UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

SCHOOL OF GEOGRAPHY

SUPER-RESOLUTION MAPPING USING THE HOPFIELD NEURAL NETWORK WITH SUPPLEMENTARY DATA

by

NGUYEN Quang Minh

Thesis for the degree of Doctor of Philosophy

June 2006

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

SCHOOL OF GEOGRAPHY

Doctor of Philosophy

SUPER-RESOLUTION MAPPING USING THE HOPFIELD NEURAL NETWORK WITH SUPPLEMENTARY DATA

by NGUYEN Quang Minh

Land cover mapping from remotely sensed imagery provides useful information for many authorities on different spatial scales and with different thematic contents. Although the spatial resolution of the remotely sensed image has been increased, the spatial resolution of the resulting land cover using conventional classifiers is still limited by the size of a pixel. This problem can be resolved partly by current super-resolution mapping techniques based on the spatial dependence maximisation. However, limitations still exist for these super-resolution mapping techniques since they are based solely on the spatial dependence assumption.

New approaches for using supplementary data such as panchromatic and fused imagery and Light Detection And Ranging (LiDAR) elevation data to increase the accuracy and spatial resolution of the thematic map were developed in this thesis. Information from the fused and panchromatic imagery was incorporated into the Hopfield neural network (HNN) model based forward and inverse models in form of reflectance functions. For the fused image, the forward and inverse models were formulated based on a linear mixture model and local end-member values. The reflectance function for the panchromatic image was derived locally based on the spectral and spatial convolutions. Visual and statistical analyses demonstrated that the use of fused and panchromatic imagery can increase accuracy of the sub-pixel image. The HNN super-resolution mapping using LiDAR elevation data is based on an optimisation process with a probability maximisation for the building class as a goal together with the goal functions and constraints of the traditional super-resolution mapping. The results showed a considerable increase in all accuracy statistics of the new technique, particularly for building objects.

Adopting the HNN model and forward model mechanism, three approaches for super-resolving of fine sub-pixel multispectral (MS) image from the coarse MS imagery were developed based on the HNN super-resolution mapping technique with the forward model and semivariogram matching. The first approach can be applied to predict the sub-pixel image based on the super-resolution of the mixed pixels. The second approach can be used to create the sub-pixel MS image with spectral features of the coarse resolution image and spatial variation at sub-pixel resolution. The aim of the third approach is to generate a smoothed sub-pixel image by maximising the spatial dependence between the sub-pixels based on a semivariance value of zero at lag $\mathbf{h} = 1$.

To my family

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The completion of this thesis would not have been possible without the support of many people. First of all, I would like to express my gratitude to my first superviser, Professor Peter Atkinson, for his useful guidence, constructive comments and criticisms, English correction, and encouragement during this work. I would also like to thank my second supervisor, Dr. Hugh Lewis, for his help with useful ideas, technical discussions and computer codes.

I am grateful to the Vietnamese Government, the Ministry of Education and Training and the Hanoi Mining and Geology University for providing me the funding to study in the UK.

Many thanks are directed to all the individuals who have helped me during this research. Also, a big thank-you goes to all my friends and colleagues for their enjoyable time together and great support.

Finally, I have to say "thanks" to my family, particularly my wife and my daughter Minh Anh.

