AUTOMATIC MAP FEATURE EXTRACTION USING ARTIFICIAL NEURAL NETWORKS

Itthi Trisirisatayawong and Mark R. Shortis
Department of Geomatics
University of Melbourne
Parkville VIC 3052 AUSTRALIA
Mark_Shortis@mac.unimelb.edu.au

ABSTRACT

This paper describes the implementation of and experimentation with multi-layer feedforward neural networks to extract particular map features from scanned topographic maps. Commercial scan-conversion systems for automatic map input exist, but their capabilities are limited to vectorisation and other post-conversion processes on clean single-theme map images. This limitation impedes their application on paper maps which are in far more common use than single-theme images. The central issue of this paper is a technique that can be used as generally as possible as a mechanism to extract single-theme data from multi-theme documents. Backpropagation neural networks can establish any complex decision boundaries and therefore can be applied in any classification problem, which makes the technique attractive as a base for the development of such a mechanism. Experiments are carried out on scanned images of two topographic maps at different scales. The results demonstrate that neural networks have the potential to be implemented for automatic map data acquisition for GIS.

INTRODUCTION

The creation of a clean digital databases is a most important and complex task, upon which the usefulness of GIS depends (Burrough, 1988, p.56). Of the number of sources that can be used as input to GIS, secondary data like maps has been the most important source because of the variety, accuracy and effectiveness of maps to convey information about real-world phenomena, and their relationships, to the users. However, although maps are very efficient stores of information, it is surprisingly difficult to obtain certain types of numeric information from them (Goodchild, 1990). To date, the tool normally offered by commercial GISs to capture map data is manual digitisation using a hand-held cursor. It is well known that data capture by this method is slow, inconsistent and error-prone, so spatial database creation is expensive. Screen-based or head-up digitisation may eliminate the inefficiency of looking back and forth between the digitising table and screen, but in order to achieve the accuracy required, it is necessary to magnify feature data. The time required to change view windows in high magnification modes often makes capture of spatial features more time-consuming with screen digitising than similar accuracies achieved by using conventional digitising tablets (Skiles, 1990).

Since the early 1970s, many commercial systems have been offering automated line-tracing, a technique for rapid generation of vector data from line maps. The line-tracing system may be controlled by special hardware devices and software, or purely by software. The core of a scan-conversion system is vectorisation, whose fundamental requirement is that clean single-theme maps such as map separates are available. This assumption leads to a very narrow, well-defined problem domain which allows commercial development. For many reasons, however, clean single-theme maps may not be available (Faust, 1987). This situation is more severe in most developing countries where map separates are strictly controlled, mainly for reasons of national security. Consequently, the application of automatic conversion systems on multi-theme map documents, which in are far more common use than map separates, are impeded by the same assumption that allows their commercial development.

Automatic map-conversion systems can be used more widely if the assumption about the availability of single-theme maps is removed. In other words, a process that can extract particular features from scanned map images and feed them to a vectorisation process is needed. The demand on this missing component will accelerate due to the fall in prices of commercial desktop scanners in recent years, since this means that most organisations now can afford to routinely capture map data in an automatic manner. Unfortunately, the level of automation of feature extraction is far behind vectorisation and other processes thereafter. This creates a situation in which single-theme data may be reproduced by manual tracing on original multi-theme documents before being scanned and vectorised. It can be clearly seen that this practice is in fact equivalent to manual digitisation and thus suffers all the same drawbacks.

There has been little reported research concerned with automatic feature extraction from images produced by scanning multi-theme maps. Fain (1987) addressed pixel classification as part of a solution to automatic data conversion, but no indications about source documents, techniques and results were given and the emphasis was on interactive editing rather than automatic methods. In the paper of Konty et al (1991), some information about test documents was available but the core work is a benchmark test of commercial systems without any attempt to evaluate the underlying techniques. Of the more technical work, Eveleigh and Potter (1989) reported a preliminary study of using the Bayes technique to classify a USGS 7.5 minute quadrangle covering a rural area. The study, however, did not indicate whether the RGB intensity values of the map image had the normal distribution assumed by the Bayes classification technique and neither was any classified result nor quantitative information revealed.

Previous research in automatic map feature extraction concentrates on the implementation of techniques that can be possibly used for particular test maps. The important issue of the generality of the employed techniques has not yet been systematically studied. This paper investigates neural network techniques, which have the potential to be implemented as a general feature extraction mechanism. The property of neural networks that makes it a powerful technique is described first, followed by discussion of the experiments carried out on test maps.

