classification (Goodman, 1988). In addition to remote sensing data, multi-channel airborne radiometric data also has also been processed to obtain a lithologic classification using a statistical approach - the minimum distance method (Zhang, 1986).

It is difficult to extract features from potential fields using statistical classification methods. It is well known that potential fields are continuous, which results from the fact that potential fields are less sensitive to the surficial geological structures and lithology than remote sensing data. However, the performance of statistical classification algorithms is very dependent on modeling of the data source; when the weighting and modeling are done properly, we can expect the statistical algorithm to perform well. In other words, the statistical classification algorithms can work well only if we know the distribution functions of the classes. In most cases, however, we do not know the distribution function.

Neural network models show great potential in pattern recognition. One of the advantages of neural networks over the
statistical classification methods is that neural networks are
distribution-free, so no prior knowledge about the statistical
distributions of classes is needed. The neural network method
also determines automatically how much weight each pattern or
class should have in the recognition. The neural network is
described by a set of weights and these weights are computed
in an iterative training procedure. The performance of neural
network models in recognition or classification is therefore
more dependent on having representative training samples.
Another important advantage of the neural network approach is
that it can provide information concerning the significance of
the components of the input feature vector, whereas other non-
parametric classifiers do not yield information of this kind.

In recent years, neural networks have been successfully
applied to many fields, including geologic and geophysical
exploration, such as classification of multi-source remote
sensing data (Benediktsson et al., 1990), fracture detection
from high resolution well bore images (Ariza et al., 1990),
spike filtering in seismic data (Mousset, 1990) and trace
editing and refraction event picking (McCormack, 1990).

There are few papers about the application of neural
networks to potential fields. The most important reason for
this is that feature recognition from potential fields using
neural networks is much more difficult than from remote
sensing data or other geological data. The purpose of this
thesis is to develop a technique which can be used to automatically extract features (lineament recognition and lithologic classification) from potential fields using back propagation neural networks. Only magnetic data were used in this research, but the same methods could be applied to gravity data.
Chapter 1
MAGNETIC FIELDS

1.1 Magnetism of rocks

We know that magnetic anomalies are entirely caused by the magnetic minerals contained in the rocks or, in particular, by their magnetic susceptibilities. Magnetic susceptibilities, which quantify the degree to which the geologic body is magnetized, are the significant variables in magnetic exploration, playing the same role as density in gravity interpretation. In other words, only if the quantity of magnetic minerals or the magnetic susceptibilities of two kinds of rock are different from each other, we can distinguish them using magnetic methods.

Table 1.1 is a list of magnetic susceptibilities of a variety of rocks. Although there is great variation in values of susceptibility even for a particular rock and a wide overlap between different types, we can distinguish the magnetic characters of different kinds of rock. For example, sedimentary rocks usually have the lowest average susceptibility, and basic igneous rocks have the highest. There are larger ranges of variation of susceptibility in metamorphic rocks. In every case, the susceptibility depends on the amount of ferromagnetic minerals present, mainly magnetite, but sometimes ilmenite or pyrrhotite.
**Table 1.1**

*Magnetic Susceptibilities of Various Rock Types*  
(from Telford et al., 1974)

<table>
<thead>
<tr>
<th>Type</th>
<th>Magnetic Susceptibility $\times 10^6$ emu</th>
<th>Type</th>
<th>Magnetic Susceptibility $\times 10^6$ emu</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range</td>
<td>Average</td>
<td>Range</td>
</tr>
<tr>
<td>Sedimentary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dolomite</td>
<td>0-75</td>
<td>10</td>
<td>Granite</td>
</tr>
<tr>
<td>Limestones</td>
<td>2-280</td>
<td>25</td>
<td>Rhyolite</td>
</tr>
<tr>
<td>Sandstones</td>
<td>0-1660</td>
<td>30</td>
<td>Dolerite</td>
</tr>
<tr>
<td>Shales</td>
<td>5-1480</td>
<td>50</td>
<td>Augite-Syenite</td>
</tr>
<tr>
<td>Av. Var. Sed. (48)</td>
<td>0-4000</td>
<td>75</td>
<td>Olivine-Diabase</td>
</tr>
<tr>
<td>Metamorphic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amphibolite</td>
<td></td>
<td>60</td>
<td>Diabase</td>
</tr>
<tr>
<td>Schist</td>
<td>25-240</td>
<td>120</td>
<td>Porphyr</td>
</tr>
<tr>
<td>Phyllite</td>
<td>10-2000</td>
<td></td>
<td>Gabbro</td>
</tr>
<tr>
<td>Gneiss</td>
<td>10-2000</td>
<td></td>
<td>Basalts</td>
</tr>
<tr>
<td>Quartzite</td>
<td>10-2000</td>
<td>350</td>
<td>Diorite</td>
</tr>
<tr>
<td>Serpentinite</td>
<td>250-1400</td>
<td></td>
<td>Pyroxenite</td>
</tr>
<tr>
<td>Slate</td>
<td>0-3000</td>
<td>500</td>
<td>Peridotite</td>
</tr>
<tr>
<td>Av. Var. Met (61)</td>
<td>0-3800</td>
<td>350</td>
<td>Andesite</td>
</tr>
<tr>
<td>Av. acid Ign.</td>
<td></td>
<td></td>
<td>Av. basic Ign.</td>
</tr>
</tbody>
</table>
Generally, the susceptibility of a rock determines only the amplitude of its magnetic anomaly. The character of the magnetic field (features of magnetic anomalies) is mainly dependent on the attributes of the geologic formations, such as the strike, depth and shape of geologic bodies; and on the inclination and declination of inducing magnetization and remanent magnetization. Fortunately, different kinds of rock do have different characters of their magnetic fields. This feature is just the prerequisite which allows us to analyze and interpret the magnetic field.

According to magnetic field theory and experience accumulated by people in magnetic exploration, we have some knowledge as a basis for distinguishing magnetic geologic bodies as follows:

* The outlines of magnetic anomalies are approximately equal to the outlines of the corresponding magnetic geologic bodies.

* The axes of magnetic anomalies correspond to the trend of the geological bodies.

* Each kind of rock has a special magnetic character, since the distribution in the subsurface and the amount and distribution of magnetic minerals is quite different between different rock types.

