UNIVERSITY OF SOUTHAMPTON

SUPER-RESOLUTION LAND COVER MAPPING FROM REMOTELY SENSED IMAGERY USING A HOPFIELD NEURAL NETWORK

ANDREW TATEM, B.Sc.

DOCTOR OF PHILOSOPHY

FACULTY OF ENGINEERING AND APPLIED SCIENCE

DEPARTMENT OF ELECTRONICS AND COMPUTER SCIENCE

MARCH 2002

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE

ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

SUPER-RESOLUTION LAND COVER MAPPING FROM REMOTELY SENSED IMAGERY USING A HOPFIELD NEURAL NETWORK

by Andrew James Tatem

Improved information on land cover is required to aid our management and understanding of the environment. Remote sensing provides, commonly, the best method for obtaining geographically and temporally detailed estimates of land cover and its changes. However, the large ground area of each pixel, relative to the size of typical land cover parcels, means that pixels often represent a mix of land covers. The presence of such pixels can adversely affect the performance of traditional per-pixel classification approaches. Soft classification techniques have been developed to estimate the class composition of image pixels, but their output provides no indication of how these classes are distributed spatially within the pixel.

The use of a Hopfield neural network to map the spatial distribution of land cover classes using prior information of pixel composition determined from soft classification is demonstrated. An approach was adopted that used the output from a soft classification to constrain a Hopfield neural network formulated as an energy minimisation tool. Energy functions were defined for various tasks, consisting of a goal and several constraints, and the network converges to a minimum of these functions. The energy minimum represents an estimated map of the spatial distribution of land cover class components in each pixel.

The Hopfield neural network design was tailored to meet specific goals, including, super-resolution target identification, land cover mapping and land cover pattern prediction. The approach was applied to synthetic data, simulated remotely sensed imagery and real Landsat Thematic Mapper imagery, with high levels of mapping accuracy and improvements over traditional techniques produced throughout. Overall, we show that the spatial resolution of satellite sensor imagery need not necessarily represent a limit to the spatial detail obtainable in land cover maps derived from such imagery. The Hopfield neural network technique is shown to be robust, efficient and simple, and results suggest that it has the potential to predict accurately land cover at the sub-pixel scale from operational satellite sensor imagery.

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The writing of this thesis would not have been possible without the help of many individuals. Principally, I'd like to thank my two supervisors, Dr. Peter Atkinson and Dr. Mark Nixon for their continual support, constructive suggestions and useful guidance. Special thanks go also to Dr. Hugh Lewis for all his help with ideas, computer code and paper writing.

I am grateful to the Engineering and Physical Sciences Research Council (EPSRC) for providing me with funding for this research. Also, thanks to all those involved in the 'Fuzzy Land Information for Environmental Remote Sensing' (FLIERS) EU project for making the full data-set available to me.

Thanks to all those people who have helped me throughout my three years, either with work related matters or just providing good drinking partners. I am very grateful to Briony for all her help in my final year and for our many philosophical discussions. Others who deserve a mention are Antony, Sally, Gary, Rich, Marie, Stuart, Steve W, Paul C, Nick H, TP, Toby, Pedro and last, but not least, Special K.

Finally, thanks to my parents, family and friends for their support.

Chapter 1: Introduction

1.1 Rationale

In this chapter, a review is presented of how remote sensing and related techniques provide the means to map land cover at the sub-pixel scale. While an overall appraisal of the topic is provided, certain issues that are key to this research project are discussed in detail. These include spatial variation in remotely sensed imagery, and techniques for soft classification.

Initially, the topic of land cover is introduced, followed by a description of how remotely sensed image parameters affect the representation of land cover provided by such imagery. Methods employed to extract land cover information from remotely sensed imagery are then presented, including techniques for pre-processing, classification and feature extraction and, finally, previous attempts and studies of land cover mapping at the sub-pixel scale are covered.

1.2 Land cover

Information on land cover is required to aid our understanding and management of the environment. Land cover represents a critical biophysical variable in determining the functioning of terrestrial ecosystems in bio-geochemical cycling, hydrological processes and the interaction between surface and atmosphere (Cihlar *et al.*, 2000). Wyatt *et al.* (1993) define land cover as 'the description of the physical nature of the land surface, for example, vegetation, buildings, water or bare soil'. Information on land cover is central to all scientific studies that aim to understand terrestrial dynamics and is required from local to global scales to aid planning, while safeguarding environmental concerns. Land cover information has many uses:

(i) *Inventories of land resources*

Land cover data inventories have the potential to provide valuable information for determining land management practices (Townshend, 1992, Veitch *et al.*, 1995).

(ii) *Monitoring change*

Land cover data from temporal sequences can be utilised to provide information on land cover dynamics for monitoring changes in land cover (Riley *et al.*, 1997).

(iii) *Understanding land use interaction*

The interaction between land use practices and environmental processes can be better understood and explained via the analysis of land cover data (Dimyati *et al.*, 1996,

O'Neill *et al.*, 1988). This increased understanding puts planners in a better position to implement land use policies, without detrimental consequences for the environment.

(iv) *Modelling and prediction*

The use of land cover data as input to predictive models serves as an aid to policy makers for future land use strategies (Thornton and Jones, 1998, Lambin, 1997).

1.3 Land cover Monitoring

In general, there exist three main sources of primary land cover data; field survey, aerial survey and satellite sensor survey (Balchin, 1984).

(i) *Field Survey*

Field survey was the dominant source of land cover data until the late 20th century. The high accuracy produced by direct observation of the land is a benefit, but the lengthy and expensive procedure of field survey proves to be a hindrance in most cases (Fuller *et al.*, 1994, Veitch *et al.*, 1995).

(ii) *Aerial Survey*

The 1960s and 1970s saw field survey replaced by aerial photography as the dominant source of land cover data. While it can be less accurate, generally it is considerably faster and cheaper (Fuller *et al.*, 1998, Osborn, 1998). More recently, airborne sensor imagery (Curran and Ojedele, 1988, Atkinson, 1993, Green *et al.*, 1998) and airborne videography (Anderson *et al.*, 1997, Um and Wright, 1999) have also been used to obtain land cover data.

(iii) *Satellite Sensor Survey*

While aerial photography still dominates surveys at the local scale, satellite sensor imagery has recently been used widely for land cover surveys at regional, national and continental scales. The spatial detail generated by satellite sensor surveys is generally less than that of aerial surveys, resulting in lower overall classification accuracy for such studies (Aniya et al., 1996). However, the lower costs, rapid data supply, convenient automated information extraction, repeatability and wide area coverage of satellite sensor surveys has meant their wider use (Fuller *et al*, 1994). The launch of increasingly advanced satellite sensors (Aplin *et al.*, 1997) also means interest in such imagery continues to grow.

1.4 Remote Sensing

Remote sensing is the examination of an object or phenomenon at a distance from it, without physical contact with it (Clark, 1990). It is the practice of deriving information about the Earth's atmosphere and surface using images acquired from an overhead perspective, using electromagnetic radiation reflected or emitted from the Earth's surface (Campbell, 1996). The electromagnetic radiation is recorded via the use of sensors mounted on aircraft or satellites (Aplin *et al.*, 1997).

1.4.1 Physical Principles

The sensors onboard satellites or aircraft are sensitive to a broad range of electromagnetic radiation, ranging from ultraviolet to radio waves (Figure 1.1).

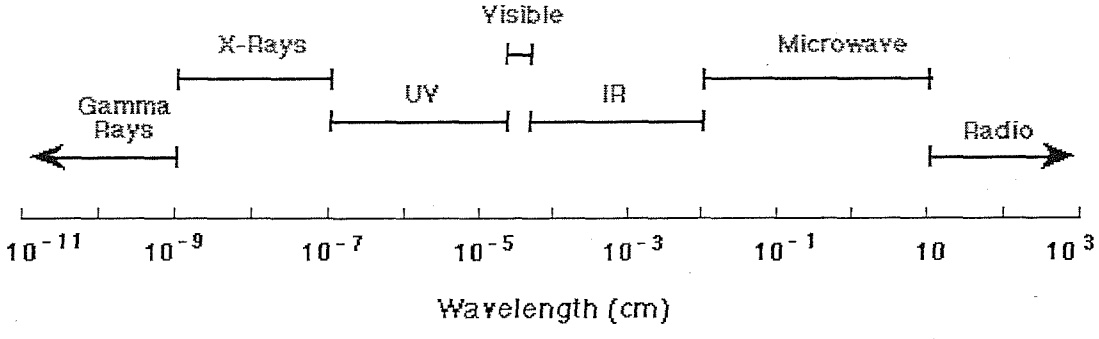


Figure 1.1 The electromagnetic spectrum

Remote sensors detect electromagnetic radiation and convert it to a data record, which can be analysed by visual or automatic means. A remote sensor detects the radiance reflected or emitted within its instantaneous field-of-view (IFOV). The magnitude of reflected or emitted radiance measured in each waveband for a single IFOV is considered to be related to the characteristics of the material forming the surface cover of the ground area corresponding to that IFOV (the Ground Resolution Element (GRE)). Therefore, the radiance recorded by a satellite sensor depends upon the sensor's characteristics, in addition to the land cover features seen by the sensor (Warner and Shank, 1997). Measurements can be taken in a variety of wavebands (or bands). For example, the Landsat Thematic Mapper (TM) sensor provides data in seven bands between 0.475 μm and 1.1 μm, with a GRE size of 30 m by 30 m. In contrast, the National Oceanographic and Atmospheric Administration Advanced Very High Resolution Radiometer (NOAA AVHRR), provides data in five wavebands between 0.58 μm and 12.5 μm, with a GRE size of 1.1 km by 1.1 km. In these regions of the electromagnetic spectrum, the ground material interacts with, and modifies the spectral

distribution of incident radiation. This means, in theory, it is possible to identify the ground surface cover type by studying the distribution of spectral radiance values recorded by the sensor. This basic principle underlies much remote sensing analysis.

1.5 Pixels

The pixel is the smallest element of a remotely sensed digital image, and represents the GRE of a sensor. The ground area represented by each pixel is determined by the characteristics of the sensor that produced the imagery. Ideally, each pixel should represent the reflectance from its own square area of ground as a square wave linear response. However, this is not the case, and several factors need to be taken into account.

1.5.1 Sensor Spatial Resolution

There exist practical limits to the level of detail that can be identified by each remote sensor. These limits are defined as the resolutions of the remote sensing system, although the image detail recorded depends also upon the character of the scene, illumination, atmospheric conditions and the ability of the image interpreter (Schowengerdt, 1997). One of the commonest measures of image characteristic used is spatial resolution. This determines the level of spatial detail depicted in an image. The IFOV of a sensor is defined as the cone angle within which incident energy is focused on the detectors (Campbell, 1996). The IFOV leads to a ground resolution element on the ground. This GRE should not be confused with the pixel, which is the output product to which a radiance value is assigned. This difference can be illustrated when studying the Landsat MSS (Multispectral Scanner Subsystem). It has a GRE of 79 m by 79 m, with an overlap of around 11 m in each scanner pass, producing the area represented by a pixel as 57 m by 79 m. The effect of the point spread function (PSF) of the sensor means that the area viewed by the sensor is larger than the nominal GRE (Schowengerdt, 1997).

1.5.2 Sensor Point Spread Function

The optical point spread function (PSF(x-u, y-v)) describes the intensity of the returned signal as a function of position (x,y) in the focal plane arising from an object which is a point source with its geometrical image at the point (u,v) in the focal plane (Cracknell, 1998). As described previously, in an ideal situation the response should be constant for all points (u,v) within the pixel, and zero for all points (u,v) outside it. Figure 1.2 shows a sketch of the PSF for the AVHRR.

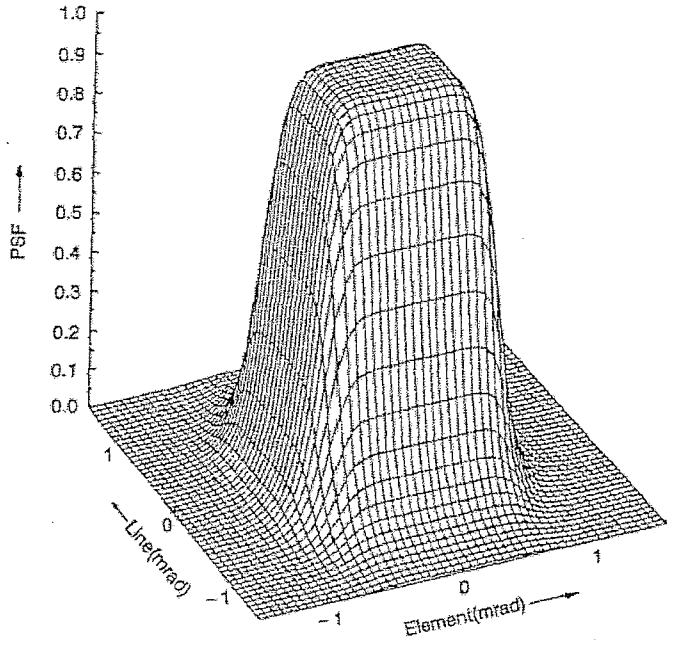


Figure 1.2 Sketch of point spread function of the AVHRR (from Cracknell, 1998).

The PSF of a sensor integrates the response from a target point and its neighbouring points, and is a result of convolution of the PSF with the scene (Forster and Best, 1994). The scale of the problem caused by the effect of the PSF depends on the scale of the surface features being studied and the size of the PSF itself. The sensor PSF means that reflectance represented by the pixels is centre-biased, such that reflectance towards the centre of the IFOV has most influence (Fisher, 1997). The shape and size of the PSF also means that in most cases overlap occurs between pixels occurs.

Craig and Labovitz (1980), Craig (1981), McGillem *et al.* (1983) and Markham (1985) have estimated the effect of the PSF on images. Tubbs and Coberly (1978) and Lewis *et al.* (2000) found the effect of the PSF to be detrimental to attempts at land cover classification from remotely sensed imagery. Many attempts have been made to remove the effect of the sensor PSF, via the use of PSF estimation then the application of deconvolution filters (e.g. Breaker, 1990, Law and Nguyen, 1995, Townshend *et al.*, 2000), with success for certain types of imagery.

1.5.3 Spatial Variation

Spatial scale is one of the key factors in the interpretation of remotely sensed data of land cover (Woodcock and Strahler, 1987). Depending on the spatial variation of the land cover being observed, and the specific terrain characteristics under consideration, the information

obtainable from remotely sensed imagery will vary greatly (Atkinson and Curran, 1995, 1997, Pax-Lenney and Woodcock, 1997). Within a remotely sensed image, one of the fundamental features is the spatial variation in land cover, dependent on the arrangement of land cover types. When examining an object-based model, the intrinsic scale of spatial variation in land cover is governed by the range of land cover features, the individual land cover features of interest, and the range of densities of these features (Cushnie, 1987). For example, spatial pattern is an important characteristic of ecological populations and in forestry, the spatial arrangement of trees within forest stands provides an insight into the allocation of above- and below-ground resources (Coops and Culvenor, 2000).

1.5.3.1 Spatial Dependence

Almost all natural and human-made phenomena exhibit spatial dependence at some scale. Points near to each other are more alike than those further apart, and the degree of dissimilarity depends on both the environment and the nature of the observations (Curran and Atkinson, 1998). These observations can be the pixels in remotely sensed images, and the assumption of spatial dependence can be used to infer relationships between these pixels (Atkinson, 1997).

1.5.3.2 Geostatistics

Spatial dependence is the assumption that underlies the field of geostatistics (Matheron, 1965, Isaaks and Srivastava, 1989, Journel and Huijbregt, 1978). In simple terms, geostatistics forms a set of techniques for predicting the local values of properties that vary in space from sample data. Crucial to geostatistics is the variogram. This is a function that relates variance to spatial separation, and therefore provides a concise description of the pattern and scale of spatial dependence (Oliver *et al.*, 1989a, 1989b). Within remote sensing, the variogram can prove a vital tool for use in the techniques of Kriging and conditional simulation (Isaaks and Srivastava, 1989, Journel and Huijbregts, 1978, Ripley, 1986). Kriging and cokriging are interpolation techniques, which provide unbiased estimates with minimum and known variance (Atkinson and Kelly, 1997). However, the results produced using these approaches could never exist in reality, as the variance of the original data is smoothed. In contrast, conditional simulation, also known as stochastic imaging (Journel, 1996, De Bruin, 2000), provides an alternative in which the variance and variogram of the original data are retained and the general pattern of spatial variation in the original data is recreated, creating a 'possible reality'. Many different realities can be conditionally

simulated, representing the inherent uncertainty in the original spatial samples (Curran and Atkinson, 1998).

1.6 Land Cover Classification

A thematic map shows the spatial distribution of identifiable surface features, providing an informational description over a given area. Image classification produces thematic maps from imagery by assigning pixels to classes (Campbell, 1996). Thematic classification of an image traditionally involves several steps:

(i) *Feature Extraction*: The multispectral image is transformed spatially or spectrally to produce a feature image. This may include, for example, a principal components transform, selection of certain bands (Schowengerdt, 1997), smoothing filters or segmentation (Tailor *et al.*, 1986).

(ii) *Training*: Pixels are extracted to train the classifier to recognize certain classes. Discriminant functions are determined in feature space, allowing the assignment of class labels to each pixel.

(iii) *Labelling*: Discriminant functions determined in the training phase are applied to the whole image, producing labels for all pixels.

(iv) *Accuracy Assessment*: The predicted pixel labels are compared to a reference map and various measures of similarity are produced.

1.6.1 Hard Classification

The majority of conventional classifiers are hard, representing land cover as a series of discrete units, whereby each pixel is associated with a single land cover class. A standard hard classification may be performed using either of two approaches: supervised or unsupervised classification. Unsupervised classification permits the classifier to aggregate land cover classes into the most frequently occurring groups in a cluster analysis. The approach usually makes no use of *a priori* information on the classes of interest, and the pixels are grouped according to their spectral reflectance values. For a supervised classification, the analyst trains the classifier with the characteristics of known land cover classes. The pixel groups are therefore associated with specific land cover classes before classification. Among the most frequently used hard classification algorithms are the following decision rules:

(i) *Minimum distance*

The spectral distance, in multi-dimensional feature space, between the position of the data pixel and the location of the mean for each land cover class is calculated. Each pixel is assigned to the closest class, based on Euclidean distance:

$$D_{xyc} = \sqrt{\sum_{i=1}^{n} (u_{ci} - X_{xyi})^2} \qquad (1.1)$$

Where D_{xyc} is the Euclidean distance from pixel (x,y) to the mean of land cover class c. n represents the number of wavebands, i is a particular band, u_{ci} is the mean of land cover class c in waveband i, and X_{xyi} is the digital number value of pixel (x,y) in waveband i.

(ii) *Parallelepiped*

The highest and lowest limits of brightness values in each waveband, defined using training data, form the dimensions for each side of a parallelepiped. The unclassified pixels are projected into data space and those that fall within an area defined by the training data parallelepiped, are allocated to the appropriate land cover class.

(iii) *Maximum likelihood*

Perhaps the most widely used approach in the literature is the Gaussian maximum likelihood classifier (MLC), which assumes that the data have a multivariate normal distribution. The MLC is based on an evaluation of the variance-covariance matrix of the spectral responses of different land cover classes in classifying an unknown pixel. Each pixel is allocated to the land cover class with which it has the highest *a posteriori* probability of membership, and an assumption is made that the *a priori* probabilities are equal for all land cover classes. It is calculated by:

$$P(X_{xyi} \mid c) = \frac{1}{(2\pi)^{n/2} |v_c|^{1/2}} \exp\left[-\frac{1}{2} (X_{xyi} - u_{ci})^T V_c^{-1} (X_{xyi} - u_{ci}) \right] \qquad (1.2)$$

Where $P(X_{xyi}|c)$ is the *a posteriori* probability of pixel (*x,y*), in waveband *i*, being a member of class *c*, *n* is the number of wavebands, u_{ci} is the mean vector for class *c* in waveband *i*, and V_c is the variance-covariance matrix for class *c*. $(X_{xyi} - u_{ci})^T v_c^{-1} (X_{xyi} - u_{ci})$ represents the Mahalanobis distance between the pixel X_{xyi} and the centre of class *c* in feature space (Campbell, 1996).

More recent work has attempted to increase the accuracy of hard classifiers by combining the statistical approaches described above with neural network (see section 1.4.3.5) techniques (Giacinto *et al.*, 2000) or soft classification (see section 1.4.3) approaches (Melgani *et al.*, 2000). In addition, new classification techniques have been introduced, such as the Kohonen self-organizing feature map (Ji, 2000) and genetic algorithms (Bandyopadhyay and Pal, 2001).

1.6.2 Mixed Pixels

Multispectral classification of remotely sensed images has previously been dominated by 'hard' algorithms, which produce one class per pixel (Campbell, 1996). This corresponds to the partitioning of feature space into mutually exclusive decision regions, leaving no possibility to label uncertainty in the resulting map. These algorithms ignore the fact that many pixels in a remotely sensed image represent a spatial average of spectral signatures from two of more surface categories. This mixing of signatures occurs as a function of:

(i) *Frequency of land cover*

The physical continuum that exists in many cases between discrete category labels, combined with the spatially mixed nature of most natural land cover classes.

(ii) *Frequency of sampling*

The spatial integration within the pixel of land cover classes due to such factors as, the sensor spatial resolution, PSF and resampling for geometric rectification.

Distinguishing which source caused mixing within an image is only possible in controlled and limited situations (Schowengerdt, 1996).

1.6.3 Soft Classification

A 'hard' classification fails to recognize or represent the existence of classes and objects which grade into one another and class boundaries at sub-pixel scales. Such simplification can be seen as a waste of the available multi-spectral information, which could be interpreted more efficiently (Wang, 1990, Foody, 1995). Therefore, since an initial approach by Horowitz *et al.* (1971), many studies have used various techniques to attempt to 'unmix' the information from mixed pixels. These produce a set of maps, one for each land cover class. Each pixel has a membership value between 0 and 1 for every class, so that each land cover map contains a continuous range of quantitative memberships (Foody, 1992, Bastin, 1997). In general, two different types of soft classification techniques exist. The most commonly used methods estimate posterior probabilities of class membership using statistical pattern recognition methods, and correlate these with area proportions. However, as Lewis *et al.* (1999b) state, posterior probabilities are measures of statistical uncertainty, and there is no causal relationship with proportions of pixels containing the class, despite their correlation. Thus, as a consequence, posterior probabilities cannot represent optimum estimates of area (Manslow *et al.*, 1999). The second type of soft classification technique directly estimates class area proportions using regression models. Lewis *et al.* (1999b) demonstrate that posterior probabilities do not represent optimum area estimates, and direct area proportion models estimate true land cover proportions more accurately.

1.6.3.1 The Fuzzy c-means Classifier

The fuzzy *c*-means classifier measures iteratively the distance of each pixel from class centroids in feature space to assign membership values (Bezdek *et al.,* 1983). It is essentially an unsupervised soft clustering algorithm in which class distributions are derived in feature space. A distance measure must be chosen from a range available, including Euclidean (see Equation 1.1) and Mahalanobis (Equation 1.2). Basically, distances calculated using Equations 1.1 or 1.2, are converted into soft memberships which sum to one for each pixel. The use of the fuzzy *c*-means technique for land cover classification has been demonstrated by Fisher and Pathirana (1990), Foody (1992) and Atkinson *et al.* (1997).

1.6.3.2 Linear Spectral Mixture Models

Underlying the linear mixing model is the basic assumption that there is no significant amount of multiple scattering between different land cover types. Each photon that reaches the sensor has interacted with just one land cover type (Settle and Drake, 1993). Therefore, the response of each pixel in any spectral wavelength is taken as a linear combination of the

responses of each component assumed to be in the mixed target. Each image pixel contains information about the proportion and the spectral response of each land cover component within the ground resolution element (Holben and Shimabukuro, 1993). The basic mixture model may be represented as:

$$\mathbf{x} = \mathbf{M}\mathbf{f} + \mathbf{e} \tag{1.3}$$

Where:

$\mathbf{x}$ = Vector of the measured sensor responses for each pixel

$\mathbf{M}$ = The matrix of endmember spectra (*nxc*; *n*=spectral bands, *c*=land cover classes)

$\mathbf{f}$ = Vector of the proportions of mixture components for each pixel

$\mathbf{e}$ = Vector of the error terms for each spectral band

Subject to:

$$\sum f = 1 \text{ and } f \geq 0. \tag{1.4}$$

The constrained least squares method estimates the proportion of each land cover component within the pixel by minimizing the sum of squares of the errors. A linear constraint is added, since the sum of the proportions for any resolution element must be one, and the proportion values must be nonnegative (Garcia-Haro *et al.*, 1996, Fletcher, 1987). The linear spectral mixture model has been used widely to generate pixel fraction images of forest cover proportions (e.g. Cross *et al.*, 1991, Shimabukuro and Smith, 1991, Holben and Shimabukuro, 1993, Wu and Showengerdt, 1993, Gong *et al.*, 1994, Garcia-Haro *et al.*, 1996, Schowengerdt, 1997, Cochrane and Souza, 1998), urban land cover proportions (e.g. Foody and Cox, 1994, Schowengerdt, 1997, Chang and Ren, 2000) and crop area proportions (e.g. Quarmby *et al.*, 1992, Puyou-Lascassies *et al.*, 1994, Gebbinck and Schouten, 1996a, 1996b, 1997). Heintz and Chang (2001) also introduced a fully constrained method for hyperspectral imagery and Oleson *et al.* (1995) used the mixture model in reverse to obtain typical reflectance curves from AVHRR data.

1.6.3.3 Support Vector Machines

Linear spectral mixture models (LSMM) rely on the linear mixing assumption, and require a single pure pixel to be identified for each class. Spectral variation occurring between these pure pixels is modelled as linear mixing and generally constraints are imposed to ensure

sensible results outside the simplex of linear mixing. Support Vector Machines (SVM) (Vapnik, 1995) by contrast, are a range of robust classification and modelling algorithms currently attracting interest due to their good empirical performance and properties (Brown and Lewis, 1998). In the simplest of forms, a linear classifier is used which maximises the distance between two class clusters, and is therefore optimal. This produces a mixture region identical to that produced by the constrained least squares Linear Spectral Mixture Model Algorithm (Brown *et al.,* 1999). In addition to this, the technique identifies automatically the pure pixels that lie on the edge of the domain from a potentially much larger data set. This means that techniques for dealing with spectral confusion and non-linear mixing margins can be applied directly, due to the equivalence between LSMM and SVM models. The use of SVMs in remote sensing image classification is discussed in Brown *et al.* (1999) and Lewis and Gunn (1999).

1.6.3.4 k-Nearest Neighbour Classifier

The *k*-nearest neighbour (*k*nn) classification scheme assigns labels to unknown pixels according to the labels of neighbouring training vectors in feature space (Schowengerdt, 1997). The majority label of the *k* nearest neighbour training pixels is chosen as the assigned label for the unknown pixel. Selection from more than one point introduces a form of feature space smoothing, allowing the classification decision not to be affected by noisy outlier points. For sub-pixel area estimation, a kernel-based smoothing algorithm is used to estimate the proportions in test data, then real (soft nearest neighbour) data. The *k* number of neighbours modifies the size of this kernel (Lewis and Brown, 1998). Each exemplar, corresponding to a single pixel, consists of an input vector of digital counts (representing the spectral signal) and a target vector of class proportions. The *k*nn model then assigns an area proportion, μ_c, of class *c*, to a new pixel by, firstly, identifying the *k* examplar input vectors having the lowest Euclidean distance, *D*, from the new pixel's input vector and by, secondly, calculating a weighted average of the *k* exemplar class area proportions, such that:

$$y_c = \frac{\sum_{i=1}^{k} w_c^i \mu_c^i}{\sum_{i=1}^{k} w_c^i} \qquad \text{for} \qquad w_c^i = 1 - \frac{D_i}{\max_{j=1}^{k}[D_i] + 0.0001} \qquad (1.5)$$

where D_i is the Euclidean distance from the ith exemplar to the new pixel's input vector, and μ_c^i is the area proportion of the ith exemplar in class c. When k=1 the outputs are simply the proportions of the nearest target exemplar. When k>1, the $k$nn model outputs, $\mathbf{y}=\{y_c\}_{c=1}^{m}$, represent a prediction of the expected (mean) values of the target area proportions (Lewis and Brown, 1998). If k exemplars have the same spectral signal, then the outputs become,

$$y_c = \frac{\sum_{i=1}^{k} \mu_c^i}{\sum_{i=1}^{k} 1} = \frac{1}{k} \sum_{i=1}^{k} \mu_c^i \tag{1.6}$$

which is the average of the target exemplar proportions. If more than k exemplars have the same spectral signal, then the algorithm selects only k of them from the distribution and computes the average as equation 1.6 shows. If fewer than k exemplars have the same spectral signal, then the average is modified to include contributions from the next nearest exemplars in such a way as to generate a smooth prediction of the mixing proportions. Application of this area estimation model to remotely sensed image classification is shown in Lewis *et al.* (1999b).

1.6.3.5 Artificial Neural Networks

Artificial neural networks (ANNs) are computer programs designed to simulate human learning processes through the establishment and reinforcement of linkages between input and output data (Campbell, 1996). These linkages form the analogy with the human learning process within the brain, in that repeated associations between input and output in the training process reinforce linkages, which are then employed to link input and output in the absence of training data. There are many types of ANN, but the most commonly used in remote sensing is the multi-layer perceptron (MLP). These are generally composed of three elements. An input layer consists of source data, in this case, the multispectral data. The output layer consists of classes defined by the analyst. Between these two layers are one or more hidden layers. During a training phase, a MLP produces association between input and output data by establishing weights within these hidden layers (Bishop, 1995). Figure 1.3 displays the general architecture of a MLP.

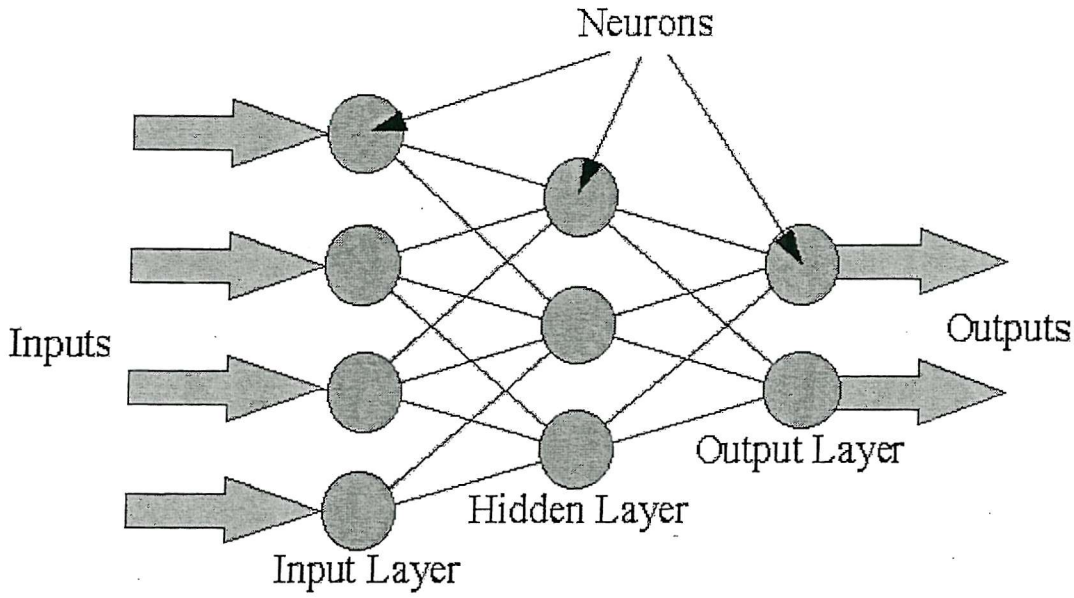


Figure 1.3 Multi-layer perceptron architecture

Repeated associations between classes and digital values strengthen the weights in the hidden layers that permit the MLP to assign correct labels when given spectral values in the absence of training values. Within each processing node there is summation and transformation (Patterson, 1996). At each hidden layer node, j, the following operation is performed on the input pattern, p, producing the output, h (w = weight):

$$\text{Hidden layer:} \quad s_j = \sum w_{ji} p_i \quad ; \quad h_j = f(s_j) \tag{1.7}$$

This is directed to each output layer node, k, where the output, o, is calculated:

$$\text{Output layer:} \quad s_k = \sum w_{jj} h_j \quad ; \quad o_k = f(s_k) \tag{1.8}$$

The most widely used of the transformation functions is the sigmoid function:

$$f(s) = 1/1 + e^{-s} \tag{1.9}$$

MLPs can be trained by back propagation. This can be considered as a retrospective examination of the links between input and output data, in which differences between expected and actual results can be used to adjust weights (Schowengerdt, 1997). This process

establishes quantitative relationships (transfer functions) between input and output layers that assign weights to emphasise effective links between the two layers.

The rapid uptake of neural approaches in remote sensing is due mainly to their ability to:

(i) Perform more accurately than other techniques, particularly when the feature space is complex and the source data has different statistical distributions (e.g., Benediktsson *et al.,* 1990).

(ii) Operate without making statistical assumptions about the input data (Campbell, 1996).

(iii) Incorporate *a priori* knowledge and physical constraints into the analysis (e.g., Foody *et al.*, 1995).

(iv) Incorporate different types of data into the analysis (e.g., Benediktsson *et al.,* 1993).

ANNs can be used within the investigation of soft sets (e.g., Fisher and Pathirana, 1990, Gopal and Woodcock, 1994). Methods of soft classification are compared with ANNs in Atkinson *et al.* (1997) and Foody *et al.* (1997). Independent networks can be applied to model the area proportions of each land cover class, so that the optimal area estimation model for each class may be applied (Lewis *et al.*, 1999b). The use of ANNs in remote sensing image classification is demonstrated by Paola and Schowengerdt (1995), Gong *et al.* (1996), Foody (1996b), Foody and Arora (1997), Atkinson and Tatnall (1997), Foody (1997), Foschi and Smith (1997), Warner and Shank (1997), Embashi (1998), Lewis *et al.* (1999b) and Lin *et al.* (2000).

1.6.4 Accuracy Assessment

Accuracy assessment can be defined as the task of comparing two maps, one based on the analysis of remotely sensed data, and another based on a different source of information. The second map is designated the reference map, assumed to be accurate, that forms the standard for comparison (Campbell, 1996). Without accuracy assessment, the user does not know the utility of a classification (Congalton, 1991, Nishii and Tanaka, 1999). Accuracy assessment can involve the simple visual comparison of two data sets, or the production of more precise empirical results, involving the comparison of a statistical sample of points throughout the study area (Gong and Howarth, 1990). Such comparison can be made on a hard or soft basis. The latter involves the comparison of the soft probabilities, or predicted area proportions of

each pixel with the land cover composition of the corresponding area in the reference data (Gopal and Woodcock, 1994, Atkinson 1999a, 1999b, Lewis and Brown, 1999). Where samples are required, each sample point requires processing and is, therefore, expensive. This means that the sample size should be kept to a tolerable level, but still be large enough to provide a statistically valid representation of the population. The accuracy of the reference data must also receive consideration. Inaccuracy can be caused by several factors (Campbell, 1996):

(i) *Measurement Error*: Features within the reference data are incorrectly classified.

(ii) *Location Error*: Image and reference data are mis-registered.

(iii) *Scale Error*: Differences in scales of measurement mean no direct comparison between reference and image features can be made.

(iv) *Timing Errors*: Reference data are acquired at a different time to image data, and land cover change has occurred in the meantime.

Errors in the reference data may be compounded at the accuracy assessment stage, thus, reducing the utility of the classification (Congalton, 1991).

Accuracy assessment results may be provided in a variety of forms. In most cases, a confusion or error matrix is used. This indicates the number of sample points assigned to each class relative to the actual land cover class, as verified using reference data. The clarity and effectiveness of such matrices in results presentation, means that they have become widely used in both hard (Congalton, 1991, Stehman, 1997) and soft classification (Gopal and Woodcock, 1994, Foody *et al.*, 1995, Lewis and Brown, 1999).

Various statistical measures can be used to express classification accuracy, each calculated from the confusion matrix:

(i) *Overall Classification Accuracy*

This is the simplest measure of accuracy, and indicates the overall proportion of area, pixels or polygons classified correctly (Campbell, 1996).

(ii) *Producer's Accuracy*

This represents, for each class, the number of pixels classified correctly, relative to the total number of reference pixels (Campbell, 1996, Zhu *et al.*, 2000).

(iii) *User's Accuracy*

This represents, for each class, the number of pixels classified correctly, relative to the total number of classified pixels (Congalton, 1991, Zhu *et al.*, 2000).

Foody (1996a) suggests that such accuracy measures are not necessarily appropriate for models that predict sub-pixel class areas. In this context, a user is typically interested in knowing that a particular class covers a certain area within a scene, and an indication of the error associated with the estimation of that area is usually sufficient (Dymond, 1992). Information about the distribution of errors can be obtained directly from the data, rather than a confusion matrix, using techniques such as those described by Foody (1996a):

(iv) *Entropy*

A different approach has been based on measures of entropy (Finn, 1993, Maselli *et al.*, 1994, Foody, 1995). Entropy is a measure of uncertainty and information formulated in terms of probability theory, which expresses the relative support associated with mutually exclusive alternative classes. Entropy is maximised when the probability of class membership is partitioned evenly between all defined classes in the classification, and minimised when it is associated entirely with one class.

(v) *Closeness*

Foody (1996a) suggests a measure, termed closeness, related to the Euclidean distance between the representation of the land cover from the classification, and reference data. This measures the separation of the two data sets, based on the relative proportion of each class in the pixel. It is calculated as:

$$S = \sum_{i=1}^{c} ({}^{1}e_i - {}^{2}e_i)^2 / c \qquad (1.10)$$

where, ${}^{1}e_i$ is the proportion of class i in a pixel from reference data, ${}^{2}e_i$ is the measure of the strength of membership to class i, taken to represent the proportion of the class in the pixel from the soft classification, and c is the total number of classes. This measure is used in subsequent chapters.

Whilst these alternatives allow the comparison of traditional classification algorithms to sub-pixel area estimation models, they are not as familiar to the users of the classified map as the measures derived from the confusion matrix, or the confusion matrix itself. Lewis and Brown (1999) and Townsend (2000) describe error matrix approaches, each derived from the generalised confusion matrix, which allow the accuracy of maps generated using area estimation models to be assessed quantitatively, and compared to the accuracies obtained from traditional classification techniques:

(vi) *Area Error Proportion*

One of the simplest measures of agreement between a set of known proportions **y**, and a set of estimated proportions **a**, is the area error proportion (*AEP*) per class,

$$AEP_q = \frac{\sum_{q=1}^{n}(y_q - a_q)}{\sum_{q=1}^{n} a_q}, \tag{1.11}$$

where, q is the class and n is the total number of pixels. This statistic informs about bias in the prediction image.

(vii) *Correlation Coefficient*

The correlation coefficient, r, represents an alternative measure of the amount of association between a target and estimated set of proportions,

$$r_q = \frac{c_{yq.aq}}{s_{yq}.s_{aq}} \quad , \qquad c_{yq.aq} = \frac{\sum_{q=1}^{n}(\bar{y}_q - y_q)\cdot(\bar{a}_q - a_q)}{n-1}, \tag{1.12}$$

where, $c_{yq.aq}$ is the covariance between y and a for class q and s_{yq} and s_{aq} are the standard deviations of y and a for class q. This statistic informs about precision in the prediction image.

(viii) *Root Mean Square Error*

The root mean square error (*RMSE*) per class,

$$RMSE_p = \sqrt{\frac{\sum_{p=1}^{n}(y_p - a_p)^2}{n}}, \qquad (1.13)$$

informs about the accuracy of the prediction (bias and precision).

Accuracy measures (vi), (vii) and (viii), are used thoughout subsequent chapters. More recently, Atkinson (1999a and 1999b) suggested comparison of a set of known class proportions with estimated proportions using a root mean square error, standardised by image variance, and Matsakis *et al.* (2000) put forward assessment techniques specifically for soft partitions in satellite imagery. Vieira and Mather (1999) also describe methods of assessment of the performance of a classifier that explicitly include the spatial pattern of classification errors, while DeFries and Chan (2000) introduce criteria that account for computational resources and noise in training data.

1.7 Previous Image Mapping at the Sub-Pixel Scale

Previous work on mapping or reconstructing images at the sub-pixel scale has taken the form of three differing approaches:

(i) *Image fusion and sharpening*
(ii) *Image reconstruction and restoration*
(iii) *Super-resolution*

1.7.1 Image Fusion and Sharpening

The basis for image fusion and sharpening stems from the fact that there will always be some trade-off within the field of remote sensing between spatial and spectral resolution. Images with high spatial resolution can locate objects with high accuracy, whereas images with high spectral resolution can be used to identify land covers. With different sensors observing over the same area, it is useful to merge the data into a hybrid product containing the useful information from both platforms. Such a hybrid image can be used to create detailed 'sharpened' images that map the abundance of various materials within a scene. Therefore, image fusion is the technique whereby images of different spatial and spectral resolutions are merged to create a high spatial resolution multispectral combination (Gross and Schott, 1998).

Generally, two approaches to fusion and sharpening exist. The first uses a two-step approach whereby firstly, the coarse spatial resolution multispectral image is used to identify the land covers in the scene via soft classification. Second, the proportion maps are combined with a fine spatial resolution panchromatic image of the same area, which serves to constrain the proportions to produce a set of 'sharpened' fine spatial resolution maps. This has been carried out for synthetic images (Gross and Schott, 1998), agricultural fields (Li *et al.,* 1998) and for a lake (Foody, 1998). Foody (1998) used a finer spatial resolution image and a simple regression based approach to sharpen the output of a soft classification of a coarser spatial resolution image, producing a sub-pixel land cover map. The results produced a visually improved representation of the lake being studied, and this was further improved by fitting class membership contours, lessening the blocky nature of the representation. However, the areal extent (soft proportions) of the lake was not maintained per-pixel, and generally, obtaining two coincident images of differing spatial resolution is difficult.

The second approach to fusion and sharpening should conceptually produce identical results to the technique described above. Image fusion is undertaken to produce a fine spatial resolution multispectral image that is then unmixed into fine spatial resolution land cover maps. There is little published work of this technique, although Robinson *et al.* (2000) attempted it with synthetic imagery and Aplin *et al.* (1999) used a similar approach on urban and agricultural imagery. Aplin (1999 and 2000) made use of sub-pixel scale vector boundary information, along with fine spatial resolution satellite sensor imagery to map land cover. By utilising Ordnance Survey Land-line vector data overlaid onto CASI imagery, and undertaking per-field rather than the traditional per-pixel land cover classification, mapping at a sub-pixel scale was demonstrated. Assessments suggested that the per-field classification technique was generally more accurate than the per-pixel classification (Aplin *et al.*, 1999). However, in most cases around the world, availability of accurate vector data sets to apply the approach will be rare, and the technique is limited to features large enough to appear on such data sets.

1.7.2 Image Reconstruction and Restoration

Digital image reconstruction refers to the process of recovering a continuous image from its samples and has fundamental importance in digital image processing, particularly in applications requiring image resampling. Its role in such applications is to provide a spatial continuum of image values from discrete pixels so that the input image may be resampled at

any arbitrary position, even those at which no data were originally supplied (Boult and Wolberg, 1993).

Whereas, reconstruction simply derives a continuous image from its samples, restoration attempts to go one step further. It assumes that the underlying image has undergone some degradation before sampling, and so attempts to estimate the original continuous image from its corrupted samples. Restoration techniques must, therefore, model the degradation and invert its effects on the observed image samples.

Both reconstruction and restoration techniques have yet to be applied fully to remotely sensed imagery due to the computational complexity involved in the calculations undertaken. However, promising results on binary images have been produced for synthetic images (Jennison and Jubb, 1988, Jubb and Jennison, 1991, Boult and Wolberg, 1993), electron micrograph images (Hitchcock and Glasbey, 1997) and astronomical images (Weir and Djorgovski, 1991). Gavin and Jennison (1997) adopted a Bayesian approach, incorporating prior information on the true image into a stochastic model that attached a higher probability to images with shorter total edge length. The model produced accurate results for synthetic and microscope images, but the complex multi-stage operation meant it was slow and only applicable to small binary images containing features larger than a pixel.

1.7.3 Super-Resolution

Super-resolution techniques are similar to the reconstruction and restoration approaches in that they attempt to recover image values at points where no data were originally supplied. However, where the reconstruction and restoration techniques attempt to recover a continuous image, the super-resolution approaches merely attempt to increase the spatial resolution of an image to a desired level. This solves the problem of the computational complexity associated with the reconstruction and restoration techniques, making the super-resolution techniques applicable to remotely sensed data. Various differing approaches to super-resolution mapping have been attempted.

Schneider (1993) introduced a knowledge-based analysis technique for the automatic localization of field boundaries with sub-pixel accuracy. The technique relies on knowledge of straight boundary features within Landsat TM scenes, and serves as a pre-processing step prior to automatic pixel-by-pixel land cover classification. With knowledge of pure pixel values either side of a boundary, a model can be defined for each 3 by 3 block of pixels. The

model uses parameters such as pure pixel values, boundary angle, and distance of boundary from the centre pixel. Using least squares adjustment, the most appropriate parameters are chosen for location of a sub-pixel boundary, dividing mixed pixels into their respective pure components. Improvements on this technique were described by Steinwendner and Schneider (1997), who used a neural network to speed up processing, and Steinwendner and Schneider (1998), Steinwendner *et al.* (1998) and Steinwendner (1999) suggested algorithmic improvements, along with the addition of a vector segmentation step. The technique represents a successful, automated and simplistic pre-processing step for increasing the spatial resolution of satellite imagery. However, its application is limited to imagery containing large features with straight boundaries at a certain spatial resolution, and the models used still have problems resolving image pixels containing more than two classes (Schneider, 1999).

Flack *et al.* (1994) concentrated on sub-pixel mapping at the borders of agricultural fields, where pixels of mixed class composition occur. Edge detection and segmentation techniques were used to identify field boundaries and the Hough transform (Leavers, 1993) was applied to identify the straight, sub-pixel boundaries. These vector boundaries were superimposed on a sub-sampled version of the image, and the mixed pixels were reassigned each side of the boundaries. By altering the image sub-sampling, the degree to which the spatial resolution was increased could be controlled. No validation was carried out, and the work was not followed up, and so the success of the technique remains unclear.

The two techniques described so far are based on direct processing of the raw imagery. Other work has focussed on using a pre-processing step where soft classification of the imagery is undertaken, and an attempt to map the location of class components within the pixels is made.

Atkinson (1997) used an assumption of spatial dependence within and between pixels, to map the location within each pixel of the proportions output from a soft classification. The assumption proved to be valid for re-creating the layout and areal coverage of the land cover, and the algorithm produced a certain degree of success for semi-natural land cover. However, the technique compared sub-pixels to pixels and so the complex mixing in the data caused the simple technique to suffer from problems.

Verhoeye and De Wulf (2000) used similar assumptions as Atkinson (1997), but formulated as a linear optimization problem. The algorithm was applied to synthetic imagery and a SPOT image of Sahelian wetlands. Compared to traditional hard classification techniques, results showed a certain degree of success, but problems were noted again due to the fact that sub-pixels were compared to pixels, rather than other sub-pixels.