NEURAL NETWORK FEATURE EXTRACTOR

Feature extraction is achievable by classification. This popular scheme in pattern recognition employs a classifier to classify image objects according to their characteristics. A classifiable characteristic is any numeric information used to distinguish one part of image from other parts. The objects may be single pixels or groups of contiguous pixels depending on the level of abstraction at which classification is performed. The derived characteristics are fed into the classifier which produces class labels for objects.

A classification technique may be categorised as supervised or unsupervised. Basically, an unsupervised classification is a data clustering technique whose fundamental idea is that different feature classes should form different clusters in characteristic space. Thus, unsupervised classification is done by grouping objects into clusters based on the criterion that the clusters should be formed as compactly or tightly grouped as possible. Other criteria or additional processing may be employed to enhance the separation between clusters. These techniques are called unsupervised because they make no use of external knowledge or the characteristics of feature classes.

Unsupervised classification is mainly used when there is little information regarding what features the scene contains. This is particularly true for satellite images which may cover inaccessible areas and, in such circumstances, it is impractical or too expensive to obtain sample data. The drawback of unsupervised techniques is that feature classes may only be marginally separable or may not form obvious clusters at all. Unsupervised classification cannot provide correct results in such situations and this problem prevents the technique from being used as a general tool.

On scanned map images, the problem of the unavailability of sample data is certainly not the case and supervised techniques are the obvious choice for the classification task. Instead of relying on compactness, which may not truly reflect the data structure, supervised techniques utilise information contained in the sample data to establish decision boundaries between feature classes.

The classification process can be geometrically interpreted as the establishment of decision boundaries in the characteristic space. The more complex the decision boundaries a classifier can establish, the more general it is. The most important of traditional supervised classifications is the Bayes technique. In the Bayes technique, the sample data is used to estimate the parameters, for instance the mean vector and covariance matrix, of the multivariate normal distribution. The bell shapes of normal probability density functions in two dimensional space mean that the decision boundaries drawn by the Bayes technique are hyper-ellipsoids in n dimensional space.

A new and increasingly important technique is backpropagation neural networks. The interest in the application of backpropagation neural networks for classification problems is driven by the analysis shown in Lippmann (1987) that a single hidden-layer backpropagation neural network can form any, possibly unbounded, convex region in n dimensional space. A simple extension to a two hidden-layer network can establish arbitrarily complex decision boundaries which can separate even meshed classes. Although the latter case is not explored in this paper, this capability means that, in theory, backpropagation neural networks can be used for virtually any classification problem regardless of the statistical distribution of data. It follows that the backpropagation neural network technique is a more general technique than the Bayes technique, which is based on the assumption that the data is normally distributed.

However, accuracy also needs to be taken into account to assess the validity of the generality. It has long been known that when the underlying assumption about the statistical distribution is met, the Bayes classifier provides an optimum result. This important property has established the Bayes technique as a benchmark against which any newly proposed technique has to measure. The next section compares performances of both techniques when applied to a scanned topographic map.

FEATURE EXTRACTION AT PIXEL LEVEL

The performance analysis of backpropagation neural network classifiers has been undertaken on remotely-sensed data by a number of reported research works, many including a comparison with the Bayes technique (Howald 1989; Hepner et al 1990; Heermann and Khazenie 1992; Bischof et al 1992). However, despite the importance of these two techniques, little attention has been paid by the mapping/GIS community for such studies on map data. The comparison of Bayes and backpropagation neural network techniques on the segmentation of real map images was probably first reported by Trisirisatayawong and Shortis (1993). The study used RGB spectral characteristics to classify a portion of 1:100,000 Australian topographic map which was scanned at a resolution of 150 dots per inch. The advantages and disadvantages of both methods in practical issues were discussed, but a detailed analysis of data and performance of each method on each feature was beyond the scope of the paper.

The analysis below is an extension of the above-mentioned work. Figure 1 illustrates the test map which is shown in grey-scale. The statistics of map features are shown in table 1. Some of the results from Bayes and one hidden-layer neural network classifiers are shown in figure 2.