1.2 General characteristics of the magnetic fields of rocks
1. Intrusive rocks

Generally, ultra-basic rocks have amplitudes of total field anomalies of over a thousand gammas. Since the magnetism of ultra-basic rocks is inhomogeneous, the features of the magnetic field appear to undulate on a rising background. Carbonization decreases magnetism and serpentinization enhances magnetism. Anomalies of hundreds of gammas are usually measured over basic rocks.

2. Volcanic rocks

There are much bigger variations in the magnetic fields of volcanic rocks than for intrusive rocks along a survey profile, and these are often uncorrelated between adjacent profiles. The magnetic character of volcanic rocks becomes weak or disappears with increasing depth of the rocks. The amplitudes of anomalies from basalt usually vary from hundreds of gammas to thousands of gammas; sometimes there are positive or negative linear anomalies due to basalt. The magnetism of andesite is often lower than basalt. Acid volcanic rocks have weak magnetism since they contain small amounts of magnetic minerals.

3. Sedimentary and metamorphic rocks

Most sedimentary rocks are non-magnetic in the sense that their susceptibilities are less than $100 \times 10^{-6}$ emu. The character of the magnetic fields of sedimentary rocks is flat and stable.
Although the average susceptibilities are lower than for intrusive rocks, the magnetic field of metamorphic rocks is more complex than that of intrusive rocks because of weathering and alteration effects. Generally, the magnetism of metamorphic rocks in the Archeozoic is higher than in the Proterozoic.

1.3 General characteristic of geologic structure from magnetic fields

The structural information from magnetic fields is based on the analysis of features of the magnetic fields. The larger faults are often associated with the activity of rock magma. In this case, we can find them easily in magnetic maps. Sometimes, there has been no activity of rock magmatism in faults, such as crushed zones. These usually present a magnetic low compared with background along the strike direction.

The lineament is one of the most important features in the structural analysis of potential fields, because other structures, such as arcs and circular anomalies, can be considered to consist of a series of short lineaments. Usually, linear magnetic anomalies can be divided into three categories: magnetic high linear anomalies, magnetic low linear anomalies and combinations of the two kinds of linear
anomalies. The first and the third are caused by the activity of magma along faults. In addition to the crushed zones, if remanent magnetization exists and the direction of remanent magnetization is opposite to the direction of earth's, the second kind of linear anomalies is produced. When the activities of magnetic magma are multi-periodic, the linear anomalies often are discontinuous. Sometimes, if the regional structural framework is complex or the scale of the fault crushed zone is large, the magnetic anomalies possess an en echelon form.
Chapter 2

NEURAL NETWORKS

2.1 General introduction to neural networks

Artificial neural networks have many names, such as connectionist models, parallel distributed processing models and neuromorphic systems. Whatever the names are, all of these attempt to achieve good performance on pattern recognition via dense interconnection of simple computational units.

Work on neural networks has a long history; their origins can be traced back to the 1950's, when scientists began to program neural network models to simulate the complex interconnections and interactions between neuronal cells in the brain. More recent work by Hopfield (1986), Rumelhart and McClelland (1986), Sejnowski (1986), Feldman (1982), Grossberg, (1982) and others has led to a new resurgence of the field. Generally, neural networks provide a technique for obtaining the required processing capacity using large numbers of simple processing elements operating in parallel. In other words, most neural network algorithms adapt connection weights over time to improve performance based on current results.

Today, there are several dozen different neural network paradigms available. Lippmann (1987) gave a taxonomy for six important neural networks types (Figure 2.1). Networks are divided into two major classes in this taxonomy: those with
Figure 2.1. A taxonomy of six types of neural networks. Classical algorithms which are most similar to the neural net models are listed along the bottom (from Lippmann, 1987).
binary and continuous valued input. Below this, networks are divided between those trained with and without supervision. Networks trained with supervision, such as the Hopfield network and the perceptron, are used as associative memories or as classifiers. Networks trained without supervision, such as Kohonen's feature-map forming networks, are used as vector quantizers or to form clusters. A further difference between networks is whether adaptive training is supported. Although all networks can be trained adaptively, the Hopfield network and Hamming network are generally used with fixed weights.

2.2 Back propagation neural networks (BPNN)

The back propagation neural network is the most popular method for performing supervised learning, and has become the most widely-used tool in the field of artificial neural networks because its performance is reasonable and it is well understood mathematically.

The BPNN consists of many processing elements organized together into "layers" with full connections between successive layers. Figure 2.2 depicts a simple three-layer BPNN showing the nodes as circles, and the interconnections with their associated weights represented by lines. Each line connecting two nodes has a unique scalar weight. There will usually be an input layer where data enters the network, and
the output layer which holds the response of the network to a
given input. Intermediate to these layers is a hidden layer
which generally acts as the synapse or junction between the
input nodes and output layer. More recent work demonstrated
that BPNN with only one hidden layer can form complex disjoint
and convex decision regions (Huang and Lippman, 1988). This
work was followed by a careful mathematical proof, which
implies that an arbitrary decision region can also be
approximated using sigmoidal non-linearities and the BPNN with
at least three layers. This proof, however, is not
constructive and does not indicate how many nodes are required
in the hidden layer (Lippman, 1989).

The BPNN is a nonlinear system which transforms an M-
dimensional input vector into an N-dimensional output vector.
The known input vectors and desired output vectors are used to
train the network. The connection weights are initially
assigned random values, and an actual input vector is fed into
the network to produce an actual output vector. The difference
between the actual and desired output values at each output
node yields an error vector which is used to modify the
connection weights between the output layer and the previous
layer. The output error vector is then recursively "back
propagated" to each of the previous layers. Figure 2.3
illustrates the principle of back propagation networks with
one layer. The examples are repeatedly presented to the
Figure 2.2. The architecture of a typical back propagation neural network with three layers.
network until the output error vector is reduced to an acceptable value. The back propagation training algorithm is described as follows:

Step 1: **Initialize weights and offsets.** Set all weights and node offsets to small random values.

Step 2: **Present inputs and desired outputs.** Present a continuous-valued input vector \(x_1, x_2, \ldots, x_n\) and specify the desired output \(d_1, d_2, \ldots, d_n\). The input could be new on each trial, or samples from a training set could be presented cyclically until weights stabilize.

