Super-resolution land cover pattern prediction using a Hopfield neural network

A.J. Tatem[a,b,*], H.G. Lewis[a], P.M. Atkinson[b], M.S. Nixon[a]

[a]Department of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
[b]Department of Geography, University of Southampton, Southampton SO17 1BJ, UK

Abstract

Landscape pattern represents a key variable in management and understanding of the environment, as well as driving many environmental models. Remote sensing can be used to provide information on the spatial pattern of land cover features, but analysis and classification of such imagery suffers from the problem of class mixing within pixels. Soft classification techniques can estimate the class composition of image pixels. However, their output provides no indication of how such classes are distributed spatially within the instantaneous field-of-view (IFOV) represented by the pixel. Techniques to provide an improved spatial representation of land cover targets larger than the size of a pixel have been developed. However, the mapping of subpixel scale land cover features has yet to be investigated. We recently described the application of a Hopfield neural network technique to super-resolution mapping of land cover features larger than a pixel, using information of pixel composition determined from soft classification, and now show how our approach can be extended in a new way to predict the spatial pattern of subpixel scale features. The network converges to a minimum of an energy function defined as a goal and several constraints. Prior information on the typical spatial arrangement of the particular land cover types is incorporated into the energy function as a semivariance constraint. This produces a prediction of the spatial pattern of the land cover in question, at the subpixel scale. The technique is applied to synthetic and simulated Landsat Thematic Mapper (TM) imagery, and compared to results of an existing super-resolution target identification technique. Results show that the new approach represents a simple, robust, and efficient tool for super-resolution land cover pattern prediction from remotely sensed imagery. © 2002 Elsevier Science Inc. All rights reserved.

1. Introduction

The landscape is a complex, hierarchically organised, spatiotemporal mosaic, where there are strong relationships coupling spatial pattern to process (Lobo, Moloney, Chic, & Chiariello, 1998). Information on land cover is required to aid understanding and management of the environment. 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 over recent years. To understand and rigorously test 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 underconstrained nature. This paper presents a technique to model land cover pattern using remotely sensed imagery and prior information on land cover distribution to constrain the predictions.

Remotely sensed imagery derived from aircraft and satellite mounted sensors can provide information on land cover. However, there exist practical limitations (Tatem et al., 2001). Perhaps the biggest drawback of obtaining land cover information from remotely sensed images relates to

* Corresponding author. Tel.: +44-23-80592215.
E-mail address: a.j.tatem@soton.ac.uk (A.J. Tatem).

scale. Spatial scale is a key factor in the interpretation of remotely sensed land cover data (Woodcock & Strahler, 1987) and the information obtainable from remotely sensed imagery can vary greatly depending on the spatial variation in the observed land cover and the specific terrain characteristics under consideration. There also exist practical limits to the level of detail that can be identified by each remote sensor and these limits are defined by the resolutions of the remote sensing system. One of the commonest measures of image characteristic used is spatial resolution, which determines the level of spatial detail depicted in an image. This measure is a function of the instantaneous field-of-view (IFOV) of a sensor, defined as the cone angle within which incident energy is focused on the detectors (Campbell, 1996). In turn, the IFOV leads to a ground resolution element (GRE) on the surface of the Earth (this GRE should not be confused with the pixel, which is the output product to which a radiance value is assigned). The pixel represents the smallest element of a digital image and has, therefore, traditionally represented a limit to the spatial detail obtainable in land cover feature extractions from remotely sensed imagery. Within remote sensing images, a significant proportion of pixels may be of mixed land cover class composition, and the presence of such mixed pixels can affect adversely the performance of image analysis and classification operations (Fisher, 1997).

Traditionally, classification approaches have focused on 'hard', one-class-per-pixel techniques (Campbell, 1996). Such approaches were found to be inaccurate, inappropriate, and an alternative approach has become more common, in the form of soft classification. Subpixel class composition is estimated through the use of techniques such as spectral mixture modelling (Garcia-Haro, Gilabert, & Melia, 1996), multilayer perceptrons (Atkinson, Cutler, & Lewis, 1997), nearest neighbour classifiers (Schowengerdt, 1997), and support vector machines (Brown, Gunn, & Lewis, 1999). These approaches allow soft proportions of each pixel to be partitioned between classes. The output of these techniques generally takes the form of a set of proportion images, each displaying the proportion of a certain class within each pixel. In most cases, this results in a more appropriate and informative representation of land cover than that produced using a hard, one-class-per-pixel classification. However, 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 accurately identifying them and using them within models.

Much previous research has been centred on attempting to extract data on subpixel scale features from remote sensing imagery. For example, information has been extracted about subpixel scale volcano vents (Bhattacharya, Reddy, & Srivastav, 1993), coal fires (Zhang, Van Genderen, & Kroonenberg, 1997), glacial features (Smith, Woodward,

Heywood, & Gibbard, 2000), and water storage ditches (Shepherd, Wilkinson, & Thompson, 2000). However, in all cases, the features have been merely detected using soft classification techniques, and no attempt at locating them has been made. This paper demonstrates that it is possible to identify subpixel scale land cover targets *and* recreate their spatial distribution across an image.

1.1. Previous super-resolution land cover target identification and mapping

Only recently has research been done on the subject of identifying and mapping land cover from remotely sensed images at the subpixel scale. Each technique put forward has produced a certain amount of success at super-resolution mapping of land cover targets larger than a pixel, but research has yet to be carried out on locating subpixel scale features. This paper attempts to build on the ideas of these existing techniques to provide a solution to the problem of subpixel scale land cover feature mapping.