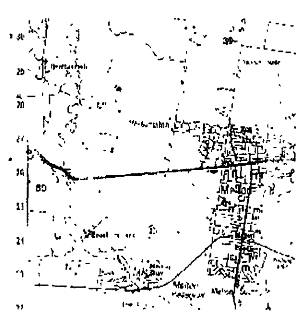


Figure 1: Test map (from the colour original at 1:100,000 scale).

Crown Copyright reproduced with the permission of the General Manager, Australian Surveying and Land Information Group, Department of the Arts and Administrative Services, Canberra, Australia

Feature	Number of Samples	Mean			Standard Deviation			Skew		
		R	G	B	R	G	B	R	G	B
Built areas	39	237	153	155	8	14	14	-0.5	0.0	-0.4
Contours	29	216	172	140	19	25	34	-0.5	0.0	0.6
Forest	50	222	233	155	21	16	36	-0.7	1.6	0.8
Roads	42	213	106	87	45	39	43	-2.2	0.2	0.3
Water	39	161	193	244	35	32	25	-0.8	-1.2	-3.7
Dark	36	80	70	83	35	24	28	1.8	1.1	0.2

Table 1: Statistics of spectral characteristics of features on the test map shown in figure 1.

Skewness, the magnitude of which indicates the degree of deviation of the data from a symmetrical distribution, of RGB characteristics is computed for each feature class. Large skewness values mean that the data is significantly skewed whereas smaller values indicate otherwise. Using skewness values as indicators, it can be seen that for the feature classes such as water, roads and forest, the sample data distributions are substantially skewed and therefore cannot be normal. It is therefore expected that the backpropagation neural network produces more accurate results, and this is evident in figure 1, particularly for the water image of rivers and lakes. On the layers of built areas and contour lines, whose spectral characteristics can be properly assumed to have normal distributions because of small skewness, the results from the two techniques are essentially similar.

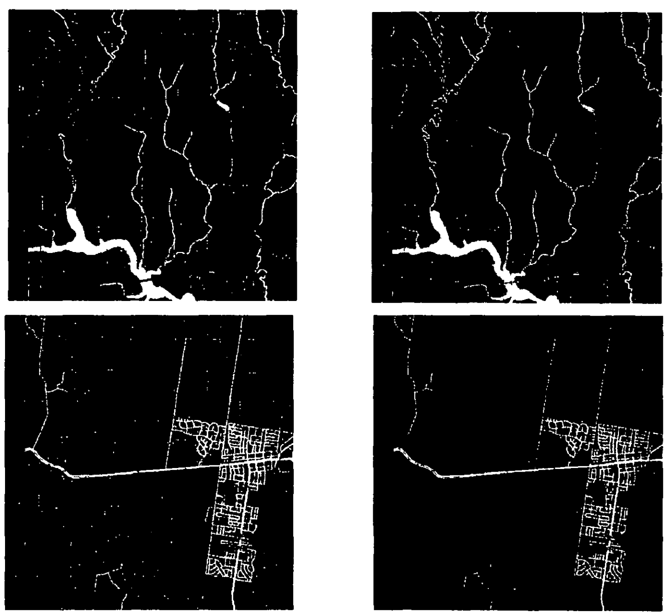


Figure 2a: From top to bottom: Classified images of the features of water bodies and roads respectively. The images on the left and right hand sides are results from the Bayes and backpropagation neural network techniques respectively.