Step 3: **Calculate actual output.** Use the sigmoid activity function \(f\) and the formula

\[
y_j = f\left(\sum_{i=0}^{M-1} w_{ij} x_j\right)
\]

to calculate outputs \(y_1, y_2, \ldots, y_n\), where \(x_j\) is either the input in the first layer or the output in the hidden layer, and \(w_{ij}\) are the connection weights between two layers.

Step 4: **Adapt weights.** Use a recursive algorithm starting at the output nodes and working back to the hidden layer. Adjust the weights by setting

\[
w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j x_i
\]

In this equation \(w_{ij}\) is the weight from hidden layer node \(i\) or
Figure 2.3. The principle of back propagation networks.
from input node i to node j at time t. The $x_i$ is either the output of node i, or is an input. The $\delta_j$ is a gain term, and $\eta$ is learning rate coefficient for node j. If node j is an output node, then

$$\delta_j = y_j(1-y_j)(d_j-y_j)$$

where $d_j$ is the desired output of node j and $y_j$ is the actual output. In fact, the term $y_j(1-y_j)$ is the derivative of the sigmoid activity function of the actual output. If node j is an internal hidden node, then

$$\delta_j = x_j(1-x_j) \sum_{k=0}^{N-1} \delta_k w_{jk}$$

where k ranges over all nodes in the layers above node j.

Step 5: Repeat from step 2.

2.3 Experiments on the effects of BPNN parameters

Back propagation neural networks form nonlinear discriminant functions using multi-layer perceptrons with sigmoid nonlinearities. They are trained with supervision, using gradient-descent training techniques, called back propagation (as described in the preceding section), which minimize the squared errors between the actual and desired
output of the networks. This is the basic principle of the BPNN.

Although the mathematical theory seems to be understood and there are a number of successful applications of BPNN, it is not easy to solve a geophysical problem in practice using BPNN because the parameters of BPNN and the forms of input and output are completely different for any given problem. There are three parameters which are most important for back propagation neural networks:

1) Number of nodes in the hidden layer.
2) Slope of the sigmoid activity function.
3) Learning rate coefficient.

2.3.1 Number of nodes in the hidden layer

The number of nodes in the input and output layers mainly depend on the size of the model pattern and form of output for representing different model patterns, which will be discussed in detail in the next two chapters.

The number of nodes in the hidden layer has a great effect on the robustness of BPNN. Although we know some criteria for choosing the number of nodes in the hidden layer, there is still a problem of how to determine the optimum value. Generally, if the hidden layer is too large, it will encourage the network to memorize the input pattern rather than generalize the input into features. This reduces the
network's ability to handle unfamiliar inputs after training is completed. On the other hand, a hidden layer that is too small will drastically extend the number of iterations required to train the network and will likely reduce the accuracy of recall.

One of the problems of applying BPNN is that there is no way to know whether the minimum found by a gradient descent is a global minimum or a local minimum. Usually, the solution is to add a momentum term to the delta rule:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j x_i + \alpha (w_{ij}(t+1) - w_{ij}(t))_{\text{prev}}.$$  

The momentum term is simply a constant $\alpha$, multiplied by the change in the weight vector of this node from the previous presentation of this input pattern. So if last weight change was in particular direction, the momentum term tries to make the next weight change in more or less the same direction. Depending on the values of parameters and the situation, the momentum term can even be sufficient to keep the network from failing into a local minimum. Another method is to increase the number of nodes in the hidden layer. Although local minima could still exist in this case, they may be much rarer with many hidden nodes than with few hidden nodes. The high dimensionality of the error surface tends to provide a downhill path along which there is no gully between the initial position and bottom of the hill (McClelland and Rummemhart, 1988). The problem is that we still do not know the
exact optimum value of momentum or number of nodes in the hidden layer. However, one of the most direct methods can be used to determine the latter approximately, especially for only one output node.

The BPNN performs a transform between the input vector and output vector using gradient descent. We say the neural network is "trained", when the trained neural network (trained weights) have learned to recognize a pattern from the input vector (model pattern). From this point of view, the neural network acts as a memory. It is true that the response of a trained network has a maximum when the input vector is the original one which was used to train this network. Generally, the larger the actual output value, the closer the new input vector is to the model pattern. The capability of resolution is one of the important properties for robustness of a neural network. In this case, the capability of resolution can be considered a kind of measurement of different values of actual output between the model pattern and checked patterns. Through analyzing a group of difference values of the response and their actual output values by means of changing the number of nodes in the hidden layer (fixing other parameters of network), we can find an approximate optimum number of hidden nodes. Table 2.1 illustrates a example of how to choose the number of nodes in the hidden layer using synthetic data for a lineament recognition test.
Table 2.1

Result of A Test of The Number of Nodes in The Hidden Layer

<table>
<thead>
<tr>
<th>Number of nodes in hidden layers</th>
<th>Model patterns (NW)</th>
<th>Checked patterns (NE)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.901575</td>
<td>0.862326</td>
<td>0.039249</td>
</tr>
<tr>
<td>10</td>
<td>0.901257</td>
<td>0.881203</td>
<td>0.020054</td>
</tr>
<tr>
<td>13</td>
<td>0.900681</td>
<td>0.887455</td>
<td>0.013226</td>
</tr>
<tr>
<td>15</td>
<td>0.902202</td>
<td>0.891723</td>
<td>0.010483</td>
</tr>
<tr>
<td>20</td>
<td>0.900399</td>
<td>0.894212</td>
<td>0.006187</td>
</tr>
</tbody>
</table>

* The number of input nodes is 49, the number of output nodes is one, and the slope of the sigmoid function is 1.0.

First, we can see that the difference of the actual output between the model pattern and checked pattern decreases with an increasing the number of nodes in the hidden layer. This implies that few nodes in the hidden layer could avoid the network being overtrained, which reduces the recognizing ability of the network; but it is not true that the least number of nodes is the best. Now, let us look the first column - the response to model pattern itself. We can see that there
is a maximum when the number of hidden nodes is 15, which is
an approximately optimum number of nodes for the hidden layer.