1.8 Summary and Research Objectives

Information on land cover is required to aid our management and understanding of the environment. Accurate mapping of land cover, as well as identification and extraction of target land cover features, are vital procedures for this. Accurate information on land cover is required for both scientific research (e.g. climate change modelling, flood prediction), and management (e.g. city planning, disaster mitigation), and remote sensing has the potential to provide this information.

Remote sensing provides perhaps the best method for obtaining geographically and temporally detailed estimates of land cover and its changes over large areas. Imagery derived from aircraft and satellite mounted sensors have attributes that make the approach suitable for land cover mapping, and the short orbit times of many sensors (of the order of 3 hours to 16 days) means a temporal sequence of images can be acquired, aiding the monitoring of specific features. Also, the large spatial coverage that can be obtained with imagery from satellite sensors (e.g. of the order of 1×10^6 km^2) provides an advantage over costly and time-consuming ground survey. Many remote sensors measure ground reflectance at a fine spectral resolution and for most mapping and target identification applications, this provides sufficient information to map accurately the land cover in question. Finally, remote sensing has the potential to provide land cover information at a variety of scales (e.g. from <1 m to 1 km). These aspects make land cover mapping from remotely sensed imagery an attractive possibility. However, there exist several practical limitations.

Perhaps the biggest drawback of land cover mapping and monitoring relates to the large ground area of each pixel, relative to the size of typical land cover parcels. In other words, the intrinsic scale of spatial variation in land cover may be finer than the scale of sampling imposed by the satellite sensor. Therefore, the pixels in such imagery may be of mixed land cover class composition. Within remotely sensed images, a significant proportion of pixels is often of mixed land cover class composition and their presence can adversely affect the performance of image mapping operations.

Traditional 'hard' techniques for classification of remotely sensed images, which are still used today, allocate each pixel to a single land cover class. This limitation within the technique inevitably causes information loss, in turn leading to unsatisfactory classification accuracy and poor information extraction. Such problems led to the development of 'soft' classification techniques.

Soft classification techniques started to come into use from the mid-1980s. They have become more popular in recent years, but only in the last couple of years have they been recognised and included in popular image processing computer packages. Soft classification allows image pixels partial and multiple land cover class membership, producing, in most cases, increases in the accuracy of land cover information extraction. Soft classification approaches include the fuzzy *c*-means classifier, linear spectral mixture models, support vector machines, soft *k*-nearest neighbour classifiers and artificial neural networks. The output of such approaches generally takes the form of a set of proportion images, each displaying the proportion of a certain class within each pixel. Results usually produce a more appropriate and informative representation of land cover than that produced using a hard, one class per-pixel classification. However, while the class composition of every pixel is estimated, the spatial distribution of these class components within the pixel remains unknown.

As the 1980s and early 1990s saw a gradual development and shift from hard classification to soft classification, so more recently are the beginnings of a further shift from soft classification to sub-pixel approaches being seen. Recent research by Atkinson (1997), Steinwendner and Schneider (1998) and Verhoeye and De Wulf (2001) has shown that it is possible to take land cover classification from remotely sensed imagery a step further and map land cover at the sub-pixel scale (super-resolution). This thesis aims to overcome the problems suffered by each of these techniques and design a robust, efficient and simple technique for super-resolution mapping of land cover from remotely sensed imagery. Such a technique would represent a great step forward in the field of land cover map production.

Chapter 2: Optimisation and the Hopfield Neural Network

2.1 Introduction

This chapter provides an overview of the Hopfield neural network as an optimisation tool, and its design for the task of super-resolution land cover mapping. Initially, the topic of optimisation is introduced, followed by a detailed description of the Hopfield neural network and previous applications. Finally, the specific network architecture and design for super-resolution land cover mapping from remotely sensed imagery is introduced.

2.2 Optimisation

Optimisation is central to many problems within computer science and engineering. Optimisation involves using efficient techniques to find minimum or maximum values of a function of many independent variables. This function is usually called the energy, error or cost function and represents a quantitative measure of the 'goodness' of some complex system (Kirkpatrick *et al.*, 1983). Optimisation techniques have been applied for problems in such areas as minimising the error function in an artificial neural network (e.g. Beale and Jackson, 1990, Atkinson and Tatnall, 1997, Foody, 1997). Minimisation of energy functions within image segmentation and feature extraction algorithms also require optimisation approaches (e.g. Zhu and Yuille, 1996, Steinwendner *et al.*, 1998).

Several practical optimisation algorithms exist. These range from simple methods such as gradient descent (Fletcher, 1987), up to the more complex methods like simulated annealing (Kirkpatrick *et al.*, 1983, Aarts and Korst, 1989).

2.3 Hopfield Neural Networks

The Hopfield neural network is a fully connected recurrent network and can be implemented physically by interconnecting a set of resistors and amplifiers with symmetrical outputs and external bias current sources (figure 2.1).

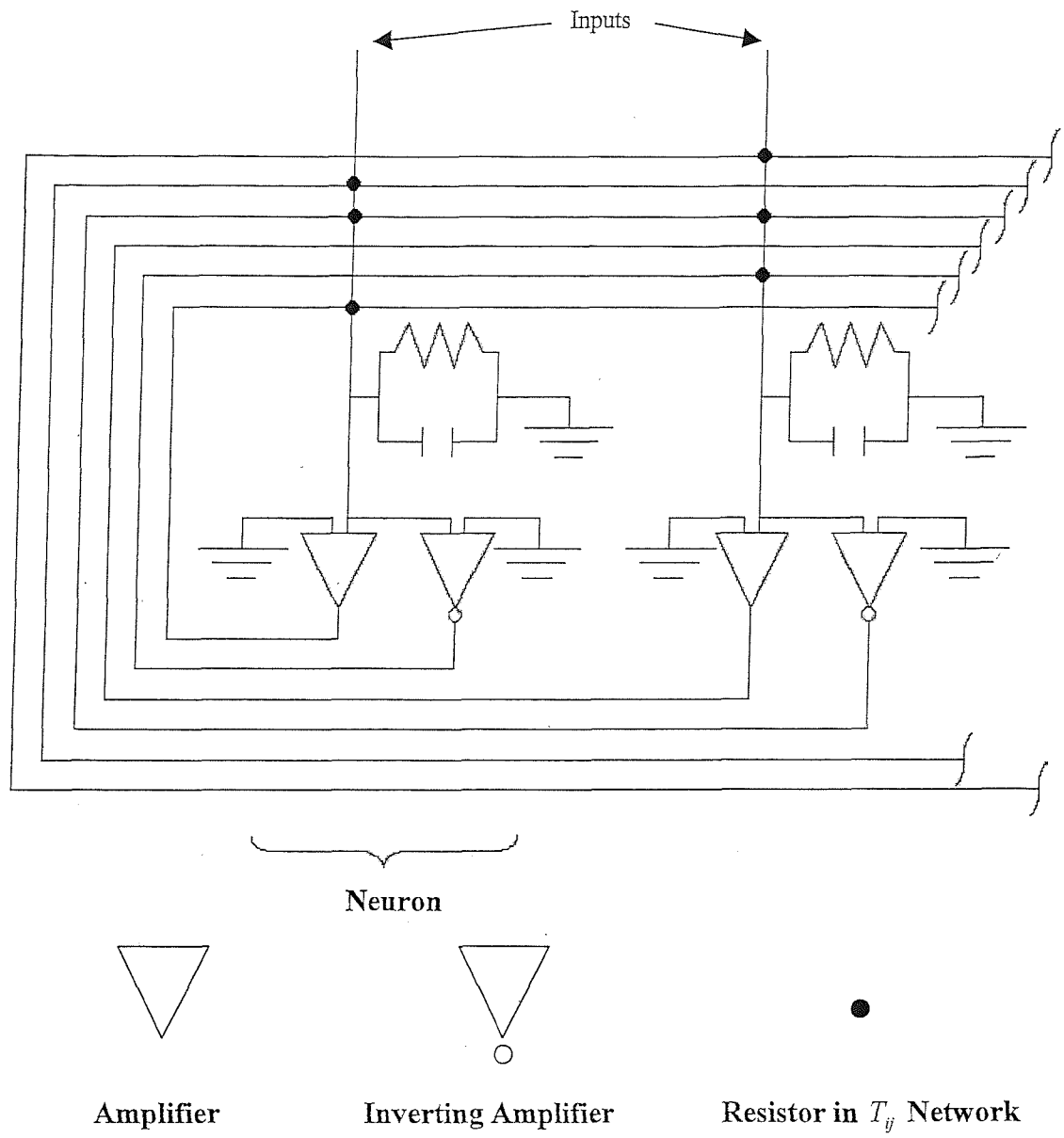


Figure 2.1. Hopfield neural network as an analogue circuit. The black circles at the intersections represent resistive connections (T_{ij}'s) between outputs and inputs. Connections between inverted outputs and inputs represent negative connections. Taken from Hopfield and Tank (1985).

The mathematical model describing the behaviour of such an array of electronic components can be derived from Kirchoff's current law (Hopfield and Tank, 1985, Cichocki and Unbehauen, 1993),

$$C_i \frac{du_i}{dt} = -\frac{u_i}{R_i} + \sum_{j=1}^{N} T_{ij} v_j + I_i \; ; i = 1 \ldots N, \tag{2.1}$$

where,

$$\frac{1}{R_i} = \frac{1}{R_{i0}} + \sum_{j=1}^{N}\left(\frac{1}{R_{ij}^{+}} + \frac{1}{R_{ij}^{-}}\right). \tag{2.2}$$

R_{ij} is the resistance between the output of amplifier j and input of amplifier i, N is the number of amplifiers, $C_i > 0$ is the capacitance of amplifier i, u_i is the internal voltage of amplifier i, and I_i is the external bias on amplifier i. T_{ij} is the conductance from amplifier j to amplifier i, where,

$$T_{ij} = \frac{1}{R_{ij}^{+}} + \frac{1}{R_{ij}^{-}}, \tag{2.3}$$

and $v_i = g_i(u_i)$ is the output voltage of amplifier i. $g_i(u_i)$ is the nonlinear activation function, defined as,

$$g_i(u_i) = \frac{1}{2}(1 + \tanh \lambda\, u_i), \tag{2.4}$$

where, λ determines the steepness of the function.

Hopfield (1984) shows how equation 2.1 can be written in a neural context for ease of interpretation, where, in this case the non-linear amplifiers correspond to neurons,

$$\tau_i \frac{du_i}{dt} = -\alpha_i u_i + \sum_{j=1}^{N} T_{ij} v_j + I_i \; ; i = 1 \ldots N, \tag{2.5}$$

where, $\tau_i = C_i$ is the time constant for neuron i, u_i is the total weighted input at neuron i, $\alpha_i = \frac{1}{R_i}$, N is the number of neurons in the network, $v_i = g_i(u_i)$ is the neural output which is a function of the input u_i, I_i is the external bias on neuron i and T_{ij} is the weight from neuron j to neuron i, which corresponds to the conductance in equation 2.1.

The set of differential equations described so far defines the time evolution of the network. Thus, from a set of initial neuron outputs, the state, v, of the network varies with time until

convergence to a stable state, where neuron output stops varying with time. Weights and biases determine the neural outputs at this stable state.

Hopfield (1984) showed that using symmetric weights with no self-connection, i.e. $T_{ji} = T_{ij}$ and $T_{ii} = 0$ is sufficient to guarantee convergence to such a stable state. Therefore, independent of its initial status, a Hopfield neural network will always reach an equilibrium state where no output variation occurs and it was also demonstrated that for high values of the gain, λ_i, the activation function $g_i(u_i)$ (equation 2.4) approaches a step function. The stable states of the network consequently correspond to the local minima of the following 'energy function' (Cichocki and Unbehauen, 1993),

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}T_{ij}v_i v_j - \sum_{i=1}^{N}v_i I_i + \sum_{i=1}^{N}\frac{\alpha_i}{\lambda}\int_0^{v_i} g_i^{-1}(v)dv, \tag{2.6}$$

where, E is the energy calculated over the whole network. For neurons where $\alpha \approx 0$ and with high values of λ, the last term becomes small and can be neglected, such that,

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}T_{ij}v_i v_j - \sum_{i=1}^{N}v_i I_i . \tag{2.7}$$

From equations 2.5 and 2.7, the equation describing the dynamics, i.e. the rate of change of neuron input of the Hopfield network can be written as,

$$\frac{du_i}{dt} = -\frac{\delta E}{\delta v_i}, \tag{2.8}$$

or

$$\frac{du_i}{dt} = -\sum_{j=1}^{N}T_{ij}v_j + I_i . \tag{2.9}$$

The Hopfield network can therefore be used for energy minimisation problems if the weights and biases are arranged such that they describe an energy function, with the minimum of energy occurring at the stable state of the network (Hopfield and Tank, 1985). By specifying different values for the weights and biases, any hypothetical energy minimisation problem can be simulated.

2.4 Problem solving using a Hopfield Neural Network

The system of differential equations given in section 2.3 completely describes the dynamics of the network and, therefore, the time evolution of each neuron composing it. The dynamics of the Hopfield neural network can be simulated on a digital computer using these equations. In addition, by specifying different values for the weights and biases, any hypothetical network can be simulated. Therefore, an energy function can be constructed from a specific set of weights and biases, and the corresponding Hopfield neural network used to find a minimum of that function.

Hopfield (1984) demonstrated the first application of a Hopfield neural network, by using it as a content-addressable memory (CAM), employing his own energy function to be minimised. CAM represents a departure from classical memory devices where items are retrieved via numbers or addresses. Within CAM, items can be retrieved from partial information about them. This was achieved with a Hopfield network by choosing a set of weights and biases in order to obtain stable network states to represent stored items. This means that a network initialised with a state containing some information about a stored item is likely to converge to the stable state representing that item. Further work by Storkey and Valabregue (2000) introduced a new learning rule specifically to improve CAM results.

Many real world problems can be formulated as the minimisation of an energy function, and this is central to the design of a Hopfield neural network formulated as an optimisation tool. The energy function used must represent the problem correctly, and reach a minimum at the solution of the problem. Once this function is designed, the weights and biases can be set, and the network is built around these.

After initialisation, the network is allowed to converge to a stable state, which is interpreted as a solution to the given problem. Most real world problems contain built-in constraints in addition to a goal, which must be considered. These constraints form a cost added to the objective within the energy function, which can then be defined as:

$$Energy = Goal + Constraints \tag{2.10}$$

If the energy function is coded in this particular way, the constraints become part of the minimisation process. This means the constraints do not need to be treated separately, just weighted by their importance to the problem. The Hopfield network process then minimises the value of the goal, while attempting to satisfy as many constraints as possible. The elegance of the Hopfield network as an optimisation tool is that constraints are coded into the energy function. This means such constraints are therefore satisfied at the same time as the goal is minimised.

Hopfield and Tank (1985) also used the network's energy minimization capability, and demonstrated solutions to complex combinatorial problems such as the travelling salesman problem. The travelling salesman problem is a NP-complete (non-deterministic polynomial time complete) problem, and the time required to solve it grows exponentially with its size. A travelling salesman must minimise the distance covered to travel between a number of cities, but must visit each city only once and cannot visit two cities at the same time. The problem is therefore to find the shortest path between cities. Hopfield and Tank formulated the problem as the minimisation of an energy function. The goal was to minimise the total distance travelled by the salesman, while adding constraints to the energy function ensured the validity of the tour. These constraints made sure that each row of a matrix representing a valid tour contained just one activated neuron (representing a visited city). Similarly, since only one city can be visited at a given time, a constraint ensured there was only one activated neuron per matrix column. Finally, the total number of activated neurons in the matrix was constrained to equal the number of cities. For a ten-city problem, Hopfield and Tank achieved convergence to a valid tour over 80% of the time. Later, work by Aiyer *et al.* (1990), Abe (1989, 1993) and Martin-Valdivia *et al.* (2000) has reduced the presence of local minima in the energy functions where constraints were not satisfied by setting various constants in a certain arrangement. This guaranteed convergence to a valid solution. Further work by Gee *et al.* (1993) outlined how effective problem mapping and modified network dynamics could improve solutions. The use of the Hopfield network means that all cities are being considered at the same time, representing the pursuit of many hypotheses in parallel. This kind of architecture distributes processing among many elements, providing robustness because of the large number of processing nodes, each responsible for a small part of the task (Côté, 1996).

Other applications include computer vision matching problems between sets of stereo or sequential images (Nasrabadi and Choo, 1992, Chang *et al.*, 1993, Li *et al.*, 1999) and colour

image segmentation (Campadelli *et al.*, 1997). Nichani (1994) used a Hopfield network to solve a stereo correspondence problem, and obtained good solution quality in addition to a fast rate of convergence. In such applications, the architecture of the network is designed as a set of rows and columns. Each row represents a feature identified in one image, and each column a feature identified in a second image. The neurons represent possible matches between features and an energy function is built, defining the problem. The network is used to find a minimum of this function, which represents a set of matches, linking similar features.

2.4.1 The Hopfield Neural Network in Remote Sensing

The Hopfield network has been used within the field of remote sensing for ice mapping, cloud motion and ocean current tracking (Côté and Tatnall, 1994, 1995, Côté, 1996, Côté and Tatnall, 1997, Lewis, 1998). For all such applications, similar matching techniques to those described above were applied. The above applications demonstrate the utility of the Hopfield network for feature tracking, the basic principle of which is to match common features in a sequence of images. The success of the method in such applications suggests it has potential to map other features, such as ocean colour or ice flows. In addition, more general feature matching problems within remote sensing could be tackled using a Hopfield network (Forte and Jones, 1999), as well as applications such as recognition or classification (Raghu and Yegnanarayana, 1996).

2.5 Super-Resolution land cover mapping using a Hopfield Neural Network

So far, this chapter has described in detail the workings and previous applications of the Hopfield neural network. This section focuses on how the network has been redesigned and adapted to tackle the research objectives set out in chapter 1. The Hopfield neural network was chosen for the task of super-resolution land cover mapping for two main reasons. Firstly, the Hopfield neural network allows for constraint based minimisation that is guaranteed to reach a minimum. Secondly, the network architecture could be adapted simply to represent an image, making neuron spatial positioning of importance. These aspects are described in detail in the following sections.

Mapping the spatial distribution of the land cover class components within each pixel of an image was formulated as a constraint satisfaction problem and an optimal solution to this problem was determined by the minimum of an energy function coded into a Hopfield neural network. The network architecture was arranged to represent a finer spatial resolution image,

and constraints within the energy function determined the spatial layout of binary neuron activations within this arrangement. The Hopfield neural network was used to find the minimum of this energy function, which corresponded to a bipolar map of class components within each pixel. This is outlined in detail in the following sections. The input data for this research took the form of class proportion images, derived either from degradation of verification data, or soft classification of satellite sensor imagery.

2.5.1 Network Architecture

In many papers on the use of Hopfield neural networks for optimisation, the spatial relations between neurons are considered irrelevant. However, for this research, the nature of the problem and the proposed solution requires the network neurons to be considered as being arranged in a regular grid, with positioning within this grid being of significance to the network design for this task (figure 2.2). Therefore, neurons will be referred to by co-ordinate notation, for example, neuron (i,j) refers to a neuron in row i and column j of the grid, and has an input voltage of u_{ij} and an output voltage of v_{ij}. The *zoom factor*, z, determined the increase in spatial resolution from the original satellite sensor image to the new fine spatial resolution image and after convergence to a stable state, the neurons represented a bipolar classification of the land cover target *at the finer spatial resolution.* Figure 2.2 shows the notation used in this paper, and how co-ordinates are transformed linearly from the image space to the network neuron space, for example, the pixel (x,y) in the satellite image is represented by $z \times z$ neurons centred at co-ordinates $[xz + int(z/2), yz + int(z/2)]$, where *int* is the integer value.

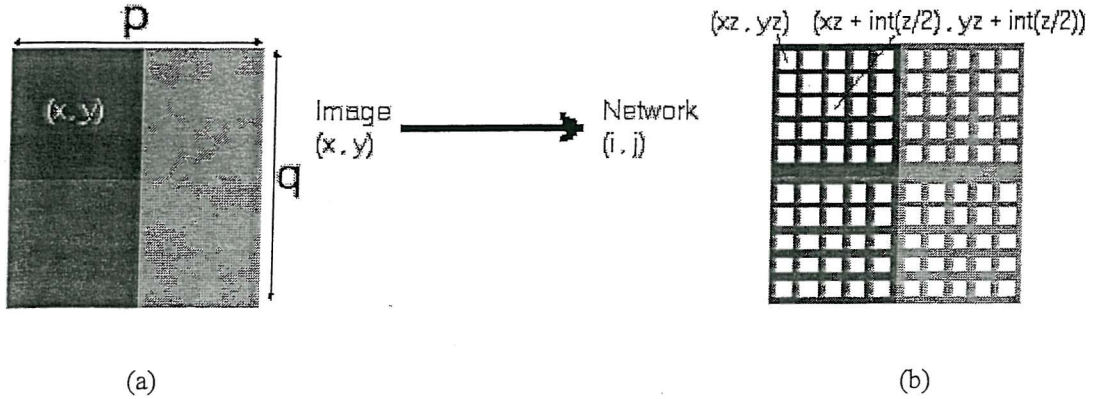


Figure 2.2. (a) 2x2 pixel image, p and q represent the image dimensions, x and y represent the image pixel co-ordinates; (b) Representation of the Hopfield network for the image in (a), i and j represent the neuron co-ordinates (int = integer value).

2.5.2 Network Initialisation

Each neuron was initialised with a starting value, u_{init}, and two strategies for initialising the network were adopted:

(i) Each set of neurons representing a pixel in the coarse spatial resolution image was identified and a proportion of this set was given randomly an output of u_{init}=0.55. This proportion was equal to the actual area proportion estimate of the class within the image pixel and the remaining neurons of the set were given an output of u_{init}=0.45. The values of 0.55 and 0.45 were chosen as the initial 'on' and 'off' outputs to speed up processing time and avoid unnecessary bias towards certain energy minimisation paths. In Hopfield and Tank (1985) and many other papers related to the use of the Hopfield network for solving the travelling salesman problem, neurons are initialised with a random value close to the central state value (0.5). This choice is justified by the fact that no initial preference should be given to any path. The small difference between the two values also enables the network to 'push' neuron outputs to 1 or 0 to represent a bipolar classification faster than if, for example, a neuron was given initially an output of 0 and had to be 'pushed' to 1 to produce an optimal solution.

(ii) The completely random initialisation of neuron outputs within the range u_{init}=[0.45, 0.55]. This allowed performance comparison with the class proportion-defined initialisation and did not introduce any possible unnecessary bias into the result, which may have occurred using (i) should estimated class proportions be inaccurate.

2.5.3 Implementation

When implemented on a computer, sets of biases and weights do not need to be determined, as the network is simulated via its equation of motion (equation 2.9) using the Euler method,

$$u_{ij}(t+dt) = u_{ij}(t) + \frac{du_{ij}(t)}{dt}dt, \tag{2.11}$$

where, dt is the time step of the iterative method and the function $\frac{du_{ij}(t)}{dt}$ is measured using $\frac{dE_{ij}}{dv}$. Equation 2.8 shows the correspondence between the two functions, and $\frac{dE_{ij}}{dv}$ was determined using the goals and constraints of the super-resolution target identification task.

Equation 2.11 was run until $\sum_{i,j}\left(u_{ij}(t+dt)-u_{ij}(t)\right)\leq du_c$, where, du_c is a sufficiently small value.

2.5.4 The Energy Function

The design of the Hopfield neural network energy function for super-resolution land cover mapping was particular to the specific type of task undertaken. Therefore, the energy function design for target identification, land cover mapping and pattern prediction, will be described in chapters 3, 4 and 5 respectively.

2.6 Summary

Hopfield neural networks are fully connected recurrent networks that can be used for optimisation problems by formulating the weights and biases such that they describe an energy function. By specifying different values for the weights and biases, any energy minimisation task can be simulated.

The Hopfield neural network simulation used in this research was designed with the task of super-resolution land cover mapping from satellite sensor imagery in mind. The network architecture was arranged to represent a finer spatial resolution image, and land cover class proportions produced from soft classification were used for initialisation. Constraints within the energy function determined neuron activations within this architecture, and the energy minimum corresponded to a bipolar map of class components within each pixel. The following chapters describe the energy function design, alongside consequent results.

Chapter 3: Super-Resolution Target Identification*

3.1 Introduction

Information on land cover features is required for management and understanding of the environment. Accurate identification, extraction and mapping of target land cover features is a vital procedure for this in many areas of work, e.g. military intelligence, agricultural planning and water resource management and remote sensing has the potential to provide this information.

The work in this chapter demonstrates that it is possible to identify land cover targets at the sub-pixel scale and produce accurate super-resolution maps of them.

3.1.2 Chapter Structure

This chapter describes an approach that uses the output from a soft classification technique to constrain a Hopfield neural network formulated as an energy minimisation tool. In section 3.2, an overview of the workings of the Hopfield neural network will be given. Section 3.2 will also introduce the modifications made to apply the Hopfield neural network to the problem of land cover target identification and mapping at the sub-pixel scale, as well as methods used to understand and improve the processes at work. In section 3.3, the results of applying the approach to synthetic imagery are used to explain and understand the performance of the network. Section 3.4 illustrates the results of applying the technique to simulated remotely sensed imagery, and section 3.5 provides analysis and explanation of these results. Finally, section 3.6 provides a summary and conclusion of the findings of this part of the research.

3.2. Using the Hopfield Neural Network for Land Cover Target Identification at the Sub-pixel Scale

The input data for the research described in this chapter were derived from aerial photography, whereby land cover targets were identified and extracted accurately from the photographs by hand, using field survey for verification. By degrading these verification images of clearly defined land cover targets to the spatial resolution of Landsat TM data using a square mean filter, accurate class proportion estimates were obtained for each pixel. These provided the input to the network, but in practice, this input could come from

* This chapter is based on Tatem *et al.* (2001a)

automated soft classification methods, such as the multi-layer perceptron. However, for the research in this chapter, the aim was to understand and test the capabilities of the Hopfield neural network technique, so any error introduced to the input data by an automated soft classification method would be detrimental to this aim.

3.2.1 Network Architecture

The network architecture for super-resolution target identification is described in section 2.5.1.

3.2.2 Network Initialisation

Network neurons were initialised using both methods ((i) and (ii)) described in section 2.5.2.

3.2.3 The Energy Function

The equations of motion were defined as,

$$\frac{dE_{ij}}{dv_{ij}} = k_1 \frac{dG1_{ij}}{dv_{ij}} + k_2 \frac{dG2_{ij}}{dv_{ij}} + k_3 \frac{dP_{ij}}{dv_{ij}} \quad . \tag{3.1}$$

Each component of equation 3.1 is described in the subsequent sections.

The goal and constraints of the sub-pixel mapping task were defined such that the network energy function was,

$$E = -\sum_i \sum_j \left(k_1 G1_{ij} + k_2 G2_{ij} + k_3 P_{ij} \right), \tag{3.2}$$

where k_1, k_2 and k_3 are constants weighting the various energy parameters, $G1_{ij}$ and $G2_{ij}$ represent the output values for neuron (i,j) of the two objective (or goal) functions (see section 3.2.3.1), and these correspond to the quadratic term in equation 2.7. P_{ij} represents the output value for neuron (i,j) of the proportion constraint (see section 3.2.3.2) which corresponds to the linear term in equation 2.7.

3.2.3.1 The Goal Functions

The goal (objective) functions were based upon an assumption of spatial dependence (Matheron, 1965). Almost all natural and human-made phenomena exhibit spatial continuity

at some scale, such that points near to each other are more alike than those further apart, and the degree of dissimilarity depends on both the environment and the nature of our observations (Curran and Atkinson, 1998). These observations can be the pixels in remotely sensed images, and the assumption of spatial dependence can be used to infer relationships between these pixels. By focusing within this research on *discrete land cover targets*, which all exhibit spatial dependence to some degree, the assumption of spatial dependence becomes particularly relevant. Therefore, by devising simple measures of spatial dependence, and incorporating each as objective functions within the Hopfield neural network to map the spatial distribution of class components within a pixel, this real world phenomenon was modelled.

In this case, the aim of the goal was to make the output of a neuron similar to that of its neighbouring neurons. Therefore, if the output of neuron (i,j) was similar to the average output of the eight neighbouring neurons, then a low energy was given. If it was different, then this represented an undesirable situation in terms of the aim of spatial dependence, and a high energy was produced. However, to produce a bipolar image, a function that just drives a neuron output to be similar to the surrounding neuron output was insufficient. Consequently, two objective functions were introduced; one to increase neuron output towards a value of 1, and another to decrease neuron output to 0, each dependent on the average output of the eight neighbouring neurons.

The first function aimed to increase the output of the centre neuron, v_{ij}, to 1, if the average output of the surrounding eight neurons, $\frac{1}{8}\sum_{\substack{k=i-1\\k\neq i}}^{i+1}\sum_{\substack{l=j-1\\l\neq j}}^{j+1} v_{kl}$, was greater than 0.55,

$$\frac{dG1_{ij}}{dv_{ij}} = \frac{1}{2}\left(1+\tanh\left(\frac{1}{8}\sum_{\substack{k=i-1\\k\neq i}}^{i+1}\sum_{\substack{l=j-1\\l\neq j}}^{j+1} v_{kl} - 0.55\right)\lambda\right)\left(v_{ij}-1\right), \tag{3.3}$$

where λ is a gain which controlled the steepness of the *tanh* function. The *tanh* function controlled the effect of the neighbouring neurons. If the averaged output of the neighbouring neurons was less than 0.5, then equation 3.3 evaluated to 0, and the function had no effect on the energy function (equation 3.2). If the averaged output was greater than 0.5, equation 3.3 evaluated to 1, and the $\left(v_{ij}-1\right)$ function controlled the magnitude of the negative gradient

output, with only $v_{ij}=1$ producing a zero gradient. A negative gradient was required to increase neuron output.

The second goal function aimed to decrease the output of the centre neuron, v_{ij}, to 0, given that the average output of the surrounding eight neurons, $\frac{1}{8}\sum_{\substack{k=i-1 \\ k\neq i}}^{i+1}\sum_{\substack{l=j-1 \\ l\neq j}}^{j+1} v_{kl}$, was less than 0.5,

$$\frac{dG2_{ij}}{dv_{ij}} = \frac{1}{2}\left(1+\left(-\tanh\left(\frac{1}{8}\sum_{\substack{k=i-1 \\ k\neq i}}^{i+1}\sum_{\substack{l=j-1 \\ l\neq j}}^{j+1} v_{kl}\right)\lambda\right)\right)v_{ij}. \tag{3.4}$$

This time the *tanh* function evaluated to 0 if the averaged output of the neighbouring neurons was more than 0.5. If it was less than 0.5, the function evaluated to 1, and the centre neuron output, v_{ij}, controlled the magnitude of the positive gradient output, with only $v_{ij}=0$ producing a zero gradient. A positive gradient was required to decrease neuron output and only when $v_{ij}=1$ and $\frac{1}{8}\sum_{\substack{k=i-1 \\ k\neq i}}^{i+1}\sum_{\substack{l=j-1 \\ l\neq j}}^{j+1} v_{kl} > 0.5$, or $v_{ij}=0$ and $\frac{1}{8}\sum_{\substack{k=i-1 \\ k\neq i}}^{i+1}\sum_{\substack{l=j-1 \\ l\neq j}}^{j+1} v_{kl} < 0.5$, was the energy gradient equal to zero, and $G1_{ij}+G2_{ij}=0$. This satisfied the objective of maximising spatial dependence, while also forcing neuron output to either 1 or 0 to produce a bipolar image.

3.2.3.2 The Proportion Constraint

While the goal functions provided the enforcement of spatial dependence, the sole use of these functions would have resulted in all neuron outputs taking the values 1 or 0. Therefore, a method of constraining the effect of those functions to the correct image areas was required. The proportion constraint, P_{ij}, aimed to retain the pixel class proportions output from the soft classification. This was achieved by adding in the constraint that the total output from the set of neurons representing each coarse spatial resolution image pixel should be equal to the predicted class proportion for that pixel. An area proportion estimate representing the proportion of neurons with an output of 0.55 or greater was calculated for all the neurons representing pixel (x,y),

$$\text{Area Proportion Estimate} = \frac{1}{2z^2}\sum_{k=xz}^{xz+z}\sum_{l=yz}^{yz+z}\left(1+\tanh\left(v_{kl}-0.55\right)\lambda\right). \tag{3.5}$$

The use of the *tanh* function ensured that if a neuron output was greater than 0.55, it was counted as having an output of 1 within the estimation of class area per pixel. Below an output of 0.55, the neuron was not counted within the estimation, which simplified the area proportion estimation procedure, and ensured that neuron output must exceed the random initial assignment output of 0.55 to be counted within the calculations.

To ensure that the class proportions per pixel output from the soft classification were maintained, the proportion target per pixel, a_{xy}, was subtracted from the area proportion estimate (equation 3.5),

$$\frac{dP_{ij}}{dv_{ij}} = \frac{1}{2z^2}\sum_{k=xz}^{xz+z}\sum_{l=yz}^{yz+z}\left(1+\tanh\left(v_{kl}-0.55\right)\lambda\right)-a_{xy}\,. \tag{3.6}$$

If the area proportion estimate for pixel (x,y) was lower than the target area, a negative gradient was produced which corresponded to an increase in neuron output to counteract this problem. An overestimation of class area resulted in a positive gradient, producing a decrease in neuron output. Only when the area proportion estimate was identical to the target area proportion for each pixel did a zero gradient occur, corresponding to $P_{ij}=0$ in the energy function (equation 3.2).

3.2.4 Advantages of the Technique

The approach described in this chapter holds several strategic advantages over those techniques mentioned in section 1.7:

- The option to choose the level of spatial resolution increase. This is essential if simulation of finer spatial resolution imagery is the aim.
- The Hopfield neural network technique has the ability to simulate any shape, rather than being restricted to straight boundaries.
- The ease by which any additional information can be incorporated within the framework to aid the target identification. Any prior information about the land cover target depicted in the input imagery can be coded easily into the Hopfield neural network as an extra constraint to increase accuracy.
- The design of the Hopfield neural network as an optimisation tool means that all constraints are satisfied simultaneously, rather than employing a multi-stage operation.

- The effect that each one of these constraints has on the final prediction image can be controlled simply via weightings.
- While the incorporation of prior information on each land cover target may increase resulting map accuracy, the Hopfield neural network technique has the benefit of being able to produce accurate, super-resolution land cover target maps from just class proportions. There is therefore no reliance on the availability of finer spatial resolution imagery, or Land-line vector data.

3.3. Interpretation

To understand and illustrate the workings of the Hopfield neural network set up in this way, several synthetic images were created. Traditionally, within the remote sensing community, there has been a reluctance to use synthetic imagery, with the application of techniques directly to real imagery being preferred. However, by breaking down the elements of real-world imagery into simplified representations, understanding an image processing technique and, in turn, making improvements to it becomes easier.

The spatial dependence exhibited in all natural and human-made landscapes is partly a function of the fact that scenes within remotely sensed imagery are composed of various shapes, for example, fields, roads and houses. Shapes can be characterised using compactness and circularity. Compactness, c, is defined as,

$$c = \frac{4\pi a}{p^2}, \tag{3.7}$$

where, p is the perimeter length, and a is the area of the shape.

Circularity, r, is defined as,

$$r = \frac{a}{\pi(\max)^2}, \tag{3.8}$$

where, a is the area of the shape, and max is the maximum distance from the shape centre to the perimeter. Figure 3.1 shows eight, 56 × 56 pixel, synthetic images of differing shapes, with their respective compactness and circularity measures. The shapes represent extremes of spatial order as well as possible land cover targets, for example (b): a road, (d): a field. In

addition, the shapes cover a wide range of compactness and circularity values, providing a useful test of the generalization capabilities of the Hopfield neural network. Figure 3.2 shows each synthetic image subsampled (using a 7 × 7 mean filter) to generate an eight by eight pixel image, which causes mixing of the two classes (white and black) at the shape boundaries, producing eight proportion images, and imitating the effect of class mixing within remotely sensed imagery.

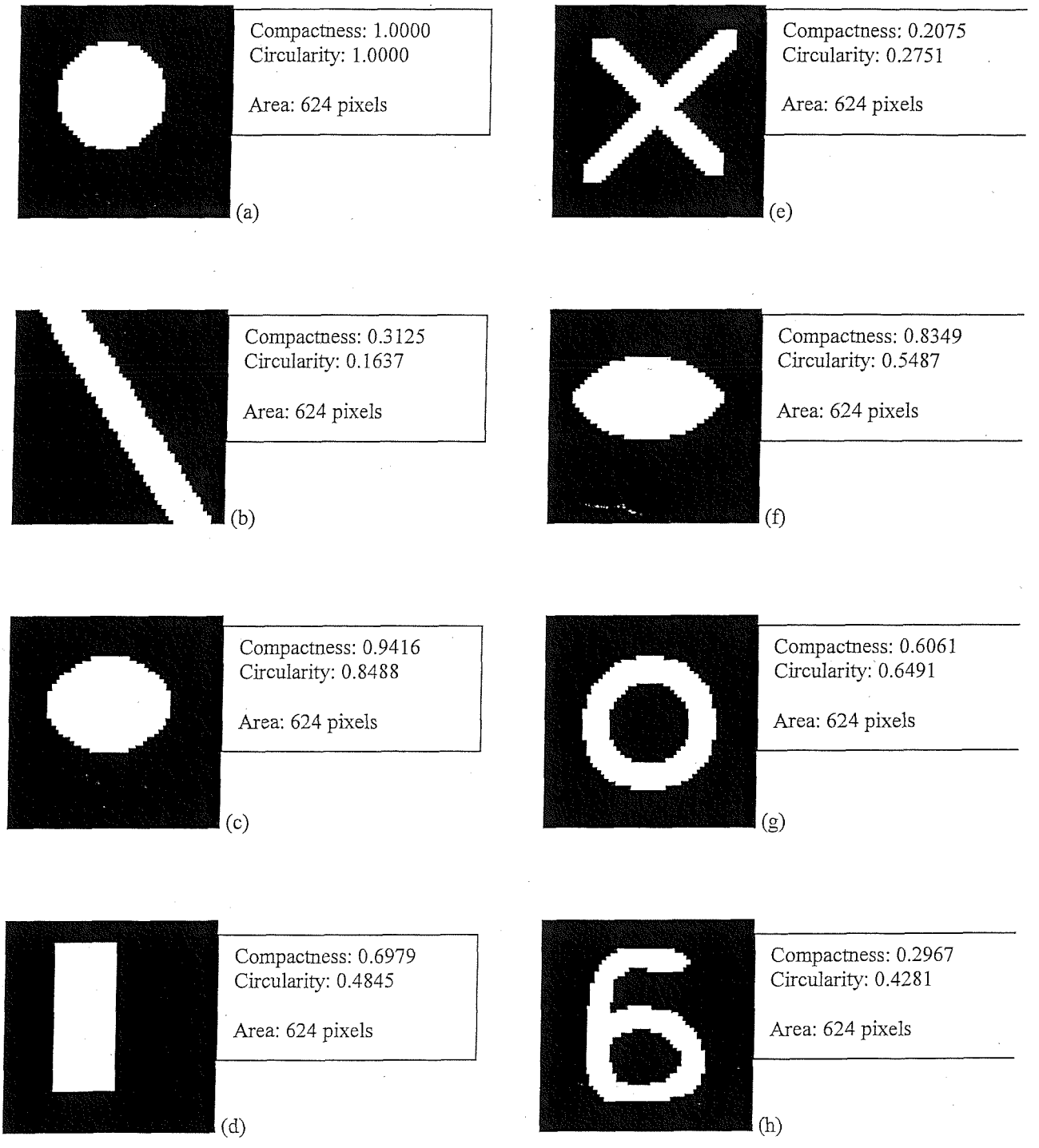


Figure 3.1. Synthetic images and the features of the shapes depicted in each.

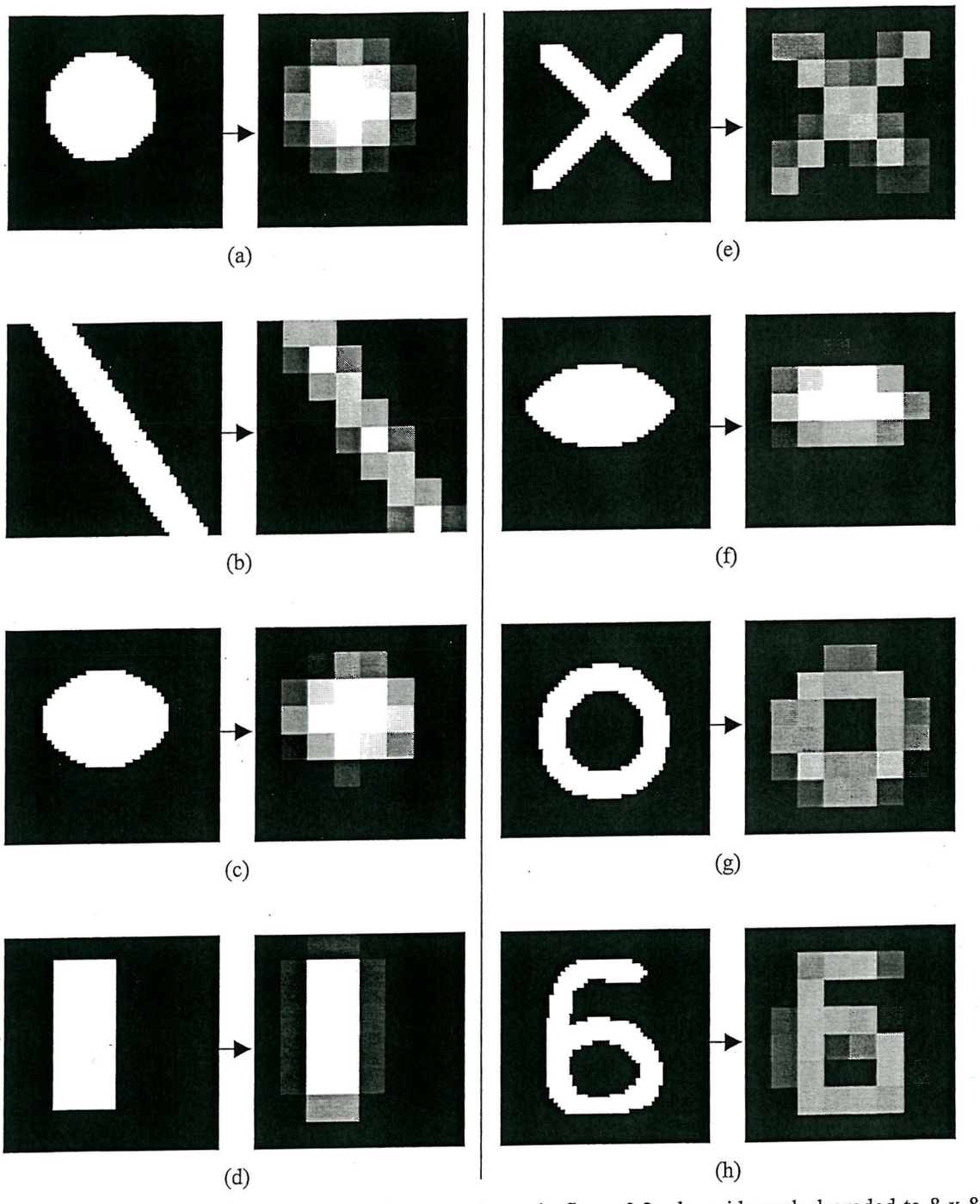


Figure 3.2. The 56 x 56 pixel synthetic imagery shown in figure 3.2, alongside each degraded to 8 x 8 pixel images.

By using these eight proportion images as inputs to the Hopfield neural network, and setting a zoom factor of 7, it should be possible to test the capabilities of the network by approximating the eight images each was derived from.

3.3.1 Setting the constraint weightings

To attempt to make predictions from the synthetic imagery, optimum values of goal and constraint weightings, k_1, k_2 and k_3, should be used, as these constants are of great importance because they control the direction of the optimisation process. Typically, finding optimum constraint values represents a notoriously difficult task, so an estimation of optimal values was made via certain assumptions and multiple network runs. For this paper, equal weightings of 150 were chosen and the justification for this decision becomes clear with the following results.

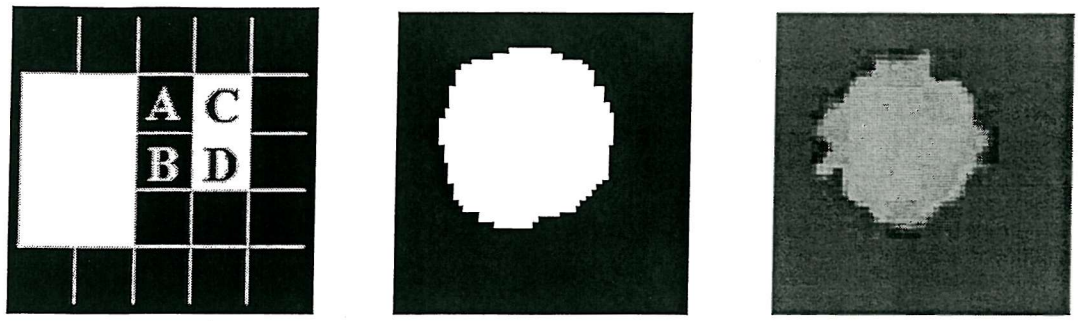
(i) Justification for equal weightings

Figure 3.3(a) shows a hypothetical situation of an image pixel with a zoom factor of 5. It should be noted that for this example, the effect of proportions within surrounding pixels is ignored for simplicity, but in practice, the distribution of class proportions within the surrounding pixels may have a significant effect on the resulting neuron activations predicted.

The target class proportion is 8/25, and using the constrained random initialisation method, this is satisfied immediately (of the 25 neurons, 8 are activated). While the proportion constraint value, P, is zero, the goal function values for this arrangement, $G1$ and $G2$, are large due to the isolated outputs of neurons C and D. To minimise $G1$ and $G2$, and keep the target proportion of 8/25, neurons A and B should have an increased output, while the outputs of C and D should be reduced.

If the proportion constraint is weighted strongly, then the arrangement of neural outputs displayed in figure 3.3(a) remains, as P is minimised, and the goal functions are not weighted strongly enough to have an effect, i.e. $k_1 G1 + k_2 G2 < k_3 P$.

If the goal functions are weighted strongly, then neurons A and B display increased output, neurons C and D stay at a high output and other neurons around these also increase in output. In such a case, the proportion constraint is not sufficiently strong to maintain target class proportions, i.e. $k_1 G1 + k_2 G2 > k_3 P$.



(a) (b) (c)

Figure 3.3. (a) Hypothetical image pixel with a zoom factor of 5 imposed. The pixel has been initialised with 8 neurons of high output (white), and 17 with low output (black); (b) Hopfield neural network prediction for the shape in figure 3.2(a), given k_1 = 150, k_2 = 150 and k_3 = 15; (c) Hopfield neural network prediction for the shape in figure 3.2(a), given k_1 = 15, k_2 = 15 and k_3 = 150.