Schneider (1993) introduced a knowledge-based analysis technique for the automatic localisation of field boundaries with subpixel accuracy. The technique relies on knowledge of straight boundary features within Landsat Thematic Mapper (TM) scenes, and serves as a preprocessing step prior to automatic pixel-by-pixel land cover classification. It represents a successful, automated, and simple preprocessing step for increasing the spatial resolution of satellite sensor 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, Gahegan, and West (1994) also concentrated on subpixel mapping at the borders of agricultural fields and used edge detection and segmentation techniques to identify field boundaries and the Hough transform to identify the straight, subpixel boundaries. These vector boundaries were superimposed on a subsampled version of the image, and the mixed pixels were reassigned to each side of the boundaries. However, no validation or further work was carried out, and so the success of the technique remains unclear. Aplin, Atkinson, and Curran (1999) utilised Ordnance Survey land line vector data, and undertaking per-field rather than the traditional per-pixel land cover classification, mapping at a subpixel scale was demonstrated. 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.

The techniques described so far are based on direct processing of the raw imagery. In other research, soft classification is first applied to the imagery, and an attempt is then made to map the location of class components within the pixels. 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. However, the complex mixing in the data caused the simple technique to suffer from problems due to the existence of subpixel scale land cover features. Unlike Atkinson, Foody (1998) made use of a higher spatial resolution image in a simple regression and contouring-based approach to sharpen the output of a soft classification of a lower spatial resolution image, producing a subpixel land cover map. However, the areal extent of the lake was not maintained using the contouring technique and generally, it is difficult to obtain two coincident images of differing spatial resolution. 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, but the complex multistage operation meant it was slow and only applicable to small images containing features larger than a pixel. Verhoeye and De Wulf (2000) used similar assumptions to Atkinson, but formulated the approach as a linear optimisation problem. Results showed a certain degree of success, but problems were noted where land cover features were smaller than a pixel.

Tatem et al. (2001) presented a technique that allowed prior information to be included in the prediction process. A Hopfield neural network was formulated as an energy minimisation tool to predict the land cover distribution within each pixel. 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 ability to incorporate prior information is extended to account for the spatial pattern of objects smaller than the ground resolution of the sensor. The long-term aim of both works is to establish a method for identifying the spatial arrangement of land cover objects at *any* scale. The results from the technique described in Tatem et al. are used for comparison with the results from the pattern prediction technique in Section 4.

The focus of each of the techniques described so far on land cover features larger than the scale of 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 paper 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 technique proposed by Tatem et al. (2001). 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. The nature of this approach should allow the production of differing, equally probable solutions each time the network is run. Therefore, the network can run multiple times and generate different realisations, each of which represents a possible land cover distribution.

2. The Hopfield neural network

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 (Fig. 1). Hopfield (1984) shows how the equation describing the behaviour of such an array of electronic components can be written in a neural context for ease of interpretation where, in this case, the nonlinear amplifiers correspond to neurons:

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

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

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

where λ determines the steepness of the function.

The set of 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 $g_i(u_i)$ (Eq. (2)) approaches a step function. The stable states of the network consequently correspond to the local minima of the following 'energy function' (Cichocki & Unbehauen, 1993),

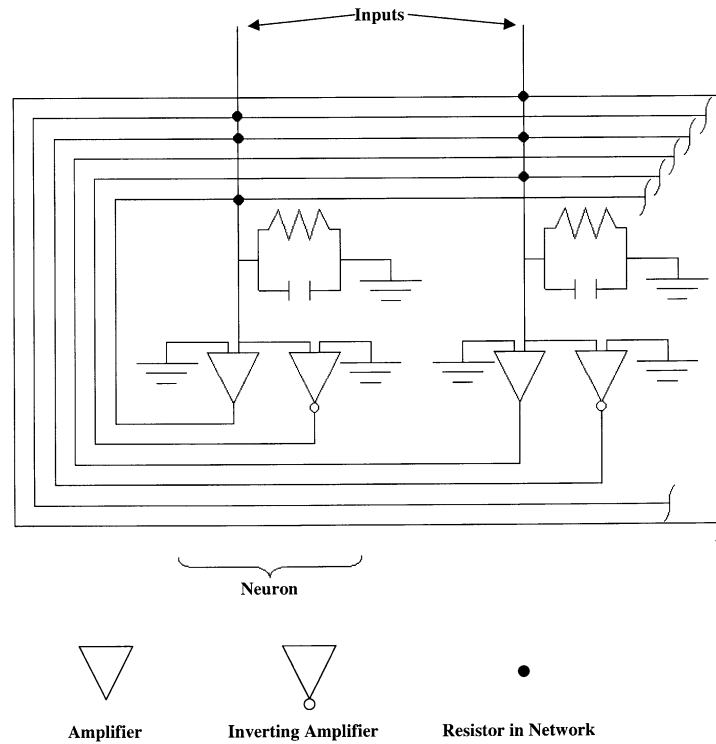$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} T_{ij} v_i v_j - \sum_{i=1}^{N} v_i I_i \qquad (3)$$

Fig. 1. Hopfield neural network as an analog circuit. The black circles at the intersections represent resistive connections between outputs and inputs. Connections between inverted outputs and inputs represent negative connections.

where E is the energy calculated over the whole network. From Eqs. (1) and (3), the equation describing the dynamics, i.e., the rate of change of neuron input of the Hopfield network can be written as:

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = -\frac{\delta E}{\delta v_i},$$
(4)

or

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = -\sum_{j=1}^{N} T_{ij}v_j + I_i.$$
(5)

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 & Tank, 1985). By specifying different values for the weights and biases, any hypothetical energy minimisation problem can be simulated.

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.

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 (Eq. (6)):

$$\text{Energy} = \text{Goal} + \text{Constraints}.$$
(6)

If the energy function is arranged in this particular way, the constraints become part of the minimisation process, which means that each does not need to be treated separately, just weighted by their importance to the problem. The Hopfield network process then finds the minimum energy that represents a compromise between the goal and the constraints.

The Hopfield network has been used within remote sensing for ice-mapping, cloud motion, and ocean current tracking (Côté & Tatnall, 1997; Lewis, 1998). These 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. In addition, more general feature matching problems have been tackled using a Hopfield network (Forte & Jones, 1999; Li, Wang, & Tseng, 1999; Nasrabadi & Choo, 1992), as well as applications such as recognition or classification (Campadelli, Medici, & Schettini, 1997; Raghu & Yegnanarayana, 1996).