In fact, the difference values of the actual output
between the model pattern and checked pattern also relate to
other parameters of the network, especially to the slope of
the sigmoid function. Table 2.2 is the result of a test of the
number of hidden nodes when the slope of sigmoid function is
equal to 0.5.

Table 2.2  
Result of A Test of The Number of Nodes in The Hidden Layer

<table>
<thead>
<tr>
<th>Number of nodes in hidden layers</th>
<th>Actual output values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model patterns (NW)</td>
</tr>
<tr>
<td>6</td>
<td>0.900245</td>
</tr>
<tr>
<td>10</td>
<td>0.901044</td>
</tr>
<tr>
<td>13</td>
<td>0.901254</td>
</tr>
<tr>
<td>15</td>
<td>0.900822</td>
</tr>
<tr>
<td>20</td>
<td>0.901398</td>
</tr>
</tbody>
</table>

* The number of input nodes is 49, the number of output nodes is one, and the slope of the sigmoid function is 0.5.
It is interesting that the maximum value of actual output occurs if the number of hidden nodes is 20, and all the difference values between the model pattern and checked patterns are about twice those for the example above.

2.3.2 Slope of the sigmoid activity function

The activation function of the neural network determines the activity or excitation level, generated in the node as a result of an input signal of a particular size. For a back propagation network, the activity function of each node should be sigmoid; that is, it must be continuous,"s"-shaped, monotonically increasing and asymptotically approach fixed values as the input approaches plus or minus infinity. The adapted activation function is

\[ f(x) = \frac{1}{1 + e^{-Cx}} \]

which has asymptotic limits of 0 and 1 as x approaches negative and positive infinity, where x is an input to the hidden layer or output layer and C is called the sharpness parameter, or slope, of the sigmoid function. A value of C greater than one essentially magnifies the weights, which will make the network converge faster; but too large a value for C increases the chance of hitting a local minimum. Figure 2.4 illustrates the curves of activity function with different
value of C: 2.0, 1.0, 0.5 and 0.3 respectively. We can see
that, from Figure 2.4, the bigger the constant C is, the
sharper the shape of the curve of the sigmoid function.

Since the features of the curves do affect the activity
of individual node, the slope of the sigmoid function is also
important to the robustness of a back propagation neural
network. Table 2.3 is the result of a test of the sigmoid
function slope.

Table 2.3
Result of A Test of The Sigmoid Function Slope

<table>
<thead>
<tr>
<th>Slope of the sigmoid function</th>
<th>Model Pattern (NW)</th>
<th>Checked pattern (NE)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.900822</td>
<td>0.882458</td>
<td>0.018374</td>
</tr>
<tr>
<td>1.0</td>
<td>0.902206</td>
<td>0.891723</td>
<td>0.010483</td>
</tr>
<tr>
<td>2.0</td>
<td>0.903217</td>
<td>0.897423</td>
<td>0.000213</td>
</tr>
</tbody>
</table>

* The number of the nodes in the three layers are 49, 15 and 1 respectively.

This is an example of a test of lineament recognition
using synthetic data. For three different slopes of the
SIGMOID ACTIVITY FUNCTION
Slope: 2.0, 1.0, 0.5, 0.3

Figure 2.4. Curves of the sigmoid activity function with different slopes C.
sigmoid function, the values of actual output of both the model pattern itself and the checked patterns increase with increasing slope; conversely, the difference values of the actual output between the model pattern and checked patterns decrease with increasing slope. Generally, we will have bigger difference values of the actual outputs using a small slope, but smaller slopes (sharpness) will result in a longer training time. In practice, the slope of the sigmoid activity function is determined by trial and error in association with choosing the number of nodes in the hidden layer.

2.3.3 Learning rate coefficient

Although the slope of sigmoid activity function and number of nodes in the hidden layer relate to convergence speed in training neural networks, the learning rate of a BPNN more directly depends on the learning rate coefficient $\eta$ in the delta rule. The learning rate coefficient, which should be in range of 0 to 1, is a measure of the speed of convergence of the weight vector to the minimum error position. Large values of $\eta$ cause faster movement to the global minimum of error but also tend to cause oscillation around the minimum. If the parameter is set to a rather small value to keep the network from oscillating, it can cause the network to take many more iterations to learn than for large values of $\eta$.

For example, If a three layer network is constructed
using 9 input nodes, 10 hidden nodes and one output node (the slope of the sigmoid function is 0.5), the number of iterations for learning increases with a decrease in the learning coefficient $\eta$, as listed on Table 2.4 ($\eta$ is same for both the hidden layer and output layer):

Table 2.4

<table>
<thead>
<tr>
<th>$\eta$</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>110</td>
</tr>
<tr>
<td>0.5</td>
<td>282</td>
</tr>
<tr>
<td>0.2</td>
<td>707</td>
</tr>
</tbody>
</table>

Although there is little difference in training time in the test above, if the sizes of input and hidden layer are larger, the training time will increase rapidly as $\eta$ decreases. However, the learning rate coefficient is not required to be the same constant in the hidden layer and output layer. In this research, variable learning rate coefficients were used in the tests of lithologic classification using synthetic data and real data. This method is based on the criterion of walking downhill. As for all gradient descent algorithms, the
initial value should be large, and can be set to 0.9. The steps gradually decrease with each iteration (walking downhill). A simple way of using a variable $\eta$ is to set the learning rate coefficient equal to the difference value between desired output and actual output, which varies from 0.9 to 0.1 with convergence of the network. For instance, in the same situation as for the test of fixed $\eta$, the number of iterations required was 924. The variable learning rate coefficient does not guarantee the fastest training rate, but it is an effective method to use to avoid local minima, especially if we have no idea how to choose the optimum fixed learning rate coefficient. The delta-bar-delta algorithm, a new method described first by Jacobs (1988), is usually effective in speeding up the convergence of a network. The delta-bar-delta principle is that each weight will have its own learning rate, and each rate will be changed according to how well the network converges (Tveter, 1991).

2.3.4 Other considerations

From 2.2, we know that the connection weights are adjusted by the formula

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j x_i$$

where $\delta_j = y_j(1-y_j)(d_j-y_j)$.