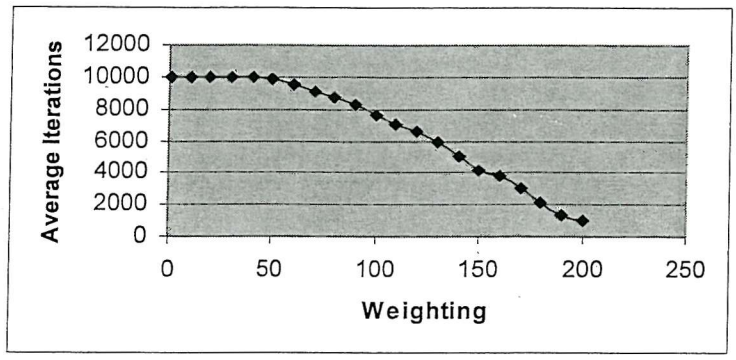
The effect of the biased weightings described above is demonstrated by running the network with a zoom factor of 7, for 10 000 iterations on the synthetic image in figure 3.2(a). Figure 3.3(b) demonstrates that the predicted shape is too large and irregular without the proportion constraint to control the positive activation effect of the goal functions. Giving P a large weight means that maintaining target class proportions becomes a priority, and the goal functions have little effect, such that the image in figure 3.3(c) is produced with a range of outputs.

By weighting the goal functions and proportion constraint equally, i.e. $k_1 = k_2 = k_3$, each affects and controls the other to minimise the overall energy of the network, and this is demonstrated in future examples.

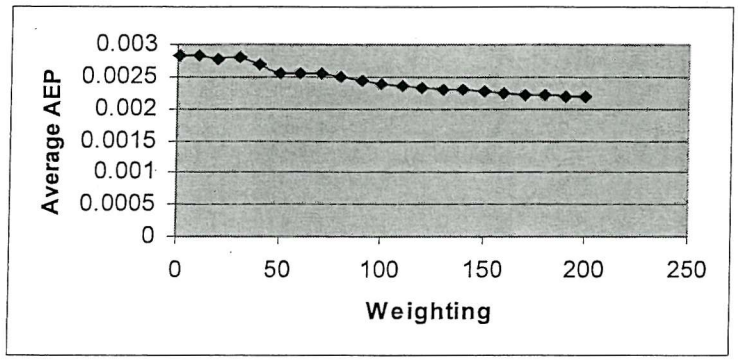
(ii) Justification for weightings of 150

While section *(i)* demonstrates why equal weightings were important, a second important choice of the value of these weightings needed to be made. This was achieved by running the Hopfield neural network on the shapes shown in figure 3.4 with a range of constraint weighting values, ensuring that

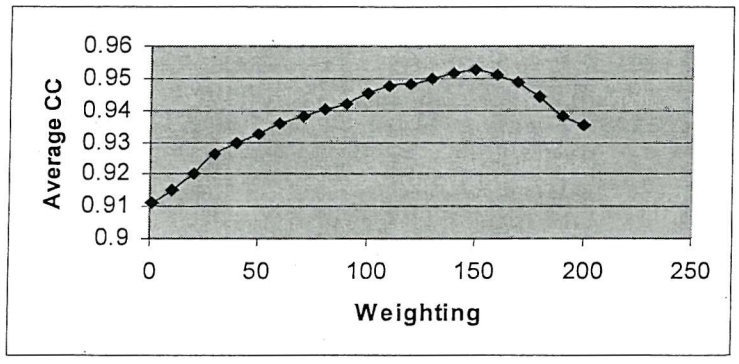
$k_1 = k_2 = k_3$. The results are shown in figures 3.4(a)-(d).



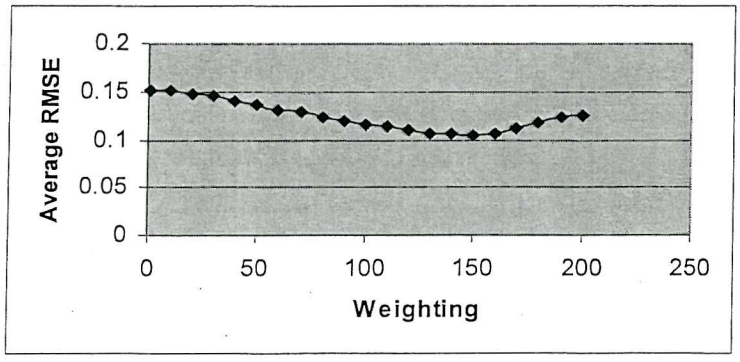
(a)



(b)



(c)



(d)

Figure 3.4. Graphs showing the relationship between constraint weighting value and (a) average number of network iterations until convergence (a) average area error proportion, (b) average correlation coefficient and (c) average RMS error.

Figure 3.4(a) shows that as the constraint weightings are increased, the average number of iterations required to reach convergence decreases (10 000 iterations was set as a maximum limit). This result is expected since increasing the effect of each constraint in the energy function produces larger steps towards an energy minimum. However, while using such larger constraint weightings may lead to quicker convergence, there is no guarantee that the result is more accurate. Therefore, the effect of varying constraint weightings was assessed using area error proportion (equation 1.11), correlation coefficient (equation 1.12) and root mean square error (equation 1.13), as described in 1.6.4.

Figure 3.4(b) shows an increase in weightings produces a small increase in the accuracy with which target class proportions are maintained, but figures 3.4(c) and 3.4(d) demonstrate more important findings. Both reveal that the prediction accuracy increases with weighting value, up to a value of approximately $k_1 = k_2 = k_3 = 150$. Beyond this value, the result suggests that the steps taken to find the energy minimum become too large, resulting in the network being unable to find the true minimum and a decrease in prediction accuracies. Consequently, weighting values of $k_1 = k_2 = k_3 = 150$ were chosen for use within the research described in this chapter.

3.3.2 Predictive ability and limitations

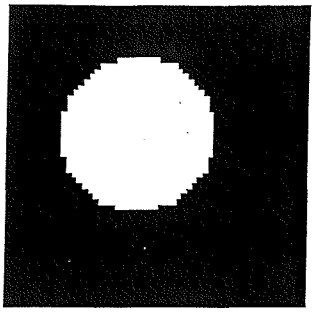
The results of the Hopfield neural network predictions from the synthetic imagery after 10 000 iterations, using values of 150 for k_1, k_2 and k_3, are shown in figure 3.5. Three measures of accuracy were calculated to assess the difference between each network prediction and the target images:

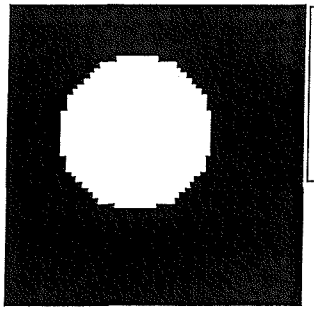
(i) Area Error Proportion (equation 1.11)
(ii) Correlation Coefficient (equation 1.12)
(iii) Root Mean Square Error (equation 1.13)

Figure 3.5(a) demonstrates the predictive ability and generalisation capabilities of the Hopfield neural network used in this chapter. From the eight by eight pixel image in figure 3.2(a), the network is able to re-create perfectly the circle it was derived from (fig 3.1(a)).

Original Image | Hopfield Network Prediction | Accuracy Statistics

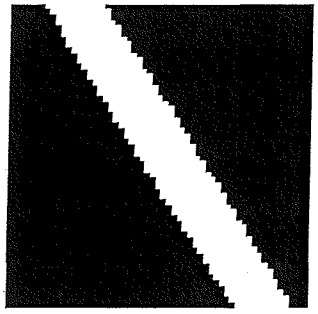
(a)

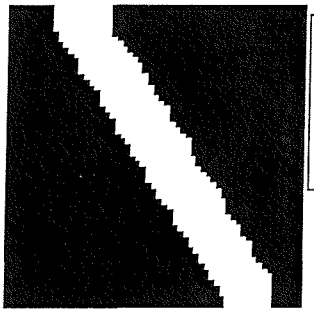




Target Area:	624
Estimated Area:	624
Area Error Proportion:	0.0000
Correlation Coefficient:	1.0000
RMS Error:	0.0000

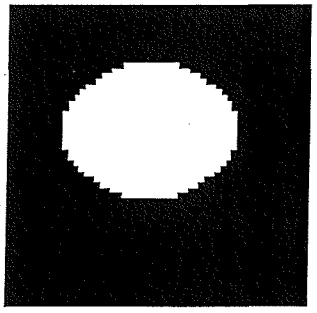
(b)

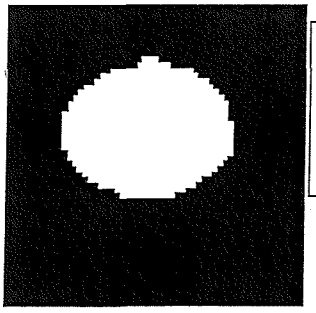




Target Area:	624
Estimated Area:	620
Area Error Proportion:	0.0013
Correlation Coefficient:	0.9599
RMS Error:	0.1129

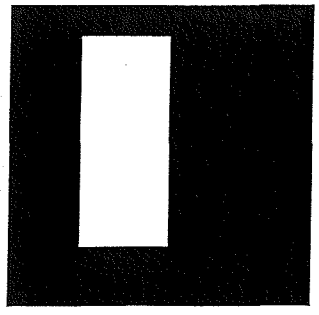
(c)

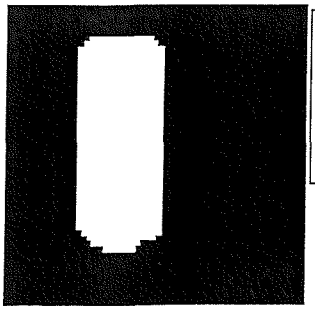




Target Area:	624
Estimated Area:	624
Area Error Proportion:	0.0000
Correlation Coefficient:	0.9860
RMS Error:	0.0668

(d)





Target Area:	624
Estimated Area:	613
Area Error Proportion:	0.0035
Correlation Coefficient:	0.9769
RMS Error:	0.0856

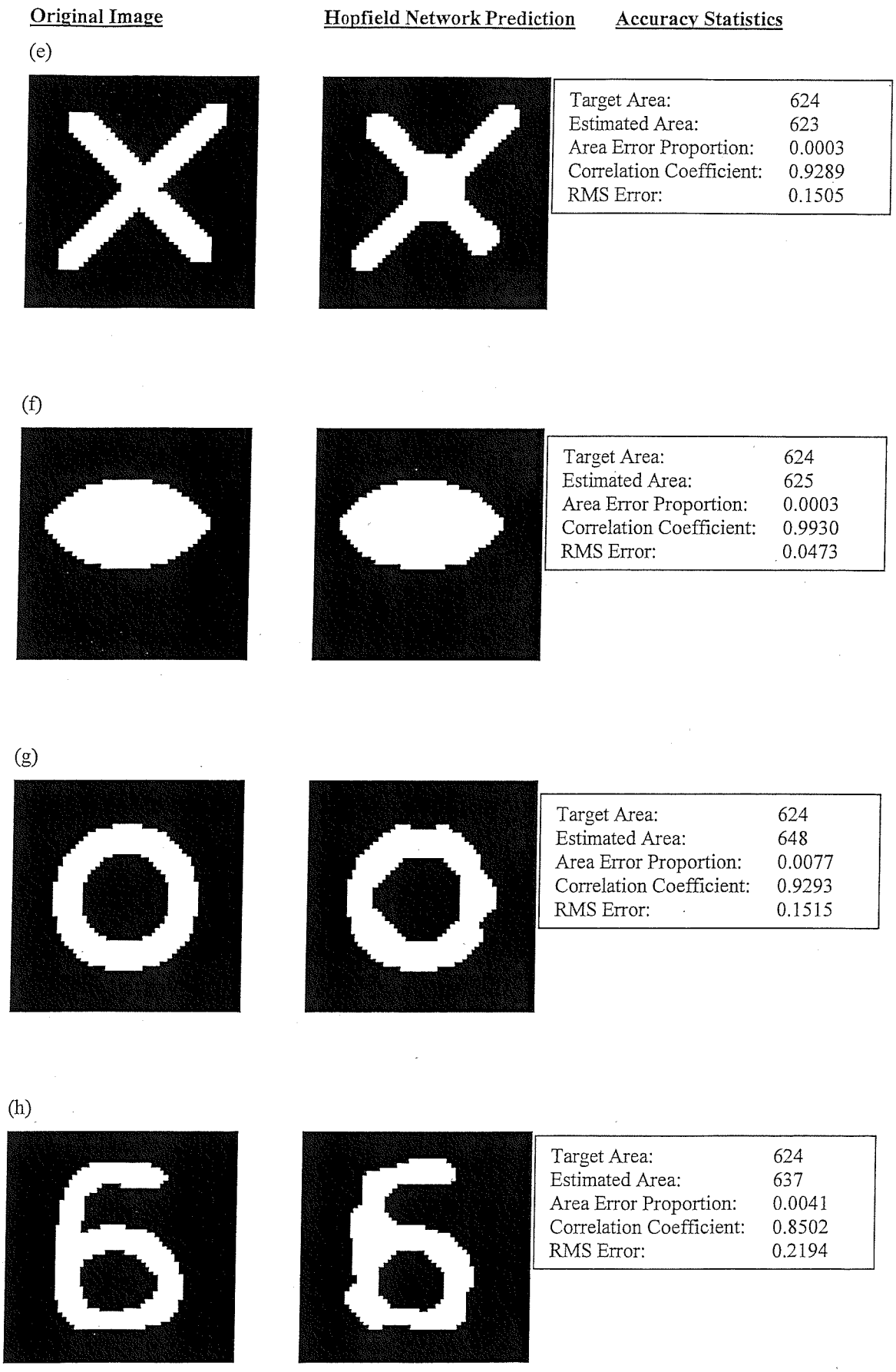


Figure 3.5. Hopfield neural network predictions and accuracy assessment, given the 8 x 8 pixel synthetic imagery shown in figure 3.2 as input.

For a set area of pixels, the circle represents the maximum spatial order attainable, and so by basing the goal functions around this assumption, the network is able to perform well on such regular shapes.

Figure 3.5(b) demonstrates the limitations that exist when class proportions lie on the edge of images. The goal functions are set up to rely on information from surrounding neurons and, consequently, the lack of such information for edge neurons means that only the proportion constraint is satisfied, leading to the low predictive accuracy at the image edge.

Figures 3.5(c) and 3.5(f) demonstrate good predictive ability in terms of class area, resulting in low area error proportion estimates. However, in both cases the predicted shapes are slightly incorrect, resulting in a larger RMS Error. This is due to certain edge formations satisfying the goal functions and proportion constraint, yet not being identical to the edges in the target image. The lack of information available to recreate such edges correctly shows that the problem is under-constrained.

Figure 3.5(d) reflects the above problem further, because as described previously, the use of spatial order as the basis for the goal functions, means that the network will almost always converge to curved, rather than sharp corners. Unless prior knowledge exists on the type of shapes that the network is aiming to recreate, this problem will remain.

Finally, figures 3.5(e), 3.5(g) and 3.5(h) demonstrate the ability of the network to cope with more complex shapes. Comparison of, for example, figure 3.5(e) with 3.5(f), shows identical area error proportion values of 0.0003, demonstrating how the class area has been predicted accurately. However, the correlation coefficients and RMS errors are significantly different, representing the difficulty in predicting the spatial distribution of class components for a complex shape such as that in 3.5(e). While the proportion constraint has ensured that class area is maintained, without prior information on the shape depicted in the input proportions, the goal functions on their own were insufficient to recreate accurately the spatial layout of the cross.

By repeating the network run on the more complex input proportions in figures 3.5(e), 3.5(g) and 3.5(h), slightly different predictions are produced each time. The use of an initialisation technique based on random neuron output, constrained by target class proportions, means that starting neuron arrangements from certain network runs produce lower energy than others.

This results in different paths of convergence along the energy surface of the network and, in turn, for the more complex shapes this results in slightly different predictions each time, again indicating the under-constrained nature of the problem.

3.3.2.1 Predictive ability versus shape type

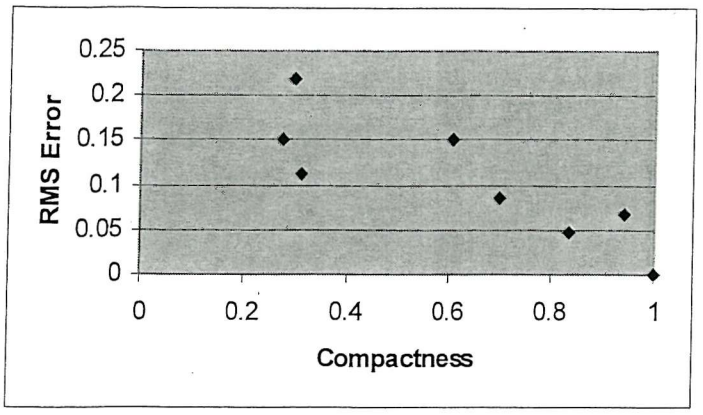
Relationships can be drawn between the shape characteristics and the predictive ability of the network. These have the potential to be used to obtain a network performance prediction, providing that a certain degree of knowledge about the input land cover shapes is known.

Figures 3.6(a) and 3.6(b) show the relationship between the various shape characteristics and the RMS error. The plots suggest increased accuracy with compactness and circularity maximised, which represents maximum spatial order, and is due to the design of the goal function as mentioned previously.

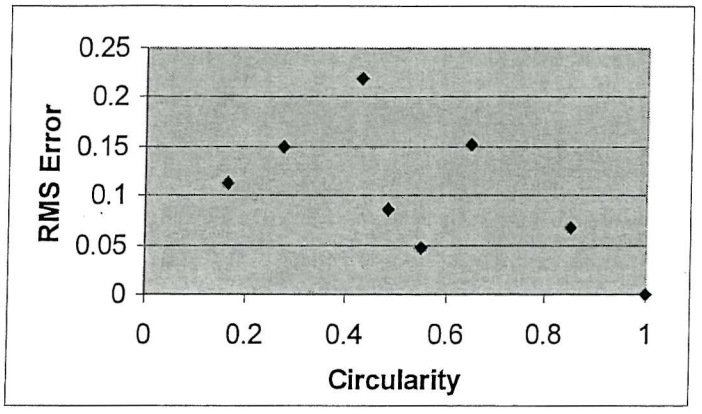
3.3.2.2 Predictive ability versus zoom factor

Figure 3.7 demonstrates the effect that different zoom factors have on the ability of the Hopfield neural network to predict accurately the spatial layout of various shapes. The network was run on the synthetic imagery shown in figure 3.2, using zoom factors of 3 to 10, and the averaged results were plotted.

Figure 3.7(a) shows that the higher zoom factors produce smaller area error proportions, indicating that the exact areal coverage of each shape can be better approximated with a larger number of pixels. Figures 3.7(b) and 3.7(c) demonstrate the effect that the zoom factor has on the precision of the prediction, because as the zoom factor is increased, the accuracy with which the target shape is recreated increases. However, above a zoom factor of 7, it is clear that further increases in accuracy are minimal, as the levelling off of figures 3.7(b) and 3.7(c) indicates. The drawback of using high zoom factors is increased computational time, shown in figure 3.7(d), meaning a balance must be found between spatial resolution and accuracy required, and computational cost. It was, therefore, decided to concentrate on zoom factors 3, 5 and 7 for the remainder of the thesis, as these showed the greatest contrast in accuracies for study, without the burden of excessive processing time.



(a)



(b)

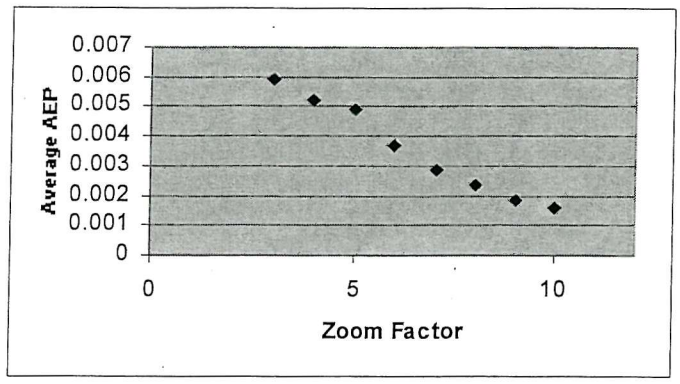
Figure 3.6. Scatterplots showing the variation in RMS error with (a) compactness and (b) circularity for the predictions in figure 3.5.

3.3.2.3 Predictive ability versus iterations

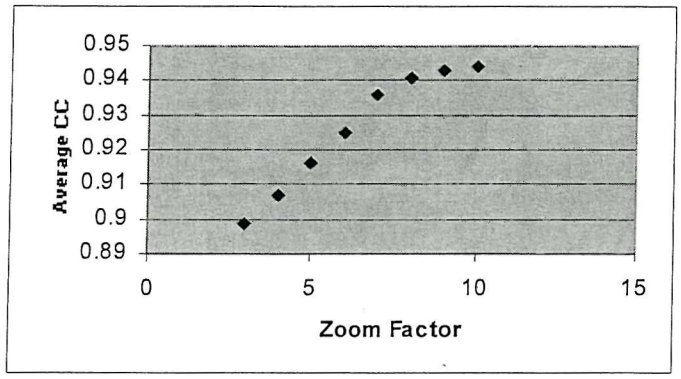
Figure 3.8 shows the performance of the network over 10 000 iterations, after initialisation using the proportion-constrained method. The network was run on the synthetic imagery shown in figure 3.2 using a zoom factor of 7, and values of 150 for k_1, k_2 and k_3, and the averaged results were plotted. The three plots demonstrate how, from the proportion-constrained random initialisation, the major spatial organisation of neuron outputs is undertaken within the first 1000 iterations. Figure 3.8(a) shows how after 2000 iterations, the area error proportion reaches a stable value, reflecting the satisfaction of the proportion constraint, and this leaves the goal functions to be minimised, reflected in the gradual convergence to stable values of figures 3.8(b) and 3.8(c).

3.3.3 Initialisation

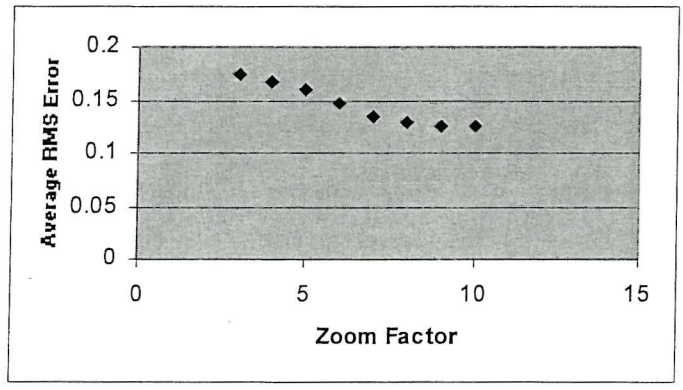
The use of just three weighting functions in determining the energy of the Hopfield neural network described in this chapter, enables the energy surface to be plotted.



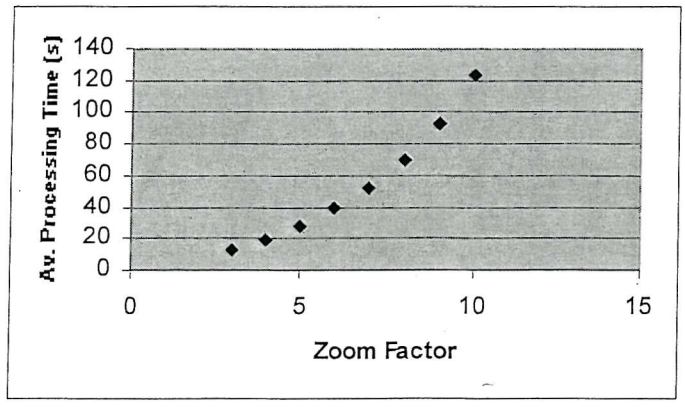
(a)



(b)

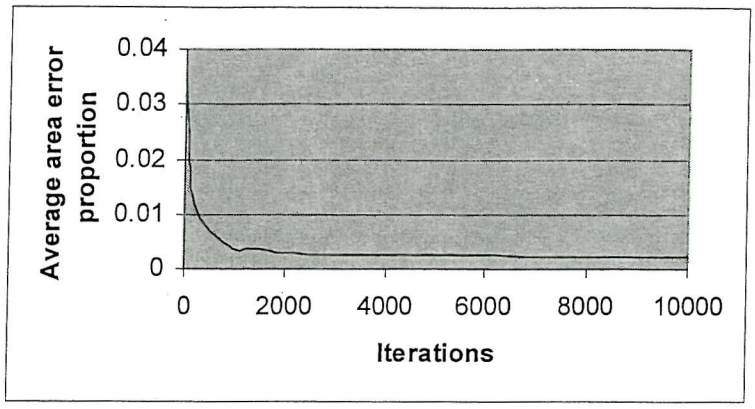


(c)

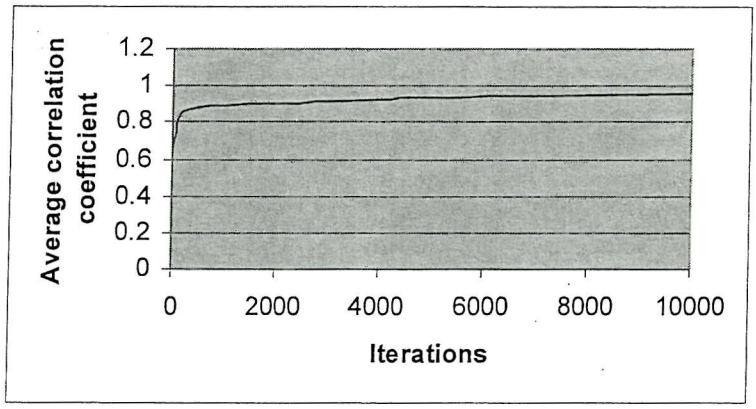


(d)

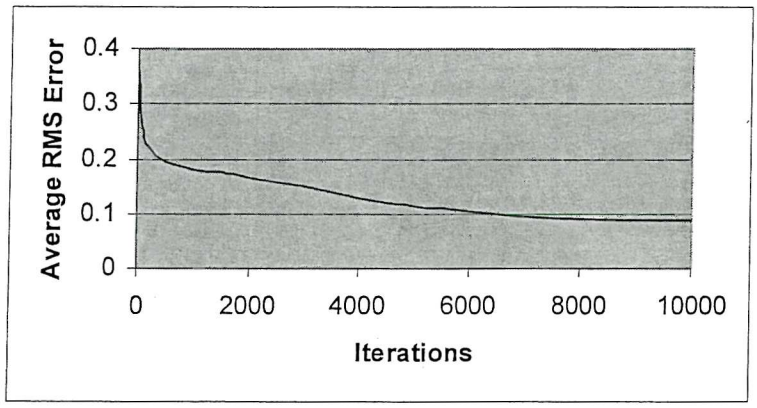
Figure 3.7. Scatterplots showing the relationship between zoom factor and (a) average area error proportion, (b) average correlation coefficient, (c) average RMS error and (d) average processing time.



(a)



(b)



(c)

Figure 3.8. Graphs showing the relationship between number of iterations and (a) average area error proportion, (b) average correlation coefficient and (c) average RMS error.

This simplicity means that the Hopfield neural network set up in this way is not only understandable, but also computationally efficient, and by studying the shape of such a surface, the workings of the network can be better understood.

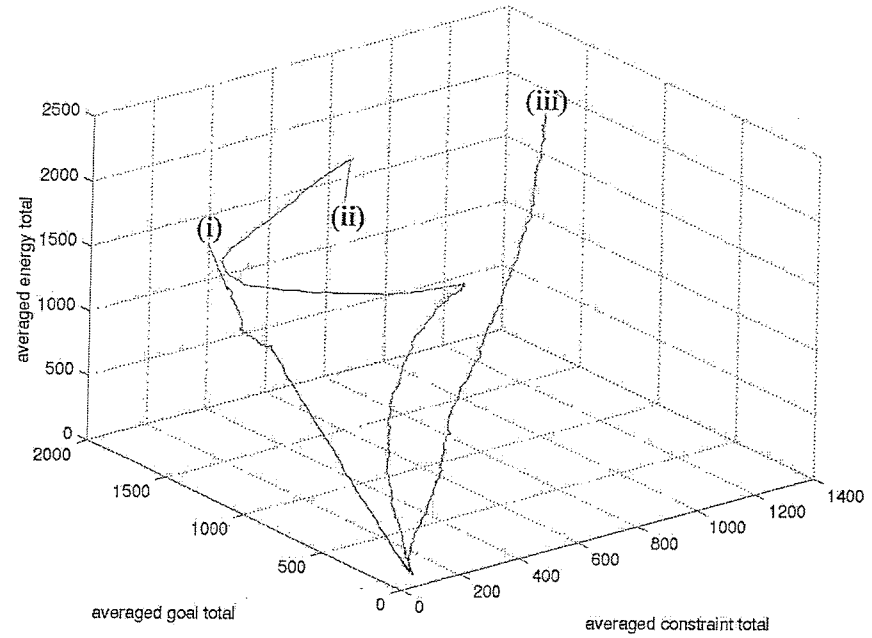
Figure 3.9(a) shows energy plots for 3 different initialisations of the network. The network was run on the 8 synthetic images shown in figure 3.2 using a zoom factor of 7, and values of 150 for k_1, k_2 and k_3, and the averaged results were plotted.

(i) represents a typical energy path when the network is initialised using image class proportions. The proportion constraint is satisfied immediately, leaving the network to minimise the total energy by altering neuron outputs to reduce the overall goal value.

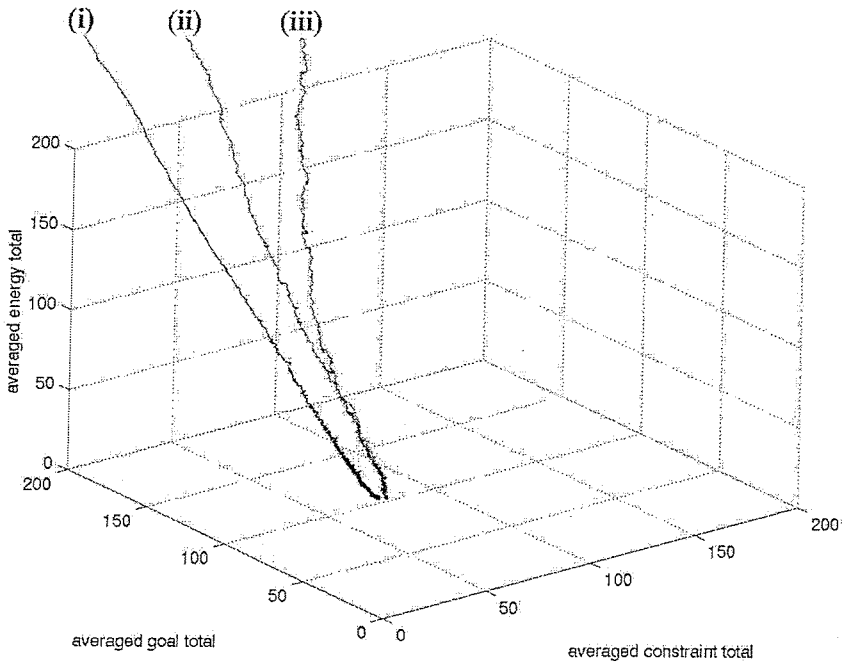
(ii) represents a typical energy path when the neurons are initialised to random outputs which are mostly less than 0.55, and the shape of the plot reflects the design of the proportion constraint. The threshold built into the proportion constraint means that, unless neuron output is greater than 0.55, it is not counted as 'on' in terms of representing class proportions. The proportion constraint, therefore, has little effect until the goal functions increase the output above 0.55 of enough neurons. At this point, the goal functions are sufficiently satisfied that the network minimises the total energy by altering neuron outputs to reduce the overall constraint function value. In satisfying the constraint function, the dominant energy-minimising force shifts towards the goal functions, which again increase neuron output, until the network finds a balance between the two functions, and converges to an energy minimum.

(iii) represents a typical energy path when the neurons are initialised to random values which are mostly greater than 0.55. This satisfies neither the goal nor the constraint functions, and the network, therefore, minimises the total energy by altering neuron outputs to find a compromise between the functions.

Figure 3.9(b) displays a close-up of the convergence of the 3 energy plots in figure 3.9(a), and shows the 3 plots converging approximately to the same point, demonstrating the effectiveness of the network at obtaining the same energy minimum, given any initialisation. Of the three runs depicted in figure 3.9(a), (i) reached the energy minimum quickest (approximately 5000 iterations), followed by (ii) (approximately 8000 iterations), then (iii) (approximately 9000 iterations).



(a)



(b)

Figure 3.9. (a) Hopfield neural network energy plots for three different initialisation settings: (i), (ii) and (iii); (b) close-up of the convergence of the 3 plots shown in (a).

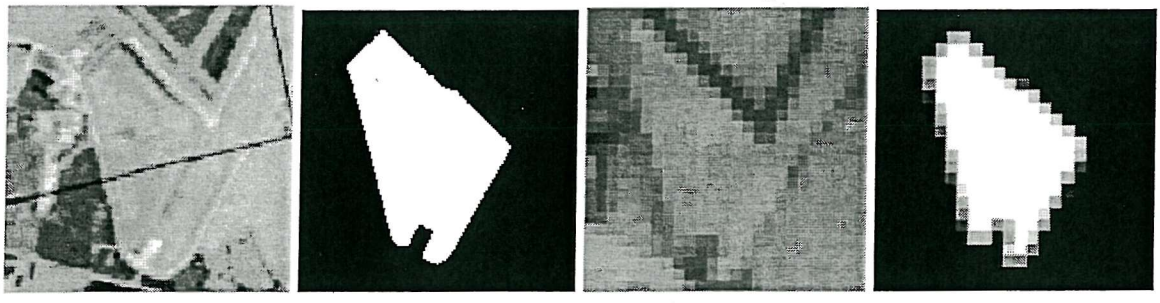
3.3.4 Performance Summary

This section has revealed several features about the workings of the Hopfield neural network that can be used to generate predictions on its performance:

- The technique will produce a more accurate prediction if the shape depicted in the input class proportions is compact and circular.
- The technique will produce a more accurate prediction if a high zoom factor is used.
- The technique will produce a more accurate prediction if the network is allowed to run for at least 1000 iterations.
- The network will converge to an accurate prediction in fewer iterations if a proportion-constrained initialisation is used.

3.4. Results

Results were produced using the Hopfield neural network run on a P2-350 computer. The network was used to identify land cover targets at the sub-pixel scale from simulated remotely sensed imagery. Landsat Thematic Mapper (TM) imagery were acquired over an agricultural area east of Leicester (Stoughton) and in seven wavebands at a spatial resolution of 30 m. Within the imagery, attention was focused on a wheat field and a section of airstrip, which provided clearly defined targets with which to evaluate the technique.

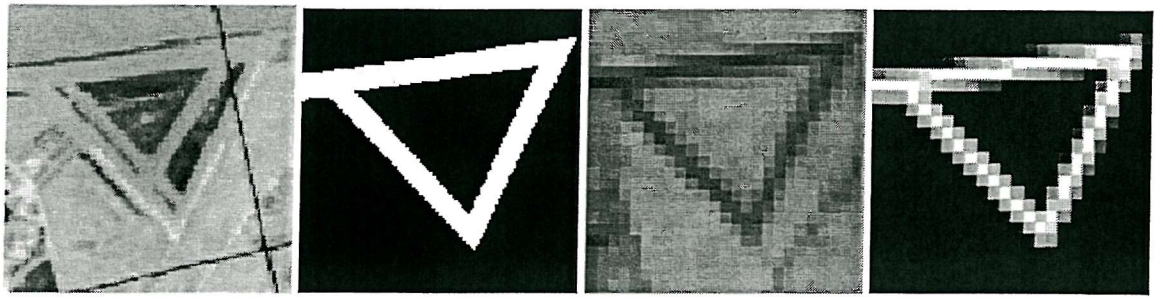


(a) (b) (c) (d)

Figure 3.10. (a) Digital aerial photograph (1 km grid overlaid); (b) Verification data derived from aerial photography and ground survey; (c) 24 x 24 pixel Landsat TM band 4 image; (d) Wheat class proportions.

Figures 3.10 and 3.11 show the various data used to initialise the network and evaluate the results produced. The verification data shown in figures 3.10(b) and 3.11(b) were derived by field survey and by hand from the 0.5 m spatial resolution digital aerial photographs shown in figures 3.10(a) and 3.11(a). The class proportion estimates shown in figures 3.10(d) and 3.11(d) were calculated from the verification data using a square mean filter which avoided

the potential problems of incorporating error from the process of soft classification of the imagery in figures 3.10(c) and 3.11(c).



(a) (b) (c) (d)

Figure 3.11. (a) Digital aerial photograph (1 km grid overlaid); (b) Verification data derived from aerial photography and ground survey; (c) 24 x24 pixel Landsat TM band 4 image; (d) Tarmac class proportions.

The network was initialised using the wheat and tarmac proportion images shown in figures 3.10(d) and 3.11(d). The proportion-constrained initialisation method was used with values for k_1, k_2 and k_3 of 150, and zoom factors of 5 and 7 were used for comparison.

After 10 000 iterations of the network at zoom factor 5 (approximately 5 minutes running time), prediction images were produced (figures 3.12(a) and (c)) with spatial resolutions 5 times higher than that of the input class proportion images in figures 3.10(d) and 3.11(d). In addition, after 10 000 iterations of the network at zoom factor 7 (approximately 10 minutes running time), prediction images were produced (figures 3.12(b) and (d)) with spatial resolutions 7 times finer than that of the input class proportion images in figures 3.10(d) and 3.11(d). The same measures of accuracy used in section 4 were calculated to assess the difference between each network prediction and the verification data.

3.5. Discussion

The accuracy shown for the results in figure 3.12, indicate that the Hopfield neural network displays great potential for accurately mapping class proportions from remotely sensed imagery at the sub-pixel scale.

The regularity and discrete nature of the wheat field enabled the network to perform well on this particular land cover target. In figure 3.12(a) and (b), the network maintained the areal coverage of the field in its prediction, while accurately predicting the field shape also, and in

both cases the nature of the goal functions meant that the network predicted rounder corners than that of the actual field.

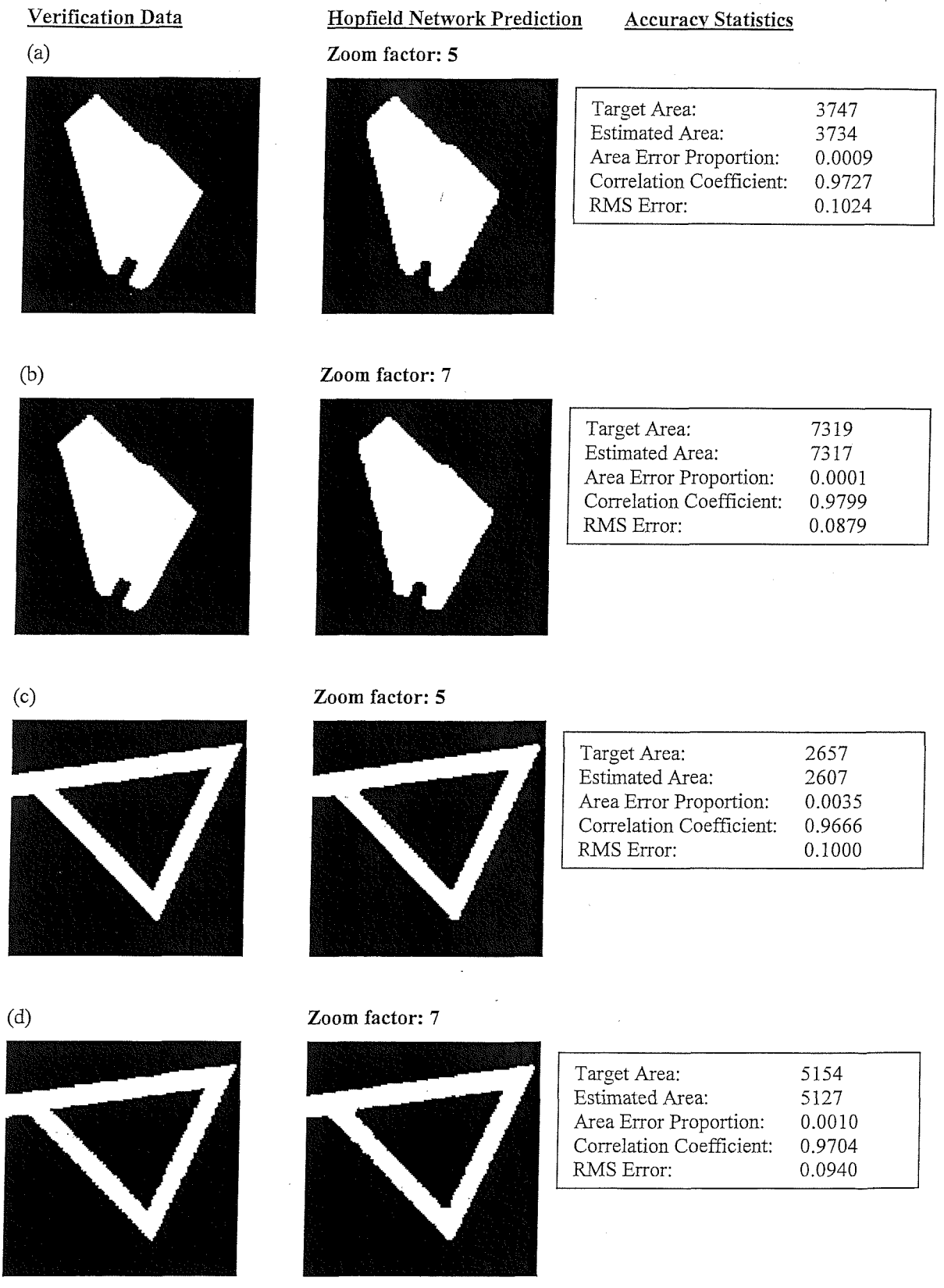


Figure 3.12. Hopfield neural network predictions and accuracy assessment, given the class proportion images shown in figures 3.10(d) and 3.11(d) as input.

With a zoom factor of 7, there was a less pronounced rounding effect, due to the finer scale that the goal functions were working on, resulting in greater accuracy. This corner-rounding problem represents the under-constrained nature of the problem, and prior information about the field shape at its corners could potentially have been built into the network as a constraint to avoid this problem.

The more complex shape of the airfield, as expected, produced results of lower accuracy than those for the wheat field. However, the statistics in figure 3.12(c) and (d) demonstrate that for both cases, the shape was predicted accurately, with an RMS error as low as 0.094 pixels using a zoom factor of 7. As predicted in section 3.3.4, more accurate results were again produced using the greater zoom factor. Again, both predictions produced rounder corners than those of the actual land cover target, but by using a zoom factor of 7, the network was able to more accurately model these corners than when it was run with a zoom factor of 5.

3.5.1 Potential Applications

Such a land cover target identification technique potentially leads to several useful applications:

(i) **To identify land cover targets at a fine spatial resolution from any remote sensing system.** While sensors on satellites such as IKONOS (up to 1 m spatial resolution) can provide sufficient spatial detail for accurate land cover target extraction, the cost and availability of such data may prohibit its use in many areas of work. By applying the developed technique to cheaper, more readily available data, for example SPOT HRV (up to 10 m spatial resolution) data, similar levels of accuracy for land cover target extraction might be achieved.

(ii) **To apply such a technique to obtain more accurate land cover metrics from remotely sensed imagery.** By increasing the spatial resolution of land cover target maps derived from medium-resolution sensors such as the Landsat Thematic Mapper or the SPOT High Resolution Visible Sensor, the potential exists to, for example, more accurately locate field boundaries or define areas of semi-natural vegetation. Such information would be of use in determining environmental or social policy, for example.

(iii) **For fine detail urban target identification.** With the advent of satellites such as IKONOS and Orbview, and the more common use of airborne remote sensing, imagery of spatial resolution less than 5 metres is becoming widely available.

Application of a technique to produce super-resolution maps of land cover targets from this source data would allow urban land cover target extraction and mapping of an unprecedented fine detail from remotely sensed imagery.

(iv) To simulate fine spatial resolution imagery from imagery of a coarser spatial resolution. Such an approach could aid decision making on future choices of imagery.

3.6. Conclusions

This chapter has shown that a Hopfield neural network can be used to estimate the location of the class proportions within pixels and produce a land cover target map of sub-pixel geometric precision. Tests using the Hopfield neural network on synthetic imagery have shown the generalisation abilities of the approach, with high accuracy predictions obtained for a range of different shapes. The Hopfield neural network represents a robust, efficient and simple technique, and results from synthetic and simulated remotely sensed data show great accuracy, suggesting that it has the potential to identify accurately land cover targets at the sub-pixel scale.

Chapter 4: Super-Resolution Land Cover Mapping*

4.1 Introduction

This chapter describes a technique for producing super-resolution land cover maps from remotely sensed imagery, incorporating any number of classes. The approach is an extension of the approach introduced in chapter 3 that uses the output from a soft classification to constrain a Hopfield neural network formulated as an energy minimisation tool.

4.2 Using the Hopfield Neural Network for Land Cover Mapping at the Sub-pixel Scale

4.2.1 The Hopfield Neural Network Design

For the task of super-resolution land cover mapping, the basic design of the Hopfield neural network described in section 2.5 was extended to cope with multiple classes. The h class proportion images output from soft classification were represented by h inter-connected layers in the network, with each neuron of high output value within these layers corresponding to a pixel in the finer spatial resolution map produced after the network had converged. Therefore, neurons will be referred to by co-ordinate notation, for example, neuron (h,i,j) refers to a neuron in row i and column j of the network layer representing land cover class h, and has an input value of u_{hij} and an output value of v_{hij}. The *zoom factor*, z, again determined the increase in spatial resolution from the original satellite sensor image to the new fine spatial resolution image. After convergence to a stable state, the output values, v, of all neurons were either 0 or 1, representing a binary classification of the land cover *at the finer spatial resolution.* The specific goals and constraints of the Hopfield neural network energy function determined the final distribution of neuron output values.

The input, u, to neuron with co-ordinates (h,i,j) was made up of a weighted sum of the outputs from every other neuron,

* This chapter is based on Tatem *et al.* (2001b) and (2001c)

$$u_{hij}(t) = \sum_{a=0}^{N} \sum_{\substack{b=0 \\ b \neq i}}^{pxz} \sum_{\substack{c=0 \\ c \neq j}}^{qxz} w_{abc,hij} v_{abc} \tag{4.1}$$

where N is the number of land cover classes, $w_{abc,hij}$ represents the weight between neuron (h,i,j) and (a,b,c), and v_{abc} is the output of neuron (a,b,c). The function describing the neural output at time t, $v_{hij}(t)$, as a function of the input was,

$$v_{hij}(t) = \frac{1}{2}(1 + \tanh u_{hij}\lambda) \tag{4.2}$$

where λ is the gain, which determines the steepness of the function. $v_{kij}(t)$ lies in the range [0,1].