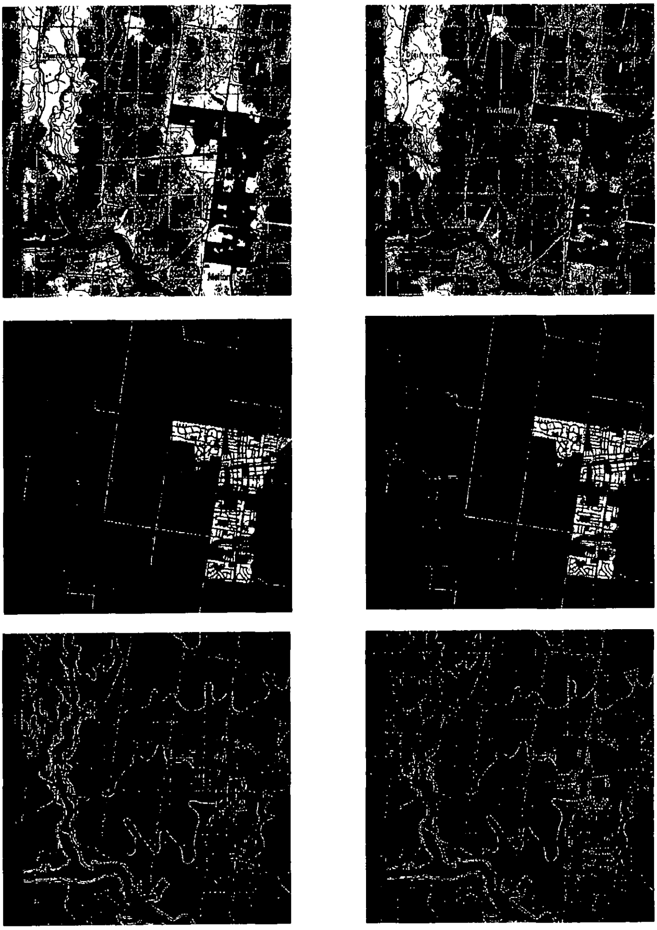
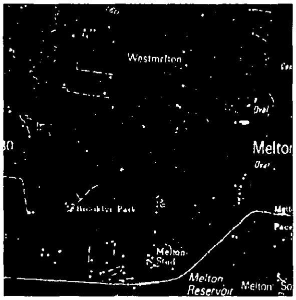


Figure 2b: From top to bottom: Classified images of the features of forest, built areas and contour lines respectively. The images on the left and right hand sides are results from the Bayes and backpropagation neural network techniques respectively.

The results indicate that a backpropagation neural network with one hidden-layer can be implemented for map-feature extraction at the pixel level. In the non-normal cases, the neural network provides better results, but even when the characteristics have a normal distribution, results provided by the Bayes bench mark are satisfactorily approximated by the neural network. Thus, the statistical distribution constraint in the Bayes technique is removed when the alternative neural network technique is applied, solving the same problem with equivalent accuracy. The proposition that neural networks are a general classifier is supported by the experimental results

However, not every feature can be extracted using a multispectral classification. Different features may have the same colour and it is not possible to differentiate one feature from another regardless of the technique used. An example of this problem is shown in figure 3 in which railway, text, house symbols and tracks are incorrectly assigned into the same class. Another example of the same situation is shown figure 4, which is an image resulted from a neural network multispectral classification of another scanned topographic map (original scale 1:50000, scanned at 300 dots per inch). The fact that these features hold the same spectral characteristics means that there is no way at pixel level to avoid the misclassification. Thus, multispectral classification provides only partial solution to the problem of map-feature extraction. Other techniques need to be utilised to resolve ambiguities resulting from the initial spectral classification.



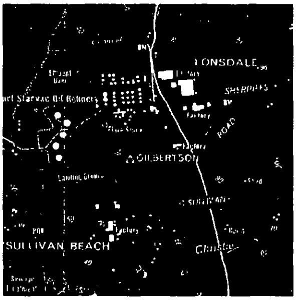


Figure 3 (left) and 4 (right): Images of mixed features resulting from a multispectral classification.

FEATURE EXTRACTION BY SHAPE ANALYSIS

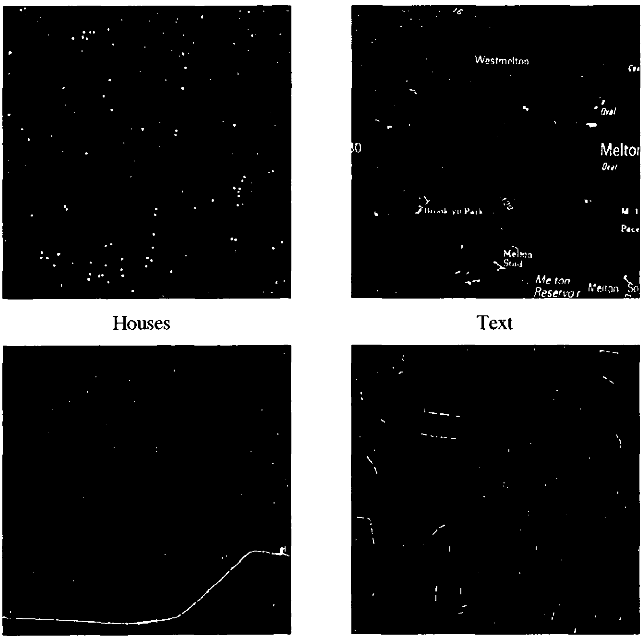
Further extraction on the image of mixed features must be carried out to produce single-theme images. If the appearances of map features are somewhat consistent, templates which define their likely pixel arrangement can be used in a classification by looking at the degree of match or similarity between the image part under consideration and the template being applied. The serious disadvantage of this simple technique is its inability to handle variations in shape and size of each feature type, which is normally the case in most maps. A very large number of templates must be defined to cover all possible occurrences and this may incur extremely heavy computational load. This may lead to an unacceptable situation in which even a high-speed processor will take several hours to locate features within the image. Another situation in which template matching techniques are not suitable is when the appearance of one feature is a part of another larger-sized feature. For example, the letter I also appears in the left portion of the letters B, D, P. Misidentifications (or false alarms) will occur when B, D, P are matched by the template of I. There is no universally effective solution for this problem.