3. Using the Hopfield network for super-resolution land cover pattern prediction

The input data for the research described in this paper 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 multilayer perceptron applied to real imagery, providing the target features to be mapped were separable enough from the other land cover classes to be accurately identified in classification. However, for the research in this paper, 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.

Mapping the spatial distribution of the class components within each pixel 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. The procedure is outlined in detail in the following sections.

3.1. Network architecture

In many papers on the use of Hopfield neural networks for optimization, the spatial relations between neurons are considered irrelevant. However, for this paper, 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 (Fig. 2). Therefore, neurons will be referred to by coordinate 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, determines the increase in spatial resolution from the original satellite sensor image to the new high-resolution image and after convergence to a stable state, the neurons represent a bipolar classification of the land cover *at the higher spatial resolution*. Fig. 2 shows the notation used in this paper, and how coordinates are transformed linearly from the image space to the network neuron space, for example, the pixel (x,y) in the satellite sensor image is represented by $z \times z$ neurons centred at coordinates $[xz + \text{int}(z/2), yz + \text{int}(z/2)]$, where int is the integer value.

3.2. Network initialisation

Each neuron is initialised with a starting value, u_{init}, and two strategies for initialising the network exist.

(i) Each set of neurons representing a pixel in the low-resolution image is identified and a proportion of this set is randomly given an output of $u_{\text{init}} = 0.55$. This proportion is equal to the actual area proportion of the class within the image pixel and the remaining neurons of the set are given an output of $u_{\text{init}} = 0.45$. 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 initially given 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 $u_{\text{init}} = [0.45, 0.55]$. This allows performance comparison with the class proportion-defined initialisation and does not introduce any possible unnecessary bias into the result, which may occur using (i) should estimated class proportions be inaccurate.

Tatem et al. (2001) demonstrated no significant differences in results from the two initialisation techniques.

3.3. Implementation

When implemented on a digital computer, sets of biases and weights do not need to be determined, as the network is simulated via its equation of motion (Eq. (5)) using the Euler method:

$$u_{ij}(t + \mathrm{d}t) = u_{ij}(t) + \frac{\mathrm{d}u_{ij}(t)}{\mathrm{d}t}\,\mathrm{d}t, \tag{7}$$

where $\mathrm{d}t$ is the time step of the iterative method and the function $(\mathrm{d}u_{ij}(t)/\mathrm{d}t)$ is measured using $(\mathrm{d}E_{ij}/\mathrm{d}v)$. Eq. (4)
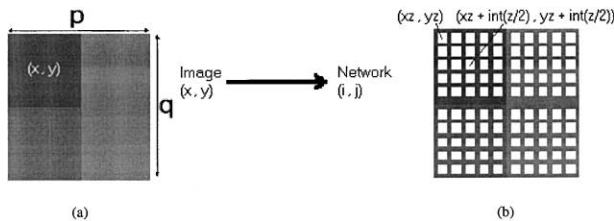

Fig. 2. (a) 2×2 pixel image, p and q represent the image dimensions, x and y represent the image pixel coordinates; (b) Representation of the Hopfield network for the image in (a), i and j represent the neuron coordinates (int = integer value).

shows the correspondence between the two functions, and $(\mathrm{d}E_{ij}/\mathrm{d}v)$ is determined using the goals and constraint of the super-resolution target identification task. Eq. (7) is run until $\sum_{i,j} u_{ij}(t + \mathrm{d}t) - u_{ij}(t) \leq \mathrm{d}u_c$, where $\mathrm{d}u_c$ is a sufficiently small value, and the equations of motion were defined as:

$$\frac{\mathrm{d}E_{ij}}{\mathrm{d}v_{ij}} = \left(\sum_{n=1}^{z} k_n \frac{\mathrm{d}Sn_{ij}}{\mathrm{d}v_{ij}}\right) + k_{z+1} \frac{\mathrm{d}P_{ij}}{\mathrm{d}v_{ij}}. \tag{8}$$

Each component of Eq. (8) is described in the subsequent sections.

3.4. The energy function

The goal and constraints of the subpixel mapping task were defined such that the network energy function for a zoom factor of 7 was:

$$E = -\sum_i \sum_j (k_1 S1_{ij} + k_2 S2_{ij} + k_3 S3_{ij} + k_4 S4_{ij}$$
$$+ k_5 S5_{ij} + k_6 S6_{ij} + k_7 S7_{ij} + k_8 P_{ij}), \tag{9}$$

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 3.4.1), and these correspond to the quadratic term in Eq. (3). 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 Eq. (3).

3.4.1. The semivariance functions

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

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1,j=1}^{N(h)} [v_{ij} - v_{i \pm h, j \pm h}]^2 \tag{10}$$

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 is calculated for z lags from an aerial photograph for example. This provides information on the typical spatial distribution of the land cover under study, which can then be used for land cover simulation from remotely sensed imagery at the subpixel scale.

By using the values of $\gamma(h)$ from Eq. (10), the output of the centre neuron, $v(c)_{ij}$, which produces a semivariance of $\gamma(h)$ can be calculated using:

$$v(c)_{ij} = \frac{1}{2a}[-b \pm \sqrt{b^2 - 4ac}] \tag{11}$$

where $a = 2N(h)$, $b = \sum_{i=1,j=1}^{N(h)} v_{1 \pm h, j \pm h}$, $c = \sum_{i=1,j=1}^{N(h)} (v_{i \pm h, j \pm h})^2$ $-2N(h)\gamma(h)$. The semivariance function value for lag 1 $(h=1)$, $S1_{ij}$, is then given by (Eq. (12)):

$$\frac{\mathrm{d}S1_{ij}}{\mathrm{d}v_{ij}} = v_{ij} - v(c)_{ij}. \tag{12}$$

If the output of neuron (i,j), v_{ij}, is lower than the target value, $v(c)_{ij}$, calculated in Eq. (11), a negative gradient is produced that corresponds to an increase in neuron output to counteract this problem. An overestimation of neuron output results in a positive gradient, producing a decrease in neuron output. Only when the neuron output is identical to the target output does a zero gradient occur, corresponding to $S1_{ij}=0$ in the energy function (Eq. (9)). The same calculations are carried out for lags 2 to 7, to produce values for $S2_{ij}$ to $S7_{ij}$.