The term $y_j(1-y_j)$ is the derivative of the sigmoid activity function. A serious problem arises here in that when
y_j is close to 1 or 0 the y_j(1-y_j) term is small and very little learning takes place. If some output unit is registering a 1 or 0, a very long time will be taken to undo this state of affairs. When this derivative term is close to 0, it is especially bad in the integer formulation because the limited precision of the arithmetic means that small values may become exactly 0 and no learning will occur at all.

A improvement was suggested by Fahlman (1988) to change the derivative term to 0.1 + y_j(1-y_j). With the addition of the extra 0.1, the term never approaches 0 and training is faster. Tjeter (1991) applied this technique to both the output layer and hidden layer units and obtained much better convergence speed. Using this improved method, the tests of the learning rate coefficient were repeated, and the number of iterations were reduced about 50 percent (Table 2.5).

Table 2.5
The Number of Iterations with Different \( \eta \) Using Improved Method

<table>
<thead>
<tr>
<th>( \eta )</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>66</td>
</tr>
<tr>
<td>0.5</td>
<td>160</td>
</tr>
<tr>
<td>0.2</td>
<td>401</td>
</tr>
<tr>
<td>variable ( \eta )</td>
<td>500</td>
</tr>
</tbody>
</table>
Chapter 3
LINEAMENT RECOGNITION

3.1 Experiments on the possibilities for lineament recognition

Structural analysis is an important part of the interpretation of potential fields. However, the structural forms in potential fields are so complex that we can not find model patterns to present their behaviors precisely for all types of structures. Fortunately, the main structural forms of the potential field consist of lineaments that can be considered to be two-dimensional anomalies; other forms can be considered composed of a number of lineaments, as was discussed in Chapter 2.

In order to explore the possibility of lineament recognition for magnetic fields using BPNN, some experiments on recognizing lineaments using synthetic data were done. Linear anomalies in a set of profiles were constructed using modified version of the CPFS program "mag2d" for a dike-like body. By choosing the dip of the magnetization, susceptibility, the width of the dike and the depth to the top of the dike, and the strike direction, different synthetic data models of a linear anomaly can be obtained (Figure 3.1).

Figure 3.2 shows one of the results of the experiments, in which the dip of magnetization was 90°, and the susceptibility was 0.002; the width of dike and the depth to
Figure 3.1. (a) 2-D linear anomaly model. (b) A model pattern of lineament with NW-trending.
Figure 3.2. Test results of lineament recognition using synthetic data.
its top were 1 and 5 units respectively for all linear anomaly models. The number of input nodes was 180 (corresponding to 6 profiles, with 30 samples on each profile), and the numbers of nodes in the hidden layer and in the output layer were 16 and 4 respectively. The slope of the sigmoid function was set to 1. The abscissa in Figure 3.2 is the strike direction of the linear anomalies. The model pattern with a N45°W (-45°) linear anomaly was used to train the BPNN; the six checked patterns contained linear anomalies with strike directions of N30°W, N15°W, NS, N15°E, N30°E and N45°E. The ordinate is the difference of the response of the network to the model pattern and the checked patterns. We can see from Figure 3.2 that the difference values increases with an increase of the angle between the strike of the model pattern and the checked pattern. This suggests that the BPNN has the capability to identify the strikes of lineaments. In other words, BPNN can recognize lineaments having different directions by picking a maximum value of the actual output from the responses to the all the model patterns.

3.2 Design and construction of model patterns

Although linear anomalies with given parameters can be recognized, trained neural networks are not suitable for real situations in which the features of magnetic anomalies are
very different from each other. We need to find model patterns that are simple and correspond to commonly-occurring situations.

The requirement that the input values of the BPNN range from zero to one prompts us to construct the model patterns in the most simple way, i.e., the model patterns are chosen to consist only of 0 and 1. The sizes of the model patterns were determined based on the general character of linear anomalies from potential fields. The sizes of the model patterns tested included 3x3, 5x5, 7x7, 15x15 and 21x21. The results indicated that if the model patterns are too small, for example 3x3, the lineaments, after recognition, look like fragments. However, if the size of the model patterns is too large, misrecognition occurs for non-linear features with small scale, and much longer computing times are needed. Finally, eight model patterns of size 7x7 were constructed, each one having a single lineament along the NW, NEE, NE, NNE, NS, NNW, NW, and NWW direction respectively (Figure 3.3). A similar approach was taken by Penn (personal communication, 1991).

The neural networks for lineament recognition were trained using the above model patterns. The number of input nodes and output nodes were 49 and 1 respectively. By trial and error, we found that we could get the best results if the sigmoid function constant is set equal to 0.5 and the number of nodes in the hidden layer is 15.
Figure 3.3. The model patterns for lineament recognition.
3.3 Normalization for lineament recognition

Generally, the linear features in the potential fields, which relate to geological structures, are not as obvious as in other geophysical data. Therefore, we need to use a method to enhance the linear features of anomalies in the magnetic field. Of course, the training of a neural network also requires the normalization of the data (translation of data into the range from 0 to 1). This increases the convergence speed of the iterations for training the neural network.

The first method we used was the shaded relief technique, which translates the potential field into the cosine of the angle between the normal vector to the potential field surface and a light source vector (Dods et al., 1988). Figure 3.4 illustrates the diagram and equation of this method. The shortcoming of this kind of normalization is that if the inclination of the light source is too low, for example lower than 30 degrees, the recognized lineaments will be somewhat shifted from their correct positions.

In order to reduce this difficulty, another method of normalization was suggested by Hansen (1991, personal communication). Using the CPFS program "analy" (type=1, order=0), one kind of analytic signal of magnetic fields, the amplitudes of the gradient of the field were obtained by the formula:
where

\[ \rho = \frac{\partial M}{\partial x}, \quad q = \frac{\partial M}{\partial y} \]

\( \lambda \) — angle between the normal vector to the 'surface' of an element and the light source vector;

\( \varphi \) — the declination or azimuth of the light source;

\( \theta \) — the inclination of the light source.
\[
\left[ \left( \frac{\partial T}{\partial x} \right)^2 + \left( \frac{\partial T}{\partial y} \right)^2 + \left( \frac{\partial T}{\partial z} \right)^2 \right]^{\frac{1}{2}}
\]

The data, using this processing technique, also needs to be normalized to 0 - 1 by a simple linear transform.