Intuitively, sets of contiguous foreground pixels displayed as regions, identified by a pixel level classifier, can be treated as individual image objects. These objects can be further classified based on the similarities and differences in shapes. Thus, the concept that feature extraction can be formulated as a classification is still applicable, provided that shape information is properly quantified.

The shape characteristics must be tolerant to transformation and uniquely defined by the objects if the problems of the template matching method are to be avoided. Basic shape characteristics are those related to size such as area, perimeter and extent. These characteristics are invariant to translation and rotation but are affected by scale. A possible way to obtain scale-invariant characteristics is by relating the given measurements of objects to some well known geometric figure such as a circle. The result is dimensionless shape measurements which are invariant under magnification or reduction. For example, a compactness ratio could be derived by dividing the area of the object by the area of the circle having the same perimeter as the object. However, although it is possible to produce characteristics which are invariant to translation, rotation and scale in this way, there is no guarantee that two different object types will not produce the same characteristics. The use of shape values which are not uniquely defined by objects prevents the implemented classification technique on a particular test map to be subsequently applied to different map images.

The theory of moment invariants can be applied to produce object characteristics that are invariant under transformation and uniquely defined by objects. This analytical method was first introduced to the classification problem by Hu (1962). Details of moment invariants is omitted here but can be found in Hu (1962). Since its introduction, moment invariants have been used in aircraft identification (Belkasim et al, 1991), detection of ships in remotely sensed images (Smith and Wright 1971), and optical character recognition (El-Dabi et al 1990). All of this research applied moment characteristics in conjunction with conventional supervised or unsupervised classifiers. The performance of the technique combining moment characteristics with neural networks has not yet been explored.

In theory, coupling a neural network classifier with moment characteristics should result in a general classification technique. Of the infinite number of moments that can be chosen, only three, namely m00 (area or number of pixels comprising an object), M1 (spread) and M2 (elongation), are employed. The selection of this subset of moments is based on the consideration that these three values carry substantial shape information and should contain discrimination power adequate for classifying objects within a map image. Statistics of the three moment characteristics of the test image of figure 3 are shown in table 2 and the classified image results from using a one hidden-layer backpropagation neural net are shown in figure 5.



Houses | Text | Railways | Dashed lines

Figure 5: Images of the object classifications of the map image shown in figure 3.

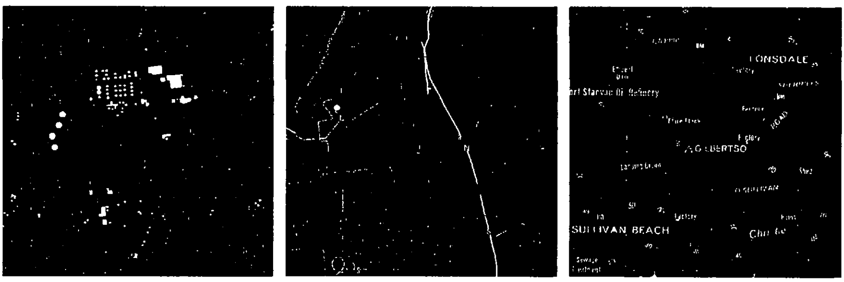
The results clearly illustrate that extraction of features from the test image is achieved with a high degree of accuracy and completeness. However, there are what seem to be misclassifications appearing in each classified image. If the classification of objects into a class is posed as the null hypothesis in statistical test, then it can be seen that most of the misclassifications are type-two errors. The classified images of house symbols and railways are free of type-one errors. In the text image, there are a few type-one errors but all of them occur from characters having similar shapes to dashed lines. This is reasonable since there is no way for the classifier to identify these characters without the help of extra information, such as context.