As in 2.5, the dynamics of the neurons were again simulated numerically by the Euler method. After a user defined time step, dt, the input to neuron (h,i,j), u_{hij}, became,

$$u_{hij}(t + dt) = u_{hij}(t) + \frac{du_{hij}(t)}{dt} dt \tag{4.3}$$

where $\frac{du_{hij}(t)}{dt}$ represents the energy change of neuron (h,i,j) at time t. The total energy change of the network was then,

$$\frac{dE}{dv} = -\sum_{h=0}^{N} \sum_{i=0}^{p \times z} \sum_{j=0}^{q \times z} \left| \frac{du_{hij}(t)}{dt} \right| \tag{4.4}$$

and when this reached zero, or the change in energy from E_t to E_{t+1} was very small, the network had converged on a stable solution.

The goal and constraints of the sub-pixel mapping task were defined such that the network energy function was,

$$E = -\sum_{h}\sum_{i}\sum_{j}(k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij}) \tag{4.5}$$

For neuron (h,i,j), $G1_{hij}$ and $G2_{hij}$ are the values of spatial clustering (goal) functions, P_{hij} is the value of a proportion constraint, and M_{hij} is the value of a multi-class constraint. The constants k_1 , k_2, k_3 and k_4 were used to decide which constraint weightings to apply to solve the problem.

4.2.1.1 The Goal Functions

The spatial clustering (goal) functions, $G1_{hij}$ and $G2_{hij}$, were again based upon an assumption of spatial dependence, as described in 3.2.3.1. For neuron (h,i,j), an average output of its neighbours for each land cover class h, $\overline{v_n}$, was calculated, which represented the target output for that neuron,

$$\overline{v_n} = \frac{1}{8}\sum_{\substack{b=i-1 \\ b\neq i}}^{i+1}\sum_{\substack{c=j-1 \\ c\neq j}}^{j+1} v_{bc} \tag{4.6}$$

The first function aimed to increase the output for each layer, h, of the centre neuron, v_{ij}, to 1, if the average output of the surrounding eight neurons, $\frac{1}{8}\sum_{\substack{b=i-1 \\ b\neq i}}^{i+1}\sum_{\substack{c=j-1 \\ c\neq j}}^{j+1} v_{bc}$, was greater than 0.5,

$$\frac{dG1_{hij}}{dv_{hij}} = \frac{1}{2}\left(1 + \tanh\left(\frac{1}{8}\sum_{\substack{b=i-1 \\ b\neq i}}^{i+1}\sum_{\substack{c=j-1 \\ c\neq j}}^{j+1} v_{bc} - 0.55\right)\lambda\right)\left(v_{hij} - 1\right), \tag{4.7}$$

where, λ is a gain which controlled the steepness of the *tanh* function. The *tanh* function controlled the effect of the neighbouring neurons in the same way as in section 3.2.3.1. If the averaged output of the neighbouring neurons was less than 0.5, then equation 4.7 evaluated to 0, and the function had no effect on the energy function (equation 4.5). If the averaged output was greater than 0.5, equation 4.7 evaluated to 1, and the $\left(v_{hij} - 1\right)$ function controlled the magnitude

of the negative gradient output, with only $v_{hij} = 1$ producing a zero gradient. A negative gradient was required to increase neuron output.

The second clustering function aimed to decrease the output for each layer, *h*, of the centre neuron, v_{ij}, to 0, given that the average output of the surrounding eight neurons, $\frac{1}{8}\sum_{\substack{b=i-1 \\ b\neq i}}^{i+1}\sum_{\substack{c=j-1 \\ c\neq j}}^{j+1} v_{bc}$, was less than 0.5,

$$\frac{dG2_{hij}}{dv_{hij}} = \frac{1}{2}\left(1 + \left(-\tanh\left(\frac{1}{8}\sum_{\substack{b=i-1 \\ b\neq i}}^{i+1}\sum_{\substack{c=j-1 \\ c\neq j}}^{j+1} v_{bc}\right)\lambda\right)\right)v_{hij}. \tag{4.8}$$

This time, the *tanh* function evaluated to 0 if the averaged output of the neighbouring neurons was more than 0.5. If it was less than 0.5, the function evaluated to 1, and the centre neuron output, v_{hij}, controlled the magnitude of the positive gradient output, with only $v_{hij} = 0$ producing a zero gradient. A positive gradient was required to decrease neuron output and only when $v_{hij} = 1$ and $\frac{1}{8}\sum_{\substack{b=i-1 \\ b\neq i}}^{i+1}\sum_{\substack{c=j-1 \\ c\neq j}}^{j+1} v_{bc} > 0.5$, or $v_{hij} = 0$ and $\frac{1}{8}\sum_{\substack{b=i-1 \\ b\neq i}}^{i+1}\sum_{\substack{c=j-1 \\ c\neq j}}^{j+1} v_{bc} < 0.5$, was the energy gradient equal to zero, and $G1_{hij} + G2_{hij} = 0$. This satisfied the objective of recreating spatial dependence, while also forcing neuron output to either 1 or 0 to produce a bipolar image.

4.2.1.2 The Proportion Constraint

In a similar way to 3.2.3.2, the proportion constraint, P_{hij}, aimed to retain the pixel class proportions output from the soft classification. This was achieved by adding in the constraint that for each land cover class layer, *h*, the total output from each pixel should be equal to the predicted class proportion for that pixel. An area proportion estimate, $\hat{a}_{hxy}$, was calculated for all the neurons representing pixel (*h*,*x*,*y*),

$$\text{Area Proportion Estimate} = \frac{1}{2z^2}\sum_{b=xz}^{xz+z}\sum_{c=yz}^{yz+z}\left(1 + \tanh\left(v_{bc} - 0.55\right)\lambda\right). \tag{4.9}$$

The *tanh* function was used in the same way as 3.2.3.2, ensuring that if a neuron output was above 0.55, it was counted as having an output of 1 within the estimation of class area per pixel. Below an output of 0.55, the neuron was not counted within the estimation, which simplified the area proportion estimation procedure, and ensured that neuron output must exceed the random initial assignment output of 0.55 to be counted within the calculations.

To ensure that the class proportions per pixel output from the fuzzy classification were maintained, the proportion target per pixel, a_{hxy}, was subtracted from the area proportion estimate (equation 4.9),

$$\frac{dP_{hij}}{dv_{hij}} = \frac{1}{2z^2} \sum_{b=xz}^{xz+z} \sum_{c=yz}^{yz+z} \left(1 + \tanh\left(v_{bc} - 0.55\right)\lambda\right) - a_{hxy} . \qquad (4.10)$$

If the area proportion estimate for pixel (h,x,y) was lower than the target area, a negative gradient was produced, which corresponded to an increase in neuron output to counteract this problem. An overestimation of class area resulted in a positive gradient, producing a decrease in neuron output. Only when the area proportion estimate was identical to the target area proportion for each pixel did a zero gradient occur, corresponding to $P_{hij} = 0$ in the energy function (equation 4.5).

4.2.1.3 The Multi-Class Constraint

The multi-class constraint, M_{hij}, aimed to ensure that the outputs from each class layer fitted together with no gaps or overlap between land cover classes in the final prediction map. This was achieved by ensuring that the sum of the outputs of each set of neurons with position (i,j) equalled one.

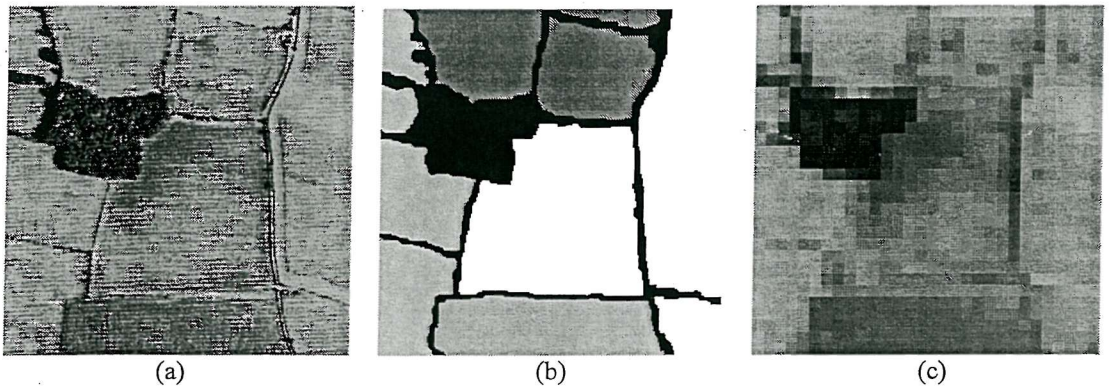
$$\frac{dM_{hij}}{dv_{hij}} = \left(\sum_{k=0}^{N} v_{hij} \right) - 1 \qquad (4.11)$$

If the sum of the outputs of each set of neurons representing pixel (i,j) in the final prediction image was less than one, a negative gradient was produced which corresponds to an increase in

neuron output to counteract this problem. A sum greater than one produced a positive gradient, leading to a decrease in neuron output. Only when the output sum for the set of neurons in question equalled one, did a zero gradient occur, corresponding to $M_{hij} = 0$ in the energy function (equation 4.5).

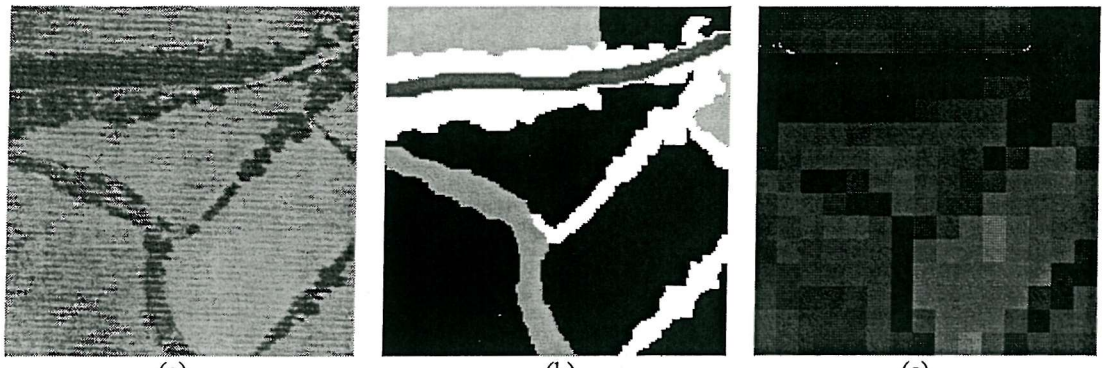
4.3 Simulated Remotely Sensed Imagery

In this section, simulated remotely sensed imagery is used to enable refinement of the technique and to give clear demonstration of the workings of the network. This avoids the uncertainty inherent in real imagery caused by the sensor point spread function, atmospheric and geometric effects, and classification error, which is introduced in section 4.4. Figures 4.1 and 4.2 show the two areas used for this study, the verification data and simulated satellite sensor images. The corresponding class proportion images are shown in figures 4.3 and 4.4. All the proportion images were derived by degrading the verification data using a square mean filter, to avoid the potential problems of incorporating error from the process of soft classification. For the study area in the city of Bath, UK, each class within the verification data was degraded to produce pixels of 20 m by 20 m, similar to that produced by SPOT HRV. For the study area in the city of Bristol, UK, the verification class images were degraded by a higher degree to produce pixels of 30 m by 30 m, to simulate those produced by Landsat TM.



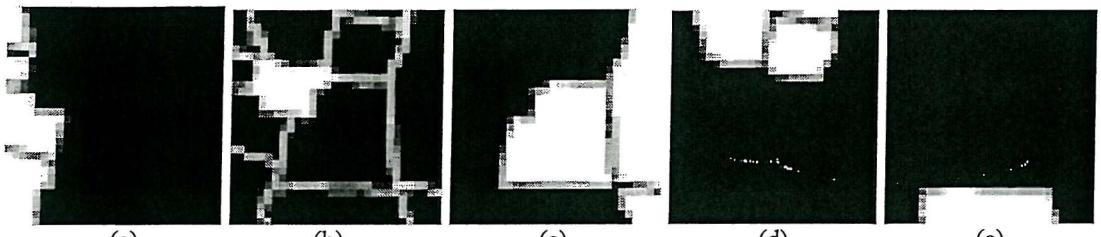
(a) (b) (c)

Figure. 4.1.
(a) Aerial photograph of an area near Bath, England.
(b) Verification image derived from digitisation of (a).
(c) Simulated SPOT HRV image (20 m spatial resolution).



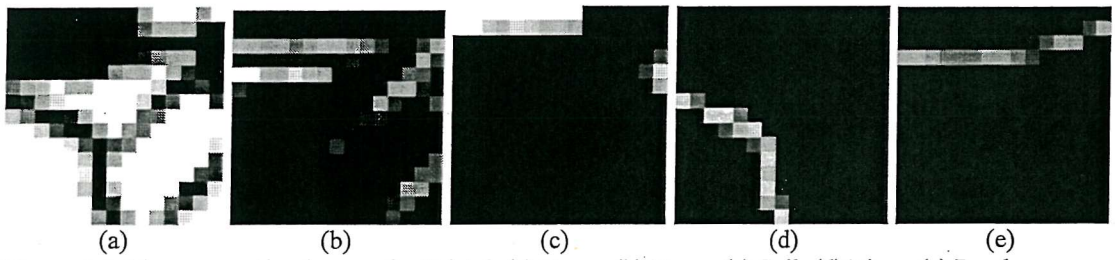
(a) (b) (c)

Figure 4.2.
(a) Aerial photograph of an area near Bristol, England.
(b) Verification image derived from digitisation of (a).
(c) Simulated Landsat TM image (30 m spatial resolution).



(a) (b) (c) (d) (e)

Figure. 4.3. Class proportion images for Bath, (a) Grass; (b) Trees; (c) Soil; (d) Wheat; (e) Weeds.



(a) (b) (c) (d) (e)

Figure 4.4. Class proportion images for Bristol, (a) Grass; (b) Trees; (c) Soil; (d) River; (e) Road.

4.3.1 Accuracy Assessment

Three measures of accuracy were estimated to assess the difference between each network prediction and the verification images. These were:

(i) Area Error Proportion (AEP) (equation 1.11)

(ii) Correlation Coefficient (CC) (equation 1.12)

(iii) Root Mean Square Error (RMSE) (equation 1.13)

4.3.2 Network Settings

The class proportion images shown in figures 4.3 and 4.4 provided the inputs to the Hopfield neural network. These proportions were used to initialise the network and provide area estimates for the proportion constraint. To produce predictions at similar scales, a zoom factor of z=5 was used on the images of Bath in figure 4.3, and z=7 on the images of Bristol in figure 4.4. This meant that from input images of the Bath area (with a spatial resolution of 20 m), a prediction map with a spatial resolution of 4 m was produced. From the input images of Bristol (with a spatial resolution of 30 m), a prediction map with a spatial resolution of 4.3 m was produced. In both cases, the constraint weightings k_1, k_2, k_3 and k_4 were set to 150, to ensure that no single function had a dominant effect on the energy function, as described in 3.3.1.

4.3.3 Hard Classification

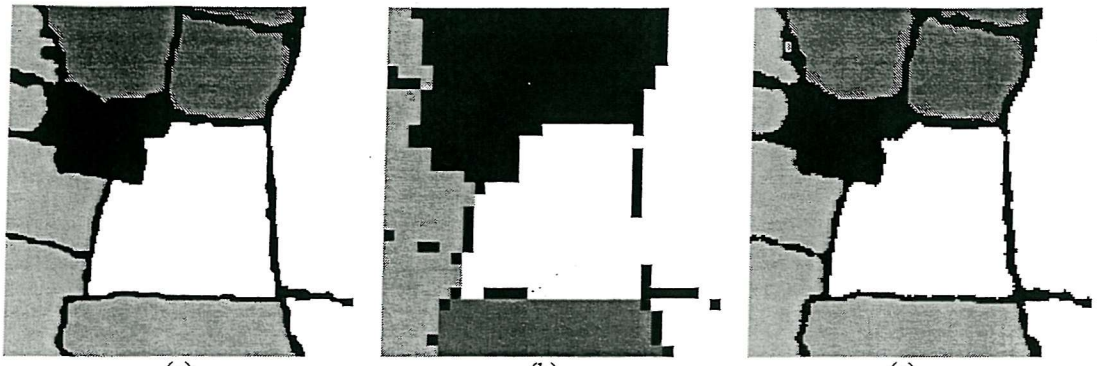
To evaluate the success of the Hopfield neural network technique, the traditional method of producing a land cover map from a satellite sensor image was also carried out for comparison. This involved undertaking a 'hard' maximum likelihood classification of the imagery, as described in 1.6.1 (iii).

4.3.4 Results

Illustrative results were produced using the Hopfield network run on a P2-350 computer.

4.3.4.1 Bath

After 10 000 iterations of the network with a zoom factor of z=5 (approximately 15 minutes running time), a prediction image was produced (figure 4.5(c)) with spatial resolution 5 times finer than that of the input class proportions in figure 4.3. In addition, a maximum likelihood classification was produced for comparison (figure 4.5(b)). Per-class (table 4.1) and overall (table 4.3(a)) accuracy statistics were also calculated to assess the difference between each map and the verification data.



(a) (b) (c)

Figure 4.5.
(a) Verification image (Bath).
(b) Traditional maximum likelihood classification.
(c) Hopfield neural network prediction (5 times increase in spatial resolution).

	Grass	Trees	Soil	Wheat	Weeds
Target Area	3420	3859	9214	3689	2318
Estimated Area	3425	3675	9375	3675	2350
Area Error Proportion	-0.0015	0.0477	-0.0175	0.0038	-0.0138
Correlation Coefficient	0.930	0.745	0.943	0.933	0.962
RMS Error	0.135	0.267	0.167	0.135	0.0838

(a)

	Grass	Trees	Soil	Wheat	Weeds
Target Area	3420	3859	9214	3689	2318
Estimated Area	3324	4044	9209	3659	2264
Area Error Proportion	0.0281	-0.0479	0.0005	0.0081	0.0233
Correlation Coefficient	0.965	0.902	0.982	0.973	0.981
RMS Error	0.0943	0.169	0.0926	0.0864	0.0589

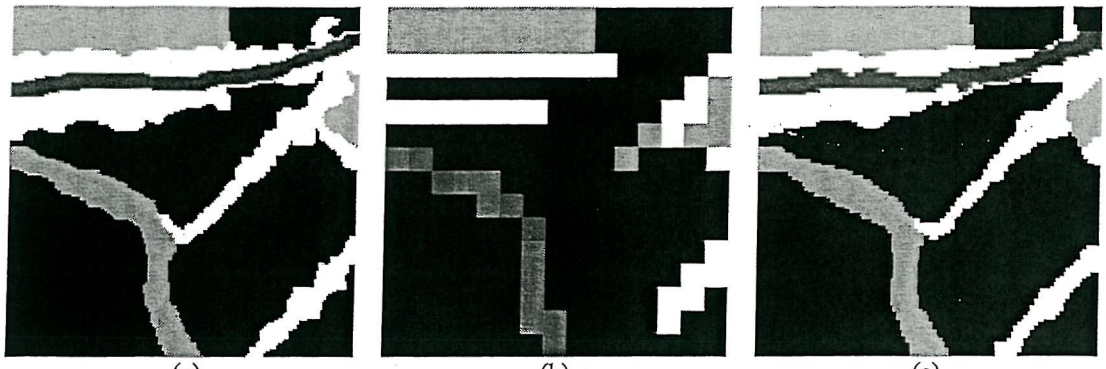
(b)

Table 4.1
(a) Per-class accuracy statistics for the hard classification of the image of Bath.
(b) Per-class accuracy statistics for the Hopfield network classification of the image of Bath.

4.3.4.2 Bristol

After 10 000 iterations of the network with a zoom factor of z=7 (approximately 30 minutes running time), a prediction image was produced (figure 4.6(c)) with spatial resolution 7 times finer than that of the input class proportions in figure 4.4. In addition, a maximum likelihood classification was produced for comparison (figure 7(b)). Per-class (table 4.2) and overall (table

4.3(b)) accuracy statistics were also calculated to assess the difference between each map and the verification data.



(a) (b) (c)

Figure 4.6.
(a) Verification image (Bristol).
(b) Traditional maximum likelihood classification.
(c) Hopfield neural network prediction (7 times increase in spatial resolution).

	Grass	Trees	Soil	River	Road
Target Area	17958	4617	2739	1901	1685
Estimated Area	18256	4042	2750	1816	2036
Area Error Proportion	-0.0166	0.125	-0.004	0.0447	-0.208
Correlation Coefficient	0.816	0.639	0.912	0.794	0.709
RMS Error	0.294	0.304	0.123	0.158	0.189

(a)

	Grass	Trees	Soil	River	Road
Target Area	17958	4617	2739	1901	1685
Estimated Area	17619	5036	2683	2045	1517
Area Error Proportion	0.0189	-0.0908	0.0204	-0.0757	0.0997
Correlation Coefficient	0.92	0.834	0.955	0.928	0.898
RMS Error	0.195	0.216	0.088	0.0959	0.104

(b)

Table 4.2
(a) Per-class accuracy statistics for the hard classification of the image of Bristol.
(b) Per-class accuracy statistics for the Hopfield network classification of the image of Bristol.

	Hard Classification	Hopfield Network Prediction
Overall Area Error Proportion	0.00352	0.00329
Overall RMS Error	0.169	0.107

(a)

	Hard Classification	Hopfield Network Prediction
Overall Area Error Proportion	0.00914	0.00779
Overall RMS Error	0.226	0.150

(b)

Table 4.3
(a) Overall accuracy statistics for the two image classifications of Bath.
(b) Overall accuracy statistics for the two image classifications of Bristol.

4.3.5. Discussion

The high accuracies shown for the results in figures 4.5 and 4.6 indicate that the Hopfield network displays potential for mapping accurately land cover class proportions within pixels, from remotely sensed imagery.

4.3.5.1 Bath

Visual assessment of figure 4.5 suggests immediately that the Hopfield network has mapped the land cover more accurately than the hard classifier. The major visible improvement attributable to the Hopfield network is the accurate mapping of the tree class, which has in turn lead to the maintenance of the fields as discrete features. Examination of the traditional maximum likelihood classification shows that, in contrast, the lines of trees have been mis-classified in many areas.

The statistics shown in table 4.1 confirm the above visual assessments. Whereas there is little significant difference between the area estimates from both techniques, the correlation coefficients and RMS errors demonstrate that the Hopfield network prediction was the more accurate. For all classes, the Hopfield network produced an increase in correlation coefficient over the hard classification, particularly for the tree class, with a value of 0.902 compared to that of 0.745. This result was mirrored by the RMS errors, with a particular decrease of around 0.1 pixels evident again for the tree class.

Table 4.3(a), which shows the overall image statistics, indicates that while both techniques maintained class area, the Hopfield network mapped the locations of these classes more accurately, with an RMS error of just 0.11 pixels.

4.3.5.2 Bristol

With input images with a spatial resolution of 30 m (similar to that of Landsat TM), mapping of the Bristol area proved a more difficult task than for the Bath area (20 m pixels). However, visual assessment of figure 4.6 suggests that the Hopfield network again mapped the location of classes within pixels accurately. The hard classification map appears visually to be even poorer than for the Bath area. The land cover appears as a series of blocky features, while the river and road classes exhibit discontinuities. In contrast, the Hopfield network prediction has maintained the approximate shapes and sizes of the discrete land cover features.

Table 4.2 confirms the above observations. The area error proportions are very similar, except for the road class where the hard classification has over-estimated the area. For all classes, the Hopfield network prediction produced a larger correlation coefficient than those for the maximum-likelihood prediction. However, the values are smaller overall than those for the Bath area, due to the coarser pixels in the input imagery. The RMS errors for the Hopfield network are less than for the hard classification, by up to 0.1 pixels, but again the results are generally less accurate than those for the Bath area.

The overall image statistics in table 4.3(b) this time show a decrease in area error proportion for the Hopfield network technique compared to the hard classification. A considerable decrease in overall RMS error is also evident. A decrease in RMS error of just 0.04 pixels (compared to that for the Bath area) may be explained by the use of a higher zoom factor, enabling more accurate recreation of the bends and corners of land cover features.

4.4 Landsat Thematic Mapper Imagery

In this section, the ability of the Hopfield neural network technique to produce accurate results from real Landsat Thematic Mapper (TM) data was explored. This enabled assessment of the network's abilities to cope with uncertainty in the form of:

(i) Uncertainty in image representation (e.g. point spread function, atmospheric effects, noise, geometric effects)

(ii) Uncertainty of classification (e.g. spectral confusion, class choices, classifier choice, classifier error)

(iii) Uncertainty in verification data (e.g. registration error, class identification, generalisation)

4.4.1 Data

Landsat TM imagery and corresponding verification data were acquired from two agricultural areas. The areas were chosen to test the capabilities of the Hopfield neural network. The first area displayed a regular arrangement of small fields with only a small number land cover classes, whereas the second area provided a far sterner test with irregular shaped large fields and a large number of classes. Both the satellite sensor imagery and verification data were acquired as part of the European Union 'Fuzzy Land Information for Environmental Remote Sensing' (FLIERS) project (contract number: ENV4-CT96). The techniques used to obtain and process these data are described in the FLIERS final report (Fisher (ed.), 2000).

(i) Sindos, Greece (small-area agriculture)

The area under study in Greece consists of many small regular fields of between 0.009 km^2 and 0.07 km^2 in area. An area of 2.04 km by 2.13 km was focused on, and in August 1997, TM imagery with a spatial resolution of 30 m was acquired in 6 wavebands, and 120 m in the thermal infra-red band. Figure 4.7 shows band 3 of the imagery. To provide co-incidental verification data, a field survey was carried out in August 1997. This field survey was used to identify land covers within a digitised cadastral map, and consequently, produce a verification map displaying 11 different land cover classes, shown in figure 4.8. By degrading these verification data to the spatial resolution of Landsat TM imagery, target class proportions were produced for use in regression based classification models (see section 4.4.5), and these are shown in figure 4.9.

Figure 4.7 Landsat TM band 3 image of Sindos, Greece.

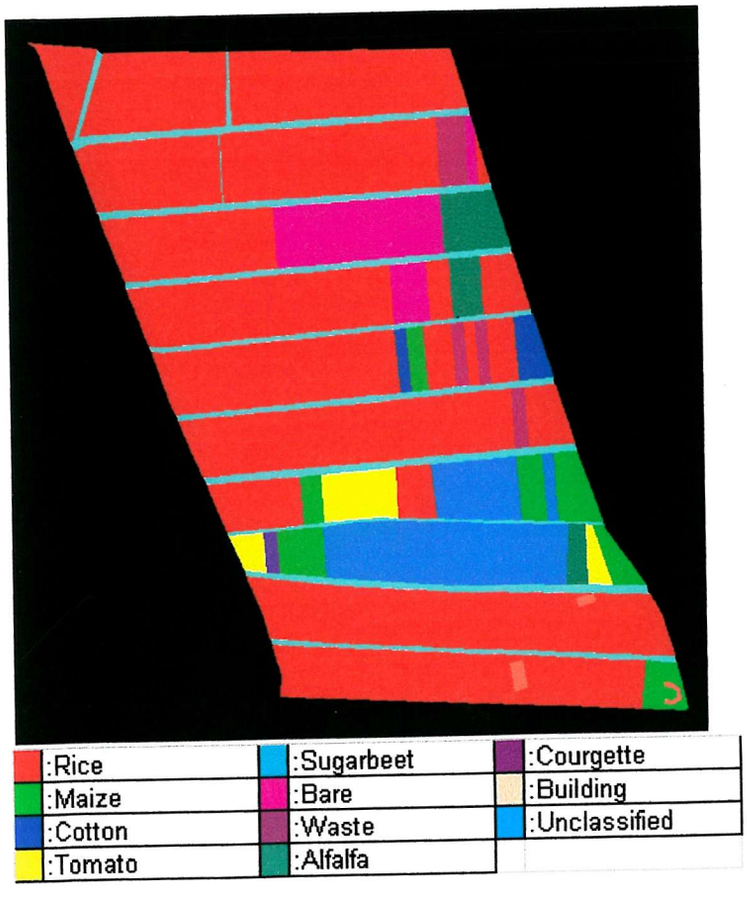


Figure 4.8 Verification map of Sindos, Greece.

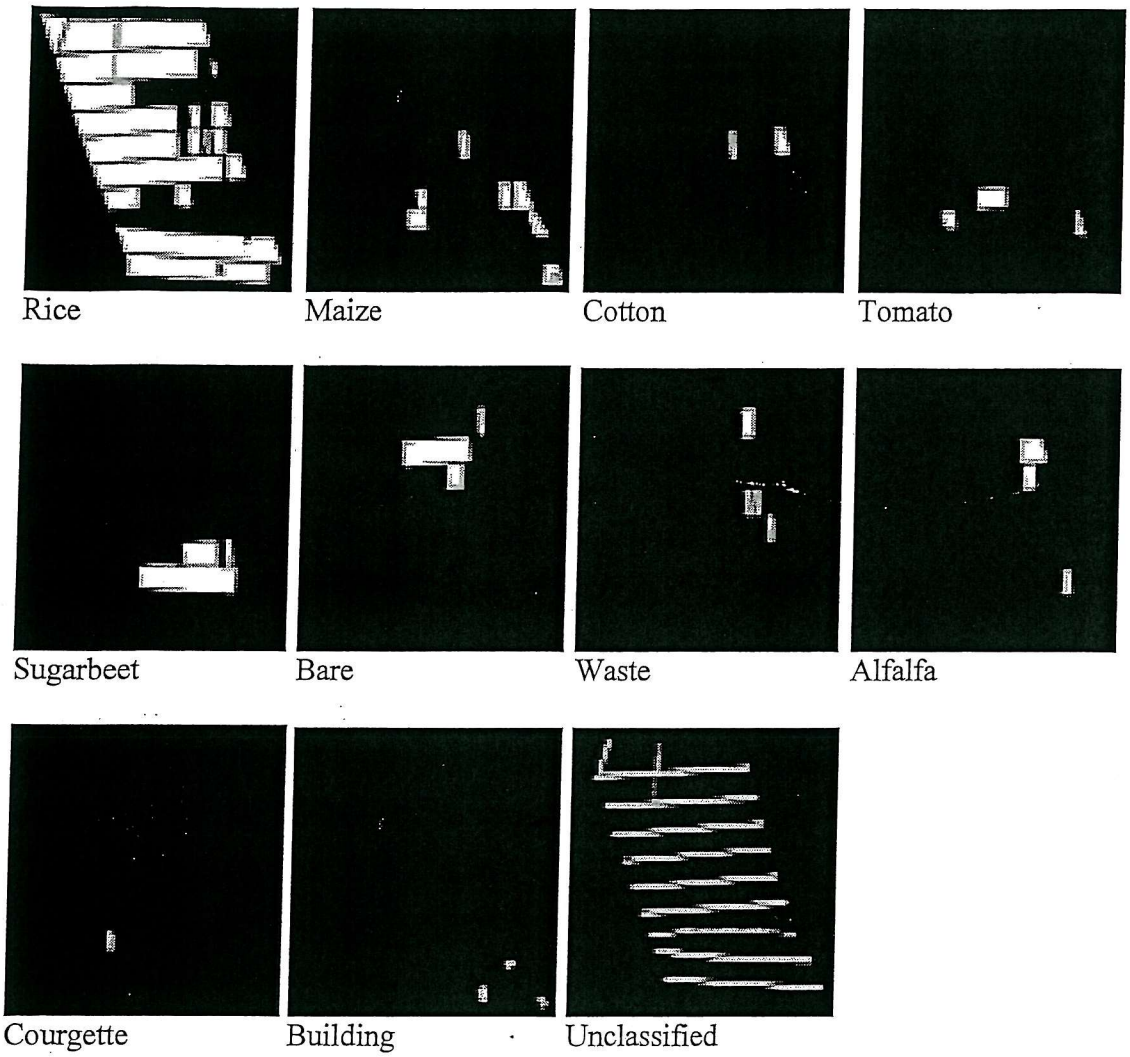


Figure 4.9 Target class area proportions for the image of Sindos, Greece.

(ii) Stoughton, UK (large-area agriculture)

The area of Stoughton is east of Leicester, UK, and consists of mainly large agricultural fields of between 0.1 km^2 and 0.12 km^2 in area. An area of 3 km by 2.73 km was focused on, and in 1998 (month unknown), Landsat TM imagery with a spatial resolution of 30 m were acquired in 6 wavebands, and 120 m in the thermal infra-red band. Figure 4.10 shows band 4 of the imagery. To provide co-incidental verification data, a field survey was undertaken in June and July 1998, and semi-orthorectified digital aerial photographs were obtained. This photography was used with field survey to identify land cover types and produce a verification map displaying 21 different land cover classes, shown in figure 4.11. By degrading these verification data to the spatial resolution of Landsat TM imagery, target class proportions were produced for use in regression based classification models (see section 4.4.5), and these are shown in figure 4.12.

Figure 4.10 Landsat TM band 4 image of Stoughton, UK.

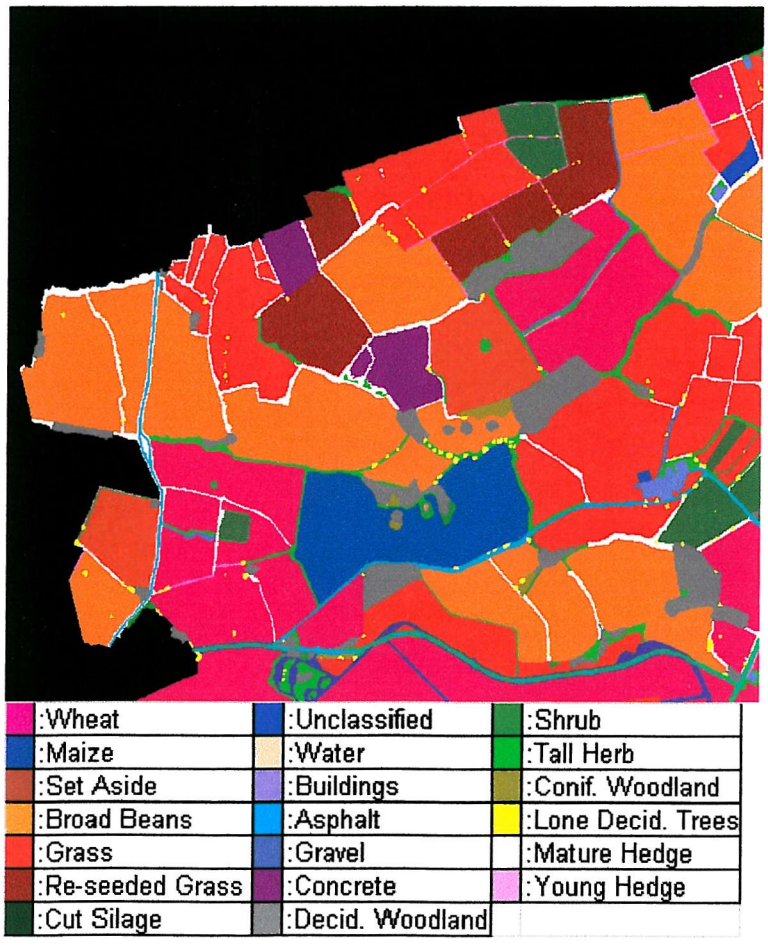


Figure 4.11 Verification map of Stoughton, UK.

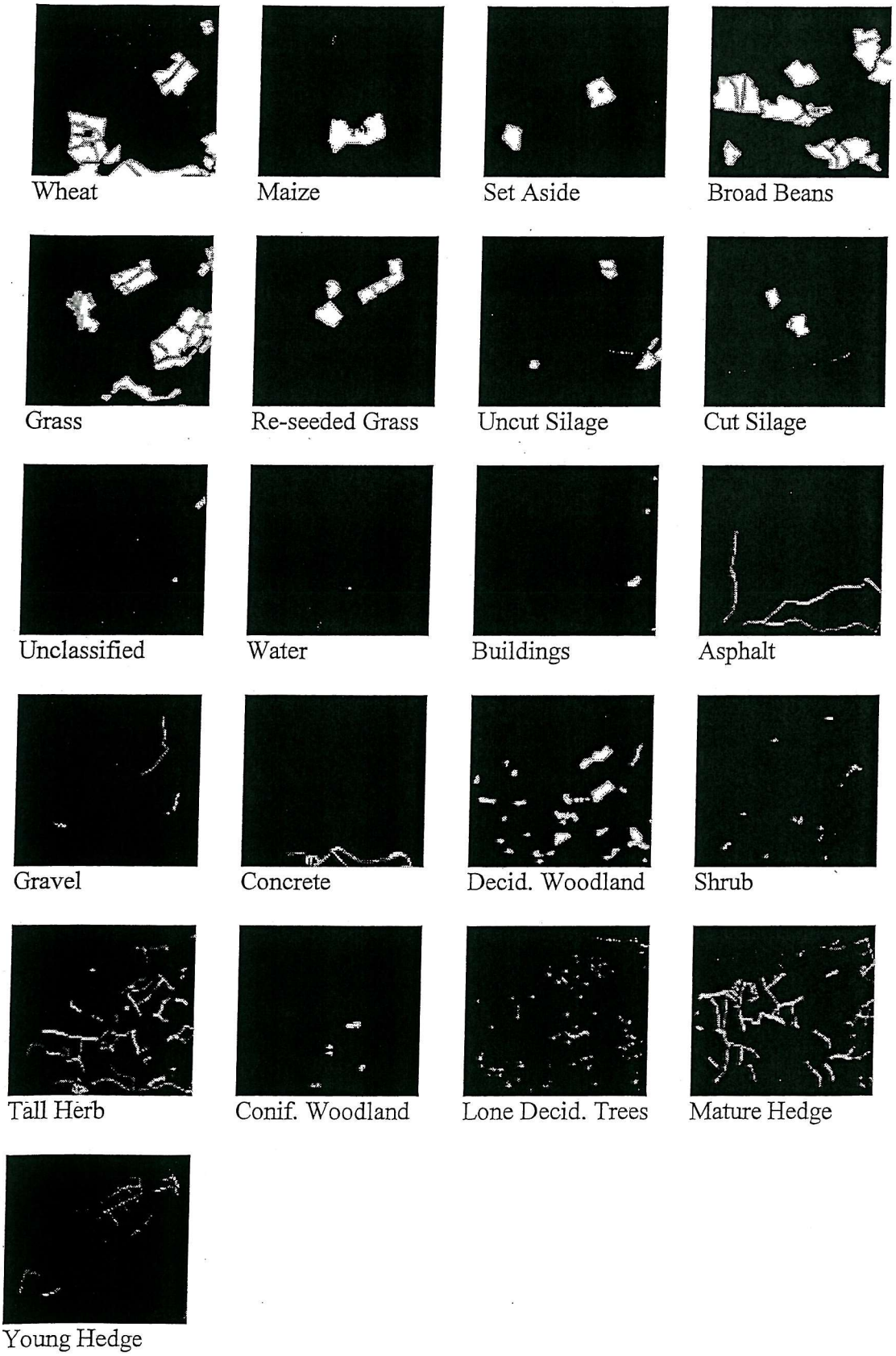


Figure 4.12 Target class area proportions for the image of Stoughton, UK.

4.4.2 Accuracy Assessment

Four measures of accuracy were estimated to assess the difference between each network prediction and the verification images. These were:

(i) Area Error Proportion (AEP) (equation 1.11)
(ii) Correlation Co-efficient (CC) (equation 1.12)
(iii) Closeness (S) (equation 1.10)
(iv) Root Mean Square Error (RMSE) (equation 1.13)

4.4.3 Network Settings

4.4.3.1 Constraint Weightings

The uncertainty and error inherent within data post-soft classification meant a change in constraint weightings could increase the accuracy of Hopfield network predictions. The effects of most of the uncertainty factors described in the introduction to section 4.4 can result in over or under estimation, for each pixel, of the predicted class area proportions output from soft classification. To assess the effect of this upon Hopfield network predictions, and formulate methods to minimise these errors, synthetic imagery was used. Specifically, the eight shape images shown in figure 3.2 were used. These were again degraded to 8 x 8 pixel images, as shown in figure 3.3, but were then altered to add error to the area proportions. For each pixel in each image, a new class proportion value was randomly drawn from a normal distribution with a mean of the original perfect class proportion, and a variance of 0.1. This led to many pixel class proportion estimates remaining unchanged with others over or under estimating class proportions. Figure 4.13 shows the new 8 x 8 pixel class proportion images, which can be compared to figure 3.3 to see the amount of error introduced.

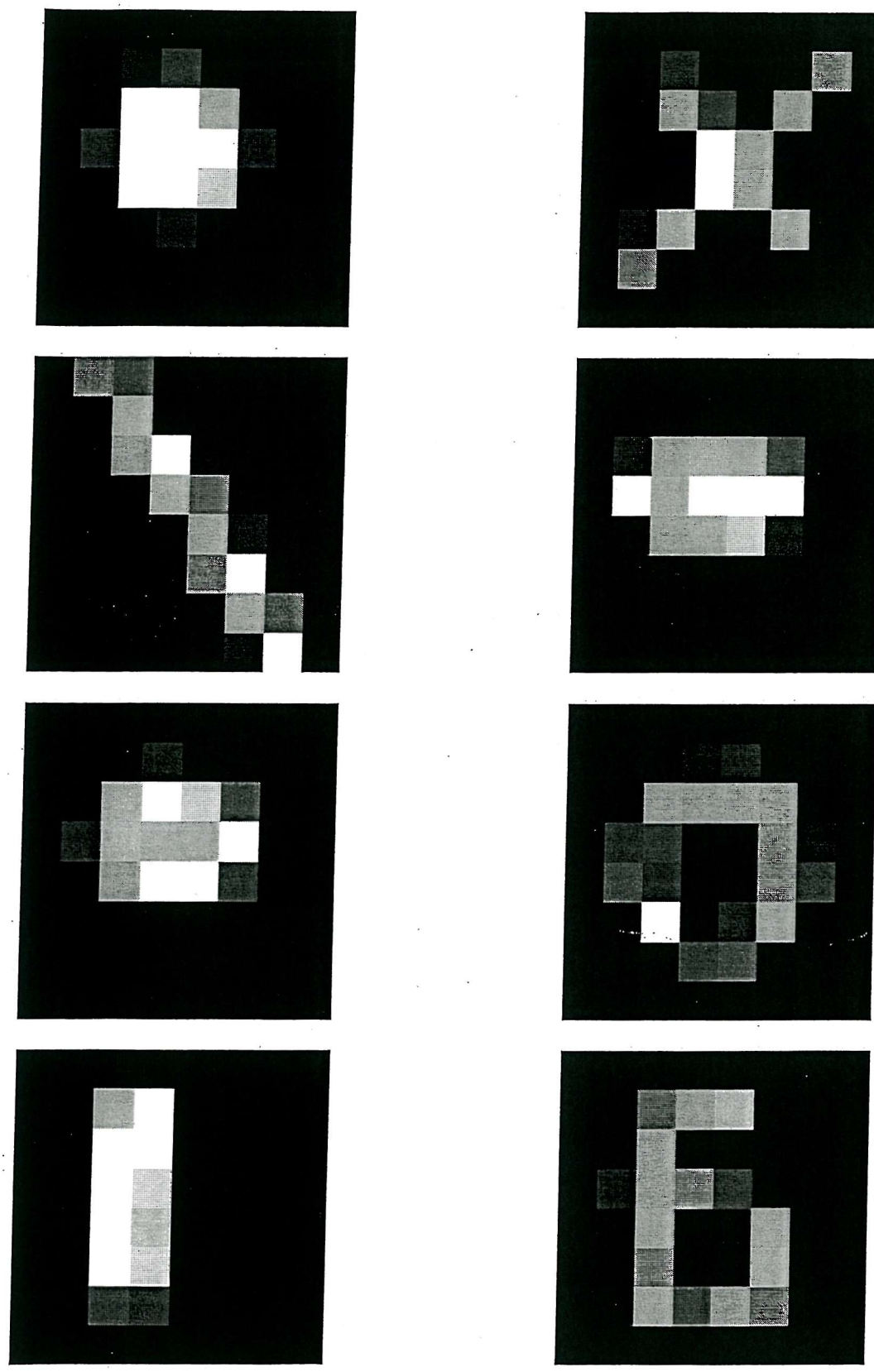
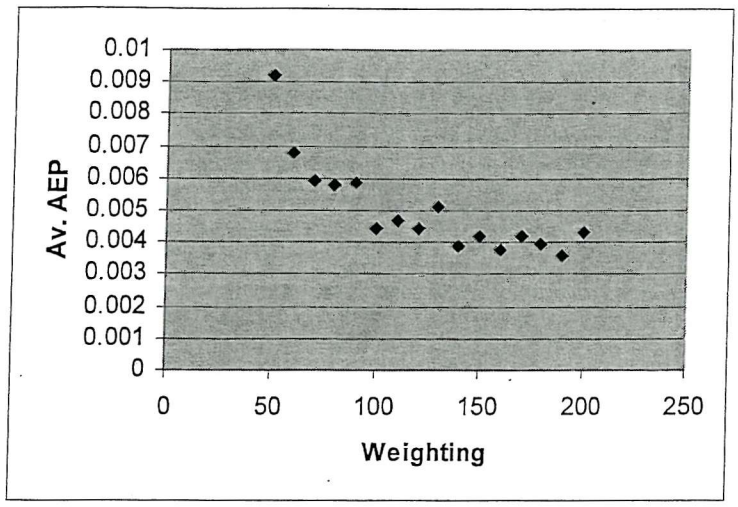


Figure 4.13 8x8 pixel class proportion images derived from the shape images in figure 3.3, with added noise.

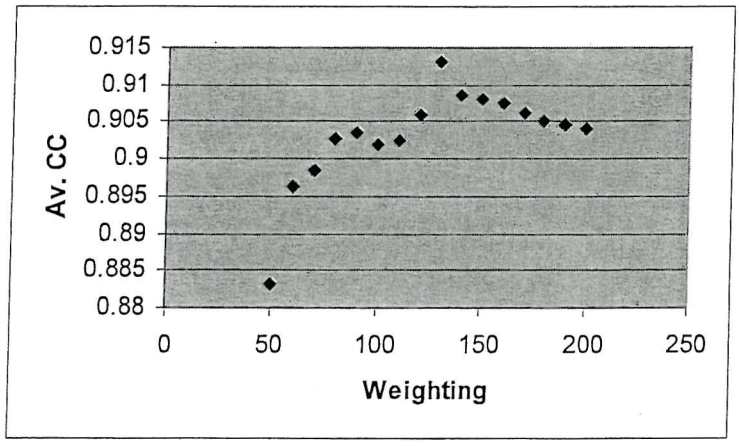
The synthetic class proportion images in figure 4.13 now displayed incorrect area proportions for the target shapes in figure 3.2. Therefore, if the network set-up used previously in this chapter, of $k_1 = k_2 = k_3 = k_4 = 150$, was applied, the network would attempt to maintain the class area proportion displayed in figure 4.13 and produce an inaccurate prediction. Consequently, it was assumed that a relaxation of the proportion constraint by reducing k_3, would produce more accurate results by allowing the goal and multi-class constraints to become the dominant functions. This assumption was tested by running the network at z=7 on each shape for k_3 values of 50 up to 200, while maintaining k_1, k_2 and k_4 at 150. Area error proportion (equation 1.11), correlation coefficient (equation 1.12) and root mean square error (equation 1.13) were calculated for each prediction image, in comparison to the target images in figure 3.2. The results for each k_3 value were averaged and plotted on graphs shown in figure 4.14.