LIST OF ABBREVIATIONS AND SYMBOLS

ARSIS - Amelioration de la Résolution Spatiale par Injection de Structures

AVHRR – Advanced Very High Resolution Radiometer

DEM – digital elevation model

DN – digital number

DSM – digital surface model

DTM – digital elevation model

GIS – geographical information system

GRE – ground resolution element

HNN – Hopfield neural network

HPF – high pass filtered

HRV – high resolution visible

IBSM – Inter-Band Structures Model

IFOV – instantaneous field-of-view

IHS – Intensity-Hue-Saturation

KIA – Kappa Index Agreement-κ

Landsat – Land Remote-sensing Satellite

LiDAR – Light Detection And Ranging

ME – mean error

MS – Multispectral

NOAA – National Oceanic and Atmospheric Administration's Advanced

PCA – principal component analysis

PAN – panchromatic

RMSE – Root Mean Square Error

SPOT – Satellite Pour l'Observation de la Terre

TM – Thematic Mapper

$\gamma_m(\mathbf{h})$ – semivariance at lag $\mathbf{h}$

λ – wavelength

Θ – fuzzy subset

Ξ – a universal of generic elements denoted by ξ

ξ – generic elements denoted by

μ – membership grade within the interval of a fuzzy subset

C_i – the covariance matrix

E – Energy

e – vector of error terms for each spectral band

$G1$, $G2$ – Goal functions

h – distance and direction vector of separation or lag

J_m – square error function

$N(\mathbf{h})$ – number of vector data a lag **h**

N – number of pixels

O – object function

P – vector of the proportions of mixture components for each pixel

P – land cover proportion

$p(w_i|x)$ – the conditional probability of a pixel x belonging to class w_i

PLC_{kj} – spatial dependence of pixel j of class k

R_λ – spectral reflectance

$\mathbf{S}_{b\cdot c}$ – the matrix of end-member spectra (b = number of spectral bands, c= number of classes)

$s_{i,j}$ – end-member spectra values

v – data value, output value of the neuron

x – vector of the measured sensor responses for each pixel

x – data value

z – zoom factor

Chapter 1: Introduction

1.1 Summary

This chapter is a review of remote sensing techniques for land cover mapping and related aspects such as data acquisition and data analysis. While a broad review of techniques are covered, the most important concepts that are relevant to the thesis are also discussed. These include land cover, hard and soft classification, super-resolution mapping techniques and image fusion techniques.

Firstly, the basic principles of remote sensing are presented, followed by a review of methods of land cover classification. These concepts aim to provide a background for the research. Secondly, definition and reviews of techniques for both super-resolution mapping and image fusion techniques are introduced. Finally, the chapter concentrates on the problems which this research aims to solve: urban land cover mapping and the rationale of using super-resolution mapping techniques as a solution.

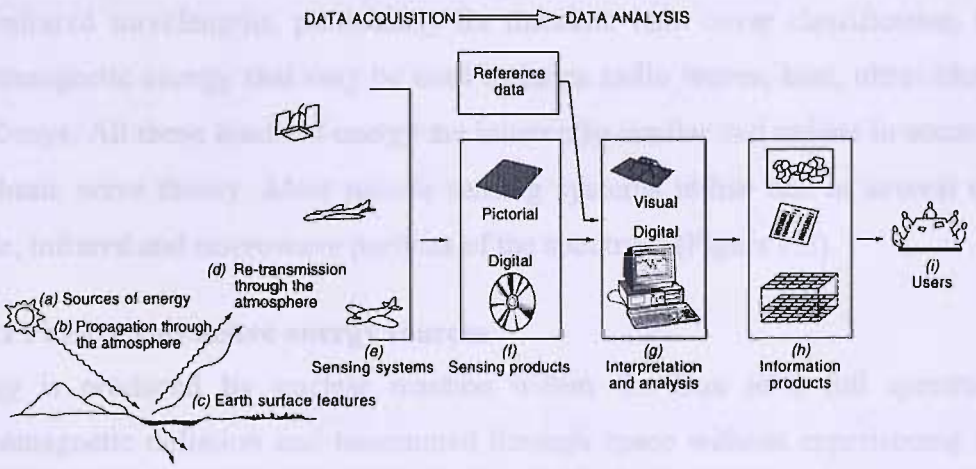


Figure 1.1. A remote sensing system (Lillesand and Kiefer (2000))

1.2 Definition of remote sensing

Remote sensing, as broadly defined by Lillesand and Kiefer (2000), "is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation". A narrower definition was provided by Campbell (1996) that "remote sensing is the practice of deriving information about the Earth's land and water surface using images acquired from an overhead perspective, using

electromagnetic radiation in one or more regions of the electromagnetic spectrum, reflected or emitted from the earth's surface". A remote sensing system is depicted in Figure 1.1.

According to the above definitions, remote sensing involves two basic processes: *data acquisition* and *data analysis*. The elements of the *data acquisition* process are energy sources, energy interactions with the atmosphere and the Earth surface, and the physical model for measurement. The introduction of these elements is the objective of Section 1.3 of this chapter. The second process, data analysis, involves interpreting the data acquired by the sensors to provide information applied in a wide range of fields. The present research is focused on the analysis of remotely sensed data for land cover extraction, which is often referred to as image classification.