Object	Mean			Standard Deviation			Skew	
	m00	M1	M2	m00	M1	M2	M1	M2
Dashed lines	49	60	57	7	13	12	-0.38	-0.42
Houses	38	16	3	5	1	2	0.00	0.25
Railway	2719	1526	1414	0	0	0	0.00	0.00
Text	111	226	21	50	16	23	0.34	0.41

Table 2: Statistics of moment characteristics of map objects in figure 3.

Every classified image suffers from a different degree of type-two errors. The most serious case is the text image. However, almost all of the errors occur from compound objects which are an incorrect aggregation of two or more objects. Considering that there is no class representing them and they are not used in the training phase, these type-two errors are not mistakes of the classifier. In fact, a visual inspection reveals that, except in one instance where two characters on the railway image are mistakenly joined by the pixel classification, all compound objects appear on the original document or are a result of the finite sampling size of the scanning process.

A similar process of classification by moment characteristics performed on the test image of figure 4 produces similar results. In this case a slight modification is made. The number of classes of line features is restricted to one only, since each line type has only a few objects. So, there are three classes representing points, lines and text with another class being assigned as a noise channel. The results are illustrated in figure 6 below.



Points Lines Text

Figure 6: Images of the object classifications of the map image shown in figure 4.

Like the previous analysis, it can be seen that most of the features have been correctly classified. Except for two point symbols being misclassified into the layer of text and three elongated characters being incorrectly assigned to the line layer, most of the problems are caused by compound objects. White noise appearing in the classified images can be simply removed by size criterion. The classification accuracy of both text and point symbols are in the high 90 percent range and this is achieved without any extra information, such as contextual information, which certainly will enhance the results.

A number of further processes are required to convert the classified objects into features appropriate for the generation of a spatial database. Aggregated objects must be separated using, for example, mathematical morphology techniques (Trisirisatayawong, 1994). Line objects must be vectorised, text must be recognised and linked to associated map features

(Shortis and Trisirisatayawong, 1994), and a final phase of attribute tagging must be conducted (Trisirisatayawong, 1994).

CONCLUSIONS

Neural networks re-formulate all problems by finding the correct internal weights, so the technique can be viewed as a black-box problem solver in which the weights have no obvious physical meaning in the context of problem. The statistical distribution of data is insignificant compared to other traditional classifiers. This means that neural networks can be universally applied to all classification problems, provided the network is properly trained by appropriate and accurate sampling.

A drawback of the neural network technique is that it is often difficult to determine whether the neural network is correctly trained. Learning error is the only information used by the neural network to indicate the degree of success. There is no guarantee that when the error has converged to a particular value that it is the global, rather than a local, minimum. So, the magnitude of error often does not truly reflect the degree of learning. One widely-used practice is to set an acceptable error threshold and the network is accepted as adequately trained once the learning error has converged to a value less than this threshold. Thus, the amount of training of the network is subjectively determined by the operator, who must specify the threshold based on experience or any other suitable guideline.

Neural networks are extremely flexible in solving a wide variety of problems. The key factor determining the accuracy of a neural network is its structure, which can be constructed as single hidden-layer, multiple hidden-layer, partial inter-layer connection, full connection or other varieties. However, it also means that different neural networks may be constructed to solve the same problem and so, in mathematical sense, the technique of neural networks does not provide a unique solution. There is no general rule to determine whether the chosen structure is optimum. The most serious problem in practice is the determination of the learning rate, the initial weights and especially the structure of neural networks. All of these factors must be pre-determined by the operator who will in general set them from prior knowledge and experience.

Nevertheless, the drawbacks of neural networks occur mostly because the technique is still a relatively young science. The problems will dissipate as the knowledge of neural computing expands. For example, some guidelines about the determination of learning rate can be found in Kung and Hwang (1988), although the final settings must still be determined on a trial and error basis. Also, research on the automation of the determination of structure and the removal of redundant elements in the network to improve efficiency are under way (Wen et al 1992). The efforts in these areas will lead to less time and frustration incurred from training neural networks in the future.

Overall, the advantages of neural networks as a general classifier outweigh the disadvantages. As the experimental results on real map data shown here firmly support the theoretical claims, it is believed that neural networks can be further developed as general map feature extraction mechanism.