The results of the latter test using real data indicated that the second method is better for recognizing the main structures and outlining 3D anomalies, but there are some distortions at the edges of the data set. The shaded relief normalization method is more sensitive to the subtle features of magnetic fields and yields a more adequate result. Generally, the inclination of the light source is chosen between 30 and 45 degrees, and the declination is chosen to be a direction which is orthogonal to the strike of main structures. If we process the data twice using the shaded relief normalization with two different azimuths and combine the two output files, i.e., overlay two lineament maps together, we can obtain more reasonable results.

3.4 Principles of lineament recognition

Figure 3.5 is the diagram of the principles of lineament recognition using BPNN. In this figure, trained networks or trained weights of eight model lineament patterns are
represented by squares. When real data are processed for lineament recognition, they are recalled sequentially to match the real data set in a moving window (the moving window is the same as that of the model patterns). In other words, the real data in the window are fed into eight trained networks respectively, then eight actual outputs of the networks are obtained. Suppose there is a lineament with a NNE strike in the window; then the response from the network of the NNE model pattern should be a maximum. If we can select this maximum, we can obtain information about the orientation of the lineament in the window. Generally, the lineament in the real data set is closer to a model pattern if the matching value is higher. The maximum shows the best match between the lineaments in the real data and in a model pattern. Finally, when the window has moved through the whole data set, we obtain a lineament map.

According to this principle, the program "rline" was written for lineament recognition. There are two types of recognition and output which can be chosen: 1) Window-by-window. The output is a file of recognized lineament coordinates, which can be displayed on the screen of an IBM RS/6000 using the Center for Wave Phenomena (CWP) program "xgraphics" or plotted on a laser printer using the CWP program "psgraphics". 2) Point-by-point. The output is a grid file, in which the different lineaments are represented by
Figure 3.5. Diagram of the principles of lineament recognition using BPNN.
different values, that can be displayed or plotted using the CWP programs "ximage" and "psimage" respectively.

Although the point-by-point technique requires a longer computation time, the lineament map is more reasonable than the window-by-window result. This further proves that one important thing for application of neural networks is representation!

The threshold is a important parameter for lineament recognition because sometimes there are no lineaments at all in a window, although we can get a maximum actual output value. If the threshold is chosen too large, real lineaments will not be identified; if it is too small, the quantity of noise will increase. By statistics and analysis for actual output values of networks an experimental formula for the threshold was obtained:

\[
    \text{THRESHOLD} = \text{MEAN} + C \times \text{STANDARD DEVIATION}
\]

Where C is a constant which ranges from 0.5 to 1. According to the mean and the standard deviation of the response of networks given by the program, we can get an appropriate threshold.
3.5 Tests of lineament recognition using real data

Figure 3.6 is a contour map of the total magnetic field data from Xinjiang Province, China. The size of data set is 100 x 100, and the area of the data set is 16x16 km². It is clear that, on the map, a group of NE linear anomalies forms the main structure in this area, and there are also three dimensional anomalies due to basic rocks in the southern part of the map. The contour interval is 50 gammas.

Figure 3.7 is one of the results of lineament recognition using the trained neural networks with shaded relief normalization and by the window-by-window method. The inclination and declination of the sun were 30 and 0 degrees (NS) respectively. If we compare this figure with the original contour map, we find that, in addition to the NE-trending main structures, some subtle features, (e.g. NW-trending linear anomalies and basic rock anomalies), have also been identified. The recognized lineaments with the analytic signal normalization (Figure 3.8) match the NE main structures well, and the basic rock anomalies are represented by a number of short lineaments. However some small linear features of the total field cannot be recognized with this normalization. The representation of lineaments using the point-by-point method is much better than the window-by-window method because some structures, such as arcs, can be displayed precisely by the
Figure 3.6. Contour map of the total magnetic field of Xinjiang, China. Contour interval 100 gammas.
Figure 3.7. Lineament map of the total magnetic field using the window-by-window method, and shaded relief normalization (inclination: 30 degrees, azimuth: NS).
Figure 3.8. Lineament map of the total magnetic field using the window-by-window method, and analytic signal normalization.
Figure 3.9. Lineament map of the total magnetic field using the point-by-point method, and shaded relief normalization (inclination: 30 degrees, azimuth:NS).
image. On the other hand, this kind of lineament map is similar to one plotted by hand. Figure 3.9 illustrates the recognized lineament, with parameters the same as those for Figure 3.7. From the new representation, we can see that the NEE linear anomaly was cut off in the center of map, which implies that there is a hidden NW fault. The subtle features are delineated in more detail also.
Chapter 4

LITHOLOGIC CLASSIFICATION

4.1 The possibility and principles of lithologic classification

The characteristics of the magnetic field are intricate. Different magnetic rocks below the surface may often give rise to anomalies with the same character. But sometimes, the character of the magnetic fields of the same kind of rock have distinctive signatures in different geological units or in different regions. In theory, almost all inverse problems for potential fields are not unique. However, we know from Chapter 1 that there are some distinct magnetic behaviors for different kinds of rocks due to the differences of susceptibility and other attributes, which provide the possibility of lithology classification using neural networks. Of course, it is difficult to classify lithology using only magnetic data because the magnetic field is affected only by the magnetic geological bodies. However, these results give us information about subsurface sources. In fact, lithologic classification from magnetic fields using neural networks is actually a classification of magnetic anomalies.

The basic idea of lithologic classification using neural networks is to train a BPNN with some model patterns which are related to different lithologies, i.e., to train the networks
to learn to recognize the characters of the anomalies we want. These trained networks can then be used to classify the whole data set. For lithologic classification using neural networks, the main difference from lineament recognition is that we cannot construct synthetic model patterns to correspond to all situations. Therefore, the model patterns have to be extracted from real data.