Figure 4.14(a), as expected, shows the AEP decreasing as k_3 is increased, due to the greater effect that the proportion constraint is exerting. A levelling off at around $k_3 = 130$ suggests that the error in the class proportions means little increases in accuracy above this value can be made. Figures 4.14(b) and (c) demonstrate more conclusive findings. Both indicate a peak in the accuracy of shape recreation at $k_3 = 130$. Below this value, accuracy levels fall off because the proportion constraint is exerting too little an influence on the network energy function. However, above $k_3 = 130$, accuracy levels also fall off because the proportion constraint is exerting a strong enough effect as to maintain the class proportions of figure 4.13, and therefore, maintain the errors introduced.

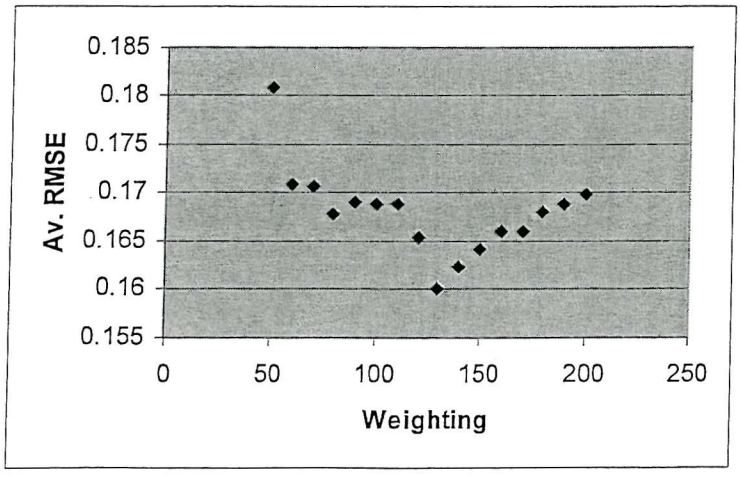
These results provide evidence to support the assumption that relaxation of the proportion constraint provides greater robustness to uncertainty and error. Therefore, values of $k_1 = k_2 = k_4 = 150$ and $k_3 = 130$ were used for the processing of Landsat TM imagery.



(a)



(b)



(c)

Figure 4.14 Graphs showing the effects of changing the Hopfield neural network proportion constraint value on (a) Area error proportion (AEP), (b) Correlation coefficient (CC), (c) Root mean square error (RMSE).

4.4.3.2 Other Network Settings

Aside from constraint weightings, similar network settings to the processing described for simulated imagery were maintained. To undertake a thorough investigation and assessment of network prediction on Landsat TM imagery, the network was run using zoom factors of z=3, z=5 and z=7, as defined in section 3.3.2.2. The network was initialised using the proportion constrained initialisation (section 3.3.3 (i)), and left to run for 10 000 iterations or until convergence, whichever came first.

4.4.4 Hard Classification

To evaluate the success of the Hopfield neural network technique, the traditional method of producing a land cover map from a satellite sensor image was carried out for comparison. This involved undertaking a 'hard' maximum likelihood classification of the imagery, as described in section 1.6.1 (iii). The verification data were used to identify representative, homogeneous areas for training data. Several individual training areas for each class were defined, where possible, containing a total of at least 100 pixels (Campbell, 1996). For the area of the Stoughton imagery under study, several classes were made up entirely of sub-pixel scale features, meaning no training pixels could be obtained. In these cases, training pixels from elsewhere in the imagery were used where homogeneous areas of each class were found.

4.4.5 Soft Classification

The soft classification of the Landsat TM imagery for both Sindos and Stoughton was undertaken using a regression based soft *k*-nearest neighbour (*k*nn) classifier, as described in section 1.6.3.4. Research undertaken for the FLIERS project (Fisher (ed), 2000) found the *k*nn regression classifier to be a robust area estimation model with generalisation abilities. In all cases, the results from the *k*nn model were found to be either comparable or better than those produced by various linear and non-linear neural network regression models.

The number of k nearest neighbours was determined by carrying out a classification of the Landsat TM imagery of both areas using k values of 2, 3, 4 and 5. The resultant area proportion predictions were then compared to class proportions derived from the verification data (figures

4.9 and 4.12), and accuracy statistics were calculated. These were area error proportion (equation 1.11), closeness (equation 1.10) and root mean square error (equation 1.13).

4.4.6 Results

4.4.6.1 Sindos

A soft *k*-nearest neighbour classification of the Sindos imagery was undertaken using *k* values of 2, 3, 4 and 5. The results of the accuracy assessment of each classification are shown in table 4.4, and indicate that *k*=3 produced the most accurate result, with the lowest closeness and RMS error values shown. This classification therefore provided the input class area proportions to the Hopfield neural network. Figure 4.15 shows the resultant proportion images, and comparison with those proportion images derived from the verification data in figure 4.9, suggests that error has been introduced by the classification process.

		k value			
		2	3	4	5
Sindos	AEP	0.00061	0.00065	0.00076	0.00081
	S	0.00327	0.00319	0.00322	0.00329
	RMSE	0.0569	0.0565	0.057	0.0573

Table 4.4 Overall accuracy statistics for the soft *k*-nearest neighbour classification of the image of Sindos, Greece, using different *k* values.

The *k*nn-derived proportions shown in figure 4.15 were input to the Hopfield neural network, which was run using settings of $k_1 = k_2 = k_4 = 150$ and $k_3 = 130$, as explained in section 4.4.3.1. After 10 000 iterations of the network with a zoom factor of *z*=3 (approximately 40 minutes running time on a P2-350 computer), a prediction map, shown in figure 4.16, was produced, with spatial resolution 3 times finer than that of the input class proportions in figure 4.15 The new map shown in figure 4.16, therefore, displays a spatial resolution of 10 m. The map was compared to the verification data (figure 4.8), and accuracy statistics were calculated, per-class and for the whole image, to assess the network performance. These are shown in table 4.5.

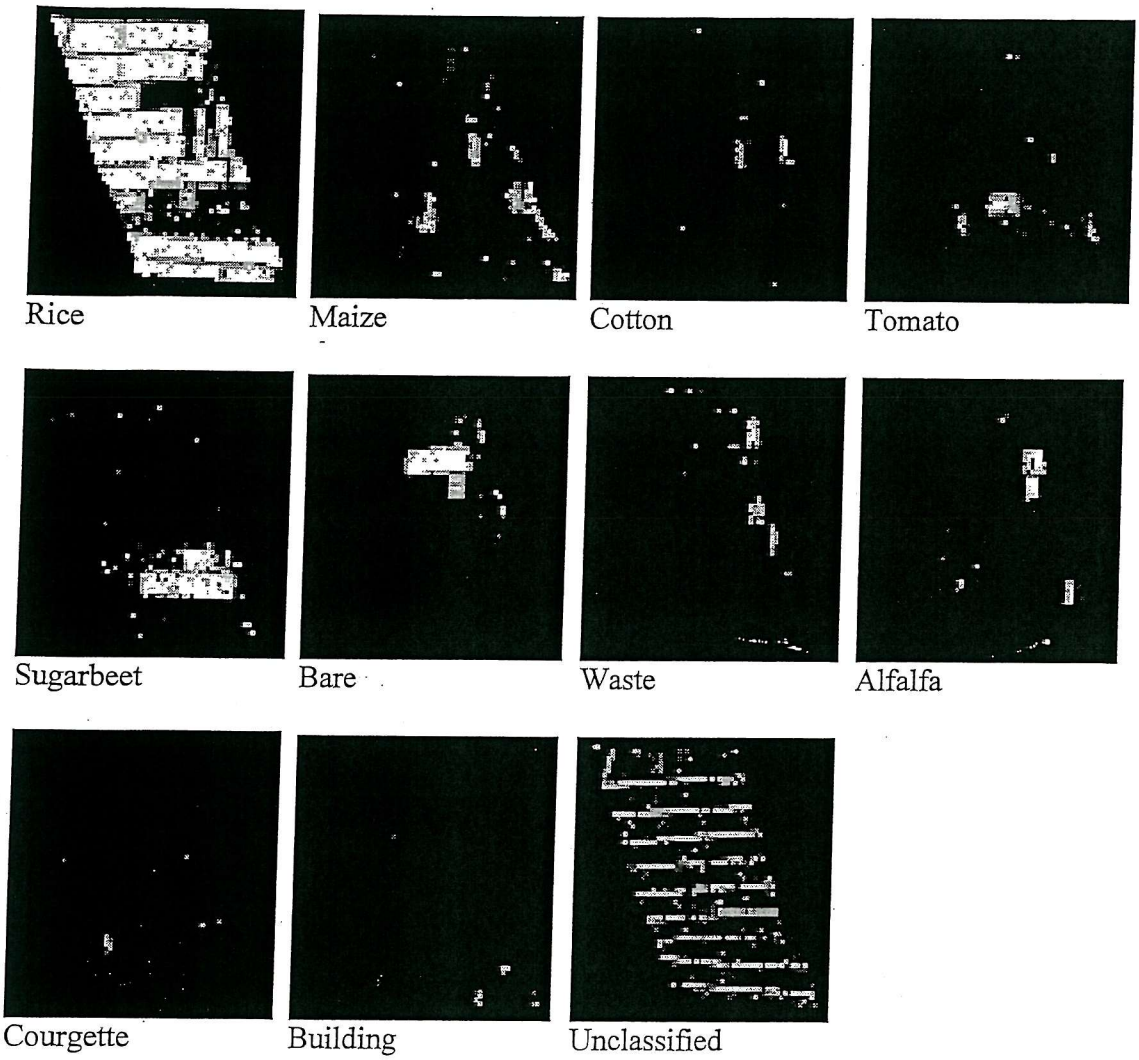


Figure 4.15 Predicted class area proportions for the image of Sindos, Greece, derived using a soft *k*-nearest neighbour classification.

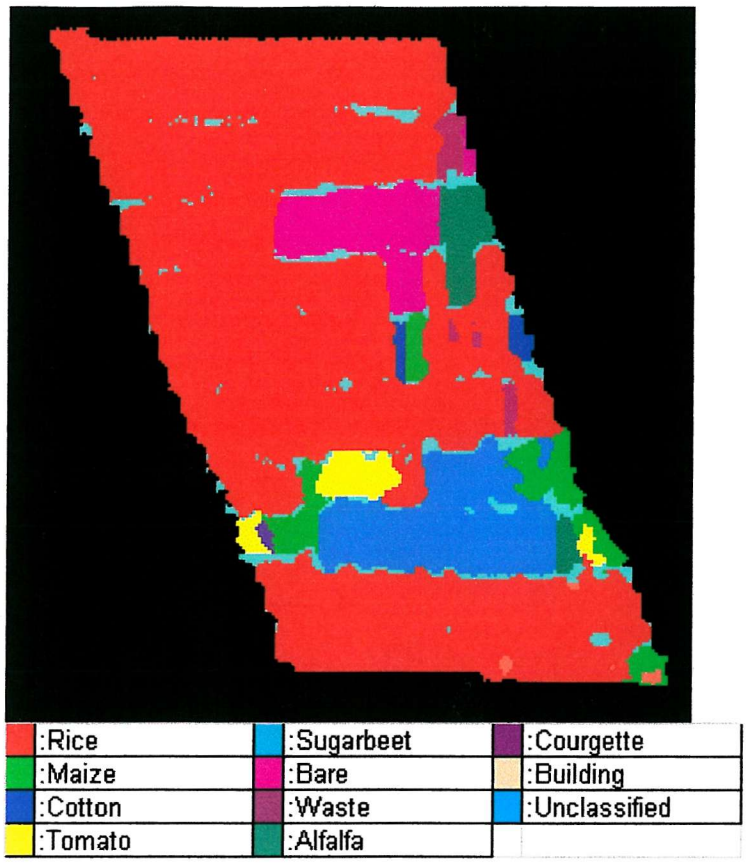


Figure 4.16 Hopfield neural network classification with z=3 for the image of Sindos, Greece (spatial resolution = 10 m).

	CC	AEP	S	RMSE
Rice	0.824	-0.0263	0.0799	0.2826
Maize	0.5412	0.2872	0.0177	0.1331
Cotton	0.5498	0.391	0.0037	0.061
Tomato	0.5427	0.2929	0.0089	0.0946
Sugarbeet	0.8682	-0.0537	0.0107	0.1033
Bare	0.8964	-0.0433	0.0053	0.0731
Waste	0.6544	0.3754	0.0052	0.072
Alfalfa	0.8145	0.1071	0.0045	0.0668
Courgette	0.5252	0.2004	0.0009	0.0303
Buildings	0.2722	0.6391	0.0016	0.0403
Unclassified	0.2466	0.6758	0.0474	0.2178
Entire Image		0.00559	0.016898	0.129993

Table 4.5 Per-class and overall accuracy statistics for the Hopfield neural network classification of the image of Sindos, Greece, with z=3.

The same procedure was undertaken for zoom factors of z=5 (approximately 100 minutes running time), and z=7 (approximately 210 minutes running time). For z=5, a map displaying spatial resolution of 6 m was produced (figure 4.17), and accuracy statistics were calculated (table 4.6). For z=7, a map of spatial resolution 4.3 m was produced (figure 4.18), and the corresponding accuracy statistics are shown in table 4.7 In addition, the maximum likelihood hard classification undertaken, produced the map shown in figure 4.19 (spatial resolution of 30 m), and corresponding accuracy statistics are shown in table 4.8. Also, the proportions output from the *k*nn classification (figure 4.15) were compared to the verification data (figure 4.8) using accuracy statistics, shown in table 4.9. This provided a benchmark to assess the success of the Hopfield neural network in locating the class proportions correctly compared to the original proportions.

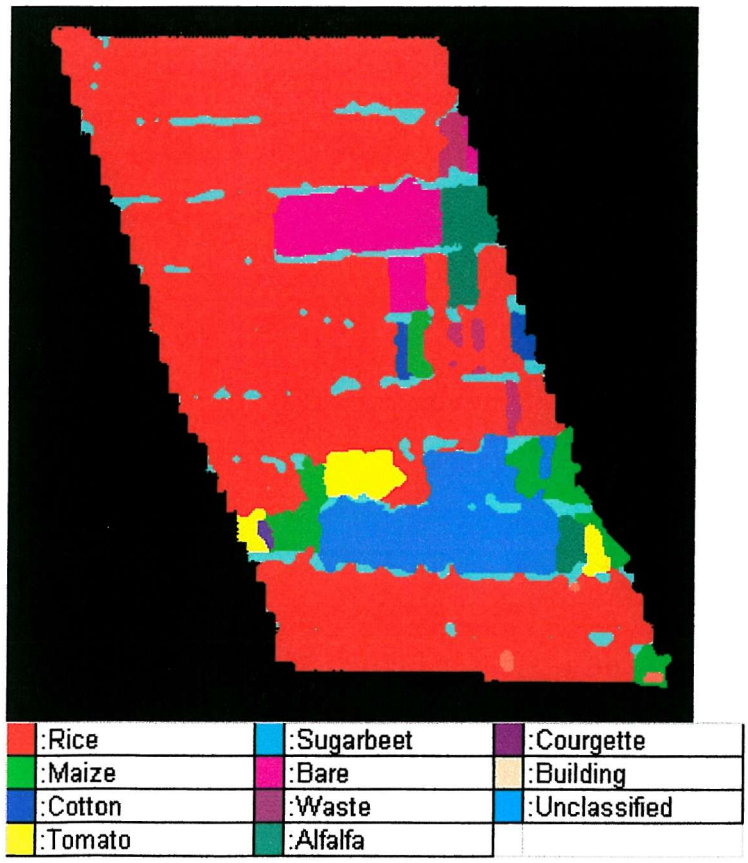


Figure 4.17 Hopfield neural network classification with z=5 for the image of Sindos, Greece (spatial resolution = 6 m).

	CC	AEP	S	RMSE
Rice	0.8252	0.0007	0.0787	0.2806
Maize	0.5317	0.2909	0.018	0.1343
Cotton	0.553	0.3673	0.0037	0.0611
Tomato	0.5251	0.2642	0.0094	0.0969
Sugarbeet	0.8614	-0.0491	0.0112	0.1058
Bare	0.9034	-0.0102	0.0049	0.07
Waste	0.701	0.2613	0.0047	0.0684
Alfalfa	0.8165	0.1157	0.0044	0.0664
Courgette	0.5393	0.161	0.0009	0.0301
Buildings	0.3576	0.4663	0.0016	0.0398
Unclassified	0.3577	0.4789	0.0452	0.2127
Entire Image		0.003721	0.016619	0.128915

Table 4.6 Per-class and overall accuracy statistics for the Hopfield neural network classification of the image of Sindos, Greece, with z=5.

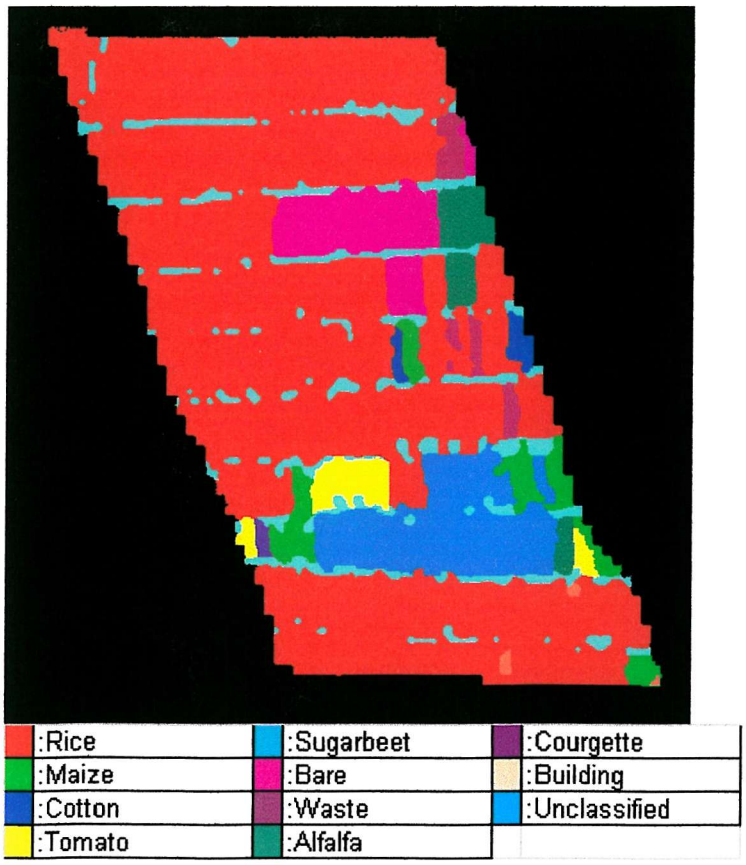


Figure 4.18 Hopfield neural network classification with z=7 for the image of Sindos, Greece (spatial resolution = 4.3 m).

	CC	AEP	S	RMSE
Rice	0.8297	0.021	0.0764	0.2764
Maize	0.6778	0.1946	0.013	0.1142
Cotton	0.5915	0.4271	0.0034	0.0579
Tomato	0.7367	0.1344	0.0056	0.0746
Sugarbeet	0.8725	-0.0258	0.0102	0.1008
Bare	0.9011	-0.0243	0.0051	0.0711
Waste	0.7257	0.1928	0.0044	0.0664
Alfalfa	0.8335	-0.0907	0.0039	0.0621
Courgette	0.6272	-3.372	0.0007	0.0262
Buildings	0.4145	0.5172	0.0014	0.0377
Unclassified	0.4328	0.3713	0.0423	0.2056
Entire Image		0.003774	0.015108	0.122913

Table 4.7 Per-class and overall accuracy statistics for the Hopfield neural network classification of the image of Sindos, Greece, with z=7.

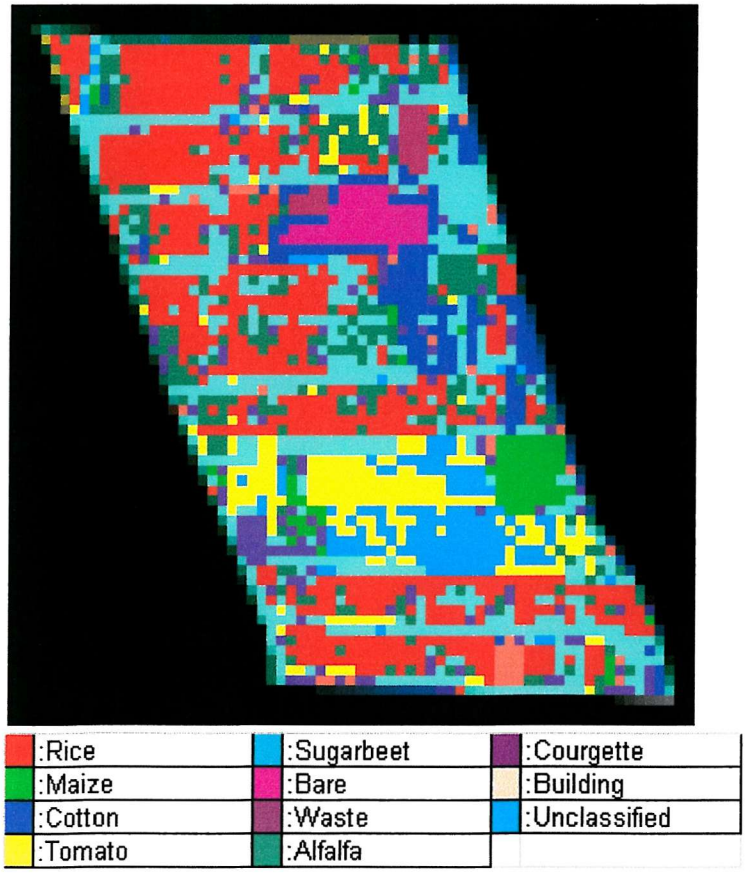


Figure 4.19 Maximum likelihood classification for the image of Sindos, Greece (spatial resolution = 30 m).

	CC	AEP	S	RMSE
Rice	0.5023	0.5731	0.2201	0.4692
Maize	0.3409	0.2726	0.0254	0.1594
Cotton	0.2038	-6.1056	0.0347	0.1862
Tomato	0.3241	-2.8633	0.0402	0.2005
Sugarbeet	0.5969	0.2542	0.0283	0.1683
Bare	0.6283	0.5329	0.0155	0.1247
Waste	-0.0076	0.2771	0.0152	0.1234
Alfalfa	0.0867	-4.1319	0.0719	0.2681
Courgette	0.1372	-26.6217	0.0289	0.1701
Buildings	0.2283	-4.0953	0.0079	0.0891
Unclassified	0.1241	-1.7094	0.1436	0.3789
Entire Image		0.041759	0.057437	0.239661

Table 4.8 Per-class and overall accuracy statistics for the maximum likelihood classification of the image of Sindos, Greece.

	CC	AEP	S	RMSE
Rice	0.2788	0.0798	0.0778	0.2788
Maize	0.5273	0.1859	0.0174	0.1317
Cotton	0.4319	0.3595	0.0043	0.0654
Tomato	0.3622	0.1126	0.0123	0.111
Sugarbeet	0.8226	-0.0048	0.0132	0.1151
Bare	0.8261	0.0534	0.0082	0.0907
Waste	0.4536	-0.0668	0.0088	0.0939
Alfalfa	0.7124	0.187	0.0064	0.0799
Courgette	0.3415	-0.3507	0.0013	0.0362
Buildings	0.1445	-0.3526	0.0028	0.053
Unclassified	0.4635	0.1896	0.0355	0.1883
Entire Image		0.004449	0.017089	0.130723

Table 4.9 Per-class and overall accuracy statistics for the *k*-nearest neighbour derived area proportions of the image of Sindos, Greece.

4.4.6.2 Stoughton

Soft classification of the Stoughton imagery was undertaken using the *k*-nearest neighbour area estimation model with *k* values of 2, 3, 4 and 5. The results of each classification are shown in table 4.10, and indicate that k=4 produced the most accurate result, with the lowest closeness and RMS error values shown. The classification produced with k=4 was, therefore, used to provide the input class area proportions to the Hopfield neural network. Figure 4.20 shows the

resultant proportion images, and comparison with the verification data-derived proportions in figure 4.12 suggests that the classification process has introduced a degree of error.

		k value			
		2	3	4	5
Stoughton	AEP	0.0004	0.0005	0.00053	0.00058
	S	0.0102	0.0095	0.0094	0.0097
	RMSE	0.1012	0.0974	0.097	0.0982

Table 4.10 Overall accuracy statistics for the soft *k*-nearest neighbour classification of the image of Stoughton, UK, using different *k* values.

The class proportions derived from the *k*-nearest neighbour classification shown in figure 4.20 were input to the Hopfield neural network, again run using settings of $k_1 = k_2 = k_4 = 150$ and $k_3 = 130$. After 10 000 iterations of the network with z=3 (approximately 90 minutes running time on a P2-350 computer), a prediction map, shown in figure 4.21, was produced with spatial resolution 3 times finer than that of the input class proportions in figure 4.20. The new map shown in figure 4.21, therefore displays a spatial resolution of 10 m. Accuracy statistics for each class, and the whole image, were calculated by comparing the map produced (figure 4.20) to the verification data (figure 4.11). These are shown in table 4.11.

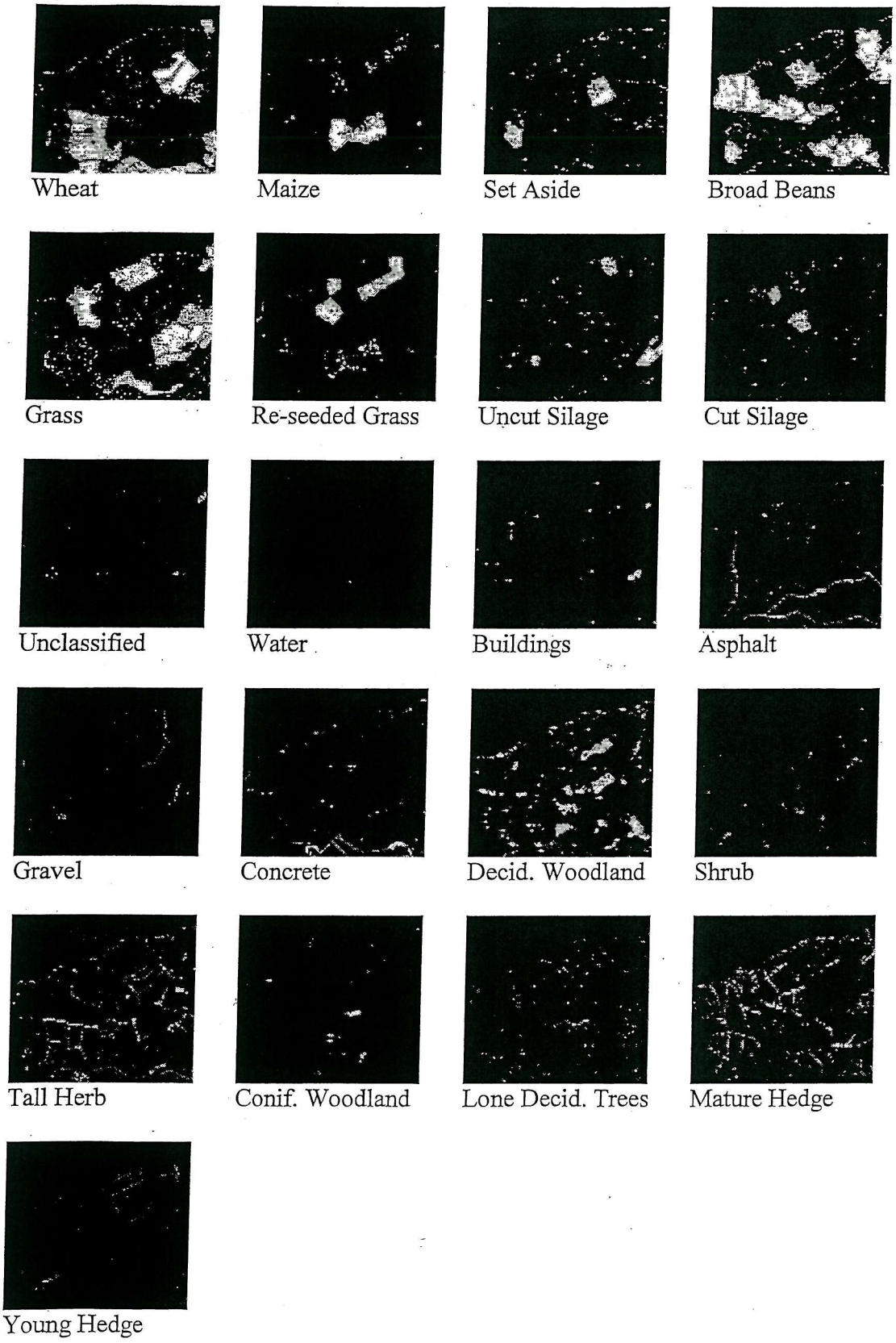


Figure 4.20 Predicted class area proportions for the image of Stoughton, UK, derived using soft *k*-nearest neighbour classification.

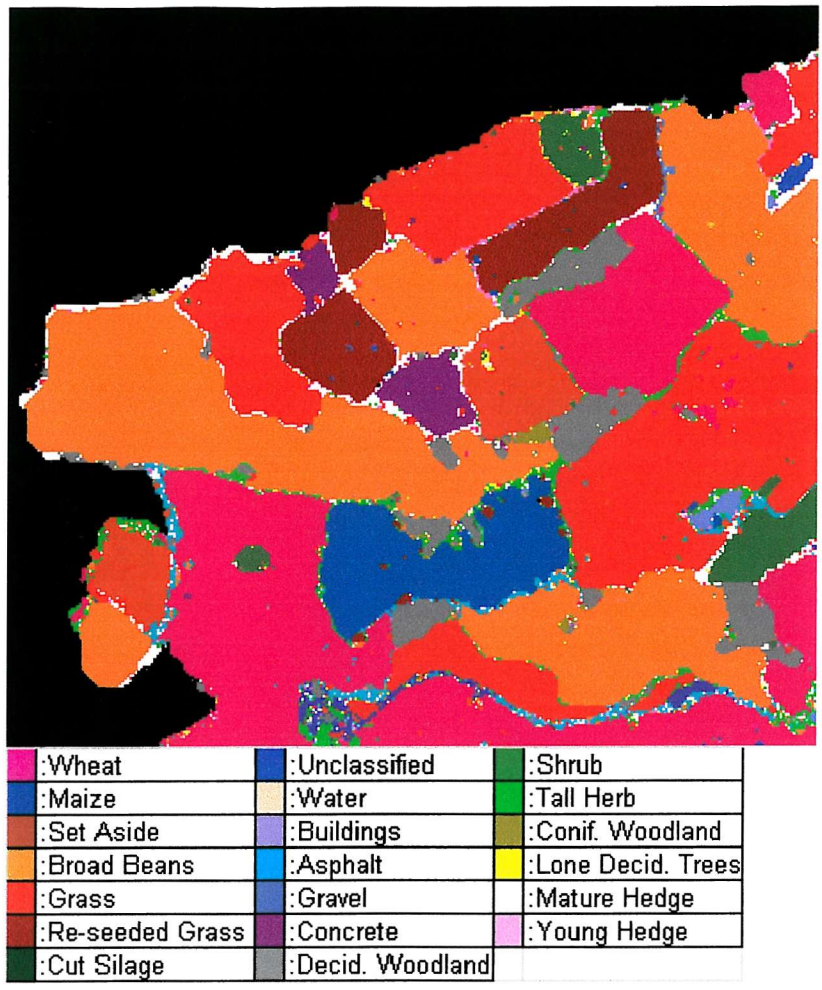


Figure 4.21 Hopfield neural network classification with z=3 for the image of Stoughton, UK (spatial resolution = 10 m).

The procedure was repeated for zoom factors of z=5 (approximately 220 minutes running time) and z=7 (approximately 510 minutes running time). For z=5, a map displaying spatial resolution of 6 m was produced (figure 4.22), and accuracy statistics were calculated (table 4.12). For z=7, a map of spatial resolution 4.3 m was produced (figure 4.23), and corresponding accuracy statistics are shown in table 4.13 In addition, figure 4.24 shows the map produced from maximum likelihood hard classification (spatial resolution of 30 m), with accuracy statistics shown in table 4.14. Also, the proportions output from the *k*nn classification (figure 4.20) were compared to the verification data (figure 4.11) using accuracy statistics, shown in table 4.15. This provided a benchmark to assess the success of the Hopfield neural network in locating the class proportions correctly compared to the original proportions.

	CC	AEP	S	RMSE
Wheat	0.8932	-0.1016	0.0258	0.1605
Maize	0.9191	-0.0215	0.0057	0.0755
Set Aside	0.876	-0.0487	0.0065	0.0805
Broad Beans	0.915	-0.0776	0.0262	0.1619
Grass	0.8822	-0.0809	0.028	0.1674
Re-seeded Grass	0.915	-0.0641	0.0065	0.0809
Uncut Silage	0.8448	0.003	0.0053	0.0728
Cut Silage	0.8402	0.0753	0.0042	0.65
Unclassified	0.7432	0.2715	0.0008	0.0284
Water	0.0816	0.3333	0.0001	0.0094
Buildings	0.6336	0.1245	0.0019	0.0432
Asphalt	0.2639	0.5548	0.0099	0.0994
Gravel	0.1022	0.6827	0.0026	0.0506
Concrete	0.3404	0.3527	0.007	0.0839
Deciduous Woodland	0.7538	0.1377	0.0159	0.1262
Shrub	0.1894	0.5397	0.0029	0.0542
Tall Herb	0.2632	0.5739	0.0227	0.1507
Coniferous Woodland	0.5399	0.2651	0.0018	0.042
Lone Deciduous Trees	0.1187	0.5475	0.0065	0.0805
Mature Hedge	0.2169	0.6978	0.0261	0.1615
Young Hedge	0.0975	0.7341	0.0049	0.0699
Entire Image		0.004541	0.010062	0.10031

Table 4.11 Per-class and overall accuracy statistics for the Hopfield neural network classification of the image of Stoughton, UK, with z=3.

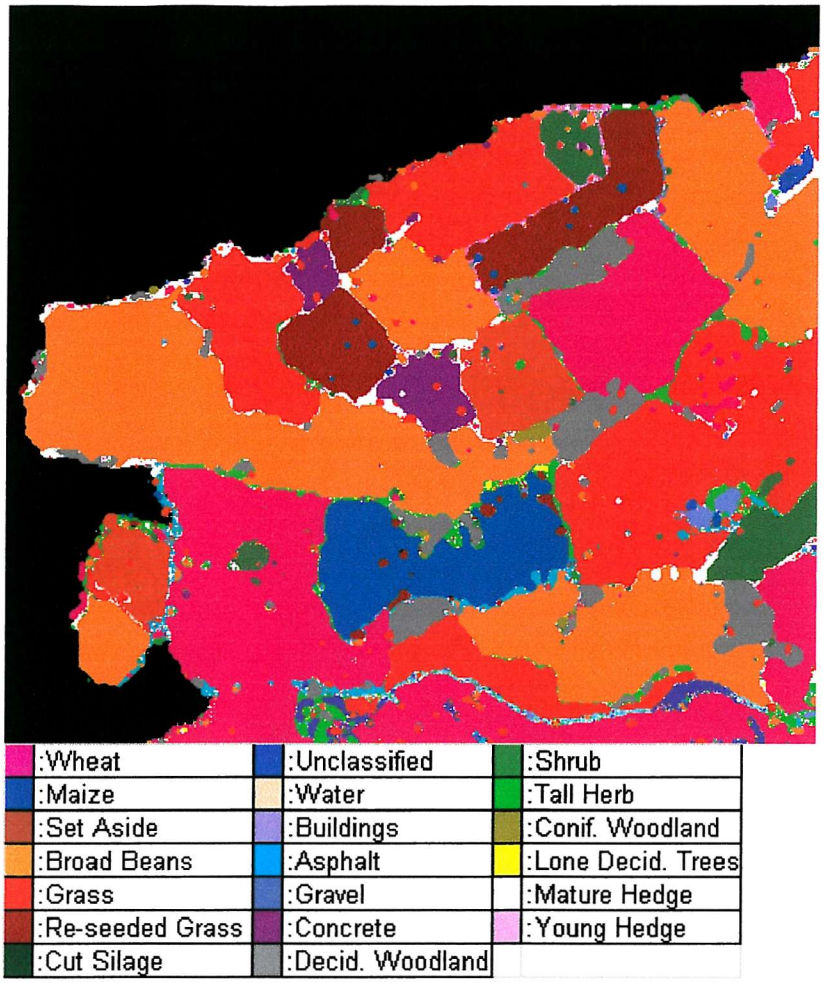
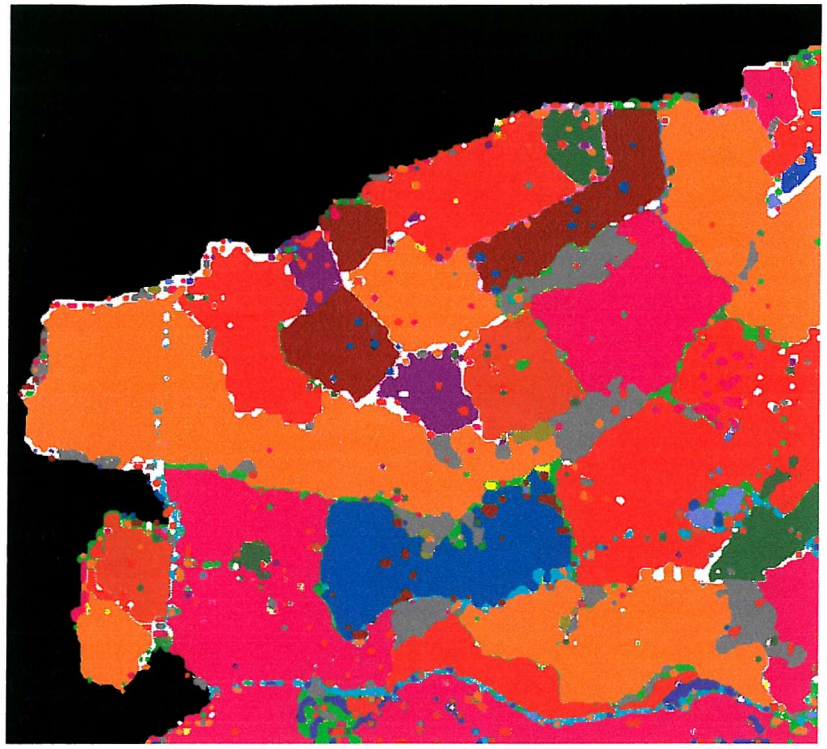


Figure 4.22 Hopfield neural network classification with z=5 for the image of Stoughton, UK (spatial resolution = 6 m).

	CC	AEP	S	RMSE
Wheat	0.8844	-0.0856	0.0276	0.1661
Maize	0.9125	-0.0091	0.0061	0.0782
Set Aside	0.8597	-0.0521	0.0073	0.0857
Broad Beans	0.9139	-0.0701	0.0264	0.1626
Grass	0.8742	-0.0725	0.0298	0.1725
Re-seeded Grass	0.8995	-0.0666	0.0077	0.088
Uncut Silage	0.8272	-0.0168	0.006	0.0772
Cut Silage	0.833	0.0561	0.0044	0.0667
Unclassified	0.7572	0.1468	0.0008	0.0282
Water	0.1885	0.5	0.0001	0.0085
Buildings	0.6247	0.0844	0.0019	0.0441
Asphalt	0.2828	0.4646	0.0101	0.1007
Gravel	0.1382	0.6845	0.0025	0.0497
Concrete	0.3914	0.2147	0.007	0.0835
Deciduous Woodland	0.7676	0.0832	0.0154	0.124
Shrub	0.284	0.4484	0.0028	0.0525
Tall Herb	0.3238	0.491	0.0219	0.1481
Coniferous Woodland	0.5951	0.1788	0.0016	0.0402
Lone Deciduous Trees	0.1366	0.5475	0.0064	0.0797
Mature Hedge	0.268	0.6173	0.0257	0.1604
Young Hedge	0.099	0.7512	0.0048	0.0695
Entire Image		0.003926	0.010303	0.101505

Table 4.12 Per-class and overall accuracy statistics for the Hopfield neural network classification of the image of Stoughton, UK, with z=5.

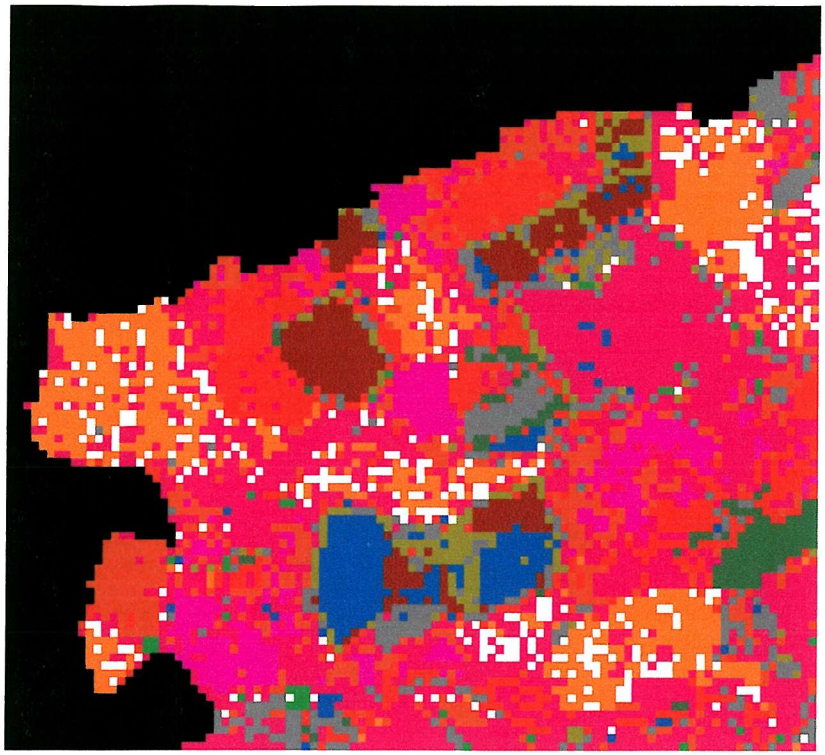


:Wheat	:Unclassified	:Shrub
:Maize	:Water	:Tall Herb
:Set Aside	:Buildings	:Conif. Woodland
:Broad Beans	:Asphalt	:Lone Decid. Trees
:Grass	:Gravel	:Mature Hedge
:Re-seeded Grass	:Concrete	:Young Hedge
:Cut Silage	:Decid. Woodland	

Figure 4.23 Hopfield neural network classification with z=7 for the image of Stoughton, UK (spatial resolution = 4.3 m).

	CC	AEP	S	RMSE
Wheat	0.8766	-0.0679	0.0291	0.1706
Maize	0.9025	-0.0069	0.0068	0.0825
Set Aside	0.8437	-0.0637	0.0082	0.0908
Broad Beans	0.9094	-0.0569	0.0276	0.1661
Grass	0.8639	-0.0616	0.032	0.1789
Re-seeded Grass	0.8914	-0.0608	0.0083	0.0913
Uncut Silage	0.82	-0.0167	0.0062	0.0787
Cut Silage	0.8248	0.0752	0.0046	0.068
Unclassified	0.7768	0.1634	0.0007	0.027
Water	0.527	0.6	0	0.0065
Buildings	0.6249	0.0408	0.002	0.0446
Asphalt	0.3198	0.4437	0.0098	0.0988
Gravel	0.2336	0.6618	0.0023	0.0477
Concrete	0.3979	0.26	0.00678	0.0821
Deciduous Woodland	0.7562	0.0822	0.0161	0.127
Shrub	0.3169	0.5659	0.0025	0.0498
Tall Herb	0.3343	0.495	0.0216	0.1469
Coniferous Woodland	0.6099	0.2153	0.0015	0.0392
Lone Deciduous Trees	0.1778	0.6892	0.0056	0.0747
Mature Hedge	0.2965	0.5859	0.0253	0.1589
Young Hedge	0.1086	0.8343	0.0045	0.0673
Entire Image		0.003678	0.010548	0.102704

Table 4.13 Per-class and overall accuracy statistics for the Hopfield neural network classification of the image of Stoughton, UK, with z=7.



:Wheat	:Unclassified	:Shrub
:Maize	:Water	:Tall Herb
:Set Aside	:Buildings	:Conif. Woodland
:Broad Beans	:Asphalt	:Lone Decid. Trees
:Grass	:Gravel	:Mature Hedge
:Re-seeded Grass	:Concrete	:Young Hedge
:Cut Silage	:Decid. Woodland	

Figure 4.24 Maximum likelihood classification for the image of Stoughton, UK (spatial resolution = 30 m).

	CC	AEP	S	RMSE
Wheat	0.3535	-0.6584	0.1934	0.4398
Maize	0.6018	0.2243	0.0251	0.1583
Set Aside	0.1731	-1.2608	0.0692	0.2631
Broad Beans	0.5766	0.541	0.1113	0.3336
Grass	0.3218	0.3941	0.1312	0.3622
Re-seeded Grass	0.6534	0.1279	0.0243	0.1558
Uncut Silage	0.4788	0.1693	0.0164	0.1281
Cut Silage	0.2838	-2.1106	0.0423	0.2057
Unclassified	N/A	1	0.0018	0.0419
Water	N/A	1	0.0001	0.0076
Buildings	N/A	1	0.0027	0.0521
Asphalt	N/A	1	0.0091	0.0954
Gravel	N/A	1	0.0021	0.0462
Concrete	N/A	1	0.0064	0.0802
Deciduous Woodland	0.2113	-0.3595	0.0638	0.2525
Shrub	0.0235	-0.8571	0.0068	0.0825
Tall Herb	0.0373	-0.0824	0.0416	0.204
Coniferous Woodland	N/A	1	0.0022	0.0467
Lone Deciduous Trees	N/A	1	0.005	0.0709
Mature Hedge	0.0135	-0.6399	0.0626	0.2502
Young Hedge	N/A	1	0.0042	0.0648
Entire Image		0.018114	0.039122	0.197792

Table 4.14 Per-class and overall accuracy statistics for the maximum likelihood classification of the image of Stoughton, UK.

LIBRARY

	CC	AEP	S	RMSE
Wheat	0.8833	-0.0072	0.0251	0.1584
Maize	0.8998	0.0039	0.0066	0.0815
Set Aside	0.8475	-0.0689	0.0073	0.0853
Broad Beans	0.914	-0.0152	0.0244	0.1562
Grass	0.8668	-0.0171	0.0284	0.1686
Re-seeded Grass	0.886	-0.0317	0.008	0.0895
Uncut Silage	0.8177	-0.0402	0.0057	0.0754
Cut Silage	0.8129	0.0209	0.0046	0.0681
Unclassified	0.7047	-0.0765	0.0009	0.0298
Water	0.5545	0.2698	0	0.0064
Buildings	0.62	-0.0772	0.0017	0.0414
Asphalt	0.4287	0.0078	0.0074	0.086
Gravel	0.4026	-0.1234	0.0018	0.0424
Concrete	0.4949	-0.1428	0.0049	0.0698
Deciduous Woodland	0.7456	0.0058	0.0153	0.1236
Shrub	0.4044	-0.0231	0.0021	0.0454
Tall Herb	0.4422	0.0031	0.0168	0.1296
Coniferous Woodland	0.6008	-0.0884	0.0014	0.0376
Lone Deciduous Trees	0.3622	-0.0986	0.0044	0.066
Mature Hedge	0.4198	0.075	0.02	0.1414
Young Hedge	0.3196	0.18	0.0038	0.0613
Entire Image		0.000731	0.009073	0.095254

Table 4.15 Per-class and overall accuracy statistics for the *k*-nearest neighbour derived area proportions of the image of Stoughton, UK.