3.4.2. The proportion constraint

While the semivariance functions provide the enforcement of a certain spatial pattern, using only these functions would result 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. 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 higher was calculated for all the neurons representing pixel (x,y) (Eq. (13)):

Area Proportion Estimate
$$= \frac{1}{2z^2} \sum_{k=xz}^{xz+z} \sum_{l=yz}^{yz+z} (1 + \tanh(v_{kl} - 0.55)\lambda). \tag{13}$$

The use of the tanh function ensures that if a neuron output is above 0.55, it is counted as having an output of 1 within the estimation of class area per pixel. Below an output of 0.55, the neuron is not counted within the estimation, which simplifies the area proportion estimation procedure, and ensures that the neuron output must exceed the random initial assignment output of 0.55 in order 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 (Eq. (14)):

$$\frac{\mathrm{d}P_{ij}}{\mathrm{d}v_{ij}} = \frac{1}{2z^2} \sum_{k=xz}^{xz+z} \sum_{l=yz}^{yz+z} (1 + \tanh(v_{kl} - 0.55)\lambda) - a_{xy}. \tag{14}$$

If the area proportion estimate for pixel (x,y) is lower than the target area, a negative gradient is produced, which corresponds to an increase in neuron output to counteract this problem. An overestimation of class area results in a positive gradient, producing a decrease in neuron output. Only when the area proportion estimate is identical to the target area proportion for each pixel does a zero gradient occur, corresponding to $P_{ij} = 0$ in the energy function (Eq. (9)).

3.5. Advantages of the technique

The approach described in this paper holds several strategic advantages over those techniques mentioned in Section 1.1:

- 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 multistage operation.
- The effect that each one of these constraints has on the final prediction image can be controlled simply via weightings.

4. 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 Tatem et al. (2001), results of both techniques were compared to highlight the need for a dual approach. To determine the success of each technique at recreating the spatial pattern exhibited in the target images, plots of lag against semivariance (a variogram; Curran & Atkinson, 1998), calculated using Eq. (10), 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 association between the various plots. This was given by:

$$\rho_{tp} = \frac{\text{Cov}(t,p)}{\sigma_t \sigma_p}, \tag{15}$$

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

In Eq. (15), 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 is also calculated to give an indication of the success of the proportion constraint.

4.1. 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 making improvements to it becomes easier.

Figs. 3 and 4 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, Fig. 3(a): sparse natural woodland or seminatural vegetation; Fig. 3(b): dense natural woodland or seminatural vegetation; Fig. 4(a): woodland plantation or housing; Fig. 4(b): intergrade between land covers.

Each 56×56 pixel synthetic image in Figs. 3 and 4 was subsampled (using a 7×7 averaging filter) to generate an 8×8 pixel image. This causes 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 Eq. (10) 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 $k_7 = 0.2$ and $k_8 = 1.4$, which gave the semivariance and proportion functions equal weighting, ensuring that neither had a dominant effect. After 1000 iterations (approximately 15 s running time), the results in Figs. 3 and 4 were produced.

Figs. 3 and 4 demonstrate the effectiveness of the pattern prediction technique in maintaining the spatial layout, while mapping subpixel scale features. Visual comparison of the target images and variograms with the predictions of the two techniques suggests the pattern prediction technique is more successful, and the variogram correlation coefficients confirm this.