Figure 4.1 is a diagram of the principles of lithologic classification using BPNN. Generally, we should first pick model patterns from real data, then train the BPNN with each model pattern and get the weights of the trained neural networks. Finally, we recall the trained BPNN in a moving window to analyze the entire data set. The size of model pattern is chosen to be 3x3 because we hope to identify a class in a minimum unit area. In other words, we hope to classify the magnetic field in as much detail as possible. It should be noted that the accuracy of the classification is directly dependent on the accuracy of the model patterns. So, we should select the model patterns carefully when picking them from real data. The rules of thumb are that the data size of the model pattern should be as large as possible and should contain only one kind of feature. Usually, the shaded relief image of the magnetic field is helpful for choosing model patterns from real data.
Figure 4.1. Diagram of the principles of lithologic classification.
The program "rlitho" is specialized for lithologic classification. It has two recognition methods: "window-by-window" and "point-by-point", as does the program "rline". Since the number of model patterns can not be fixed in advance for different data sets, we need to enter each file name of trained weights sequentially. The output files are grid files for both recognition methods, which can be displayed or plotted on the screen or on the laser printer, respectively.

4.2 Parameters of BPNN and normalization

The parameters of BPNN for lithologic classification are quite different from those for lineament recognition. Two of the most important parameters are the slope of the sigmoid function and the number of nodes in the hidden layer. For normalized data, if the number of input nodes is 9 (the size of model pattern is 3 by 3), the results of tests show the sigmoid slope should be about 0.05 and the number of nodes in the hidden layer about 10 (notice that this is more than the number of input nodes).

Normalization is also important to lithologic classification. It can be used either to convert the data into a scale from zero to one or to enhance useful information. Tests for lithologic classification using synthetic data without normalization were carried out. Although the trained
networks could achieve convergence, their recognition ability was rather weak and unstable. Comparing several normalization methods, we found that the analytic signal method with the parameters type=1 and order=2 generates the best recognition results (CPFS program "analy"). The equations of normalization are as follows:

\[ A_i = \left[ \left( \frac{\partial^3 T}{\partial^2 z \partial x} \right)^2 + \left( \frac{\partial^3 T}{\partial^2 z \partial y} \right)^2 + \left( \frac{\partial^3 T}{\partial^2 z \partial z} \right)^2 \right]^{\frac{1}{2}} \]

\[ N_i = \frac{A_i - A_{\text{Min}}}{A_{\text{Max}} - A_{\text{Min}}} \]

4.3 Tests on synthetic data

Table 4.1 lists the range and average values of susceptibilities of sedimentary, metamorphic and igneous (including acid and basic igneous) rocks extracted from Table 1.1. Using these values, a synthetic data set consisting of a grid file which contains random numbers with different means (75, 350, 650 and 2600) and different standard deviations (5, 20, 50 and 100) was constructed (Figure 4.2).
Table 4.1

Range and Average Values of Susceptibility of Four Kinds of Rocks (from Telford et al., 1974)

<table>
<thead>
<tr>
<th></th>
<th>Susceptibility x 10^-6 emu</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range</td>
</tr>
<tr>
<td>Sedimentary</td>
<td>0-4000</td>
</tr>
<tr>
<td>Metamorphic</td>
<td>0-5800</td>
</tr>
<tr>
<td>Acid Igneous</td>
<td>3-6530</td>
</tr>
<tr>
<td>Basic Igneous</td>
<td>44-9710</td>
</tr>
</tbody>
</table>

From the results after recognition (Figure 4.3), we can see that the four kinds of classes are identified reasonably well except for the horizontal boundaries between the four classes. The reason for this mis-recognition is that the moving window crosses at least two kinds of classes at boundaries. In this case, the analytic signal normalization was not used and the parameters of the neural networks are actually different from the situation for real data. However, the tests on synthetic data indicate that BPNN really has the capability to recognize patterns and shapes in two dimensions.
Figure 4.2. Image of synthetic data for tests of lithologic classification.
Figure 4.3. Image of the result of lithologic classification with synthetic data using BPNN.
4.4 Tests of lithologic classification using real data

Based on the experiments using synthetic data, a real data set (aeromagnetic data from Xinjiang, China; Figure 4.4) was processed as a test of lithology classification. Four kinds of magnetic fields (i.e., four classes) were picked from the original data set. Pattern 1 and pattern 2, which both belong to gentle and stable fields, differ only slightly. In this way, we can check the potential of BPNN for identifying classes with small differences. The average amplitude of pattern 3 is about 150 gammas. Pattern 4 is one of the typical anomalies of basic rocks, with average amplitude about 300 - 400 gammas. The following step was to train the BPNN with the model patterns; then the weights of the trained networks were extracted from the neural network and used to create a subroutine which represented the trained network.

Figure 4.5 is a classification map after recognition using the window-by-window method. The magnetic data are classified into three parts. Pattern 4 represents a class (3.7 % of total area) which relates to the anomalies of the basic rocks. Pattern 3 corresponds to a magnetic high with lower amplitude (48.58 %). pattern 2 (20.66 %) and pattern 1 (27.65 %) are the responses of the background magnetic field. Figure 4.6 is an image of lithologic classification using the point-by-point method. The results seem to be better than
Figure 4.4. Contour map the aeromagnetic field from Xinjiang, China. Contour interval 100 gammas.
Figure 4.5. Classification map of the aeromagnetic field using the window-by-window method.
Figure 4.6. Classification map of the aeromagnetic field using the point-by-point method.
those of Figure 4.5 for extracting lithologic information according to grey levels.

Although the classification map looks somewhat rough and simplified, it demonstrates an important fact: magnetic anomalies in a grid can be recognized by neural networks. Results of the real data tests indicate that if we use reasonable parameters for the neural networks and an effective normalization method, especially if we can pick precise model patterns, we can classify the magnetic fields into several parts which are related to the lithology. This method is very useful to recognize one or two classes, (e.g., special kinds of magnetic anomalies).
Chapter 5
FEATURE RECOGNITION FROM THE AEROMAGNETIC FIELD
OF THE PUEBLO QUADRANGLE, COLORADO

In order to evaluate the capability of this technique for lineament recognition and lithologic or textural classification, the aeromagnetic data of the Pueblo Quadrangle, Colorado was processed for lineament recognition and lithologic classification. The reasons for choosing this area are that high quality aeromagnetic data is available (Figure 5.1) and the geology in this area is well known.

5.1 Geology

Figure 5.2 shows the geologic map of the Pueblo Quadrangle, Colorado (Tweto, 1979). The survey area is a rectangle with sides 200 km by 120 km, which is depicted on the index map in the lower-right corner of Figure 5.2.