4.4.7 Discussion

4.4.7.1 Sindos

Results produced for the Landsat TM imagery of Sindos suggest that the Hopfield neural network super-resolution technique provides greater mapping accuracy than traditional approaches, in addition to producing finer spatial resolution land cover maps.

Figure 4.19 shows some of the problems commonly associated with traditional supervised hard classification. Although training pixels were carefully chosen to represent homogeneous areas, using the verification data as a guide, a large number of pixels were still mis-classified. The problem mainly lies in the assumption that each pixel contains one 'pure' land cover class with a homogenous reflectance across it. This is of course a false assumption since, within every 30 m

x 30 m instantaneous field of view, either a mixture of classes will occur, or if the area does contain one class, within-class variation may make it spectrally different from the training data for that class. The majority of mis-classification within figure 4.19 occurred at the boundaries between classes where large numbers of mixed pixels existed. Because the Sindos scene is made up of a patchwork of small fields, the majority of pixels in the imagery were boundary pixels, resulting in a high overall RMS error of 0.24, and closeness of 0.57. Table 4.8 shows that the most accurate results in the maximum likelihood classification were for the larger classes (e.g., Rice, Sugarbeet, Bare), with correlation coefficients of above 0.5 due to these classes having a larger number of 'pure' pixels compared to boundary pixels of mixed composition.

Whereas figure 4.19 and table 4.8 highlight the problems of attempting to assign each pixel to a single land cover class, particularly in imagery of the nature of Sindos, table 4.15 shows that accounting for mixed pixels increases accuracy greatly. The result of classifying the Landsat TM imagery using a soft *k*-nearest neighbour regression model, and comparing to the verification data, is an overall RMS error of 0.13, a decrease of 0.11 over the hard classification. The overall AEP value of 0.004 also shows how land cover class proportions have been predicted with increased accuracy, compared to the maximum likelihood prediction of 0.042. Table 4.15 can now be used as a benchmark to assess how effectively the Hopfield neural network approach can actually map these class proportions within pixels.

Figures 4.16, 4.17 and 4.18 show the results of the Hopfield neural network predictions for zoom factors of 3, 5 and 7 respectively. Visual comparison with the verification map in figure 4.8, clearly suggests that using z=7 (figure 4.18) produces more accurate results. Whereas all three maps (figures 4.16, 4.17 and 4.18) appear to have mapped the fields accurately, the linear features of the 'unclassified' class become more accurately represented as the zoom factor is increased. This fact is clearly reflected in the accuracy statistics of tables 4.5, 4.6 and 4.7 As the zoom factor is increased, the precision with which the unclassified class is mapped also increases, with a correlation coefficient of 0.25 for z=3, 0.36 for z=5 and 0.43 for z=7. Although such large increases in mapping precision are not shown for the other classes in tables 4.5, 4.6 and 4.7, as zoom factor is increased, all classes show a decrease in closeness and RMS error values, reflected in the overall image values.

Comparison of table 4.9 with tables 4.5, 4.6 and 4.7 shows that all three zoom factor settings provide an increase in mapping accuracy over the class proportions used to initialise the Hopfield neural network. This indicates that the Hopfield neural network is actually locating the class proportions in the correct places within pixels. A comparison of table 4.9 with the best Hopfield neural network prediction (table 4.7), shows that in all classes except 'unclassified', the network provides a decrease in RMS error, and overall, a decrease of 0.008. Although this is only a small increase in overall accuracy, it should be noted that the soft class proportions assessed in table 4.9 can only be viewed as 11 separate maps, whereas the Hopfield neural network prediction provides a single thematic map. This is far more useful and acceptable to decision makers, GIS input, and industry as a whole. Traditional techniques of single thematic map production from satellite sensor imagery involve hard classification, and examination of figures 4.18 and 4.19 suggests that the Hopfield neural network provides a more accurate approach. Tables 4.7 and 4.8 confirm this, as the network prediction shows overall closeness and RMS error values of 0.015 and 0.123 respectively, compared to 0.057 and 0.24 for the maximum likelihood result. In addition to this, the AEP results confirm how accurate class proportion estimates from the *k*nn classifier have been maintained within the network prediction (0.0038 in table 4.7). In contrast, by forcing each 30 m x 30 m pixel to represent a single land cover class, the hard classification produced an overall AEP value of just 0.0418.

4.4.7.2 Stoughton

Results produced for the Landsat TM imagery of Stoughton suggest that the Hopfield neural network super-resolution technique produces increased mapping accuracy over traditional approaches, but suffers from problems when presented with complex data. Throughout the various classification procedures it was clear that the nature of the imagery provided a stern test of the capabilities of the classifiers and Hopfield neural network. Figure 4.12 shows that over half of the land cover classes identified through field survey and aerial photograph interpretation consisted of partly or entirely sub-pixel scale features. Matters were complicated further by the fact that many of these classes were located in close proximity to one another (e.g. the Asphalt, Concrete and Tall Herb classes), producing many pixels at field borders representing mixtures of up to six classes. For certain purposes it may have been sensible to merge some classes to simplify the image processing task, but as this chapter focuses on an exploration of the capabilities of the Hopfield neural network approach, it was decided to leave the situation unchanged.

Figure 4.24 shows how the hard classification map appears to perform well in mapping the relatively 'pure' pixels in the large fields, but suffers at field boundaries. This is an expected result as, for example, when attempting to assign a pixel containing a mixture of 6 different land cover classes to a single class, error will inevitably occur. This mapping error is reflected in the accuracy statistics in table 4.14. The first point to note is that nine of the classes made up of mainly sub-pixel scale features have not been identified in the maximum likelihood classification. This is due to the fact that within the Stoughton imagery under study, none of the features from these classes covered a large enough area of a pixel to warrant assigning a whole pixel to. Elsewhere in figure 4.24 the maximum likelihood classifier mapped the larger classes more precisely, with correlation coefficients of above 0.6 for the Maize and Re-seeded Grass classes. However, the complex nature of the imagery meant that the majority of pixels represented a mixed response. As described in section 4.4.7.1 for the Sindos image, this is not ideal for obtaining accurate maps via hard classification, producing high overall closeness and RMS error values of 0.039 and 0.198 respectively for the hard classification of the Stoughton imagery.

Whereas figure 4.24 and table 4.14 highlight the problems of attempting to assign mixed pixels to a single land cover class, table 4.15 shows again that accounting for mixed pixels produces an increase in classification accuracy. By classifying the same Landsat TM image of Stoughton using a soft *k*nn classifier, an overall decrease in RMS error of 0.105 over the hard classification is produced. Table 4.15 shows that the *k*nn classifier mapped precisely the larger classes with a correlation coefficient of over 0.8 in most cases. Also, whereas the maximum likelihood classifier failed to identify the classes made up of sub-pixel features, the *k*nn classifier identified them. In the case of these sub-pixel classes, the high precision measures should not be looked upon too seriously, as table 4.15 represents the results of a comparison of *soft class proportions* with the hard verification map shown in figure 4.11. The function of the Hopfield neural network technique is to process these proportions to predict their locations within each pixel, thereby increasing mapping accuracy.

Figures 4.21, 4.22 and 4.23 show the results of the Hopfield neural network predictions for zoom factors of 3, 5 and 7 respectively. Visual comparison, as with the Sindos results, appear to show the fields being mapped accurately, but with more detail at the field boundaries being

identified as the zoom factor is increased. However, whereas the Sindos results showed increased accuracy with increased zoom factor, tables 4.11, 4.12 and 4.13 show that overall, the opposite is true for Stoughton. As the zoom factor is increased, the RMS error increases from 0.1 for z=3, to 0.102 for z=5, to 0.103 for z=7. However, these statistics only tell part of the story of the effect of the zoom factors. Comparison of tables 4.11 and 4.13 reveals that for the classes made up of large fields, the network prediction with z=7 produces the lowest closeness and RMS error values. This shows a similar result to the Sindos imagery. Fields with dimensions larger than a pixel dominated the Sindos imagery, and as the zoom factor was increased, the Hopfield neural network could better approximate each field shape at the finer spatial resolution. This is also true for the classes representing large fields in the Stoughton imagery, e.g., Wheat, Maize, Set Aside, Broad Beans etc. all show increases in accuracy as the zoom factor is increased. However, the classes made up of sub-pixel scale features, e.g., trees or linear features with sub-pixel width, appear to present problems to the network, as all show a decrease in accuracy as zoom factor is increased. It can be seen from figure 4.23 that at z=7 more of these sub-pixel scale features have been included in the map, but the statistics in table 4.13 suggest that the majority of these features have been located incorrectly. This result means that overall, the more generalised map of figure 4.21 which has failed to map much of the field boundary detail shown in the verification map (fig 4.11), produces the most accurate map overall.

Comparison of the overall image statistics of table 4.15 with tables 4.11, 4.12 and 4.13 shows that all three Hopfield neural network predictions have produced lower accuracy predictions than the soft class proportions that each processed. Examination of the per-class statistics shows that for the large field classes, the Hopfield neural network approach has produced an increase in mapping accuracy over the class area proportions. This shows that for these classes, the network has been successful in predicting where, within each pixel, the classes are located. However, for those classes with sub-pixel dimensions, the network was not accurate in this result, which is understandable due to the nature of the goal functions. As described in section 4.2.1.1, the goal functions rely on utilising information from surrounding pixels to infer the location of class proportions at the finer spatial resolution and maximise spatial dependence. The classes made up of sub-pixel scale features in the Stoughton imagery have no such information available, leaving the Hopfield neural network to cluster the proportions in the centre of each pixel, a kind of 'blind guess', which as expected, has been shown to be inaccurate. A merging of such sub-pixel classes would have produced a much more manageable

problem with mostly features larger than the size of a pixel. This should have meant the network would then have produced a map of increased accuracy over the soft proportions, but as described before, this section was aimed at testing the network's capabilities. As it is, single thematic maps have been produced of very similar accuracy levels to the class area proportions each was derived from. Such thematic maps, as described previously, are more acceptable in the world of decision making and GIS.

Finally, comparison of the best Hopfield neural network prediction map in figure 4.21 with the hard classification in figure 4.24 shows that, despite suffering from problems due to the complex class make-up of the area, the network has produced a more accurate map. By comparing table 4.11 with table 4.14, the statistics show a large increase in accuracy for all classes, in addition to the overall figures. The AEP results also confirm how the accurate *k*nn-derived area proportion estimates have been maintained through the network processing, with an overall value of 0.0045. This compares favourably to the same statistics for the hard classification where, forcing each 30 x 30 m pixel to represent a single land cover class, resulted in a value of just 0.0181.

4.4.8 Applications

The success of the Hopfield network in mapping varying land cover features accurately from differing spatial resolution images, suggests future potential applications for the technique. The technique proposed here has the potential to be used for land cover mapping from imagery that has been derived from any remote sensing system. For example, this may be IKONOS MSS data with a spatial resolution of 4 m for fine detail urban feature mapping, hyper-spectral MERIS or MODIS data with a spatial resolution of 250-300 m, or even NOAA AVHRR data with a spatial resolution of 1.1 km.

4.5 Conclusions

This study has shown that a Hopfield neural network can be used to predict the location of class proportions within each pixel, to produce a sub-pixel scale land cover map. The technique presented here represents a robust, efficient and simple technique. Results from simulated and Landsat TM remotely sensed data show accurate prediction, suggesting that the Hopfield network has the potential to predict accurately land cover at the sub-pixel scale from operational satellite sensor imagery.

Chapter 5: Super-Resolution Land Cover Pattern Prediction Using a Hopfield Neural Network*

5.1. Introduction

The landscape is a complex, hierarchically organised, spatio-temporal mosaic, where there are strong relationships coupling spatial pattern to process (Lobo *et al.*, 1998). Land cover represents a critical biophysical variable that affects the functioning of terrestrial ecosystems in bio-geochemical cycling, hydrological processes and the interaction between surface and atmosphere (Cihlar *et al.*, 2000). It is therefore central to all scientific studies that aim to understand terrestrial dynamics at any scale. Spatial patterns of cover types in landscape systems are the result of an interaction among dynamical processes operating across a range of spatial and temporal scales (Lobo *et al.*, 1998). Such spatial issues have interested ecologists for a long time, and have been receiving increasing attention in recent years. To understand and test rigorously the effects of landscape pattern on ecological processes, models of spatial pattern have been developed. Such models attempt to capture a set of constraints dictating landscape pattern and assign the remaining pattern to a purely random process (Keitt, 2000). These were originally introduced to generate spatial patterns of land cover in the absence of any structuring process and can suffer from inadequacies due to their under-constrained nature. This chapter presents a technique to model land cover pattern using remotely sensed imagery and prior information on land cover distribution to constrain the predictions.

The spatial detail obtainable from many remote sensors means that within the imagery produced, some land cover features (e.g. individual trees or buildings) are smaller than a pixel. Consequently, while these features can be detected within pixels by soft classification techniques, there exist many difficulties in identifying them accurately and using them within models. Much previous research has centred on attempting to extract data on sub-pixel scale features from remote sensing imagery. For example, information has been extracted about sub-pixel scale volcano vents (Bhattacharya *et al.*, 1993), coal fires (Zhang *et al.*, 1997), glacial features (Smith *et al.*, 2000) and water storage ditches (Shepherd *et al.*, 2000). However, in all cases, the features have been merely detected using soft classification techniques, and no

* This chapter is based on Tatem *et al.* (2001d)

attempt at locating them has been made. This chapter demonstrates that it is possible to identify sub-pixel scale land cover targets *and* to recreate their spatial distribution across an image.

Chapters 3 and 4 presented techniques that allowed prior information to be included in the sub-pixel mapping process. By utilising information contained in surrounding pixels, the land cover within each pixel was mapped using a simple spatial clustering function coded into a Hopfield neural network. In that work, the prior information was representative of spatial coherence, i.e. the property of objects within natural landscapes to be similar to neighbouring objects. Here, the prior information is extended to the spatial pattern of objects smaller than the ground resolution of the sensor. The results from the technique described in chapter 3 are used for comparison with the results from the new technique described in this chapter.

The focus of each of the techniques described in chapters 3 and 4 on land cover features larger than a pixel (e.g. agricultural fields), enables the utilisation of information contained in surrounding pixels. However, this source of information is unavailable when examining imagery of land cover features that are smaller than a pixel (e.g. trees in a forest). Consequently, while these features can be detected within a pixel by soft classification techniques, surrounding pixels hold no information for inference of spatial relationships to aid their mapping. Therefore, the technique presented in this chapter attempts to overcome this problem and to present a novel and effective solution to super-resolution land cover pattern prediction from remotely sensed images, as well as an extension to the techniques introduced in chapters 3 and 4. This method is based on prior information on the spatial arrangement of land cover. A simple function to match land cover distribution within each pixel to this prior information is coded into a Hopfield neural network.

5.2. Using the Hopfield Neural Network for Super-Resolution Land Cover Pattern Prediction

The input data for the research described in this chapter were derived from aerial photography, whereby land cover targets were identified and extracted accurately from the photographs by hand. By degrading these verification images of clearly defined land cover targets to the spatial resolution of Landsat TM data using a square averaging filter, perfect class proportions were obtained for each pixel. These provided the input to the network. In practice, the input could

come from automated soft classification methods, such as the multi-layer perceptron applied to real imagery, providing the target features to be mapped were separable enough from the other land cover classes to be identified accurately in classification. However, for the research in this chapter, the aim was to understand and test the capabilities of the Hopfield network technique. Any error introduced to the input data by automated soft classification would be detrimental to this aim.

5.2.1 Network Architecture

The network architecture for super-resolution pattern prediction is described in section 2.5.1.

5.2.2 Network Initialisation

Network neurons were initialised using method (i), described in section 2.5.2.

5.2.3 The Energy Function

The equations of motion were defined as,

$$\frac{dE_{ij}}{dv_{ij}} = \left(\sum_{n=1}^{z} k_n \frac{dSn_{ij}}{dv_{ij}} \right) + k_{z+1} \frac{dP_{ij}}{dv_{ij}}. \qquad (5.1)$$

Each component of equation 5.1 is described in the subsequent sections.

The goal and constraints of the pattern prediction task were defined such that the network energy function for a zoom factor of 7 was,

$$E = -\sum_i \sum_j \left(k_1 S1_{ij} + k_2 S2_{ij} + k_3 S3_{ij} + k_4 S4_{ij} + k_5 S5_{ij} + k_6 S6_{ij} + k_7 S7_{ij} + k_8 P_{ij} \right), \quad (5.2)$$

where, k_1 to k_8 are constants weighting the various energy parameters, $S1_{ij}$ to $S7_{ij}$ represent the output values for neuron (i,j) of the seven semivariance (objective) functions (see section 5.2.3.1), and these correspond to the quadratic term in equation 2.7. P_{ij} represents the output value for neuron (i,j) of the proportion constraint (see section 3.4.2) which corresponds to the linear term in equation 2.7.

5.2.3.1 The Semivariance Functions

The semivariance (objective) functions, $S1_{ij}$ to $S7_{ij}$, aimed to model the spatial pattern of each land cover at the sub-pixel scale. Prior knowledge about the spatial arrangement of the land cover in question was utilised, in the form of semivariance values, calculated by:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1,j=1}^{N(h)} \left[v_{ij} - v_{i\pm h, j\pm h} \right]^2 \tag{5.3}$$

where, $\gamma(h)$ is the semivariance at lag h, and $N(h)$ is the number of pixels at lag h from the centre pixel (i,j). This was calculated for z lags from an aerial photograph, for example. This provided information on the typical spatial distribution of the land cover under study, which could then be used for land cover simulation from remotely sensed imagery at the sub-pixel scale.

By using the values of $\gamma(h)$ from equation 5.3, the output of the centre neuron, $v(c)_{ij}$, which produced a semivariance of $\gamma(h)$ could be calculated using:

$$v(c)_{ij} = \frac{1}{2a} \left[-b \pm \sqrt{b^2 - 4ac} \right] \tag{5.4}$$

where, $a = 2N(h), b = \sum_{i=1,j=1}^{N(h)} v_{i\pm h, j\pm h}, c = \sum_{i=1,j=1}^{N(h)} (v_{i\pm h, j\pm h})^2 - 2N(h)\gamma(h)$. The semivariance function value for lag 1 ($h = 1$), $S1_{ij}$, was then given by:

$$\frac{dS1_{ij}}{dv_{ij}} = v_{ij} - v(c)_{ij} \tag{5.5}$$

If the output of neuron (i,j), v_{ij}, was lower than the target value, $v(c)_{ij}$, calculated in equation 5.4, a negative gradient was produced that corresponded to an increase in neuron output to

counteract this problem. An overestimation of neuron output resulted in a positive gradient, producing a decrease in neuron output. Only when the neuron output was identical to the target output, did a zero gradient occur, corresponding to $S1_{ij} = 0$ in the energy function (equation 5.2). The same calculations were carried out for lags 2 to 7, to produce values for $S2_{ij}$ to $S7_{ij}$.

5.2.3.2 The Proportion Constraint

While the semivariance functions provided the enforcement of a certain spatial pattern, using only these functions would have resulted in the entire image displaying the same regular pattern. Therefore, a method of constraining the effect of those functions to the correct image areas was required. The proportion constraint, P_{ij}, aimed to retain the pixel class proportions output from the soft classification, and is described in section 3.2.3.2.

5.2.4 Advantages of the Technique

The approach described in this chapter holds several strategic advantages over those techniques mentioned in chapter 1:

- The technique has the ability to map land cover features smaller than a pixel, rather than being restricted to large features.
- The option to choose the level of increase in spatial resolution. This is essential if land cover pattern prediction of higher spatial resolution imagery is the aim.
- The ease by which any additional information can be incorporated within the framework to aid the pattern prediction. Any prior information about the land cover depicted in the input imagery can be coded easily into the Hopfield network as an extra constraint to increase accuracy.
- The design of the Hopfield network as an optimisation tool means that all constraints are satisfied simultaneously, rather than employing a multi-stage operation.
- The effect that each one of these constraints has on the final prediction image can be controlled simply via weightings.

5.3. Results

Illustrative results were produced using the Hopfield network run on a P2-350 computer. As the pattern prediction technique represents an extension of the method described in chapter 3, results of both techniques were compared to highlight the need for a dual approach. To

determine the accuracy of each technique at recreating the spatial pattern exhibited in the target images, plots of lag against semivariance (a variogram (Curran and Atkinson, 1998)), calculated using equation 5.3, were used. In addition to visual comparison of the target image variograms with the prediction variograms, a correlation coefficient was calculated to provide a statistical measure of similarity between the various plots. This was given by,

$$\rho_{tp} = \frac{Cov(t,p)}{\sigma_t \sigma_p}, \tag{5.6}$$

where, $-1 < \rho_{tp} \leq 1$, and, $Cov(t,p) = \frac{1}{n}\sum_{q=1}^{n}(t_q - \mu_t)(p_q - \mu_p)$.

In equation 5.6, t is the target semivariance value, p is the predicted value, n is the number of semivariance values, σ is the standard deviation of each set of values and μ is the mean of each set of values. A correlation coefficient of 1 represents a perfect match between target and predicted semivariance values, whereas a value of –1 corresponds to a perfect negative correlation. The class area shown in each image was also calculated to give an indication of the success of the proportion constraint.

5.3.1 Synthetic Imagery

To understand and illustrate the workings of the Hopfield network set up in the above way, several synthetic images were created. By breaking down the elements of real-world imagery into simplified representations, understanding such an image processing technique and improving it becomes easier.

Figures 5.1 and 5.2 show the four images used to test the prediction capabilities of the Hopfield network. The images were designed to represent a range of possible landscape features, so that the generalisation ability of the network could be examined, for example, 5.1(a): sparse natural woodland or semi-natural vegetation, 5.1(b): dense natural woodland or semi-natural vegetation, 5.2(a): woodland plantation or housing, 5.2(b): intergrade between land covers.

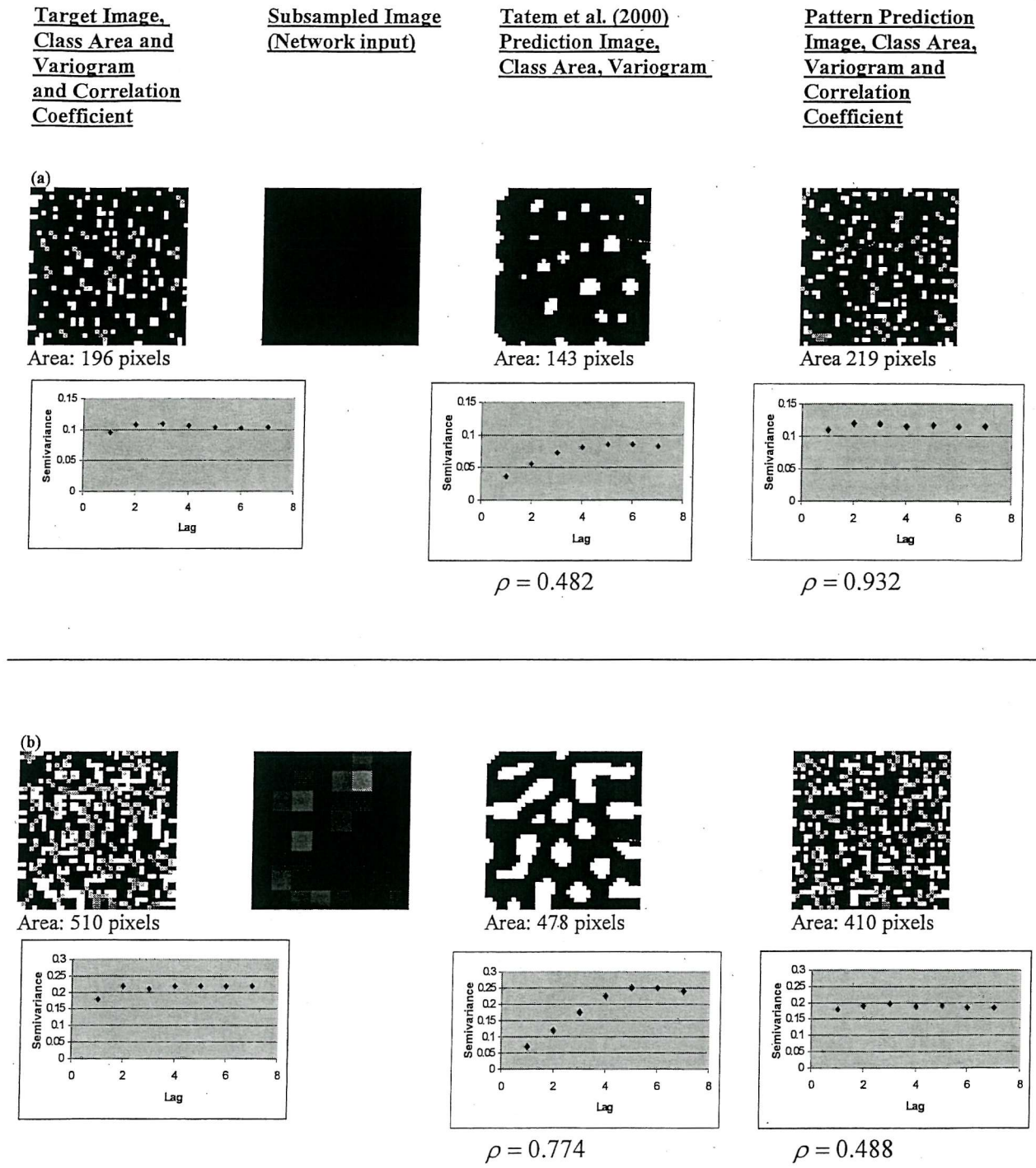


Figure 5.1. Synthetic imagery, each 56x56 pixels, and results of the Hopfield neural network predictions.

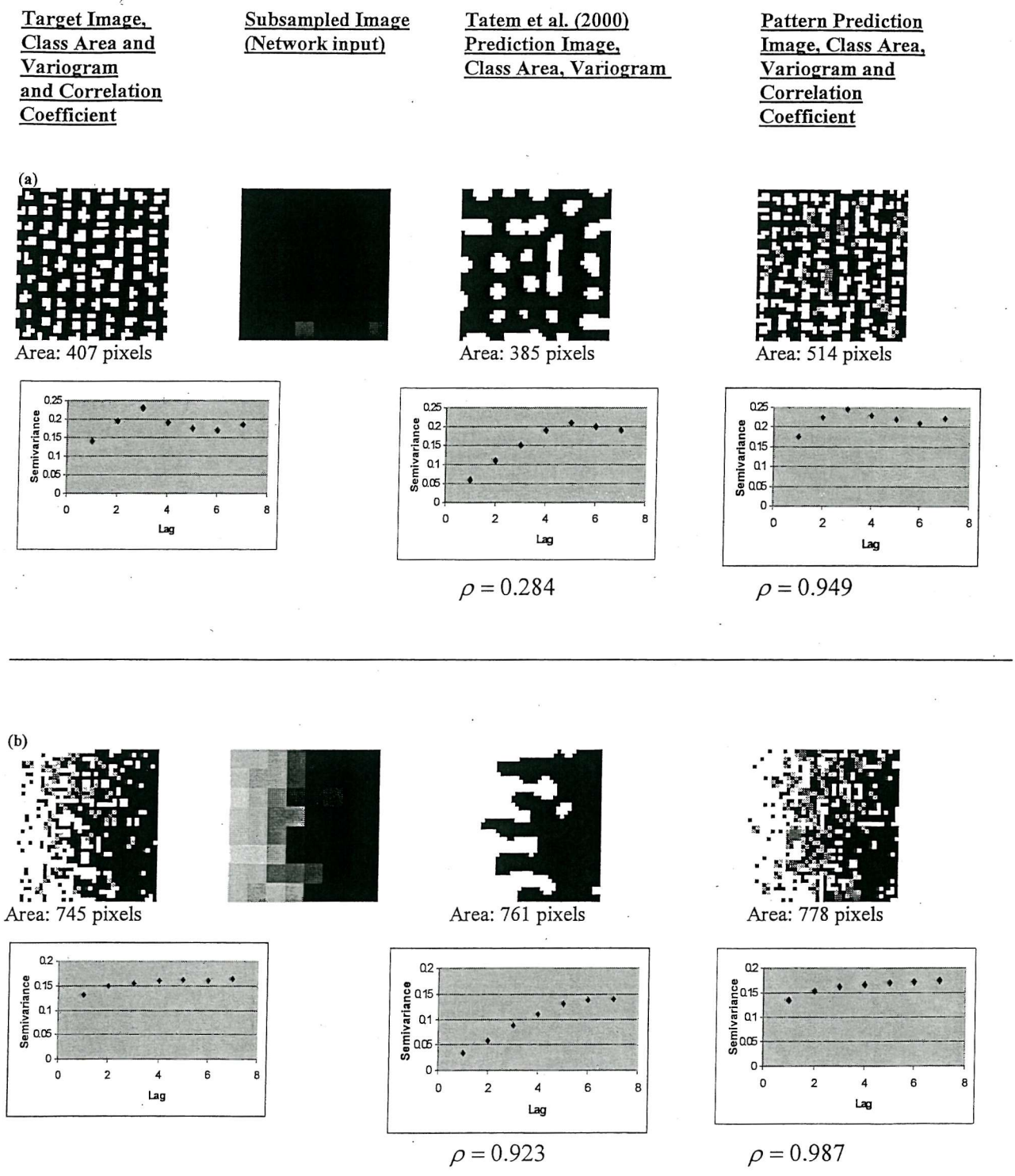


Figure 5.2. Synthetic imagery, each 56x56 pixels, and results of the Hopfield neural network predictions.

Each 56x56 pixel synthetic image in figures 5.1 and 5.2 was subsampled (using a 7x7 averaging filter) to generate an 8x8 pixel image. This caused mixing of the two classes (white and black) at the object boundaries, producing four proportion images, and imitating the effect of class mixing within remotely sensed imagery. By using these four proportion images as inputs to the Hopfield network, and setting a zoom factor of 7, it should be possible to test the capabilities of the techniques by approximating the four images each was derived from. Semivariance was calculated using equation 5.3 at lags 1 to 7 from the target images, and used as input to the semivariance functions.

The network was run using settings of k_1 to $k_7 = 20$ and $k_8 = 140$, which gave the semivariance and proportion functions equal weighting, ensuring that neither had a dominant effect. After 1000 iterations (approximately 15 seconds running time) the results in figures 5.1 and 5.2 were produced.

Figures 5.1 and 5.2 demonstrate the effectiveness of the pattern prediction technique in maintaining the spatial layout, while mapping sub-pixel scale features. Visual comparison of the target images and variograms with the predictions of the two techniques suggests the pattern prediction technique is more precise, and the variogram correlation coefficients confirm this.

The earlier spatial clustering based technique (chapter 3) groups the class proportions into features larger than a pixel in most cases, indicated by the shape of each variogram. In contrast, the pattern prediction technique described in this paper maintains a similar spatial structure to that of the target image. In terms of class area, apart from 5.1(a), the clustering technique proves more accurate, although the over or under-estimation of the pattern prediction technique is at most 107 pixels, representing just 3% of the total number of pixels. Comparison of the various variogram plots indicates that the relative semivariance between lags has been maintained more accurately using the pattern prediction technique. The images resulting from each pattern prediction technique match the target image better than any by the clustering technique, which tend to exhibit the same shape, regardless of input image. The variogram correlation coefficients for figures 5.1(a) and 5.2(a) show significant differences between the two techniques in approximating the target semivariance values, while 5.2(b) shows the highest correlation of

0.987 for the pattern prediction technique. Only in 5.1(b) does the clustering technique show a closer variogram correlation with the target image.

5.3.2 Simulated Remotely Sensed Imagery

Figures 5.3 and 5.4 show the two areas used for this study, and the imagery obtained. Both class proportion images were derived from aerial photographs to avoid the potential problems of incorporating error from the process of soft classification.

Both super-resolution techniques were again run for the proportions in figures 5.3(c) and 5.4(c). After 1000 iterations (approximately 80 seconds running time) at a zoom factor of 7, the results in figures 5.5 and 5.6 were produced.

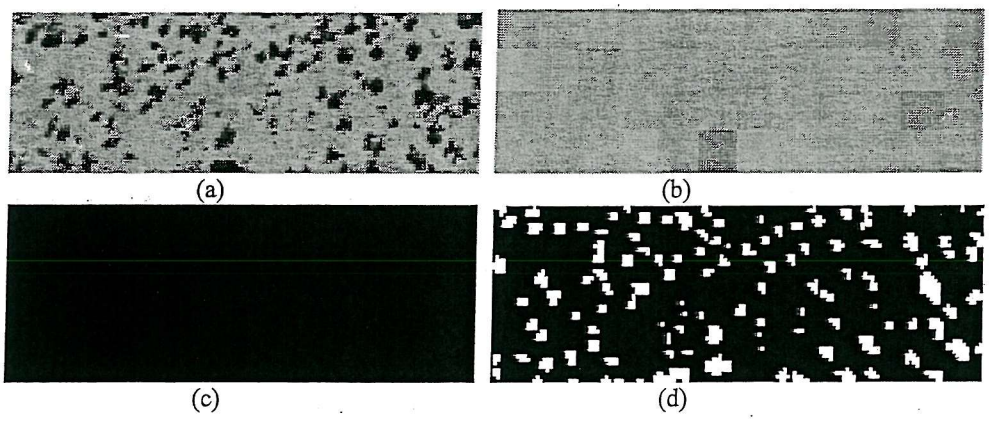


Figure 5.3. (a) Digital aerial photo of an area of the New Forest, UK; (b) 12x4 pixel simulated Landsat TM image (for illustration only); (c) Tree class proportions derived from verification data; (d) Verification image derived from aerial photography.

Figures 5.5 and 5.6 again demonstrate the effectiveness of the pattern prediction technique in mapping sub-pixel scale features while maintaining their spatial distribution across the image. Visual comparison of the target images and variograms with the predictions of the two techniques indicates the pattern prediction technique is more precise, and the variogram correlation coefficients confirm this.

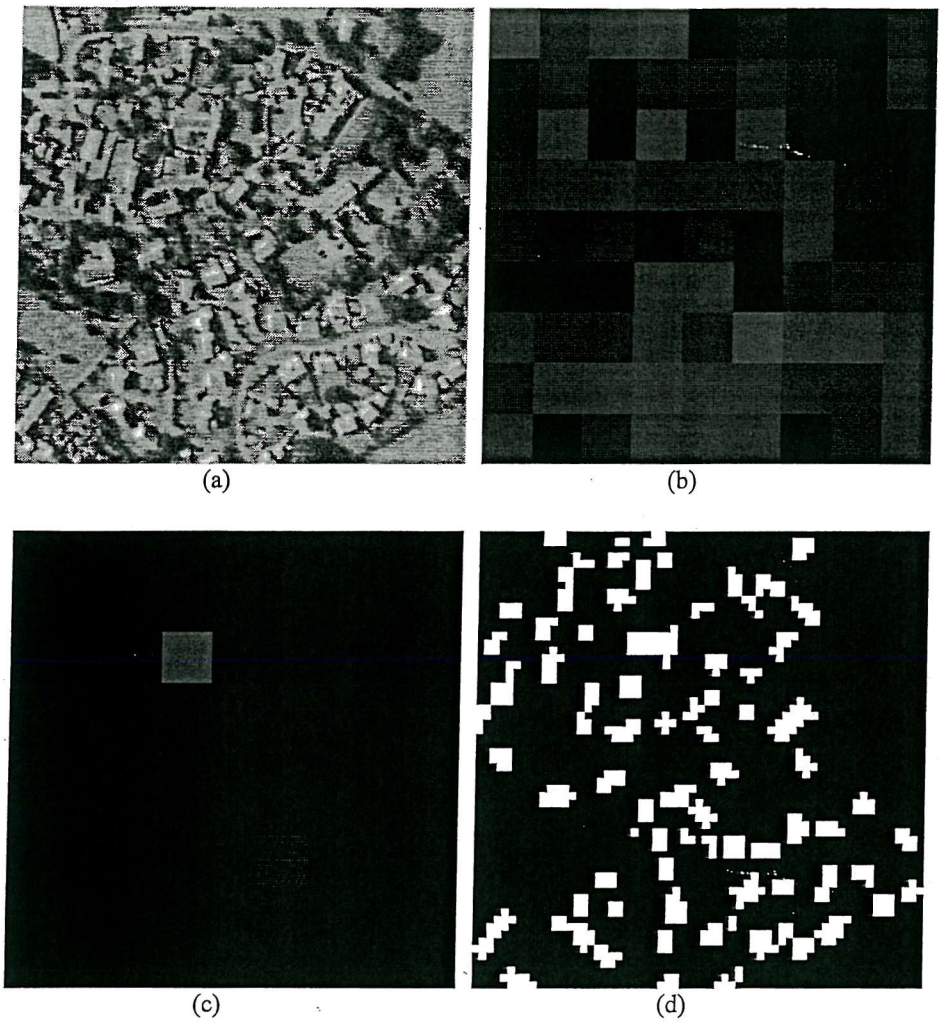


Figure 5.4. (a) Digital aerial photo of an area of Bath, UK; (b) 12x4 pixel simulated Landsat TM image (for illustration only); (c) Building class proportions derived from verification data; (d) Verification image derived from aerial photography.

In the same way as for figure 5.1, both figures 5.5 and 5.6 show that the spatial clustering technique produces a map where the land cover features are much larger and fewer than those in the target image. In contrast, the technique described in this chapter appears to maintain a similar spatial structure to that of the target images. Visual comparison shows that in figure 5.5, the pattern prediction image contains more and slightly smaller trees than the target image, and in figure 5.6 the technique predicts fewer and slightly larger buildings than those depicted in the target image.

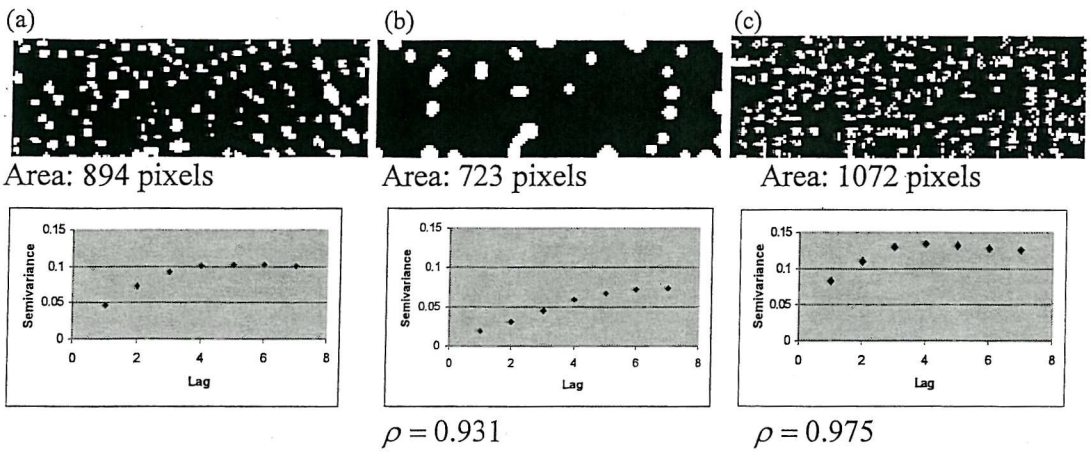


Figure 5.5. (a) Verification image, class area and variogram; (b) Hopfield neural network prediction image, class area, variogram and variogram correlation coefficient (Chapter 3 technique); (c) Hopfield neural network prediction image, class area, variogram and variogram correlation coefficient (pattern prediction technique).

Class area comparisons show that, as with the synthetic imagery, the spatial clustering technique proves more accurate, with the pattern prediction technique demonstrating overestimation on both occasions. However, comparison of the variogram plots in figures 5.5 and 5.6 indicates that the relative semivariance between lags has been maintained more accurately using the pattern prediction technique. The shapes of each pattern prediction technique plot in figures 5.5 and 5.6 match the target plots better than any of the clustering technique plots, and the correlation coefficients confirm this. Figure 5.5 shows that while the correlation of the clustering technique plot with the target plot is 0.931, the correlation for the pattern prediction plot is 0.975, demonstrating a larger correlation. The precision of the pattern prediction technique in matching the semivariance of the target image is mirrored in figure 5.6, with a correlation coefficient of 0.968, significantly larger than that of 0.843 for the spatial clustering technique.

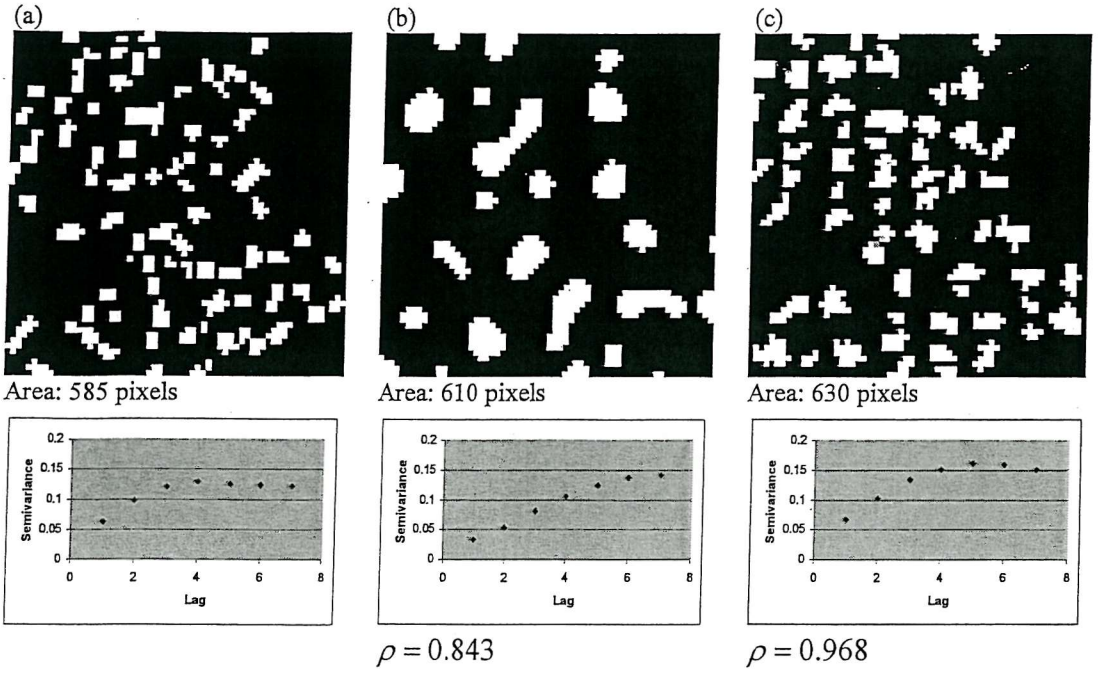


Figure 5.6. (a) Verification image, class area and variogram; (b) Hopfield neural network prediction image, class area, variogram and variogram correlation coefficient (Chapter 3 technique); (c) Hopfield neural network prediction image, class area, variogram and variogram correlation coefficient (pattern prediction technique).

5.4. Discussion

5.4.1 Results Analysis

The results shown in section 5.3 indicate that the Hopfield neural network technique described in this chapter displays excellent potential for super-resolution land cover pattern prediction. Whereas the spatial clustering approach described in chapter 3 failed to recreate the spatial characteristics of land cover features that were smaller than a pixel, the pattern prediction approach introduced here is relatively accurate. This suggests that a combination of the two approaches could potentially identify the spatial arrangement of land cover for objects of any reasonable size in relation to the spatial resolution of the imagery. It should be noted that while the focus of the results of this paper is features such as trees or buildings that are smaller than a pixel in Landsat imagery, the technique could be equally applied to AVHRR imagery, to predict the spatial pattern of sub-pixel scale features such as lakes, tree stands or villages.

The result of applying the technique to synthetic imagery highlights the generalisation capabilities of the approach. Visual inspection of figures 5.1 and 5.2 shows a wide range of synthetic land cover patterns, and comparison of these with the results of the pattern prediction technique in figures 5.1 and 5.2 suggests good generalisation ability. This is of course important if the approach is to be applied in an automated fashion. Throughout the testing of the technique, the various weighting parameters, k_1 to k_8, remained at the same values, and the maintenance of accurate results for each spatial arrangement indicate the ability of the approach to generalise. The synthetic imagery shows that the technique produces surprisingly accurate results for random patterns, as well as the more difficult challenge of a regular grid pattern, while the results from aerial photos demonstrates its potential success in dealing with real world problems. Figures 5.5 and 5.6 demonstrate clearly that the proportion constraint has ensured the class proportions occur in the correct areas of the images, while the semivariance functions ensure these proportions are arranged in a similar spatial pattern to those in the target images.

While the pattern prediction approach was extremely precise at recreating the spatial arrangement of class proportions from coarse spatial resolution imagery, irrespective of the type of pattern in the target images, in many cases there were inaccuracies in recreating the target class area. In all cases in figures 5.1, 5.2, 5.5 and 5.6, except for 5.1(b), there was an overestimation of class area by the pattern prediction technique. This suggests a failure by the proportion constraint, P_{ij}, as it is the role of this function to maintain the target proportions given in the input coarse spatial resolution image. A possible solution could be to increase the weighting of this function, k_8, within the energy function, making it a dominant factor. However, this is likely also to hinder the effect of the semivariance functions, lessening the precision with which the target spatial pattern is recreated. This highlights one of the difficulties with such an optimisation technique, in that there is no ideal method for determining optimal constraint weightings. However, while the difficult choice of weightings can be a drawback, the inclusion of such weightings means that the technique can be tailored to the task in hand. If the maintenance of class area is of prime importance then the proportion constraint can be weighted strongly, whereas if recreation of spatial pattern is the objective then the semivariance functions can be given priority. This flexibility makes the use of the Hopfield neural network an attractive technique.

In all cases except one in figures 5.1, 5.2, 5.5 and 5.6, the pattern prediction technique produced higher correlation coefficients than the other Hopfield neural network technique described in chapter 3 in recreating the spatial pattern of land covers. The only example where the spatial clustering technique produced a larger variogram correlation coefficient is in figure 5.13(b). However, visual inspection of the images and variogram shapes suggest that the pattern prediction technique has produced a more realistic and accurate result, and it is only the fact that the semivariance is underestimated at small lags by the clustering technique that produces the larger correlation coefficient. While the super-resolution spatial clustering technique has been shown to be accurate for mapping land cover targets larger than a pixel described in chapter 3, when features are of sub-pixel scale, the lack of prior information on their layout produces low accuracy, indicating that the work in this chapter is a necessary extension to the technique.

5.4.2 Applications

The precision of the technique demonstrated in this chapter leads to potential application in fields of work where recreation of the spatial pattern of land cover is more important than attempting accurate mapping.