The earlier spatial clustering-based technique (Tatem et al., 2001) 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 Fig. 3(a), the clustering technique proves more successful, although the over- or underestimation 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 more accurately maintained 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 tends to exhibit the same shape, regardless of input image. The variogram correlation coefficients for Figs. 3(a) and 4(a) show significant differences between the two techniques in approximating the target semivariance values, while Fig. 4(b) shows the highest correlation of .987 for the pattern prediction technique. Only in Fig. 3(b) does the clustering technique show a better variogram correlation with the target image.
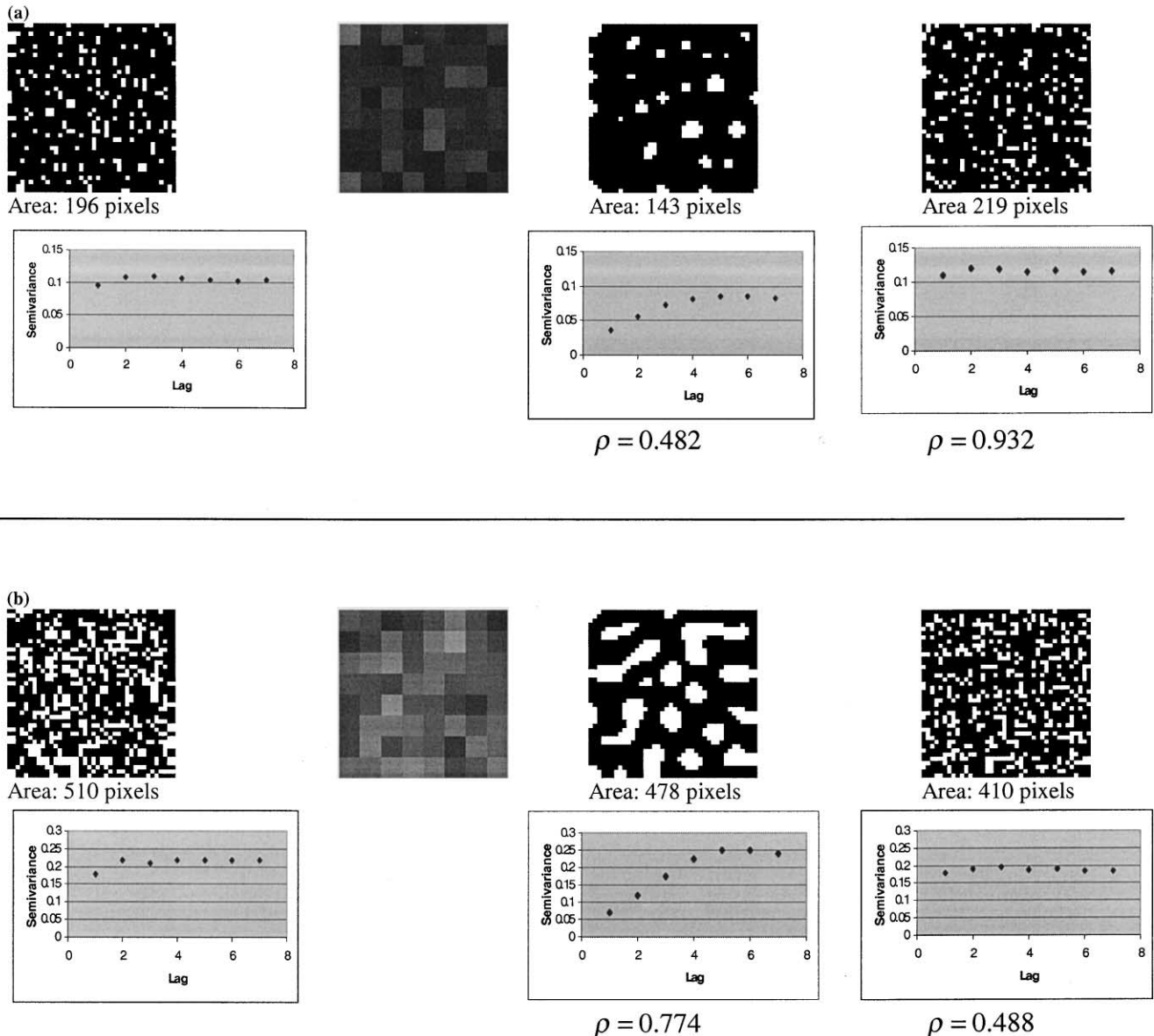

Fig. 3. Results of the Hopfield neural network predictions on synthetic imagery.

4.2. Simulated remotely sensed imagery

Figs. 5 and 6 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 Figs. 5(c) and 6(c). After 1000 iterations

(approximately 80 s running time) at a zoom factor of 7, the results in Figs. 7 and 8 were produced.

Figs. 7 and 8 again demonstrate the effectiveness of the pattern prediction technique in mapping subpixel 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 that the pattern prediction technique is more

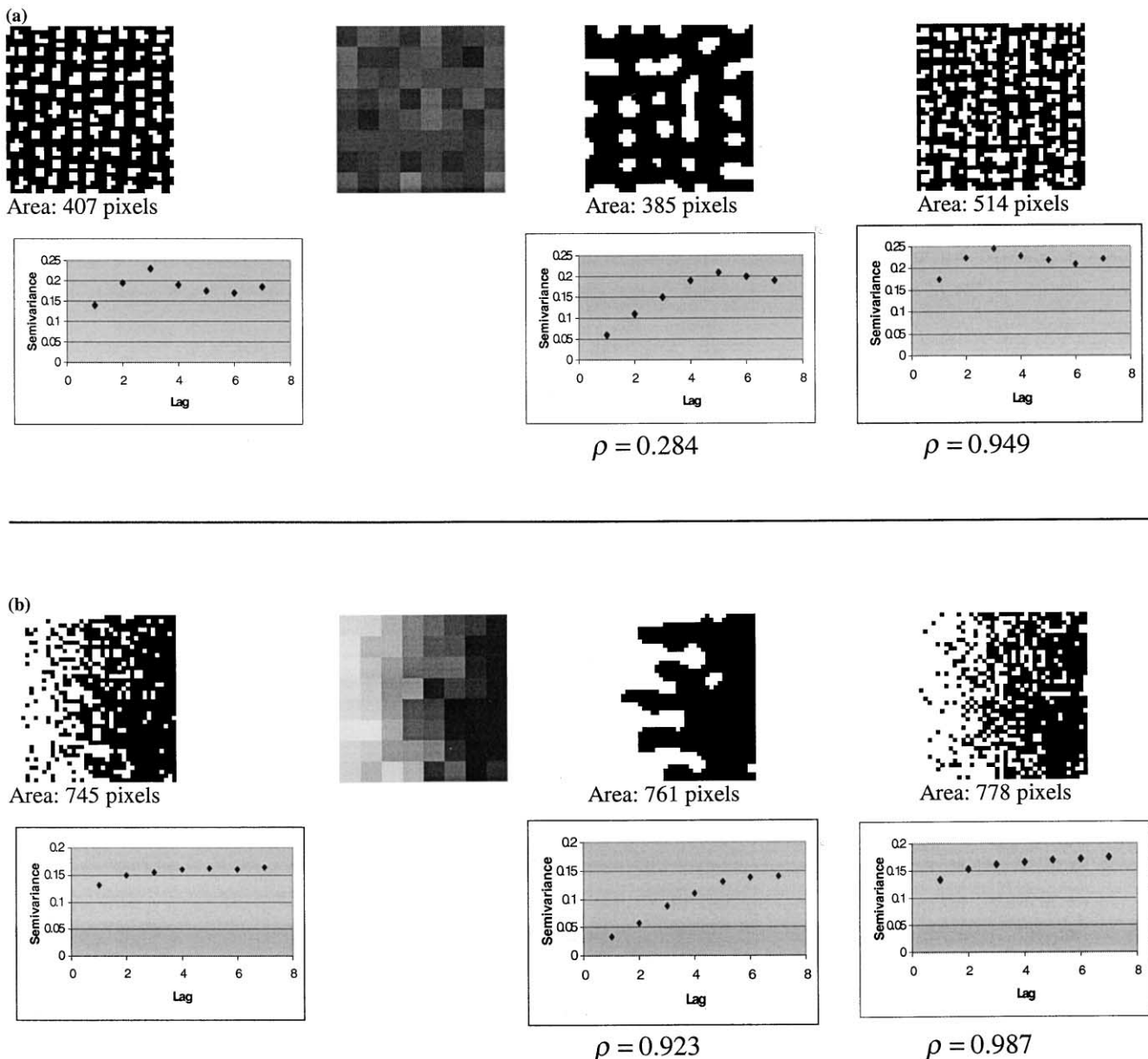| **Target Image, Class Area and Variogram**	**Subsampled Image (Network input)**	**Tatem et al. (2001) Prediction Image, Class Area, Variogram and Correlation Coefficient**	**Pattern Prediction Image, Class Area, Variogram and Correlation Coefficient**