1.3 Data acquisition in remote sensing

1.3.1 Energy sources

The most widely used electromagnetic energy in remote sensing exists at visible and near-infrared wavelengths, particularly for thematic land cover classification. Other electromagnetic energy that may be used includes radio waves, heat, ultraviolet rays and X-rays. All these kinds of energy are inherently similar and radiate in accordance with basic wave theory. Most remote sensing systems utilise one or several of the visible, infrared and microwave portions of the spectrum (Figure 1.2).

1.3.1.1 Passive and active energy sources

Energy is produced by nuclear reaction within the Sun in a full spectrum of electromagnetic radiation and transmitted through space without experiencing major changes. Some remote sensing systems detect energy that is emitted or reflected

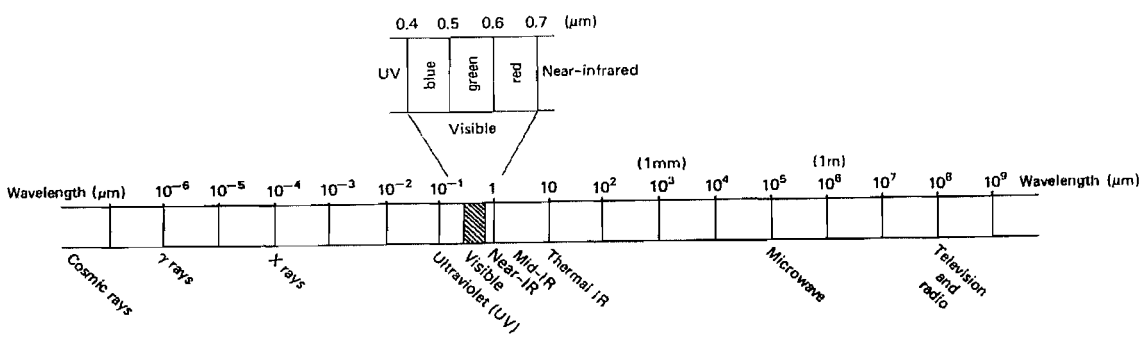


Figure 1.2 Electromagnetic spectrum

naturally from the ground. Other systems use instruments that generate their own energy, and then record the reflection from the Earth's surface. The two kinds of systems are termed *passive* and *active* systems, respectively.

1.3.2 Energy interactions of the Earth surface

As electromagnetic energy reaches the Earth's surface, it interacts with the surface in three basic ways: transmission, absorption and reflection. The proportions of light attributed for each process depend on the properties of the surface, the energy's wavelength and the angle of illumination (Campbell, 1996). The state of energy interactions of wavelength λ can be expressed as follows:

$$E_I(\lambda) = E_R(\lambda) + E_A(\lambda) + E_T(\lambda) \tag{1.1}$$

$E_T(\lambda)$ represents the part of the energy transmitted by the surface. Its magnitude depends on the transmittance property of the surface. $E_A(\lambda)$ represents the absorbed amount of incident energy $E_I(\lambda)$ and $E_R(\lambda)$ is the reflected energy.

1.3.3 Reflectance spectra as a function of ground properties

Among the three interaction properties of the surface, reflectance is the most important because all the remote sensing systems operate using reflected energy. The reflectance of each surface feature is determined as function of wavelength and is called *spectral reflectance* R_λ, which is a percentage, defined as

$$R_\lambda = \frac{E_R(\lambda)}{E_I(\lambda)} \tag{1.2}$$

where $E_I(\lambda)$ and $E_R(\lambda)$ are defined as in (1.1). Thematic information can be extracted from the remotely sensed data based on the *spectral reflectance* of the Earth surface features.