REFERENCES

Belkasim S.O., Shridhar M., and Ahmadi M., 1991, “Pattern Recognition with Moment Invariants: A Comparative Study and New Results”, *Pattern Recognition*, Vol. 24, No. 12, pp. 1117-1138.

Bischof H., Schneider W., and Pinz A.J., 1992, “Multispectral Classification of Landsat-Images Using Neural Networks”, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, No. 3, pp. 482-490.

Burrough P.A., 1988, *Principles of Geographical Information Systems for Land Resources Assessment*, Reprinted, Oxford University Press, UK.

El-Dabi S.S., Ramsis R., and Kamel A., 1990, “Arabic Character Recognition System: A Statistical Approach for Recognizing Cursive Typewritten Text”, *Pattern Recognition*, Vol. 23, No. 5, pp. 485-495.

Eveleigh T.J., and Potter K.D., 1989, "Spectral/Spatial Exploitation of Digital Raster Graphic Map Products for Improved Data Extraction", *Proceedings of Auto-Carto 9*, Baltimore, Maryland, pp. 348-356.

Fain M.A., 1987, "Optical Scanning and Interactive Color Graphics: An Integrated Approach to Automated Data Capture", *Technical Papers of the 1987 ASPRS-ACSM Annual Convention*, Baltimore, Maryland, Vol. 2, pp. 323-326.

Faust N.L., 1987, "Automated Data Capture for Geographic Information System", *Photogrammetric Engineering and Remote Sensing*, Vol. 53, No. 10, pp. 1389-1390.

Goodchild M.F., 1990, "Geographic Information Systems and Cartography", *Cartography*, Vol. 19, No. 1, pp. 1-13.

Heermann P.D., and Khazenie N., 1992, "Classification of Multispectral Remote Sensing Using a Back-Propagation Neural Network", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, No. 1, pp. 81-88.

Hepner G.F., Logan T., Riter N., and Bryant N., 1990, "Artificial Neural Network Classification Using a Minimal Training Set: Comparison to Conventional Supervised Classification", *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 4, pp. 469-473.

Howald K.J., 1989, "Neural Network Image Classification", *Technical Papers of ASPRS-ACSM Fall Convention*, Cleveland, pp. 207-215.

Hu M.K., 1962, "Visual Pattern Recognition by Moment Invariants", *IRE Transaction on Information Theory*, Vol. 8, No. 2, pp. 179-187.

Konty L., and Gross G., 1991, "Scanning for GIS Data Conversion: Perils, Pitfalls & Potentials", *URISA Proceedings*, San Francisco, Vol. 2, pp. 192-197.

Kung S.Y., and Hwang J.N., 1988, "An Algebraic Projection Analysis for Optimal Hidden Units Size and Learning Rates in Back-Propagation Learning", Proceedings of IEEE International Conference on Neural Network, San Diego, Vol. 1, pp. 363-370.

Lippmann R.P., 1987, "An Introduction to Computing with Neural Nets", *IEEE ASSP Magazine*, Vol. 4, No.2, pp. 4-22.

Shortis, M. R. and Trisirisatayawong, I., 1994, "Automatic text recognition and the potential for text to feature association on scanned maps", *Australian Journal of Geodesy, Photogrammetry and Surveying*, (in press).

Skiles J.M., 1990, "Using Scanning, Automated Conversion, and Raster Data to Speed GIS Implementation and Lower Cost", *Proceedings of GIS/LIS'90*, Anaheim, California, pp. 476-483.

Smith F.W., and Wright M.H., 1971, "Automatic Ship Photo Interpretation by the Method of Moments", *IEEE Transactions on Computers*, Vol. 20, No. 9, pp.1089-1095.

Trisirisatayawong I., 1994, "Automatic Feature Extraction from Scanned Topographic Maps", unpublished Ph.D. thesis, The University of Melbourne, Australia, 152 pages.

Trisirisatayawong I., and Shortis M.R., 1993, "A Comparison of Two Feature Classification Methods for Scanned Topographic Maps", *Cartography*, Vol. 22, No. 1, pp. 1-14.

Wen W.X., Liu H., and Jennings A., 1992, "Self-Generating Neural Networks", Proceedings of the International Joint Conference on Neural Networks, Baltimore, Maryland, Vol. 2, pp. 111-117.