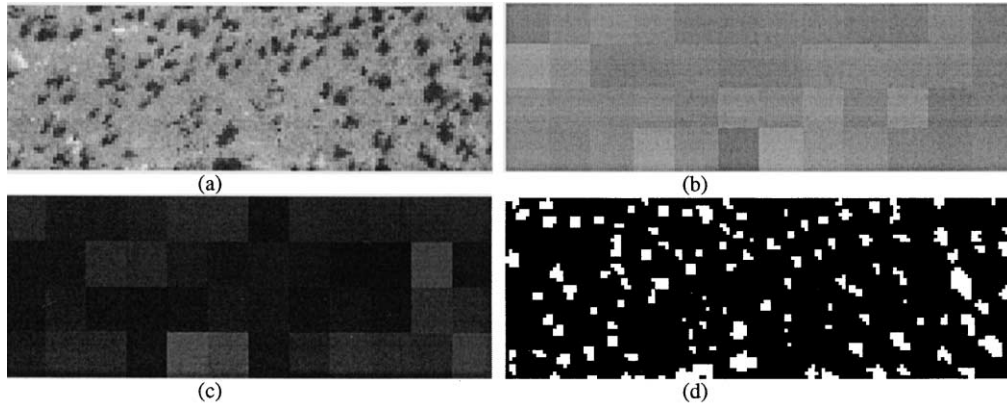Fig. 4. Results of the Hopfield neural network predictions on synthetic imagery.

Fig. 5. (a) Digital aerial photo of an area of the New Forest, UK; (b) 12 × 4 pixel simulated Landsat TM image (for illustration only); (c) Tree class proportions derived from verification data; (d) Verification image derived from aerial photography.

successful, and the variogram correlation coefficients confirm this.

In the same way as for Fig. 3, both Figs. 7 and 8 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 paper appears to maintain a similar spatial structure to that of the target images. Visual comparison shows that in Fig. 7, the
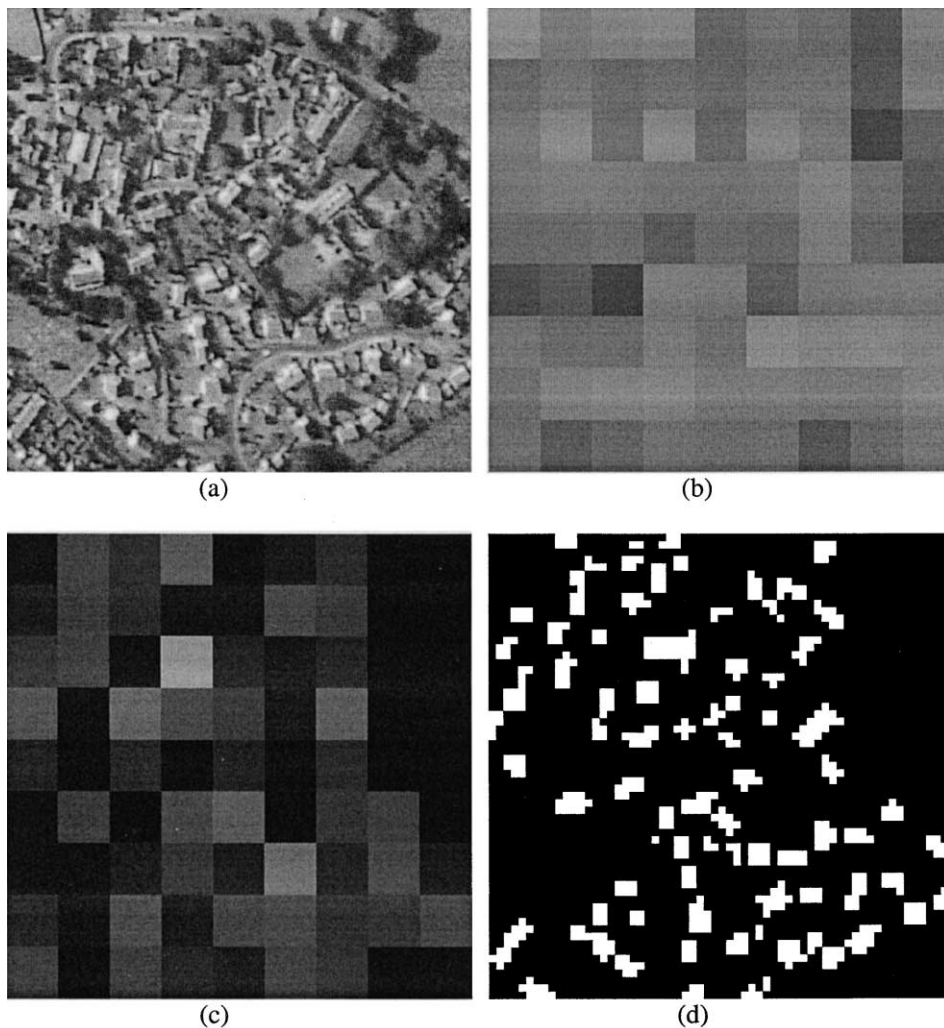