1.3.4 Physical model of the remote sensing system

Electromagnetic radiation is detected by sensors and recorded in pictorial formats for analysis by human operators or by computer automatically. At one instant in time, a sensor collects energy in a certain waveband from a ground area (ground resolution element-GRE) within its instantaneous field-of-view (IFOV- used area measurement unit) and transforms the detected energy to a digital value. The recorded digital value depends on the energy interactions with the features of the Earth surface area covered

by the IFOV and the sensor's characteristics (Warner and Shank, 1997). Remote sensing systems are often operated in several wavebands (bands). For example, the Land Remote-sensing Satellite (Landsat) Enhanced Thematic Mapper (ETM) and acquires data in seven bands between 0.574 μm to 12.50 μm, while IKONOS data consist of four Multispectral (MS) bands between 0.4 μm to 0.85 μm and a panchromatic (PAN) band.

1.3.4.1 Pixels

Pixels are the smallest picture elements or cells that make up photographs or images. In remote sensing images, each pixel covers an area on the ground. The size of the ground covered by a pixel varies between remote sensing systems. The quantity of energy reflected or emitted from a GRE determines the digital value or brightness of the pixels. However, the GRE should not be confused with the pixel. The difference can be illustrated by Landsat MSS with the GRE size of 57 m by 79 m (overlap of around 11 m in each scanner pass) and the pixel size of 57 m by 57 m. Within a GRE, the properties of the Earth surface are usually considered homogeneous, although this is rarely the case in practice.

i) Spatial Resolution

Spatial resolution determines the spatial detail depicted in an image. The ground dimension of the GRE defines the spatial resolution of remotely sensed data. The spatial resolution is influenced primarily by the sensor characteristics and altitude above the Earth. High spatial resolution data such as provided by IKONOS are produced at 1 m in PAN mode and 4 m in multispectral mode. In contrast, coarse spatial resolution data of 1.1 km are produced by the National Oceanic and Atmospheric Administration's Advanced (NOAA) Advanced Very High Resolution Radiometer (AVHRR).

ii) Radiometric Resolution

Radiometric resolution is determined by the number of bits used to record a brightness value within a digital image. This shows the sensitivity of the sensor as well as the capability for recording and transmitting data.

iii) Spectral Resolution

Spectral resolution is the ability to discriminate fine spectral differences. The spectral resolution of a remote sensing system is determined by the number of spectral bands over the electromagnetic spectrum and the spectral range of each band.

1.3.4.2 Spatial variation

Spatial variation is the variation in data values across space. In remote sensing, spatial variation is exhibited through the change of data values across images. Spatial dependence is the likelihood that observations close in space are more alike than those far apart. This observation was confirmed by Goodchild (1987). The spatial dependence assumption is true for most geographic phenomena at different scales. In remotely sensed data, the ubiquity of spatial dependence was detected by Atkinson and Curran (1995). Spatial dependence can be described by the covariance function, which represents the variation between pairs of data at specific distance and direction vectors of separation or lags **h** (Atkinson and Tate, 2000). The spatial variation can also be represented by the semivariogram, which is a function relating semivariance γ_m at lag **h** as

$$\gamma_m(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{1}^{N(\mathbf{h})} [v_i - v_{i+\mathbf{h}}]^2 \qquad (1.3)$$

where v_i is the data value at point (i,j) and $v_{i+\mathbf{h}}$ is the data value at point $(i,j)+\mathbf{h}$, and $N(\mathbf{h})$ is the number of data values.

1.3.4.3 Geostatistics

Geostatistics applies the theories of stochastic processes and statistical inference to geographic phenomena. In simple terms, geostatistics is a set of techniques for estimating the local values of properties that vary in space from sample data (Atkinson and Curran, 1998). Geostatistics has been employed for numerous remote sensing tasks such as defining the optimal pixel size (Marceau et al., 1994, Atkinson and Curran, 1995), land cover classification (Carr and Mirranda, 1998, de Bruin, 2000) and estimation of continuous variables (Davidson and Watson, 1995, Dungan, 1995).

1.4 Classification of land cover

Information on land cover is important for a variety of subjects at local, national and international levels (Campbell, 1996). This information can be acquired from remote sensing imagery automatically based on a process commonly known as land cover

classification. The process closely relates to the concepts presented in the following sections.

1.4.1 Basic concepts in land cover classification

1.4.1.1 Land use versus land cover

Land use can be defined as the use of land by humans, usually with emphasis on the functional role of land in economic activities. In contrast, land cover often relates to natural or human-made features of the surface (Lillesand and Kiefer, 2000).