(i) **To provide spatial information for local or regional scale environmental models.**

The technique has potential future application in providing data for input to local or regional scale process models where spatial pattern is important. For example, local scale flood-routing models require maps containing information on the spatial pattern of individual trees within an area of forest. The pattern prediction approach described here may provide this information from coarse spatial resolution imagery. Previously, where input maps of sufficient spatial detail have been unavailable, the pattern prediction technique can utilise widely available coarse spatial resolution imagery to produce a prediction of land cover at the spatial resolution required. This may be imagery of the scale produced by MERIS or MODIS (300 m spatial resolution), Landsat Thematic Mapper (30 m spatial resolution) or even up to the newly launched IKONOS sensor (1 to 4 m spatial resolution).

(ii) **To provide spatial information for global scale environmental models**

In an age when global scale models are being increasingly used to study the world's climate, oceans and ecosystem interactions, input spatial data with sufficient coverage and spatial resolution is a valuable resource. The spatial pattern of land cover,

meteorological or ocean features is often vital input information to such models, and in many cases the satellite sensor imagery used to obtain large spatial coverage (e.g., AVHRR, 1.1 km spatial resolution) has insufficient spatial detail to identify such patterns. Therefore, application of the super-resolution technique described in this chapter could potentially solve this problem by providing spatial pattern predictions from coarse spatial resolution satellite sensor imagery.

(iii) **Environmental management**

In many cases, information on the spatial pattern of a certain land cover class is far more useful for its management than attempting to obtain an accurate map. For example, in forestry, the spatial arrangement of trees within forest stands provides an insight into the allocation of above and below ground resources. The spatial distribution reflects stand history, microclimate differences, climate, sunlight factors as well as competition between individuals and the chance of success of different species over time (Coops and Culvenor, 2000). The approach described in this chapter has the potential to provide maps of tree stands from coarse spatial resolution imagery, which maintain the spatial pattern of the trees and, therefore, make such information available for interpretation.

(iv) **Visualisation**

Often, within certain fields of geographical research and environmental or urban management, the typical spatial arrangement of a certain land cover feature will be known, e.g., 1960s housing, savannah trees, animal habitats, without having a method for visualising such an arrangement. The approach described in this chapter potentially has applications in solving this problem by providing a visualisation method. For example, a town planner may have knowledge on the typical pattern of housing and a coarse spatial resolution satellite sensor image of such an area of housing. By applying the pattern prediction technique described here, a possible realistic map of the housing could be produced to aid planning.

5.5. Conclusions

This study has shown that a Hopfield neural network can be used to predict the location and spatial pattern of class proportions within each pixel. When examining complex, disperse land covers composed of sub-pixel scale features, the super-resolution pattern prediction technique

based on recreating pre-defined semivariance measures provides an accurate and realistic mapping approach.

The unique Hopfield neural network application presented here represents a robust, efficient and simple technique. Results from synthetic and simulated remotely sensed data show high precision, suggesting that it has potential to predict accurately land cover patterns at the sub-pixel scale.

Chapter 6: Further Work

6.1 Introduction

This chapter introduces areas of research that display potential for more extensive future research than that undertaken here. Section 6.2 describes how the two approaches to super-resolution mapping described in chapters 4 and 5 were combined to produce a multiple-scale mapping approach. Section 6.3 presents initial results from applying the super-resolution technique described in chapter 4 to simulated imagery of a new high spatial resolution sensor, IKONOS. Section 6.4 examines a potential extension to the mapping approaches of chapters 3 and 4, which reduces error by allowing uncertainty to be incorporated in the process. Finally, section 6.5 looks at potential future directions for the research detailed in this thesis.

6.2 Super-Resolution Mapping of Multiple-Scale Land Cover Features*

Previous work in chapters 4 and 5 focused on two separate areas of super-resolution mapping: 1. Land cover features larger than a pixel, 2. Land cover features smaller than a pixel. The focus of chapters 3 and 4 on super-resolution mapping of features larger than the scale of a pixel (e.g. agricultural fields in SPOT HRV imagery), enables the utilisation of information contained in surrounding pixels. However, this source of information is unavailable when examining imagery of land cover features that are smaller than a pixel (e.g. houses in Landsat TM imagery). Therefore, chapter 5 presented a technique that attempted to overcome this problem using a Hopfield neural network.

Most remotely sensed images contain both sub-pixel features and those larger than a pixel. Therefore, this section presents the initial results of combining the super-resolution mapping of features larger (as described in chapter 4) and smaller (as described in chapter 5) than a pixel, into a single approach. The approach should be capable of super-resolution land cover mapping from imagery of *any* spatial resolution, containing *any* scale of feature. This was undertaken by combining the functions described in chapters 3, 4 and 5, into a single energy function (equation (1)), and weighting their influence on certain classes.

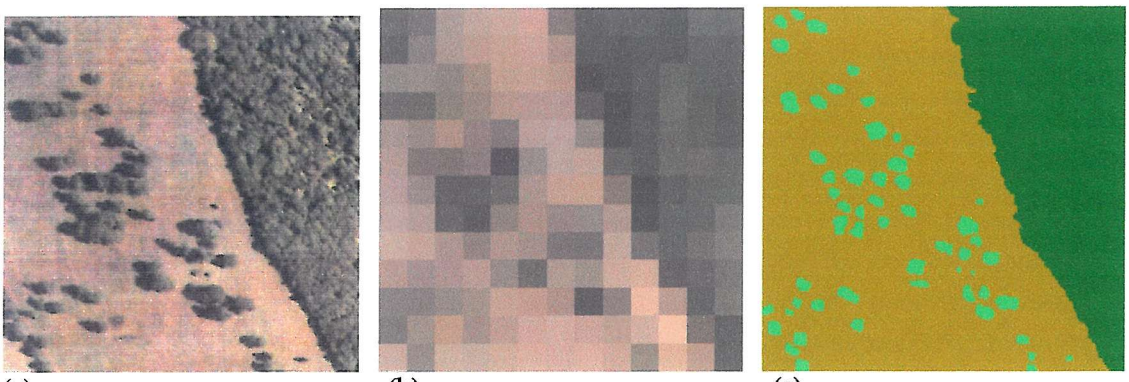
$$E = -\sum_{h}\sum_{i}\sum_{j}\left(\left(\sum_{n=1}^{z} k_n Sn_{hij}\right) + k_{z+1}G1_{hij} + k_{z+2}G2_{hij} + k_{z+3}P_{hij} + k_{z+4}M_{hij}\right) \qquad (6.1)$$

* This section is based on Tatem *et al.* (2001e)

For example, the clustering functions ($G1_{hij}$ and $G2_{hij}$) were weighted highly for features larger than the pixel size, whereas, the semivariance functions (Sn_{hij}) were given more influence when sub-pixel scale features were dominant in a class. Throughout, P_{hij} and M_{hij} were given the largest weightings to ensure correct class proportions were maintained, without gaps or overlaps in the final map.

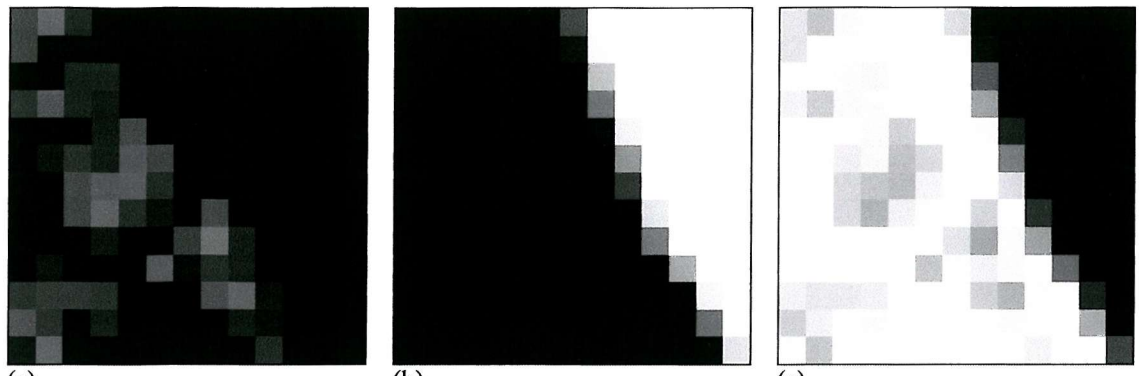
6.2.1 Results

The new network set-up was tested on simulated SPOT HRV imagery. Figure 6.1(a) shows an aerial photograph of the chosen test area, which contained both a large area of woodland, and lone trees amongst grassland. The verification image in figure 6.1(c) was degraded, using a square mean filter, to produce three class proportion images, shown in figure 6.2, that provided input to the Hopfield neural network. In addition, a variogram (figure 6.3) was calculated from a small section of figure 6.1(c) to provide the prior spatial information on the lone tree class, required by the *SV* function. After 1000 iterations of the network with z=7, the map shown in figure 6.4(c) was produced, and a traditional hard classification was undertaken for comparison. Both maps were compared to the verification data, and accuracy statistics calculated. These included the correlation coefficient between classes (CC, equation 1.12), area error proportion (AEP, equation 1.11) and root mean square error (RMSE, equation 1.13), all shown in tables 6.1 and 6.2. Figure 6.5 shows the variogram calculated from the network prediction in figure 6.4(c).



(a) (b) (c)

Figure 6.1(a) Aerial photograph of an area near Bath, UK; (b) Simulated SPOT HRV image of the area shown in (a); (c) Verification image derived from aerial photography.



(a) (b) (c)

Figure 6.2 Class proportion images; (a) lone trees; (b) woodland; (c) grass.

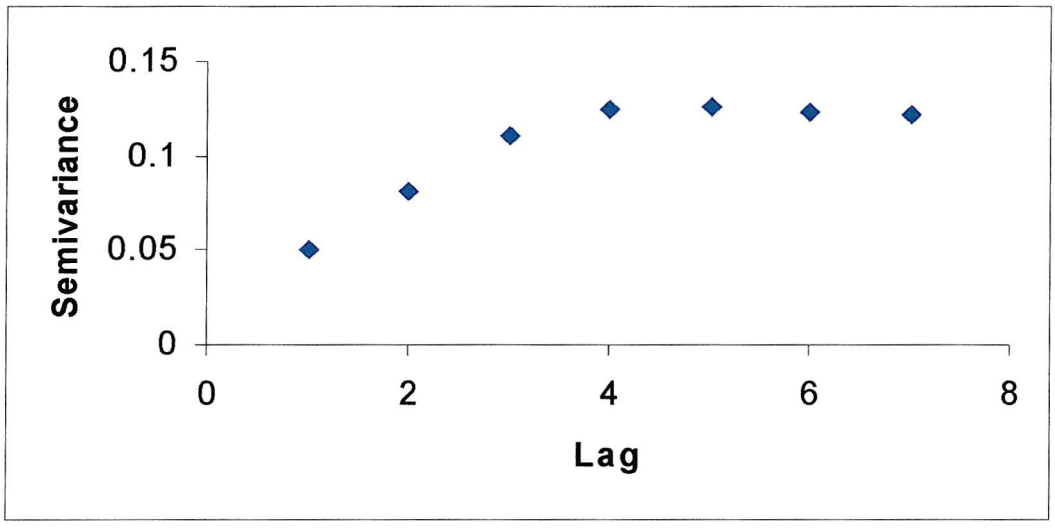
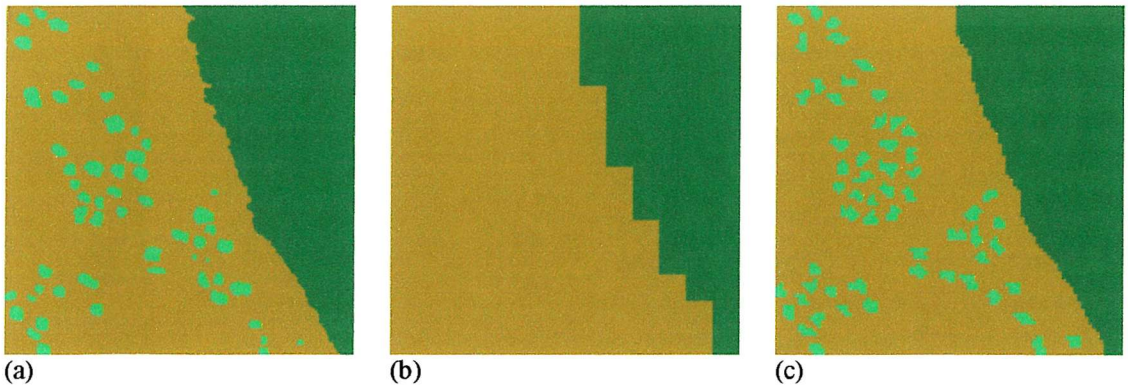


Figure 6.3 Variogram of the typical spatial arrangement of the lone tree class.



(a) (b) (c)

Figure 6.4(a) Verification image; (b) Traditional hard classification; (c) Hopfield neural network prediction.

	CC	AEP	RMSE
Lone Trees	N/A	1	0.251
Woodland	0.887	0.0095	0.218
Grass	0.761	-0.101	0.331
Entire Image		0.0434	0.271

Table 6.1 Accuracy statistics for the hard classification in figure 6.4(b)

	CC	AEP	RMSE
Lone Trees	0.43	-0.161	0.269
Woodland	0.985	0.0085	0.081
Grass	0.831	0.0136	0.28
Entire Image		0.00711	0.229

Table 6.2 Accuracy statistics for the Hopfield neural network prediction in figure 6.4(c)

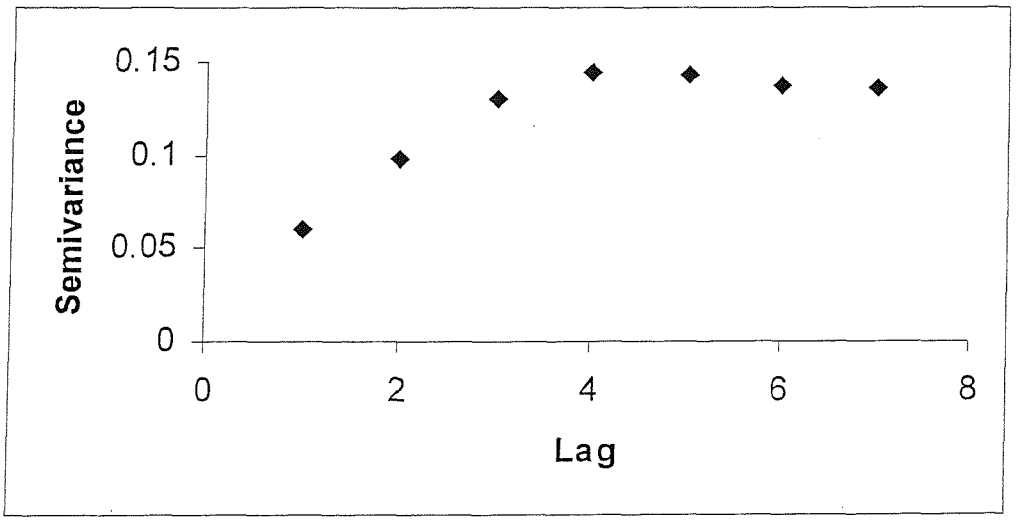


Figure 6.5 Variogram of the spatial arrangement of the lone tree class in the Hopfield neural network prediction map.

6.2.2 Discussion

The results show clearly that the super-resolution technique provides an increase in accuracy over traditional hard classification. Visual inspection of figure 6.4 reveals that hard classification (figure 6.4(b)) has failed to identify the lone tree class, and produced an uneven woodland boundary. In contrast, the Hopfield neural network prediction appears to have identified and mapped both features correctly. This is confirmed after inspection of the accuracy statistics and variograms. While there is little difference between the AEP values in tables 6.1 and 6.2, showing that both techniques maintained class area to a similar degree, the other statistics show how accurate the Hopfield neural network was. The woodland class was mapped accurately, with a correlation coefficient of 0.985, compared to just 0.887 using hard classification. Overall image results also show an increase in accuracy, with a RMSE value of just 0.229. The only low accuracy result is for the lone tree class, with a correlation coefficient of just 0.43. However, as the aim of the *SV* function is to recreate the spatial arrangement of sub-pixel scale features, rather than map accurately their locations, this is expected. The best way to test the performance of this function is therefore, to compare the

shape of the variograms in figures 6.3 and 6.5, which do confirm that a similar spatial arrangement of trees to that of the verification image has been recreated.

6.2.3 Conclusions

This study has shown that the techniques for super-resolution mapping described in chapters 3, 4 and 5 can be combined to map features both larger and smaller than a pixel. The Hopfield neural network used in this way represents a simple, robust and efficient tool for super-resolution mapping of multiple-scale land cover features, and future research should be directed at adapting and applying the approach to more complex imagery.

6.3 Super-Resolution Mapping of Urban Scenes from IKONOS Imagery*

The recent launch of the satellite IKONOS has led to many new areas of research being examined. The unprecedented availability of multispectral satellite sensor imagery with spatial resolution of up to 4 m potentially leads to new advances in the field of land cover mapping from remotely sensed data. One obvious usage of such detailed imagery is in urban land cover mapping.

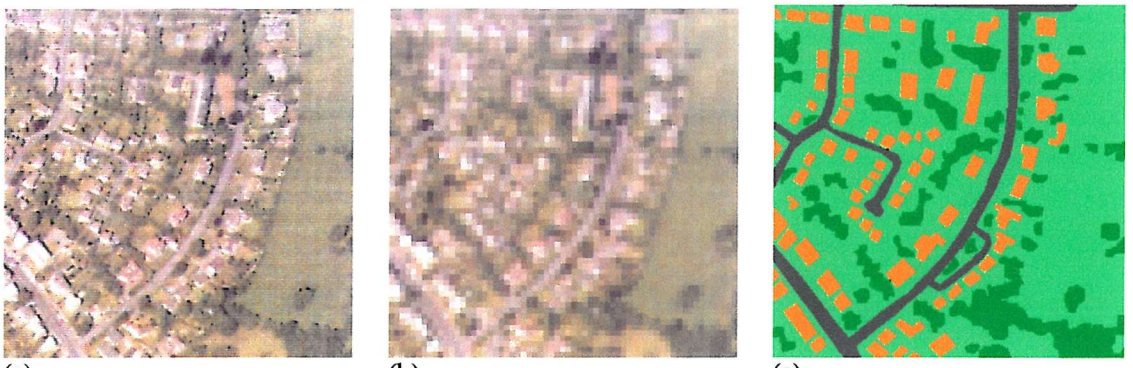
Previously, the coarse spatial resolution of existing satellite sensors has prevented the detailed mapping of urban scenes. Identifying and classifying the small objects (e.g. buildings, roads, trees) that make up the complex mosaic of an urban scene was impossible with sensors such as Landsat TM (30 m spatial resolution), but there exist possibilities when imagery of 4 m spatial resolution is available. However, when examining multispectral IKONOS imagery, it is clear that a large amount of class mixing still occurs. This indicates that a traditional 'hard' classification is not appropriate for accurate mapping of urban scenes from IKONOS imagery. The application of a 'soft' classifier enables the contents of each pixel to be estimated, resulting in a more informative classification. However, the location of land cover class components within each pixel still remains unknown, hindering the production of accurate urban maps from IKONOS imagery.

This section describes initial results from the application of the Hopfield neural network technique, described in chapter 4, to map the location of urban class components at the sub-pixel scale, for simulated IKONOS images of Bath, UK.

* This section is based on Tatem *et al.* (2001f)

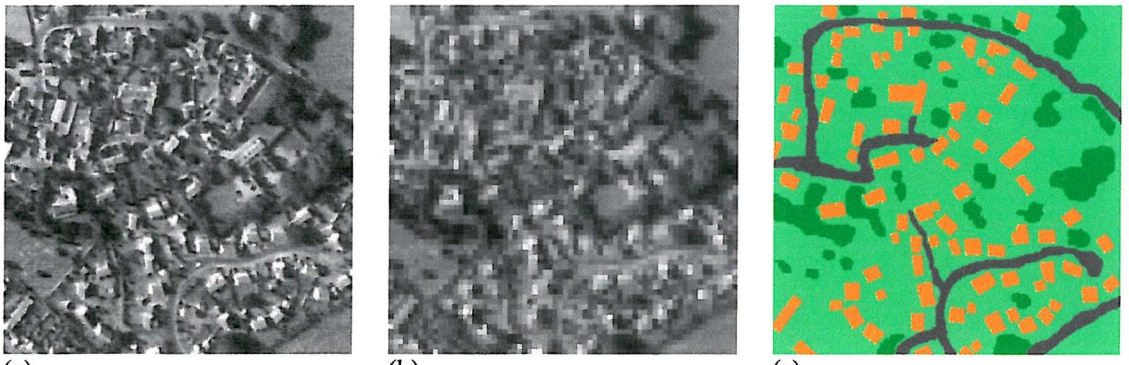
6.3.1 Data

Figures 6.6 and 6.7 show the study areas in Bath, UK, used for this section, simulated IKONOS imagery of the same areas and the verification imagery derived from aerial photography. The two verification images were degraded using a square mean filter to produce the class proportion images shown in figures 6.8 and 6.9. These images represent the expected output from soft classification, ignoring the effects of the sensor point spread function and classifier error, to demonstrate clearly the capabilities of the Hopfield neural network. The images in figures 6.8 and 6.9 were the sole input to the network.



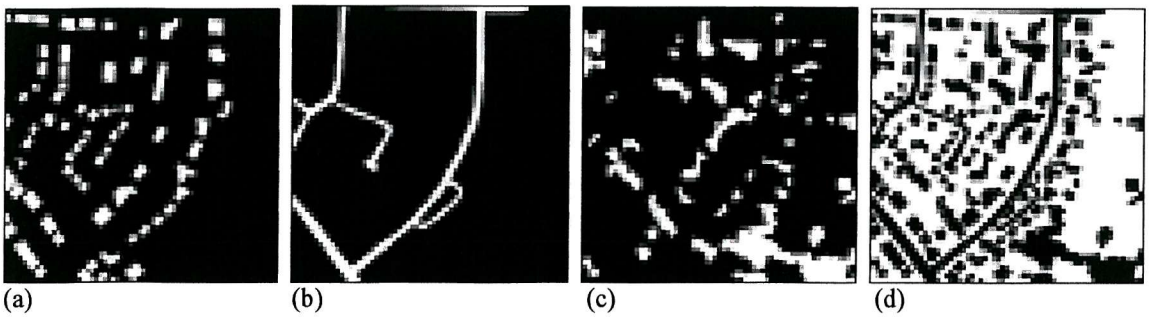
(a) (b) (c)

Figure 6.6(a) Aerial photograph of an area of Bath, UK; (b) Simulated IKONOS image of the area shown in (a); (c) Verification image derived from aerial photography.



(a) (b) (c)

Figure 6.7(a) Aerial photograph of an area of Bath, UK; (b) Simulated IKONOS image of the area shown in (a); (c) Verification image derived from aerial photography.



(a) (b) (c) (d)

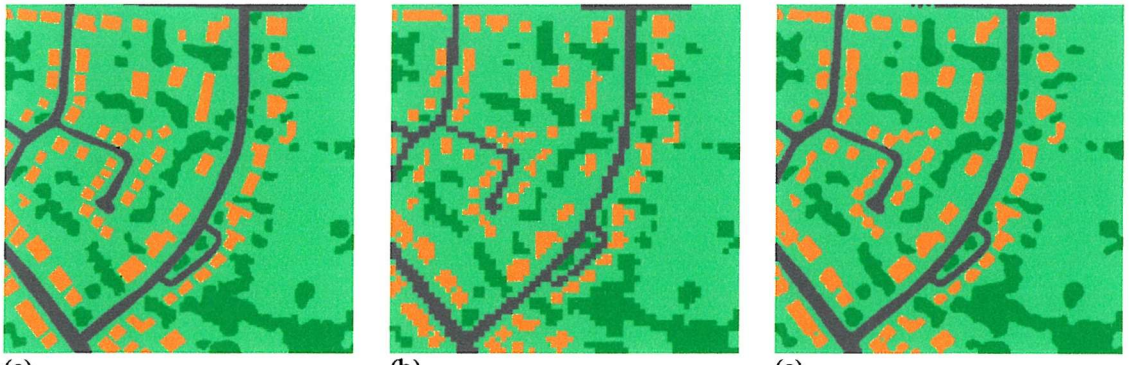
Figure 6.8 Class proportion images; (a) buildings; (b) roads; (c) trees; (d) grass.

(a) (b) (c) (d)

Figure 6.9 Class proportion images; (a) buildings; (b) roads; (c) trees; (d) grass.

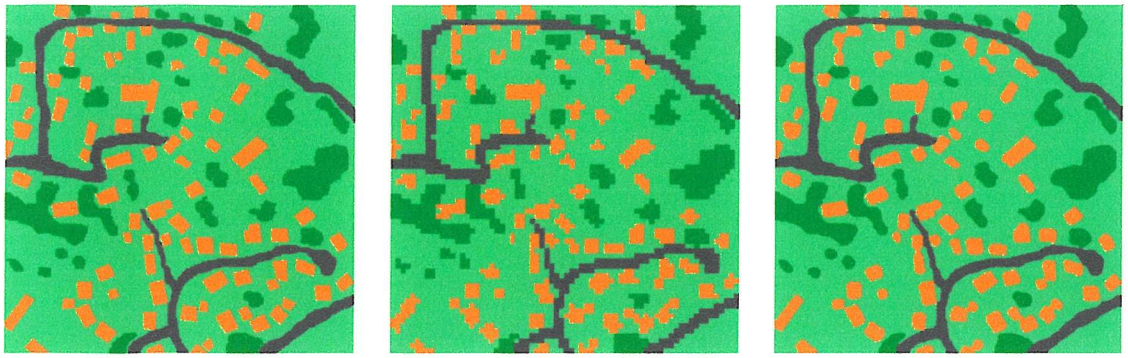
6.3.2 Results

After approximately 1000 iterations of the network at z=7, the network converged to solutions, and figures 6.10(c) and 6.11(c) show the resultant maps. Traditional hard classifications were also undertaken for comparison, with the results shown in figures 6.10(b) and 6.11(b). Both maps were compared to the verification images in figures 6.10(a) and 6.11(a), and accuracy statistics calculated. These included the correlation coefficient between classes (CC, equation 1.12), area error proportion (AEP, equation 1.11) and root mean square error (RMSE, equation 1.13). Tables 6.3 and 6.4 show the assessment results for the maps shown in figure 6.10, and tables 6.5 and 6.6 show assessment statistics for the predictions in figure 6.11.



(a) (b) (c)

Figure 6.10 (a) Verification image; (b) Traditional hard classification; (c) Hopfield neural network prediction.



(a) (b) (c)

Figure 6.11(a) Verification image; (b) Traditional hard classification; (c) Hopfield neural network prediction.

	CC	AEP	RMSE
Building	0.791	0.0308	0.213
Road	0.836	0.0274	0.156
Tree	0.842	0.0154	0.204
Grass	0.764	-0.003	0.33
Entire Image		0.00261	0.235

Table 6.3 Accuracy statistics for the hard classification in figure 6.10(b)

	CC	AEP	RMSE
Building	0.907	-0.0239	0.144
Road	0.951	-0.0272	0.086
Tree	0.931	-0.0196	0.136
Grass	0.9	0.013	0.215
Entire Image		0.00415	0.152

Table 6.4 Accuracy statistics for the Hopfield neural network prediction in figure 6.10(c)

	CC	AEP	RMSE
Houses	0.813	0.042	0.213
Road	0.859	-0.0036	0.153
Trees	0.874	0.0097	0.157
Grass	0.797	-0.0009	0.303
Entire Image		0.00202	0.215

Table 6.5 Accuracy statistics for the hard classification in figure 6.11(b)

	CC	AEP	RMSE
Houses	0.935	-0.003	0.127
Road	0.95	-0.035	0.092
Trees	0.95	-0.006	0.01
Grass	0.915	-0.012	0.195
Entire Image		0.00295	0.135

Table 6.6 Accuracy statistics for the Hopfield neural network prediction in figure 6.11(c)

6.3.3 Discussion

The results show clearly that the Hopfield neural network technique provides an increase in accuracy over traditional hard classification. Visual inspection of figures 6.10 and 6.11 show that hard classification has resulted in both maps displaying broken sections of road, blocky

areas of trees and merged buildings. In contrast, the Hopfield neural network prediction appears to have mapped most features correctly. This is confirmed after inspection of the accuracy statistics in tables 6.3 to 6.6. Comparison of tables 6.3 and 6.4 shows that while both sets of AEP values are similarly low, indicating that both techniques maintained class area to a similar degree, the other statistics reveal how accurate the super-resolution technique was. Whereas correlation coefficients in table 6.3 range from just 0.76 to 0.84, those for the Hopfield neural network technique are all above 0.9. These levels of accuracy are also reflected in the overall image statistics with table 6.4 showing a RMSE value of just 0.152, significantly higher than the corresponding value for hard classification. Tables 6.5 and 6.6 show very similar results to those in tables 6.3 and 6.4. Both tables show similarly low AEP values, indicating that areas estimated from soft classification have been maintained accurately, but differences arise with the other statistics. The correlation coefficients are all above 0.9 again for the Hopfield neural network prediction in table 6.6, reflecting the high precision with which features have been mapped. This again compares favourably to those values for the hard classification in table 6.5. The individual RMS errors show how each class was mapped more accurately using the Hopfield neural network approach, and the overall decrease from 0.215 for the hard classification to 0.135 for the Hopfield neural network prediction demonstrates this further.

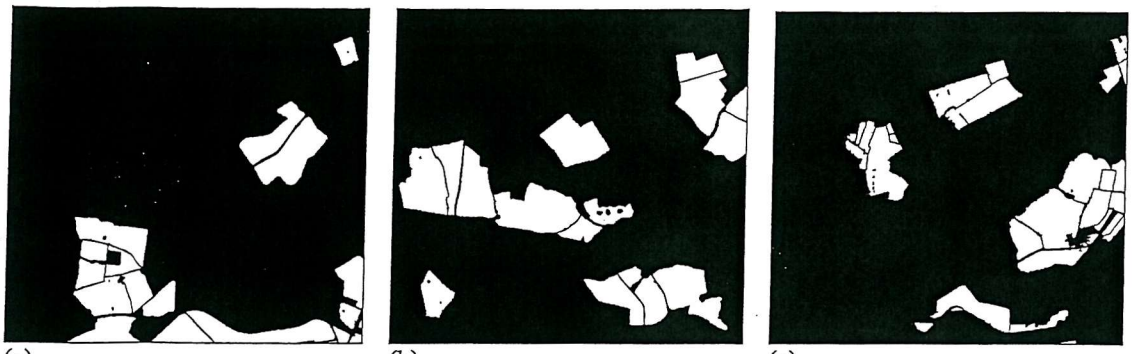
6.3.4 Conclusions

This study has shown that a Hopfield neural network can be used to estimate the location of urban land cover class proportions within pixels from simulated IKONOS imagery. Results suggest that the use of this simple, robust and efficient technique can significantly increase mapping accuracy. In addition, this work demonstrates that the 4 m spatial resolution of IKONOS imagery does not represent a limit to the spatial detail obtainable in classification of such data.

6.4 Dealing with Uncertainty in Super-Resolution Land Cover Mapping

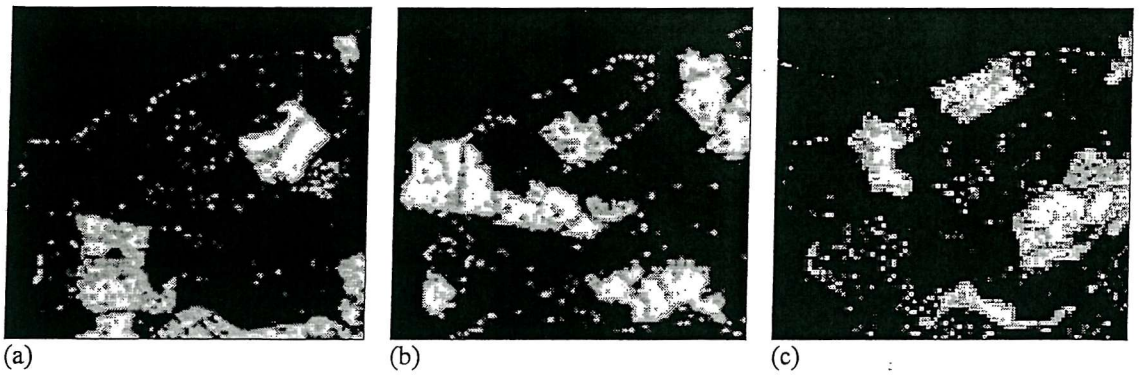
Chapter 4 presented results from applying the Hopfield neural network mapping technique to Landsat TM imagery. Application of the technique to real imagery enabled assessment of the network's abilities to cope with uncertainty (section 4.4 lists potential sources of uncertainty). Section 4.4.3.1 showed how relaxation of the proportion constraint provided greater robustness to uncertainty and error. However, problems still remained. The existence of isolated small groups of pixels producing a spotted appearance to figures 4.21, 4.22 and 4.23

is evidence that the relaxed proportion constraint was insufficient to eliminate isolated mis-classified pixels in the input imagery (figure 4.20).



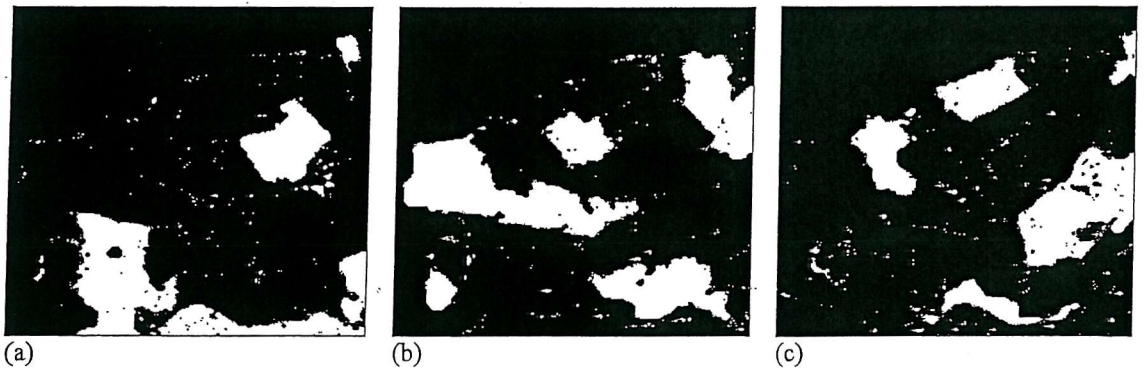
(a) (b) (c)

Figure 6.12 Stoughton verification imagery; (a) wheat; (b) broad beans; (c) grass.



(a) (b) (c)

Figure 6.13 Soft *k*-nearest neighbour classification results; (a) wheat; (b) broad beans; (c) grass.



(a) (b) (c)

Figure 6.14 Hopfield neural network predictions from chapter 4; (a) wheat; (b) broad beans; (c) grass.

Figures 6.12, 6.13 and 6.14 illustrate the problem tackled in this section in detail. Figure 6.12 shows the verification data for the three classes within the Stoughton imagery with the largest numbers of isolated mis-classified pixels after soft *k*-nearest neighbour classification (Wheat, Broad Beans and Grass). Figure 6.13 shows the results of this classification, and comparison to figure 6.12 reveals the problems caused by error and uncertainty within the imagery and classification process. When the three images in figure 6.13 were input to the Hopfield neural

network at z=7 and using the proportion constraint relaxation settings described in 4.4.3.1, the predictions shown in figure 6.14 were produced. It can be seen from these predictions that a small amount of the error from the soft classification has been eliminated, but a large amount also remains. This section describes the design and implementation of a new windowed proportion constraint that allows for error and uncertainty to alleviate these problems and increase mapping accuracy.

6.4.1 The Energy Function

The equations of motion were defined as,

$$\frac{dE_{ij}}{dv_{ij}} = k_1 \frac{dG1_{ij}}{dv_{ij}} + k_2 \frac{dG2_{ij}}{dv_{ij}} + k_3 \frac{dWP_{ij}}{dv_{ij}}, \tag{6.2}$$

and the goal and constraints of the sub-pixel mapping task were, therefore, defined such that the network energy function was,

$$E = -\sum_i \sum_j (k_1 G1_{ij} + k_2 G2_{ij} + k_3 WP_{ij}) \tag{6.3}$$

where, k_1, k_2 and k_3 are constants weighting the various energy parameters, $G1_{ij}$ and $G2_{ij}$ represent the output values for neuron (i,j) of the two objective (or goal) functions (see section 3.2.3.1). WP_{ij} represents the output value for neuron (i,j) of the windowed proportion constraint, described in the subsequent section.

6.4.2 The Windowed Proportion Constraint

Section 3.2.3.2 described the design of the proportion constraint used throughout chapters 3, 4 and 5, in addition to sections 6.2 and 6.3 in this chapter. The proportion constraint, P_{ij}, aimed to retain the pixel class proportions output from the soft classification by constraining the total output from each set of neurons representing a coarse spatial resolution image pixel to be equal to the predicted class proportion for that pixel. The windowed proportion constraint, WP_{ij}, also aimed to retain the pixel class proportions output from the soft classification, but allows for uncertainty and error in each single pixel area estimation. This was achieved by operating a similar constraint to that described in 3.2.3.2, over a centre-weighted square window of 3x3 pixels. Therefore, instead of ensuring that the output of each

set of neurons representing one pixel was equal to the predicted class proportion for that pixel, the centre-weighted constraint was extended to ensure that the output of each set of neurons representing 3x3 pixels was equal to the centre-weighted predicted class proportion for a window of 3x3 pixels. As this processing was undertaken for all pixels, the windows overlapped, effectively producing linkages across the entire image, and simulating the uncertainty-inducing point spread function (PSF) effect of integrating responses from surrounding pixels. By weighting the neuron outputs from the centre pixel of each 3x3 window, the PSF-induced effects were simulated further. This also allowed the goal functions to be more influential, and a degree of uncertainty and flexibility to be incorporated into the prediction of class location. Whereas previously the location of class components would have been constrained by P_{ij} to lie within a pixel, components could now migrate across pixel boundaries, reversing the sensor PSF effect, and still satisfy the replacement proportion constraint. The neuron outputs for the centre pixel of each 3x3 window were weighted more strongly than the surrounding 8 pixels. This weighting provided a simple approximation of the sensor PSF that introduces uncertainty into the image being processed (see section 1.5.2). By simulating its form in the proportion constraint, a degree of image smoothing was introduced, eliminating isolated erroneous area predictions.

An area proportion estimate representing the proportion of neurons with an output of 0.55 or greater was calculated for all the neurons representing pixel (x, y) and the surrounding eight pixels.

$$\text{Area Estimate} = \frac{1}{18z^2} \sum_{k=xz-z}^{xz+2z} \sum_{l=yz-z}^{yz+2z} \left(1 + \tanh\left(v_{kl} - 0.55\right)\lambda\right) \qquad (6.4)$$

The tanh function was used in the same way as described in section 3.2.3.2, ensuring that if a neuron output was above 0.55, it was counted as having an output of 1.0 within the estimation of class area per 3x3 pixel window. Below an output of 0.55, the neuron was not counted within the estimation, which simplified the area proportion estimation procedure, and ensured that neuron output must exceed the random initial assignment output of 0.55 to be counted within the calculations. Once these neuron outputs were calculated, the outputs of neurons representing the centre pixel were multiplied by 2, and those representing the surrounding eight pixels were multiplied by 0.5. This produced the centre weighted area estimate required.

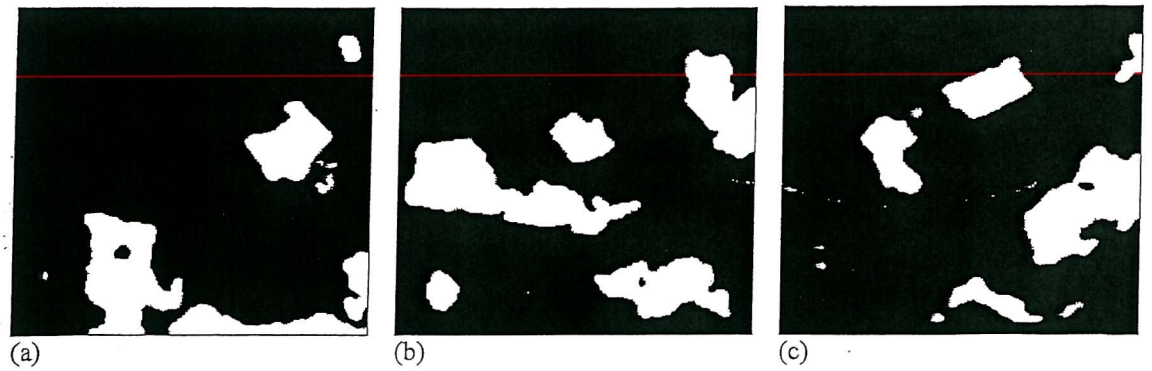
To ensure that the class proportions per 3x3 pixel window output from the soft classification were maintained, the proportion target per 3x3 pixel window, $\sum_{p=x-1}^{x+1}\sum_{q=y-1}^{y+1} a_{pq}$, centre-weighted in the same way as above, was subtracted from the area proportion estimate (equation 6.4):

$$\frac{dWP_{ij}}{dv_{ij}} = \frac{1}{18z^2}\sum_{k=xz-z}^{xz+2z}\sum_{l=yz-z}^{yz+2z}\left(1+\tanh(v_{kl}-0.55)\lambda\right) - \sum_{p=x-1}^{x+1}\sum_{q=y-1}^{y+1} a_{pq} \qquad (6.5)$$

If the weighted area proportion estimate for the 3x3 pixel window centred on pixel (x, y) was lower than the weighted target area, a negative gradient was produced that corresponded to an increase in neuron output to counteract this problem. An overestimation of class area resulted in a positive gradient, producing a decrease in neuron output. Only when the weighted area proportion estimate was identical to the weighted target area proportion for each 3x3 pixel window did a zero gradient occur, corresponding to $WP_{ij} = 0$ in the energy function (equation 6.3).

6.4.3 Results

The new network set-up was tested on the three soft k-nearest neighbour derived class proportion images shown in figure 6.13. After 10 000 iterations of the network with z=7 and $k_1 = k_2 = k_3 = 150$, the prediction maps shown in figure 6.15 were produced. These were compared, along with the predictions from chapter 4 in figure 6.14, to the verification data in figure 6.12, and accuracy statistics calculated. These included correlation co-efficient between classes (CC, equation 1.12), area error proportion (AEP, equation 1.11) and root mean square error (RMSE, equation 1.13), all shown in tables 6.7, 6.8 and 6.9.



(a) (b) (c)

Figure 6.15 Hopfield neural network predictions using the new windowed proportion constraint; (a) wheat; (b) broad beans; (c) grass.

Constraint	CC	AEP	RMSE
P_{ij}	0.893	0.0133	0.161
WP_{ij}	0.909	0.0073	0.146

Table 6.7 Accuracy statistics for the wheat class using two types of proportion constraint.

Constraint	CC	AEP	RMSE
P_{ij}	0.909	0.0103	0.166
WP_{ij}	0.926	0.0081	0.149

Table 6.8 Accuracy statistics for the broad beans class using two types of proportion constraint.

Constraint	CC	AEP	RMSE
P_{ij}	0.882	0.0106	0.167
WP_{ij}	0.9	0.0043	0.153

Table 6.9 Accuracy statistics for the grass class using two types of proportion constraint.

6.4.4 Discussion

The results show clearly that the use of the windowed proportion constraint, WP_{ij}, produces an increase in accuracy over the proportion constraint, P_{ij}, used throughout this thesis. Visual inspection of figures 6.14 and 6.15 shows that whereas the use of P_{ij} has lead to the small isolated and erroneous areas of predicted classes being maintained, the use of WP_{ij} has eliminated these errors, leaving just the large fields in all three cases. Such visual assessments are confirmed after inspection of the accuracy statistics in tables 6.7, 6.8 and 6.9. All three tables show decreases in area error proportion using WP_{ij}. This shows how the use of this constraint accepts area estimates which differ from those output by the *k*-nearest neighbour classifier to be accepted, consequently producing area estimates closer to reality, as the uncertainty and error inherent in the *k*-nearest neighbour estimates is reduced.

6.4.5 Conclusions

This section has shown that the accuracy of predictions produced from Landsat TM imagery in chapter 4 can be increased further using an alternative proportion constraint. The use of a centre-weighted, windowed proportion constraint allows uncertainty and error within the input imagery to be accommodated in the mapping process. Future work should focus on applying the technique to a full set of classes, quantification of any increases in accuracy and examinations of further approaches to account for and reduce uncertainty.

6.5 Future Work

This section introduces areas of potential future research that have not been explored in the thesis.

6.5.1 Multiple Scale Super-Resolution Mapping

Section 6.2 has shown the potential that exists for combining the mapping and pattern prediction techniques, described in chapters 4 and 5, into a single tool. While being only a simple synthetic image, the results from section 6.2 demonstrate how features both larger and smaller than a pixel can be mapped at the sub-pixel scale by a Hopfield neural network. Future work should focus on applying the technique to more complex synthetic imagery to understand more fully the processes at work, and consequently, improve them. Once refinement of the approach has been achieved, the network should be tested on real imagery containing features of a range of scales.

6.5.2 Super-Resolution Mapping from Fine Spatial Resolution Sensors

Section 6.3 has shown the potential that exists in using the Hopfield neural network for super-resolution mapping from IKONOS imagery. Although the approach was only tested on simulated IKONOS imagery, the results showed that the 4 m spatial resolution does not represent a limit to the spatial detail obtainable in classification of such data. Future work should be directed towards achieving similarly successful results for real IKONOS imagery, and also applying the technique to imagery from other fine spatial resolution sensors.

6.5.3 Super-Resolution Mapping from Coarse Spatial Resolution Sensors

Throughout this thesis, the Hopfield neural network technique has been applied to simulated and real imagery of the more popular, medium spatial resolution sensors, Landsat TM and Spot HRV (30 m and 20 m spatial resolution respectively). While the approach has proved very accurate in achieving accurate super-resolution mapping from this scale of imagery, it has not been tested on coarser spatial resolution imagery. This may include AVHRR imagery (1.1km spatial resolution), or imagery produced by the new hyper-spectral sensors of MERIS (300m spatial resolution) and MODIS (250m-1000m spatial resolution). Such imagery may present a whole new set of challenges to the Hopfield neural network with class mixing within pixels occurring at a whole new set of spatial scales. To explore the potential application of the Hopfield neural network to these types of imagery, simulated imagery should again provide the starting point, enabling understanding and development of the

approach. Once refinement of the technique has been undertaken on simulated imagery, it should then be applied to real imagery.

6.5.4 Uncertainty in Super-Resolution Mapping

Results produced from real Landsat TM imagery in chapter 4 demonstrate the problems encountered due to the error and uncertainty in the remote sensing process, from image acquisition through to soft classification. Such uncertainty propagation has been the focus of much research in the past, and the Hopfield neural network technique adds in another process that increases uncertainty further. In an attempt to reduce this problem, both chapter 4 and section 6.4 have introduced changes to the super-resolution mapping procedure that attempt to account for, and accommodate uncertainty. While these additions have increased the accuracy of results, there still remains much potential for exploration, quantification and accommodation of uncertainty within super-resolution mapping.