Fig. 6. (a) Digital aerial photo of an area of Bath, UK; (b) 12 × 4 pixel simulated Landsat TM image (for illustration only); (c) Building class proportions derived from verification data; (d) Verification image derived from aerial photography.
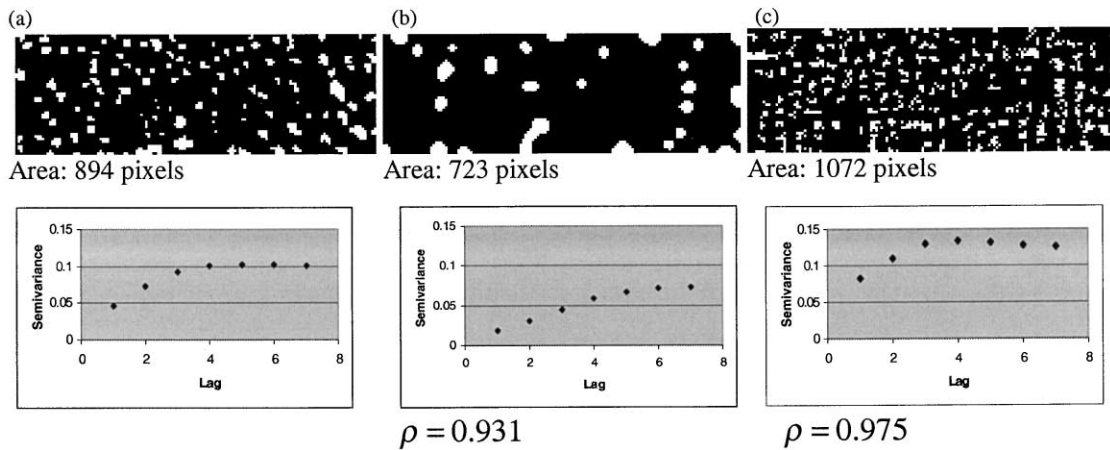
Fig. 7. (a) Verification image, class area, and variogram; (b) Hopfield network prediction image, class area, variogram, and variogram correlation coefficient (Tatem, Lewis, Atkinson, & Nixon, 2001 technique); (c) Hopfield network prediction image, class area, variogram, and variogram correlation coefficient (pattern prediction technique).

pattern prediction image contains more and slightly smaller trees than the target image, and in Fig. 8, the technique predicts fewer and slightly larger buildings than those depicted in the target image. 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 Figs. 7 and 8 indicates that the relative semivariance between lags has been

more successfully maintained using the pattern prediction technique. The shapes of each pattern prediction technique plot in Figs. 7 and 8 match the target plots better than any of the clustering technique plots, and the correlation coefficients confirm this. Fig. 7 shows that while the clustering technique plot correlates with the target plot to a level of .931, the pattern prediction plot improves on this with .975, demonstrating a strong correlation. The success of the pattern prediction technique in matching the semivariance of the
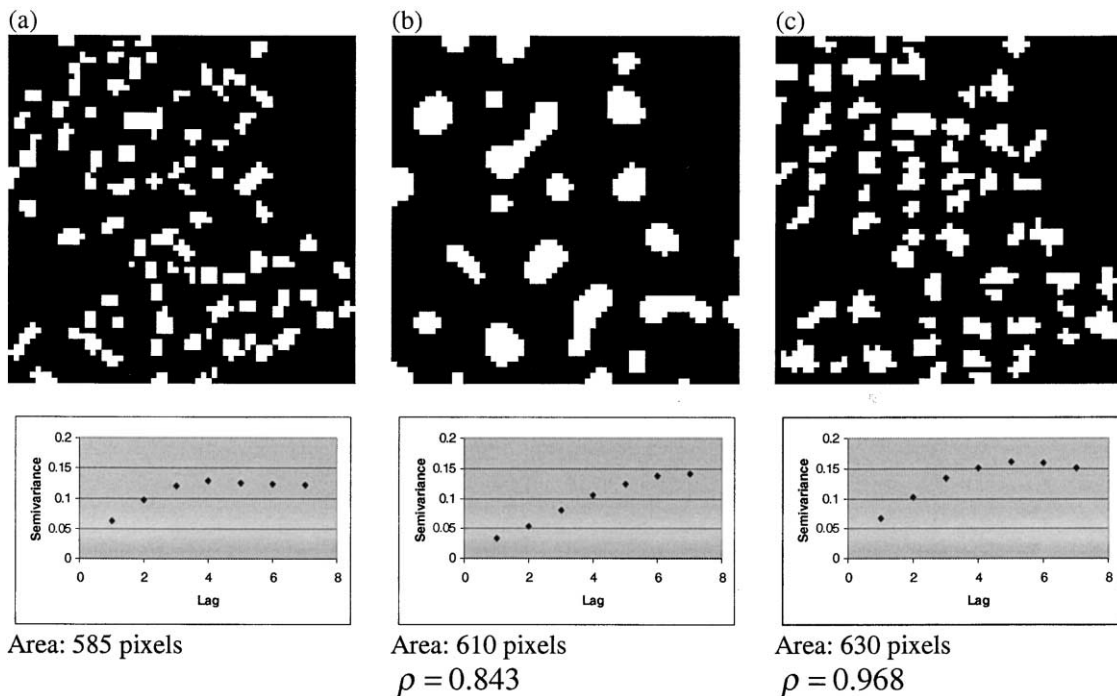

Fig. 8. (a) Verification image, class area, and variogram; (b) Hopfield network prediction image, class area, variogram, and variogram correlation coefficient (Tatem et al., 2001 technique); (c) Hopfield network prediction image, class area, variogram, and variogram correlation coefficient (pattern prediction technique).

target image is mirrored in Fig. 8, with a correlation coefficient of .968, significantly larger than that of .843 for the spatial clustering technique.