The mountains of the Front Range region are composed primarily of Precambrian crystalline rocks, and the foothills and plains are underlain mostly by Phanerozoic sedimentary rocks. The contact between them is a series of high-angle reverse faults along the Front Range from Golden to Colorado Springs (Lee, 1989).

Superimposed on the Precambrian basement are two sets of high-angle faults. The northwest-trending faults frequently
Figure 5.1. Contour map of the aeromagnetic field of the Pueblo Quadrangle, Colorado. Contour interval 200 gammas.
GEOLGIC MAP OF THE PUEBLO QUADRANGLE

Figure 5.2. Geologic map of the Pueblo Quadrangle, Colorado.
show Neogene offset, whereas the northeast faults usually do not. The northeast faults are restricted to a northeast-trending zone more-or-less coincident with the offset of the Front Range. There are northwest-trending and east-west trending faults in the Tertiary igneous rocks and sedimentary rocks, along which basaltic intrusive rocks and associate tuff are present.

Precambrian rocks are of three ages: the oldest rocks are 1.7 Ga migmatites at the Royal Gorge, augen gneisses at Cripple Creek, and mica schists in High Park and Cripple Creek. Quartz monzonites (1.4 Ga) intrude the Boulder Creek metamorphic rocks at Cripple Creek, and these in turn are intruded by the Pikes Peak granite (1.0 Ga) north and east of Cripple Creek. Phanerozoic rocks of the area include sedimentary rocks and volcanic rocks. Lower Paleozoic rocks are shelf-type marine sedimentary rocks, with massive red arkoses of Pennsylvanian (Late Carboniferous) age. Mesozoic fluvial sequences give way to Cretaceous marine sedimentary rocks. Tertiary rocks consist of ash-flow tuffs and lahars, with occasional gravels preserved in down-faulted areas. Quaternary alpine glacial deposits occur in the high mountains, with pediment alluvial along the foothills (Lee, 1989).
5.2 Lineament recognition

Since the main structural trend in this region is northwest, the azimuth of the light source for shaded-relief normalization was chosen to be northeast. The entire data was processed using the program "rline" with the point-by-point method. In order to extract the lineaments with northeast trends, and obtain information about buried faults, the aeromagnetic data was processed again using shaded-relief normalization with the declination of light source north-south. Figure 5.3 illustrates the final composite result of the two lineament maps.

Comparing Figure 5.3 with the geologic map (Figure 5.4), we can see that a number of lineaments fit the geologic map well, for example the oval structure (I) surrounding Pikes Peak; the mountain front (II), i.e., the boundary between the rocks of Pikes Peak batholith (Yp, 1,000 my. age group) and the granitic rocks of 1,400 my. age group (Yg); the northwest fault zone (III) in the Precambrian metamorphic rocks; and the metamorphic dikes (VI) on the Pennsylvanian sedimentary rocks.

In addition to the lineaments above, there are some lineaments which can not be found on the geologic map. It is believed that these lineaments relate to buried faults, for instance, a discontinuous northeast linear structure (V) in the lower-left corner of Figure 5.3, and at least two parallel
Figure 5.3. Composite lineament map from the aeromagnetic field of the Pueblo Quadrangle, Colorado using BPNN (with shaded-relief normalization both northeast-southwest and north-south sun angles).
Figure 5.4 Comparison between the geologic map and lineament map.
lineaments (VI, VII) with east-west trends in the north of the Canon City embayment.

5.3 Lithologic classification

To avoid the influence of near-surface geologic noise, the aeromagnetic data were pre-processed by upward continuation 0.8 km. Using the CPFS program "analy" and the subroutine "norm", the original data set was normalized into values from 0.0 to 1.0.

The boxes Cl - C5 on Figure 5.1 are the model patterns we want to pick from the data; the CPFS program "wind" was used to extract these model patterns from the normalized data. The next step was to train the neural networks with the model patterns using the programs "prep" and "bp1". Finally, using the program "rlitho" and the trained networks or weights, the whole data set was classified. Figure 5.5 illustrates the results of classification. This classification map looks somewhat simplified and is not completely satisfactory; for example class 4 and class 5 cannot be distinguished from each other since the magnetic characters of model patterns 4 and 5 are so close. However, we can still easily distinguish Precambrian metamorphic rocks, igneous rocks, Tertiary sedimentary and Cretaceous sedimentary according to the grey level and their patterns.
Figure 5.5. Lithologic classification map from aeromagnetic data of the Pueblo Quadrangle, Colorado using BPNN.

The diagram shows a map with various geological classifications including Precambrian igneous, Precambrian metamorphic, Cretaceous sedimentary, and Tertiary sedimentary. The map includes a scale for 50 Km.
DISCUSSION AND CONCLUSIONS

Results of a number of the experiments using synthetic data and real data in this dissertation demonstrate that artificial neural networks can be applied to automatically extract information about lineaments and magnetic textures from potential fields.

The neural networks used to recognized features from potential fields (generally, the same as those from other geophysical data) depend on

1) Selection of parameters of neural networks. For example, BPNN parameters: number of nodes in the hidden layer, learning rate and slope of the sigmoid activity function.

2) Design and construction of suitable model patterns.

3) Construction of a reasonable representation.

The example of feature recognition from the aeromagnetic data of the Pueblo Quadrangle, Colorado using BPNN is a successful application. The structural lineament map and the lithologic classification map after recognition show that lineaments recognized using BPNN technique correspond to those which would be recognized by a skilled interpreter. Textures or patterns of magnetic fields can be classified using BPNN, which also can be used as a effective tool for picking one class of magnetic anomalies. The lithologic or textural classification from magnetic fields actually is one kind of recognition for features or characteristics of magnetic
fields, which relate indirectly to lithology in geology.

This technique can be applied to any kind of gridded data. Feature recognition with multiple geophysical data types using neural networks can be expected to achieve geologically reasonable results.
REFERENCE


Jacobs, R. A. 1988, Increased rates of convergence through learning rate adaptation: Neural Networks, v. 1, p. 295-307


Mousset, E. 1990, Spike filtering in seismic data with neural networks, in Conference on artificial intelligence in petroleum exploration & production: Texas A&M University, College Station, p. 115-123.


Tweto, O., 1979, Geologic map of Colorado: U.S. Geological Survey, Reston, VA.