6.5.5 Building in New Constraints

The beauty of the Hopfield neural network as an optimisation technique is that new constraints can be built in easily, and the network will find a compromise between all the functions. The number of constraint functions used depends on the amount of prior knowledge available on the scene to be mapped. Chapters 3 and 4 demonstrated the workings of the network without specific prior knowledge, showing good generalisation abilities. However, mapping accuracy could have been increased had prior knowledge of the scenes been utilised. Chapter 5 gave an example of how such prior knowledge can be utilised. In that case, prior knowledge of sub-pixel class spatial arrangement in the form of a variogram, was used to increase mapping accuracy. Future work should focus on developing extra constraints to increase accuracy in this way. This may include ensuring that the prediction for a certain class displays straight edges and corners for each feature, should it be known that a building class is being mapped. This could be achieved using simple shape descriptor measures, or even the Hough transform. Such a constraint would avoid the erroneous predictions shown in section 6.3 for example, where building corners were rounded. Other such constraints may involve ensuring continuous class linkage across an image for a road class, or the use of directional variograms for more complex spatial arrangement recreation.

6.5.6 Change Detection

At the spatial resolution of SPOT HRV imagery (20 m) or coarser, small changes in land cover, boundaries or feature migration may not be detected using conventional classification

techniques. By obtaining a sequence of satellite sensor images of the same area, taken at different times, and applying the Hopfield neural network approach to each, small changes not visible using traditional techniques could potentially be identified. Future work should focus on examining the potential of the Hopfield neural network for small-scale change detection from satellite sensor imagery, and defining the scale limits and robustness to noise and error of the technique.

6.5.7 New Measures of Accuracy

Throughout this thesis, traditional statistical measures of accuracy have been used to assess the performance of the Hopfield neural network technique over traditional approaches. These have proved sufficient for the aims of this research. However, future research may require new measures to be developed. Within this thesis, pixel-by-pixel comparison methods have been used to assess the accuracy with which the Hopfield neural network approach has assigned each pixel to the correct class. However, where the aim is to identify boundaries accurately, for example, where the verification data is in vector form, such measures of accuracy are inappropriate. Therefore, alternative measures based on boundary positioning need to be developed.

6.5.8 Data Fusion

The use of the zoom factor, z, within the design of the Hopfield neural network enables spatial resolution increase by any integer value specified. Such a feature makes the approach attractive potentially for fusing datasets of differing spatial resolution. Future work should examine the potential of super-resolution mapping from coarse spatial resolution images to match the spatial resolution of a finer resolution image, and merge the two. Such operations could provide useful and more informative GIS data layers for example.

6.5.9 Application to New Fields of Research

Great potential exists in applying the Hopfield neural network approach outside the field of land cover mapping from remotely sensed imagery. The technique has potential application in any area of research where observations or measurements are spatially aggregated. Examples of research fields where the Hopfield neural network could be applied are:

- Increasing the spatial resolution of radar, or lidar imagery for target mapping and change detection.
- Increasing the accuracy of areal rainfall estimates

- Dis-aggregation of spatially aggregated sociological data, for example, health or crime statistics within enumeration districts or wards.
- Increasing the spatial resolution of close circuit television imagery for target identification.
- Increasing the spatial resolution of medical imagery for feature identification, mapping and change detection.
- Providing measurements with increased accuracy in close-range metrology.

6.5.10 New Approaches to Optimisation

Section 2.5 gave the reasons behind the choice of the Hopfield neural network for undertaking the optimisation task. However, while the Hopfield neural network does present benefits for use in the task of super-resolution land cover mapping, there exist alternative optimisation techniques that could be implemented. Future work should focus on testing alternative techniques, such as those described in section 2.2, to assess whether improvements could be made.

6.5.11 New Approaches to Super-Resolution Land Cover Mapping

As has been described in 1.7.3, alternative approaches to super-resolution land cover mapping from remotely sensed imagery do exist, but they are few and far between, and each suffers from drawbacks. Therefore, there exists great potential for research into alternative approaches to understanding the problem that has been tackled with the Hopfield neural network. Future work could focus on developing techniques along the lines of these approaches:

- Super-resolution image segmentation. The combination of sub-pixel scale edge detection operators with traditional image segmentation algorithms has the potential to be used for super-resolution target mapping.
- Model-based techniques. By applying pixel-splitting models at the local scale within an image, a complete image can potentially be segmented at the sub-pixel scale. This approach was adopted by Steinwendner and Schneider (1997), but suffered from drawbacks, as outlined in section 1.7.3, which could be overcome.
- Geostatistically based techniques. Incorporation of prior image information into a stochastic model that attaches higher probabilities to e.g. targets with shorter edge length, as was undertaken by Gavin and Jennison (1997). This model could then be optimised via

simulated annealing that would place an increasing probability at the globally optimum state.

6.6 Conclusions

This chapter has both shown and described the potential that the Hopfield neural network approach has for future research. Section 6.2 shows how techniques described in previous chapters for super-resolution mapping of features larger and smaller than a pixel can be combined successfully into a single tool. Section 6.3 demonstrates the potential that exists in applying the Hopfield neural network for super-resolution mapping from IKONOS imagery and fine spatial resolution imagery in general. Section 6.4 indicates that acknowledging and accommodating the uncertainty inherent in all remotely sensed data can produce increases in super-resolution mapping accuracy. Finally, section 6.5 presents a variety of potential future research paths, which work could take after this thesis. The broad range of future options reveals how under-explored the field of super-resolution mapping actually is, and hopefully, how the tool developed in this thesis could prove useful across a multitude of disciplines.

Chapter 7: Conclusions

7.1 Introduction

In this chapter, an attempt is made to draw together the findings of the various approaches to super-resolution land cover mapping described in previous chapters. Initially, the research is summarised, including project objectives, results, findings and implications. Finally, concluding remarks are provided.

7.2 Summary

The research undertaken in this thesis is summarised in five parts: (i) background, (ii) objectives, (iii) development and analysis, (iv) findings and (v) implications.

7.2.1 Background

Land cover information provides an important aid to our management and understanding of the environment. In many cases, remote sensing represents the only feasible and affordable means of obtaining such information, especially when large area studies are required. However, the coarse spatial resolution provided by satellite sensor imagery has proved detrimental to the success of such studies. The size of image pixels, relative to the land cover features of interest, leads to large numbers of image pixels representing a mixed response. Such mixed pixels consequently produce error in land cover maps derived from traditional hard classification techniques. Recent years have seen such problems reduced by the introduction of soft classification approaches, producing more informative land cover maps. Nevertheless, while such classifications predict the class composition of each pixel, there are few techniques for mapping the location of these classes within the pixels.

7.2.2 Objectives

The overall aim was to design and develop a robust, efficient and simple technique for super-resolution mapping of land cover from remotely sensed imagery. The approach would aim to map the location of class components output from soft classification within each pixel, and overcome the problems suffered by previous techniques.

7.2.3 Development and Analysis

The Hopfield neural network computer model was developed as an energy minimisation tool. The task of mapping the spatial distribution of land cover class components within each image pixel was formulated as a constraint satisfaction problem, with an optimal solution

determined by the minimum of an energy function. This energy function consisted of a set of component functions that differed depending on the super-resolution task.

The Hopfield neural network architecture represented an image of finer spatial resolution than the original satellite sensor image being processed. The nature of the problem and proposed solution required the network neurons to be considered as being arranged in a regular grid, with positioning within this grid being of significance to the network design. The architecture, processing steps and energy function of the Hopfield neural network were all developed and refined via the use of synthetic imagery. Initial application of the Hopfield neural network to simple synthetic imagery allowed the processes at work to be understood and, consequently, design improvements and additions to be made.

Energy functions were designed specifically for the application of the Hopfield neural network to various tasks: (i) Super-resolution land cover target identification; (ii) Super-resolution land cover mapping; (iii) Super-resolution land cover pattern prediction; (iv) Super-resolution land cover mapping of multiple-scale features; (v) Super-resolution mapping of urban scenes from IKONOS imagery; (vi) Dealing with uncertainty in super-resolution land cover mapping. Each of these specific network designs were tested and applied on simulated remotely sensed imagery from a range of sensors. In addition, the Hopfield neural network set-up for super-resolution land cover mapping was tested on Landsat TM imagery of small and large-area agriculture.

7.2.4 Findings

Overall, results showed accurate classifications, suggesting that the Hopfield neural network has potential to predict accurately land cover at the sub-pixel scale from operational satellite sensor imagery. Application of the Hopfield neural network designed as a target identification tool to synthetic and simulated imagery produced high levels of accuracy. These results also enabled findings on network performance to be made. Such findings include:

- The technique produced a more accurate prediction if the shape depicted in the input class proportions was compact and circular.
- The technique produced a more accurate prediction if a high zoom factor was used.
- The technique produced a more accurate prediction if the network was allowed to run for at least 1000 iterations.

- The network converged to an accurate prediction in fewer iterations if a proportion-constrained initialisation was used.

Conversion of the Hopfield neural network into a tool for super-resolution *land cover mapping* resulted in the production of land cover maps of greater accuracy and finer spatial resolution than those produced by traditional hard classification approaches. The results from applying the Hopfield neural network to simulated Landsat TM and SPOT HRV imagery was similar high levels of accuracy to those obtained for the target identification set-up. However, the uncertainty inherent in real satellite sensor imagery provided more of a challenge. Uncertainty in image representation (e.g. point spread function, atmospheric effects, noise, geometric effects), uncertainty of classification (e.g. spectral confusion, class choices, classifier choice, classifier error) and uncertainty in the verification data (e.g. registration error, class identification, generalisation) meant that the output Landsat TM imagery from a soft *k*-nearest neighbour classification were inaccurate. Therefore, relaxation of the constraint controlling maintenance of input proportion predictions was introduced. This allowed small differences away from the input class area prediction to be accepted and, therefore, increased prediction accuracy for synthetic imagery. Consequently, this was applied to the processing of real Landsat TM imagery of large and small-area agriculture. Results again were more accurate than traditional approaches (hard maximum likelihood classification) and maintained the accuracy levels obtained using soft classification. Problems were encountered when the network was tested on unusually complex imagery containing 21 classes, but the prediction was still more accurate than that obtained by the hard classification.

Conversion of the Hopfield neural network into a tool for super-resolution prediction of the spatial distribution of sub-pixel size features produced results of high accuracy. As with the target identification and mapping approaches, the technique was tested on synthetic and simulated satellite sensor imagery and the accuracy increased over that for hard classification. The approach showed that when examining complex, disperse land covers composed of sub-pixel scale features, the Hopfield neural network technique, based on recreating pre-defined semivariance measures, provided an accurate and realistic mapping approach.

Initial results from the combination of super-resolution mapping and pattern prediction approaches into a single tool were accurate and, as with previous work, produced increases in accuracy over that for traditional hard classification. Application of the Hopfield neural

network used as a super-resolution land cover mapping tool to simulated IKONOS imagery produced accurate results, which were, again, more accurate than those produced by hard classification. Finally, it was found that the introduction of a new constraint to control class area maintenance increased mapping accuracy when dealing with imagery containing error and uncertainty.

7.2.5 Implications

In most cases, remote sensing represents the only realistic means of obtaining fast and affordable land cover information at any scale. However, traditional 'hard' approaches for converting the satellite sensor imagery to land cover maps have, in the past, been limited by the relatively coarse spatial resolution of the available imagery.

Over the last decade, soft classification approaches have become more widely available, enabling image analysts to produce more informative land cover maps from satellite sensor imagery. But limitations still exist, as such maps give no indications of how land cover classes are distributed spatially within each pixel. In addition, the output of soft classification produces a map for each class, with no definite boundaries between classes (which decision makers, planners and GIS require). Therefore, the introduction of a technique that produces hard land cover maps from satellite sensor imagery, but is not limited by spatial resolution, has huge implications for the field of map production from remotely sensed imagery.

Providing users are willing to accept the assumptions made by the Hopfield neural network approach, results from this thesis show that in addition to providing finer spatial resolution maps, the accuracy of such maps is greater than those produced by hard classification, and potentially, as good as or better than soft classification. Such results imply that the spatial resolution of land cover maps are no longer limited by the spatial resolution of the satellite sensor imagery that each was derived from. Although there is much work to be done to refine, apply and test the Hopfield neural network approach, results produced in this thesis demonstrate that there is great potential for the widespread use of the technique for operational processing of satellite sensor imagery.

7.3 Conclusions

This thesis has shown that the spatial resolution of satellite sensor imagery need not necessarily represent a limit to the spatial detail obtainable within land cover maps derived from such imagery. Given reasonable assumptions about the nature of spatial correlation (at

the sub-pixel scale) in the scene, a soft classification can be transformed into a hard categorical variable at the sub-pixel scale. That is, multivariate proportion information can be transformed into a single categorical variable with an increase in spatial resolution. Moreover, by making these assumptions, the accuracy of such land cover maps can be greater than those produced using traditional techniques.

This thesis has shown that a Hopfield neural network can be used to predict the location of class proportions within each pixel, to produce sub-pixel scale maps of land cover. The Hopfield neural network represents a robust, efficient and simple technique, and predictions from simulated and real remotely sensed data are accurate. This suggests that the Hopfield neural network has potential to predict accurately land cover at the sub-pixel scale from operational satellite sensor imagery and future research should be directed at examining this possibility further.

References

Aarts E. and Korst J., 1989, *Simulated Annealing and Boltzmann Machines*, Wiley, Chichester.

Abe, S., 1989, Theories on the Hopfield neural networks, *Proceedings of the International Joint Conference on Neural Networks*, **1**, 557-564.

Abe, S., 1993, Global convergence and suppression of spurious states of the Hopfield neural networks, *IEEE Transactions on Circuits and Systems*, **40**, 246-257.

Aiyer, S.V.B., Niranjan, M., Fallside, F., 1990, A theoretical investigation into the performance of the Hopfield model, *IEEE Transactions on Neural Networks*, **1**, 204-215.

Anderson, J.E., Desmond, G.B., Lemeshewsky, G.P. and Morgan, D.R., 1997, Reflectance calibrated digital multispectral video: a test-bed for high spectral and spatial resolution remote sensing, *Photogrammetric Engineering and Remote Sensing*, **63**, 224-229.

Aniya, M, Sato, H., Naruse, R., Skvarca, P. and Carassa, G., 1996, The use of satellite and airborne imagery to inventory outlet glaciers of the Southern Patagonia icefield, South America, *Photogrammetric Engineering and Remote Sensing*, **62**, 1361-1369.

Aplin P., Atkinson P.M. and Curran P.J., 1997, Fine spatial resolution satellite sensors for the next decade, *International Journal of Remote Sensing*, **18**, 3873-3881.

Aplin, P., 1999, Fine spatial resolution satellite sensor imagery for per-field land cover classification, *Ph.D Thesis*, University of Southampton, Southampton.

Aplin, P., Atkinson, P.M. and Curran, P.J., 1999, Fine spatial resolution simulated satellite imagery for land cover mapping in the United Kingdom, *Remote Sensing of Environment*, **68**, 206-216.

Aplin, P., 2000, Comparison of simulated IKONOS and SPOT HRV imagery for land cover classification, *Proceedings of the 26th Annual Conference of the Remote Sensing Society*, Leicester, U.K.

Atkinson, P.M., 1993, The effect of spatial resolution on the experimental variogram of airborne MSS imagery, *International Journal of Remote Sensing*, **14**, 1005-1011.

Atkinson, P.M. and Curran, P.J., 1995, Defining an optimal size of support for remote sensing investigations, *IEEE Transactions on Geoscience and Remote Sensing*, **33**, 768-776.

Atkinson, P.M., 1997, Mapping sub-pixel boundaries from remotely sensed images, In Kemp, Z. (ed), *Innovations in G.I.S. 4*, Taylor and Francis, London, 166-180.

Atkinson, P.M. and Curran, P.J., 1997, Choosing an appropriate spatial resolution for remote sensing investigations, *Photogrammetric Engineering and Remote Sensing*, **63**, 1345-1351.

Atkinson, P.M. and Kelly, R.E.J., 1997, Scaling-up point snow depth data in the U.K. for comparison with S.S.M./I. imagery, *International Journal of Remote Sensing*, **18**, 437-443.

Atkinson, P.M. and Tatnall, A.R.L., 1997, Neural networks in remote sensing, *International Journal of Remote Sensing*, **18**, 699-709.

Atkinson, P.M., Cutler, M.E.J. and Lewis, H.G., 1997, Mapping sub-pixel proportional land cover with A.V.H.R.R. imagery, *International Journal of Remote Sensing*, **18**, 917-935.

Atkinson, P.M., 1999a, Assessing uncertainty in fuzzy land cover maps obtained by remote sensing, In *Proceedings of Geocomputation '99*, Fredericksburg, USA.

Atkinson, P.M., 1999b, Assessing accuracy in fuzzy land cover maps, In *Proceedings of the 25th Remote Sensing Society Conference,* Taylor and Francis, London, 79-86.

Balchin, W.G.V., 1984, Data sources for land use survey, *Land Use Policy*, **1**, 4-13.

Bandyopadhyay, S. and Pal, S.K., 2001, Pixel classification using variable string genetic algorithms with chromosome differentiation, *IEEE Transactions on Geoscience and Remote Sensing*, **39**, 303-307.

Bastin L., 1997, Comparison of fuzzy c-means classification, linear mixture modelling and M.L.C. probabilities as tools for unmixing coarse pixels, *International Journal of Remote Sensing*, **18**, 3629-3648.

Beale, R. and Jackson, T., 1990, *Neural Computing – an introduction*, I.O.P. Publishing, Bristol.

Benediktsson, J.A., Swain, P.H. and Ersoy, O.K., 1990, Neural network approaches versus statistical methods in classification of multisource remote sensing data, *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 540-552.

Benediktsson, J.A., Swain, P.H. and Ersoy, O.K., 1993, Conjugate-gradient neural networks in classification of multisource and very-high-dimensional remote sensing data, *International Journal of Remote Sensing*, **14**, 2883-2903.

Bezdek, J.C., Ehrlich, R. and Full, W., 1983, F.C.M.: The fuzzy c-means clustering algorithm, *Computers and Geosciences*, **10**, 191-203.

Bhattacharya, A., Reddy, C.S. and Srivastav, S.K., 1993, Remote sensing for active volcano monitoring in Barren Island, India, *Photogrammetric Engineering and Remote Sensing*, **59,** 1293-1297.

Bishop, C.M., 1995, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford.

Boult, T.E and Wolberg, G., 1993, Local image reconstruction and subpixel restoration algorithms, *CVGIP: Graphical Models and Image Processing*, **55**, 63-77.

Breaker, L.C., 1990, Estimating and removing sensor-induced correlation from advanced very high resolution radiometer satellite data, *Journal of Geophysical Research*, **95c**, 9701-9711.

Brown, M. and Lewis, H.G., 1998, Support vector machines and linear spectral unmixing for remote sensing, *Proceedings of the International Conference on Advances in Pattern Recognition*, Plymouth.

Brown, M., Gunn, S.R. and Lewis, H.G., 1999, Support vector machines for optimal classification and spectral unmixing, *Ecological Modelling*, **120**, 167-179.

Campadelli, P., Medici, D. and Schettini, R., 1997, Color image segmentation using Hopfield networks, *Image and Vision Computing*, **15**, 161-166.

Campbell, J.B., 1996, *Introduction to Remote Sensing; 2nd Edition*, Taylor and Francis, New York

Chang, C.I. and Ren, H., 2000, An experiment-based quantitative and comparative analysis of target detection and image classification algorithms for hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, **38**, 1044-1063.

Chang, J.Y., Lee, S.W. and Horng, M.F., 1993, Image sequence correspondence via a Hopfield neural network, *Optical Engineering*, **32**, 1531-1538.

Cichocki, A. and Unbehauen, R., 1993, *Neural networks for optimization and signal processing*, Wiley, Stuttgart.

Cihlar, J, Latifovic, R., Chen, J., Beaubien, J., Li, Z. and Magnussen, S., 2000, Selecting representative high resolution sample images for land cover studies. Part 2: application to estimating land cover composition, *Remote Sensing of Environment*, **72**, 127-138.

Clark, A.N., 1990, *The Penguin Dictionary of Geography*, Penguin, London.

Cochrane, M.A. and Souza, C.M., 1998, Linear mixture model classification of burned forests in the eastern Amazon, *International Journal of Remote Sensing*, **19**, 3433-3440.

Congalton, R.G., 1991, A review of assessing the accuracy of classifications of remotely sensed data, *Photogrammetric Engineering and Remote Sensing*, **37**, 35-46.

Coops, N. and Culvenor, D., 2000, Utilizing local variance of simulated high spatial resolution imagery to predict spatial pattern of forest stands, *Remote Sensing of Environment*, **71**, 248-260.

Côté, S., 1996, Measurement of sea-surface velocities from satellite sensor images using the Hopfield neural network, *Ph.D Thesis*, University of Southampton, Southampton.

Côté, S. and Tatnall, A.R.L., 1994, Estimation of sea surface velocities from space using neural networks, *Proceedings of the SPIE*, **2315**, 444-451.

Côté, S. and Tatnall; A.R.L., 1995, Estimation of ocean surface currents from satellite imagery using a Hopfield neural network, *Proceedings of the 3rd Thematic Conference on Remote Sensing for Marine and Coastal Environments,* Seattle, Washington, 538-549.

Côté, S. and Tatnall, A.R.L., 1997, The Hopfield neural network as a tool for feature tracking and recognition from satellite sensor images, *International Journal of Remote Sensing*, **18**, 871-885.

Cracknell, A.P., 1998, Synergy in remote sensing – what's in a pixel?, *International Journal of Remote Sensing*, **19**, 2025-2047.

Craig, R.G. and Labovitz, M.L., 1980, Sources of variation in landsat autocorrelation, *Proceedings of the 14th International Symposium on Remote Sensing of the Environment*, Ann Arbor, 1755-1717.

Craig, R.G., 1981, Precision in the evaluation of landsat autocorrelation: The Terrain Effect, *Proceedings of the 15th International Symposium on Remote Sensing of the Environment*, Ann Arbor, 1305-1313.

Cross, A.M., Settle, J.J., Drake, N.A. and Paivinen, R.T.M., 1991, Sub-pixel measurement of tropical forest cover using A.V.H.R.R. data, *International Journal of Remote Sensing*, **12**, 1119-1129.

Curran, P.J. and Atkinson, P.M., 1998, Geostatistics and remote sensing, *Progress in Physical Geography*, **22**, 61-78.

Curran, P.J. and Ojedele, B., 1988, Airborne MSS for land cover classification, *Geocarto International*, **4**, 67-72.

Cushnie, J.L., 1987, The interactive effect of spatial resolution and degree of internal variability within land-cover types on classification accuracies, *International Journal of Remote Sensing*, **8**, 15-29.

De Bruin, S., 2000, Predicting the areal extent of land-cover types using classified imagery and geostatistics, *Remote Sensing of Environment*, **74**, 387-396.

DeFries, R.S. and Chan, J.C., 2000, Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data, *Remote Sensing of Environment*, **74**, 503-515.

Dimyati, M., Mizuno, K., Kobayashi, S. and Kitamura, T., 1996, An analysis of land use/cover change using the combination of MSS landsat and land use map – a case study in Yogyakarta, Indonesia, *International Journal of Remote Sensing*, **17**, 931-944.

Dymond, J.R., 1992, How accurately do image classifiers estimate area?, *International Journal of Remote Sensing*, **13**, 1735-1742.

Embashi, M.R.M., 1998, Mapping sub-pixel variation in land cover at the global scale using NOAA AVHRR imagery, *Ph.D. Thesis*, University of Southampton, Southampton.

Finn, J.T., 1993, Use of the average mutual information index in evaluating classification error and consistency, *International Journal of Geographical Information Systems*, **7**, 349-366.

Fisher, P., 1997, The pixel: a snare and a delusion, *International Journal of Remote Sensing*, **18**, 679-685.

Fisher, P. (ed), 2000, *Final Report: Fuzzy Land Information from Environmental Remote Sensing*, European Union Environment and Climate Programme, Contract No. ENV4-CT96.

Fisher, P. and Pathirana, S., 1990, The evaluation of fuzzy membership of land cover classes in the suburban zone, *Remote Sensing of Environment*, **34**, 121-132.

Flack, J., Gahegan, M. and West, G., 1994, The use of sub-pixel measures to improve the classification of remotely sensed imagery of agricultural land, *Proceedings of the 7th Australasian Remote Sensing Conference,* Melbourne, 531-541.

Fletcher, R., 1987, *Practical Methods of Optimisation: 2nd Edition*, Wiley, Tiptree.

Foody, G.M., 1992, A fuzzy sets approach to the representation of vegetation continua from remotely sensed data: an example from lowland heath, *Photogrammetric Engineering and Remote Sensing*, **58**, 221-225.

Foody, G.M. and Cox, D.P., 1994, Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions, *International Journal of Remote Sensing*, **15**, 619-631.

Foody, G.M, 1995, Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data, *Photogrammetric Engineering and Remote Sensing*, **61**, 2-12.

Foody, G.M., McCulloch, M.B. and Yates, W.B., 1995, Classification of remotely sensed data by an artificial neural network: issues related to training data characteristics, *Photogrammetric Engineering and Remote Sensing*, **61**, 391-401.

Foody, G.M., 1996a, Approaches for the production and evaluation of fuzzy land cover classifications from remotely sensed data, *International Journal of Remote Sensing*, **17**, 1317-1340.

Foody, G.M., 1996b, Relating the land-cover composition of mixed pixels to artificial neural network classification output, *Photogrammetric Engineering and Remote Sensing*, **62**, 491-499.

Foody, G.M., 1997, Fully fuzzy supervised classification of land cover from remotely sensed imagery with an artificial neural network, *Neural Computing and Applications*, **5**, 238-247.

Foody, G.M. and Arora, M.K., 1997, An evaluation of some factors affecting the accuracy of classification by an artificial neural network, *International Journal of Remote Sensing*, **18**, 799-810.

Foody, G.M., Lucas, R.M., Curran, P.J. and Honzak, M., 1997, Non-linear mixture modelling without end-members using an artificial neural network, *International Journal of Remote Sensing*, **18**, 937-953.

Foody, G.M., 1998, Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution, *International Journal of Remote Sensing*, **19**, 2593-2599.

Forster, B.C. and Best, P., 1994, Estimation of SPOT P-mode point spread function and derivation of a deconvolution filter, *ISPRS Journal of Photogrammetry and Remote Sensing*, **49**, 32-42.

Forte, P. and Jones, G.A., 1999, Posing structural matching in remote sensing as an optimisation problem, in Kanellopoulos, I., Wilkinson, G. and Moons, T. (eds.), *Machine Vision and Advanced Image Processing in Remote Sensing*, Springer, London.

Foschi, P.G. and Smith, D.K., 1997, Detecting subpixel woody vegetation in digital imagery using two artificial intelligence approaches, *Photogrammetric Engineering and Remote Sensing*, **63**, 493-500.

Fuller, R.M., Sheail, J. and Barr, C.J., 1994, The land of Britain, 1930-1990: a comparative study of field mapping and remote sensing techniques, *The Geographical Journal*, **160**, 173-184.

Fuller, R.M., Groom, G.B., Mugisha, S., Ipulet, P., Pomeroy, D., Katende, A., Bailey, R. and OgutuOhwayo, R., 1998, The integration of field survey and remote sensing for biodiversity assessment: a case study in the tropical forests and wetlands of Sango Bay, Uganda, *Biological Conservation*, **86**, 379-391.

Garcia-Haro, F.J., Gilabert, M.A. and Melia, J., 1996, Linear spectral mixture modelling to estimate vegetation amount from optical spectral data, *International Journal of Remote Sensing*, **17**, 3373-3400.

Gavin, J. and Jennison, C., 1997, A subpixel image restoration algorithm, *Journal of Computational and Graphical Statistics*, **6**, 182-201.

Gebbinck, M.S.K. and Schouten, T.E., 1996a, Accurate area estimation by data-driven decomposition of mixed pixels, Computing Science Institute, Nijmegan, http://www.cs.kun.nl/csi/reports/

Gebbinck, M.S.K. and Schouten, T.E., 1996b, Application of data-driven decomposition to landsat T.M. images for crop area estimation, Computing Science Institute, Nijmegan, http://www.cs.kun.nl/csi/reports/

Gebbinck, M.S.K. and Schouten, T.E., 1997, Area estimation with subpixel accuracy for industrial imaging systems, Computing Science Institute, Nijmegan, http://www.cs.kun.nl/csi/reports/

Gee, A.H., Aiyer, S.V.B. and Prager, R.W., 1993, An analytical framework for optimising neural networks, *Neural Networks*, **6**, 79-97.

Giacinto, G., Roli, F. and Bruzzone, L., 2000, Combination of neural and statistical algorithms for supervised classification of remote-sensing images, *Pattern Recognition Letters*, **21**, 385-397.

Green, E.P., Mumby, P.J., Edwards, A.J., Clark, C.D. and Ellis, A.C., 1998, The assessment of mangrove areas using high resolution multispectral airborne imagery, *Journal of Coastal Research*, **14**, 433-443.

Gross, H.N. and Schott, J.R., 1998, Application of spectral mixture analysis and image fusion techniques for image sharpening, *Remote Sensing of Environment*, **63**, 85-94.

Gong, P. and Howarth, P.J., 1990, An assessment of some factors influencing multispectral land-cover classification, *Photogrammetric Engineering and Remote Sensing*, **56**, 597-603.

Gong, P., Miller, J.R. and Spanner, M., 1994, Forest canopy closure from classification and spectral unmixing of scene components – multisensor evaluation of an open canopy, *IEEE Transactions on Geoscience and Remote Sensing*, **32**, 1067-1080.

Gong, P., Pu, R. and Chen, J., 1996, Mapping ecological land systems and classification uncertainties from digital elevation and forest-cover data using neural networks, *Photogrammetric Engineering and Remote Sensing*, **62**, 1249-1260.

Gopal, S. and Woodcock, C., 1994, Theory and methods for accuracy assessment of thematic maps using fuzzy sets, *Photogrammetric Engineering and Remote Sensing*, **60**, 181-188.

Heintz, D.C. and Chang, C.I., 2001, Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, **39**, 529-545.

Hitchcock, D. and Glasbey, C.A., 1997, Binary image restoration at subpixel resolution, *Biometrics*, **53**, 1040-1053.

Holben, B.N. and Shimabukuro, Y.E., 1993, Linear mixing models applied to coarse spatial resolution data from multispectral satellite sensors, *International Journal of Remote Sensing*, **14**, 2231-2240.

Hopfield, J.J., 1984, Neurons with graded response have collective computational properties like those of two-state neurons, *Proceeding of the National Academy of Sciences*, **81**, 3088-3092.

Hopfield, J.J. and Tank, D.W., 1985, "Neural" computation of decisions in optimization problems, *Biological Cybernetics*, **52**, 141-152.

Horowitz, H.M., Nalepka, R.F., Hyde, P.D. and Morgenstern, J.P., 1971, Estimating the proportions of objects within a single resolution element of a multispectral scanner, *Proceedings of the 7th International Symposium on Remote Sensing of the Environment*, 1307-1320.

Isaaks, E.H. and Srivastava, R.M, 1989, *An Introduction to Applied Geostatistics*, Oxford University Press, Oxford.

Jennison, C. and Jubb, M., 1988, Statistical image restoration and refinement, *Information Processing in Medical Imaging*, 255-262.

Ji, C.Y., 2000, Land-use classification of remotely sensed data using Kohonen self-organizing feature map neural networks, *Photogrammetric Engineering and Remote Sensing*, **66**, 1451-1460.

Journel, A.G. and Huijbregts, C.J., 1978, *Mining Geostatistics*, Academic Press, London.

Journel, A,G., 1996, Modelling uncertainty and spatial dependence: stochastic imaging, *International Journal of GIS*, **10**, 517-522.

Jubb, M. and Jennison, C., 1991, aggregation and refinement in binary image restoration, *Spatial Statistics and Imaging*, **20**, 150-162.

Keitt, T.H., 2000, Spectral representation of neutral landscapes, *Landscape Ecology*, **15**, 479-493.

Kirkpatrick, S., Gelatt, C.D. and Vecchi M.P., 1983, Optimisation by simulated annealing, *Science*, **220**, 671-680.

Lambin, E.F., 1997, Modelling and monitoring land-cover change processes in tropical regions, *Progress in Physical Geography*, **21**, 375-393.

Law, N.F. and Nguyen, D.T., 1995, Multiple frame projection based blind deconvolution, *Electronics Letters*, **31**, 1734-1735.

Leavers, V.F., 1993, Which Hough transform, *CVGIP – Image Understanding*, **58**, 250-264.

Lewis, H.G., 1998, The use of shape, appearance and the dynamics of clouds for satellite image interpretation, *Ph.D. Thesis*, University of Southampton, Southampton.

Lewis, H.G. and Brown, M., 1998, *FLIERS Software Framework and Toolkit Version 2.0; User Guide*, University of Southampton.

Lewis, H.G. and Brown, M., 1999, A generalized confusion matrix for assessing area estimates from remotely-sensed data, *International Journal of Remote Sensing*, In Press.

Lewis, H.G. and Gunn, S.R., 1999, Estimation of sub-pixel land cover using support vector methods, In *Proceedings of the 25th Remote Sensing Society Conference*, Taylor and Francis, London, 395-402.

Lewis, H.G., Nixon, M.S., Tatnall, A.R.L. and Brown, M., 1999a, Appropriate strategies for mapping land cover from satellite imagery, In *Proceedings of the 25th Remote Sensing Society Conference,* Taylor and Francis, London, 717-724.

Lewis, H.G., Nixon, M.S., Tatnall, A.R.L., Edwards, M. and Hughes, M., 1999b, On the role of uncertainty in empirical mixture modelling, In *Proceedings of the 25th Remote Sensing Society Conference,* Taylor and Francis, London, 709-716.

Lewis, H.G., Brown, M. and Tatnall, A.R.L., 2000, Incorporating uncertainty in land cover classification from remote sensing imagery, *Advances in Space Research*, **26**, 1123-1126.

Li, R., Wang, W. and Tseng, H., 1999, Detection and location of objects from mobile mapping image sequences by Hopfield neural networks, *Photogrammetric Engineering and Remote Sensing*, **65**, 1199-1205.

Li, X., Kudo, M., Toyama, J. and Shimbo, M., 1998, Knowledge-based enhancement of low spatial resolution images, *IEICE Transactions on Information and Systems*, **81**, 457-463.

Lin, C., Lee, Y. and Pu, H., 2000, Satellite sensor image classification using cascaded architecture of neural fuzzy network, *IEEE Transactions on Geoscience and Remote Sensing*, **38**, 1033-1043.

Lobo, A., Moloney, K., Chic, O. and Chiariello, N., 1998, Analysis of fine-scale spatial pattern of a grassland from remotely-sensed imagery and field collected data, *Landscape Ecology*, **13**, 111-131.

Manslow, J., Brown, M. and Nixon, M., 1999, On the probabilistic interpretation of area based fuzzy land cover mixing proportions, In Lek, S. and Guegan, J. (eds): *Artificial neural networks in ecology and evolution*, Springer Verlag, Berlin.

Markham, B.L., 1985, The landsat sensors' spatial responses, *IEEE Transactions on Geoscience and Remote Sensing*, **23**, 864-875.

Martin-Valdivia, M., Ruiz-Sepulveda, A. and Triguero-Ruiz, F., 2000, Improving local minima of Hopfield networks with augmented Lagrange multipliers for large scale TSPs, *Neural Networks*, **13**, 283-285.

Maselli, F., Conese, C. and Petkov, L., 1994, Use of probability entropy for the estimation and graphical representation of the accuracy of maximum likelihood classifications, *ISPRS Journal of Photogrammetry and Remote Sensing*, **49**, 13-20.

Matheron, G., 1965, Les variables régionalisées et leur estimation, Masson, Paris.

Matsakis, P., Andrefoet, S and Capolsini, P., 2000, Evaluation of fuzzy partitions, *Remote Sensing of Environment*, **74**, 516-533.

McGillem, C.D., Yu, K.B., Anuta, P.E. and Malaret, E., 1983, Estimation of a remote sensing system point spread function from measured imagery, *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, Lafayette Ind., 62-68.

Melgani, F., Al Hashemy, B.A.R. and Taha, S.M.R., 2000, An explicit fuzzy supervised classification method for multispectral remote sensing images, *IEEE Transactions on Geoscience and Remote Sensing*, **38**, 287-295.

Nasrabadi, N.M. and Choo, C.Y., 1992, Hopfield network for stereo vision correspondence, *IEEE Transactions on Neural Networks*, **3**, 5-13.

Nichani, S., 1994, Solving the correspondence problem using a Hopfield network, *Proceedings of the IEEE international conference on neural networks,* ICNN 94, June 1994, Orlando, FL, 4107-4112.

Nishii, R. and Tanaka, S., 1999, Accuracy and inaccuracy assessments in land-cover classification, *IEEE Transactions on Geoscience and Remote Sensing*, **37**, 491-498.

Oliver, M., Webster, R. and Gerrard, J., 1989a, Geostatistics in Physical Geography, Part 1: theory, *Transactions of the Institute of British Geographers*, **14**, 259-269.

Oliver, M., Webster, R. and Gerrard, J., 1989b, Geostatistics in physical geography, Part 2: applications, *Transactions of the Institute of British Geographers*, **14**, 270-286.

Oleson, K.W., Sarlin, S., Garrison, J., Smith, S., Privette, J.L. and Emery, W.J., 1995, Unmixing multiple land-cover type reflectances from coarse spatial resolution satellite data, *Remote Sensing of Environment*, **54**, 98-112.

O'Neill, R.V., Krummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., Deangelis, D.L., Milne, B.T., Turner, M.G., Zygmunt, B., Christensen, S.W., Dale, V.H. and Graham, R.L., 1988, Indices of landscape pattern, *Landscape Ecology*, **1**, 153-162.

Osborn, K.J., 1998, United States Mexico transboundary aerial photography and mapping initiative, *Photogrammetric Engineering and Remote Sensing*, **64**, 1085-1088.

Paola, J.D. and Schowengerdt, R.A., 1995, A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery, *International Journal of Remote Sensing*, **16**, 3033-3058.

Patterson, D.W., 1996, *Artificial Neural Networks: theory and applications*, Prentice Hall, London.

Pax-Lenney, M. and Woodcock, C.E., 1997, The effect of spatial resolution on the ability to monitor the status of agricultural lands, *Remote Sensing of Environment*, **61**, 210-220.

Puyou-Lascassies, P., Flouzat, G., Gay, M. and Vignolles, C., 1994, Validation of the use of multiple linear regression as a tool for unmixing coarse spatial resolution images, *Remote Sensing of Environment*, **49**, 155-166.

Quarmby, N.A., Townshend, J.R.G., Settle, J.J., Milnes, M., Hindle, T.L. and Silleos, N., 1992, Linear mixture modelling applied to AVHRR data for crop area estimation, *International Journal of Remote Sensing*, **13**, 415-425.

Raghu, P.P. and Yegnanarayana, B., 1996, Segmentation of Gabor-filtered textures using deterministic relaxation, *IEEE Transactions on Image Processing*, **5**, 1625-1636.

Riley, R.H., Phillips, D.H., Schuft, M.J. and Garcia, M.C., 1997, Resolution and error in measuring land cover change: effects on estimating net carbon release from Mexican terrestrial ecosystems, *International Journal of Remote Sensing*, **18**, 121-137.

Ripley, B.D., 1986, *Stochastic Simulation*, Wiley, Chichester.

Robinson, G.D., Gross, H.N. and Schott, J.R., 2000, Evaluation of two applications of spectral mixing models to image fusion, *Remote Sensing of Environment*, **71**, 272-281.

Schneider, W., 1993, Land use mapping with subpixel accuracy from landsat TM image data, *Proceedings of the 25th International Symposium on Remote Sensing and Global Environmental Change*, Michigan, USA, 155-161.

Schneider, W., 1999, Land cover mapping from optical satellite images employing subpixel segmentation and radiometric calibration, in Kanellopoulos, I., Wilkinson, G. and Moons, T. (eds.), *Machine Vision and Advanced Image Processing in Remote Sensing*, Springer, London.

Schowengerdt, R.A., 1996, On the estimation of spatial-spectral mixing with classifier likelihood functions, *Pattern Recognition Letters*, **17**, 1379-1387.

Schowengerdt, R.A., 1997, *Remote Sensing: Models and Methods for Image Processing*, Academic Press, San Diego.

Settle, J.J. and Drake, N.A., 1993, Linear mixing and the estimation of ground cover proportions, *International Journal of Remote Sensing*, **14**, 1159-1177.

Shepherd, I., Wilkinson, G. and Thompson, J., 2000, Monitoring surface water storage in the North Kent Marshes using Landsat TM images, *International Journal of Remote Sensing*, **21**, 1843-1865.

Shimabukuro, Y.E. and Smith, J.A., 1991, The least-squares mixing models to generate fraction images derived from remote sensing multispectral data, *IEEE Transactions on Geoscience and Remote Sensing*, **29**, 16-20.

Smith, G.R., Woodward, J.C., Heywood, D.I. and Gibbard, P.L., 2000, Interpreting Pleistocene glacial features from SPOT HRV data using fuzzy techniques, *Computers and Geosciences*, **26**, 479- 490.

Stehman, S.V., 1997, Selecting and interpreting measures of thematic classification accuracy, *Remote Sensing of Environment*, **62**, 77-89.

Steinwendner, J. and Schneider, W., 1997, A neural net approach to spatial subpixel analysis in remote sensing, in Burger, W and Burge, M. (eds.), *Pattern Recognition 1997, Proceedings of the 21st OAGM Workshop,* Halstatt, Austria, OCG-Shriftenreihe.

Steinwendner, J. and Schneider, W., 1998, Algorithmic improvements in spatial subpixel analysis of remote sensing images, *Proceedings of the 22nd workshop of the Austrian Association of Pattern Recognition*, 205-213.

Steinwendner, J., Schneider, W. and Suppan, F., 1998, Vector segmentation using spatial subpixel analysis for object extraction, *International Archives of Photogrammetry and Remote Sensing*, **32**, 265-271.

Steinwendner, J., 1999, From satellite images to scene description using advanced image processing techniques, *Proceedings of the 25th Remote Sensing Society Conference,* Taylor and Francis, London, 865-872.

Storkey, A.J. and Valabregue, R., 1999, The basins of attraction of a new Hopfield learning rule, *Neural Networks*, **12**, 869-876.

Tailor, A., Cross, A., Hogg, D.C. and Mason, D.C., 1986, Knowledge-based interpretation of remotely sensed images, *Image and Vision Computing*, **4**, 67-83.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S., 2001a, Super-resolution target identification from remotely sensed images using a Hopfield neural network, *IEEE Transactions on Geoscience and Remote Sensing*, **39**, 781-796.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S., 2001b, Land cover mapping at the sub-pixel scale using a Hopfield neural network, *International Journal of Applied Earth Observation and Geoinformation*, in press.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S., 2001c, Increasing the spatial resolution of satellite imagery for agricultural land cover mapping, In *Proceedings of the 1st Annual Meeting of the Remote Sensing and Photogrammetry Society,* The Remote Sensing Society, Nottingham.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S., 2001d, Super-resolution land cover pattern prediction using a Hopfield neural network, *Remote Sensing of Environment*, in press.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S., 2001e, Super-resolution mapping of multiple scale land cover features using a Hopfield neural network, In *Proceedings of the International Geoscience and Remote Sensing Symposium*, IEEE, Sydney.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S., 2001f, Super-resolution mapping of urban scenes from IKONOS imagery using a Hopfield neural network, In *Proceedings of the International Geoscience and Remote Sensing Symposium*, IEEE, Sydney.

Thornton, P.K. and Jones, P.G., 1998, A conceptual approach to dynamic agricultural land-use modelling, *Agricultural Systems*, **57**, 505-521.

Townsend, P.A., 2000, A quantitative fuzzy approach to assess mapped vegetation classifications for ecological applications, *Remote Sensing of Environment*, **72**, 253-267.

Townshend, J.R.G., 1992, Land cover, *International Journal of Remote Sensing*, **13**, 1319-1328.

Townshend, J.R.G., Huang, C., Kalluri, S.N.V., Defries, R.S and Liang, S., 2000, Beware of per-pixel characterization of land cover, *International Journal of Remote Sensing*, **21**, 839-843.

Tubbs, J.D. and Coberly, W.A., 1978, Spatial correlation and its effect upon classification results in landsat, *Proceedings of the 12th International Symposium on Remote Sensing of the Environment*, 775-81.

Um, J.S. and Wright, R., 1999, The analog-to-digital transition and implications for operational use of airborne videography, *Photogrammetric Engineering and Remote Sensing*, **65**, 269-275.

Vapnik, V.N., 1995, *The Nature of Statistical Learning Theory*, Springer-Verlag, London.

Veitch, N., Treweek, J.R. and Fuller, R.M., 1995, The land cover map of Great Britain – a new data source for environmental planning and management, in Danson, F.M. and Plummer, S.E. (eds.), *Advances in Environmental Remote Sensing*, John Wiley, Chichester, 157-170.

Verhoeye, J. and De Wulf, R., 2000, Land cover mapping at the sub-pixel scale using linear optimisation techniques, *Remote Sensing of Environment*, in press.

Vieira, C.A.O. and Mather, P.M., 1999, Assessing the accuracy of thematic classifications using remotely sensed data, Proceedings of the *25th Remote Sensing Society Conference,* Taylor and Francis, London, 897-903.

Wang F., 1990, Improving remote sensing image analysis through fuzzy information representation, *Photogrammetric Engineering and Remote Sensing*, **56**, 1163-1169.

Warner, T.A. and Shank, M., 1997, An evaluation of the potential for fuzzy classification of multispectral data using artificial neural networks, *Photogrammetric Engineering and Remote Sensing*, **63**, 1285-1294.

Weir, N. and Djorgovski, S., 1991, A subpixel deconvolution method for astronomical images, *Fundamental Theories of Physics*, **43**, 275-283.

Woodcock, C.E. and Strahler, A.H., 1987, The factor of scale in remote sensing, *Remote Sensing of Environment*, **21**, 311-332.

Wu, P.H. and Schowengerdt, R.A., 1993, Improved estimation of fraction images using partial image restoration, *IEEE Transactions on Geoscience and Remote Sensing*, **31**, 771-778.

Wyatt, B.K., Greatorex-Davies, J.N., Hill, M.O., Parr, T.W., Bunce, R.G.H. and Fuller, R.M., 1993, Comparison of land cover definitions, *Institute of Terrestrial Ecology report to the Department of the Environment*, Institute of Terrestrial Ecology, Cambridge.

Zhang, X.M., Van Genderen, J.L. and Kroonenberg, S.B., 1997, A method to evaluate the capability of Landsat-5 TM band 6 for sub-pixel coal fire detection, *International Journal of Remote Sensing*, **1**, 3279-3288.

Zhu, S.C. and Yuille, A., 1996, Region competition: unifying snakes, region growing and Bayes/M.D.L. for multiband image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**, 884-900.

Zhu, Z., Yang, L., Stehman, S.V. and Czaplewski, R.L., 2000, Accuracy assessment for the US Geological Survey regional land-cover mapping program: New York and New Jersey Region, *Photogrammetric Engineering and Remote Sensing*, **66**, 1425-1435.