5. Discussion

5.1. Results analysis

The results shown in Section 4 indicate that the Hopfield neural network technique described in this paper displays excellent potential for super-resolution land cover pattern prediction. Whereas the spatial clustering approach described in Tatem et al. (2001) failed to recreate the spatial characteristics of land cover features that were smaller than a pixel, the pattern prediction approach introduced here is relatively successful. This suggests that a combination of the two approaches could potentially identify the spatial arrangement of land cover objects at any scale. It should be noted that while the focus of the results of this paper are 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 subpixel 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 Figs. 3 and 4 shows a wide range of synthetic land cover patterns, and comparison of these with the results of the pattern prediction technique in Figs. 3 and 4 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 successful results for each spatial arrangement indicates the ability of the approach to generalise. The synthetic imagery shows that the technique produces surprisingly good results for random patterns, as well as the more difficult challenge of a regular grid pattern, while the results from aerial photos demonstrate its potential success in dealing with real-world problems. Figs. 7 and 8 demonstrate clearly that the proportion constraint has ensured that the class proportions occur in the correct areas of the images, while the semivariance functions ensure that these proportions are arranged in a similar spatial pattern to those in the target images.

While the pattern prediction approach was extremely successful 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 (Figs. 3, 4, 6, and 7), except for Fig. 3(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 to also hinder the effect of the semivariance functions, lessening the success 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 (Figs. 3, 4, 6, and 7), the pattern prediction technique outperformed the other Hopfield neural network technique described by Tatem et al. (2001) in recreating the spatial pattern of land covers. The only example where the spatial clustering technique produced a larger variogram correlation coefficient is in Fig. 3(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 successful for mapping accurately land cover targets larger than a pixel by Tatem et al., when features are of subpixel scale, the lack of prior information on their layout produces poor performance by the technique, indicating that the work in this paper is a necessary extension to the technique.

5.2. Applications

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

5.2.1. 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 can 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 TM (30 m spatial resolution) or even up to the newly launched Ikonos sensor (1–4 m spatial resolution).

5.2.2. 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 are valuable resources. 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 get 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 paper could potentially solve this problem by providing spatial pattern predictions from coarse resolution satellite sensor imagery.

5.2.3. 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 & Culvenor, 2000). The approach described in this paper 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.

5.2.4. 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 paper 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 1960s 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.

5.3. Future work

The pattern prediction technique described in this paper provides an extension to the spatial clustering method described in Tatem et al. (2001). Techniques now exist to carry out super-resolution mapping of land cover targets both larger and smaller than a pixel, and future work will be directed towards examining the possibility of combining the two approaches. A large proportion of remotely sensed images of land cover will contain features both smaller and larger than the pixel size, so combination of both techniques is required for accurate super-resolution land cover mapping. To cope with complex scenes, the extension of both techniques to multiple land cover classes must also be examined. However, the successful results shown in this paper, along with the potential applications described in Section 5.2, indicate great potential for such a super-resolution land cover pattern prediction technique.

6. 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 subpixel scale features, the super-resolution pattern prediction technique based on recreating predefined 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 good performance, suggesting that it has the potential to predict accurately land cover patterns at the subpixel scale. The combination of the technique with an existing super-resolution land cover mapping approach, aimed at features larger than a pixel, potentially leads to a super-resolution land cover prediction technique capable of mapping at *any* scale.

References

Aplin, P., Atkinson, P. M., & 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.

Atkinson, P. M. (1997). Mapping sub-pixel boundaries from remotely sensed images. In: Z. Kemp (Ed.), *Innovations in GIS IV* (pp. 166–180). London: Taylor and Francis.

Atkinson, P. M., Cutler, M. E. J., & Lewis, H. (1997). Mapping sub-pixel proportional land cover with AVHRR imagery. *International Journal of Remote Sensing*, 18, 917–935.

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

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

Campadelli, P., Medici, D., & 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 ed.). New York: Taylor and Francis.

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

Cihlar, J., Latifovic, R., Chen, J., Beaubien, J., Li, Z., & 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.

Coops, N., & 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., & 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.

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

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

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

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.

Forte, P., & Jones, G. A. (1999). Posing structural matching in remote sensing as an optimisation problem. In: I. Kanellopoulos, G. Wilkinson, & T. Moons, (Eds.), *Machine vision and advanced image processing in remote sensing* (pp. 12–22). London: Springer.

Garcia-Haro, F. J., Gilabert, M. A., & 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., & Jennison, C. (1997). A subpixel image restoration algorithm. *Journal of Computational and Graphical Statistics*, *6*, 182–201.

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

Hopfield, J., & Tank, D. W. (1985). Neural computation of decisions in optimization problems. *Biological Cybernetics*, *52*, 141–152.

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

Lewis, H. G. (1998). *The use of shape, appearance and the dynamics of clouds for satellite image interpretation*. PhD Thesis, University of Southampton, Southampton, UK.

Li, R., Wang, W., & 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.

Lobo, A., Moloney, K., Chic, O., & 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.

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

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

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

Schneider, W. (1999). Land cover mapping from optical satellite images employing subpixel segmentation and radiometric calibration. In: I. Kanellopoulos, G. Wilkinson, & T. Moons (Eds.), *Machine vision and advanced image processing in remote sensing*. London: Springer.

Schowengerdt, R. A. (1997). *Remote sensing: models and methods for image processing*. San Diego: Academic Press.

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

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

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

Verhoeye, J., & De Wulf, R. (2000). Land cover mapping at the sub-pixel scale using linear optimization techniques. *Remote Sensing of Environment* (submitted for publication).

Woodcock, C. E., & Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, *21*, 311–332.

Zhang, X. M., Van Genderen, J. L., & 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*, *18*, 3279–3288.