1.4.1.2 Land cover classification

Image classification procedures are used to categorise all pixels in an image into land cover classes or themes. Each procedure employs one algorithm or combines several algorithms, which are called *classifiers*. For a specific task, the analyst selects the best classifier based on the characteristics of each image. Classified land cover data can be used directly or reorganised by geographical information system (GIS) tools to produce map-like land cover images. Most of classification procedures (except unsupervised classification methods) usually involves three stages as follows:

i) Training

Training is a stage in which the classifier is taught to recognise pixels belonging to certain classes. This stage is often implemented prior to the classification allocation stage.

ii) Allocation

The allocation stage is a second stage of classification in which a land cover class is assigned to all pixels in an image based on the training data and the digital values of the pixels.

iii) Accuracy assessment

Accuracy assessment is a third stage in which classified pixels are compared with reference data to evaluate the accuracy of the classification. With conventional classification, accuracy assessment is mainly based on statistics such as the confusion matrix, overall accuracy, user's and producer's accuracies, commission and omission errors (Congalton, 1991, Campbell, 1996), Kappa Index Agreement-κ (KIA) (Lillesand and Kiefer, 2000). KIA is accuracy statistics which uses all information in

the confusion matrix in order that the chance allocation of labels can be taken into consideration (Tso and Mather, 2001).

1.4.2 Hard classification

Hard classification is a conventional method, in which each pixel is assigned unambiguously to a single land cover class (Foody, 2002). Two approaches may be used to perform the hard classification: supervised and unsupervised classification. Supervised classification can be defined as the process of using reference data to classify pixels of the whole image. Unsupervised classification, in contrast, aggregates pixels into natural groups or clusters and then uses reference data to label those classified groups. In the following sections, the most widely used conventional classifiers are represented.

1.4.2.1 Parallelpiped classifier

The upper and lower significant bounds of brightness values in all wavebands are defined for each class from training data. Together, the range in all bands describes a multidimensional box or *parallelepiped* in the feature space. If, on classification, pixels are found to lie in a parallelepiped they are labelled as belonging to that class.

1.4.2.2 Minimum distance classifier

The minimum Euclidean distance in multi-dimensional feature space between the position of the pixels to be classified and the mean of each class is used as a criterion to identify a pixel belonging to a class.

1.4.2.3 Maximum likelihood classifier

The maximum likelihood classifier (MLC) is a supervised statistical approach to pattern recognition. In this classifier, the conditional probability of a pixel belonging to a class is calculated for all classes, and a pixel is assigned to a class if the conditional probability is the highest. This rule can be expressed as

$$x \in w_i \quad \text{if} \quad p(w_i|x) > p(w_j|x) \quad \text{for all } j \neq i, \tag{1.4}$$

where $p(w_i|x)$ is the conditional probability of a pixel x belonging to class w_i.

The conditional probability $p(w_i|x)$ can then be computed based on the Bayesian formula:

$$p(w_i|x) = \frac{p(x|w_i)p(w_i)}{p(x)}, \tag{1.5}$$

where $p(x|w_i)$ is the occurrence of events x given w_i, $p(x)$ and $p(w_i)$ are prior probabilities of events x and w_i. The $p(x)$ is the same for all pixels and, $p(w_i)$ is usually set to be uniformly distributed. The probability $p(x|w_i)$ is calculated as follows:

$$p(x|w_i) = \frac{1}{\sqrt{2\pi^{\rho}\sqrt{|C_i|}}} \exp\left(-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_i)^T C_i^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_i)\right), \tag{1.6}$$

where C_i is the covariance matrix of class w_i with the dimension ρ, μ_i is the mean vector of class w_i, C_i^{-1} is the inverse matrix of C_i and $|.|$ *is* the determinant.

1.4.3 Soft-classification

In hard classification, the land cover within a pixel is assumed to belong to a single class. However, the real landscape is not always composed of elemental squares or pixels (Fisher, 1997). Remotely sensed images, in fact, are commonly dominated by pixels that are composed of two or more land cover properties, called mixed pixels. As a consequence, the assumption of discrete, mutually exclusive classes and pure pixels may often be unsatisfied. Since hard classification is not appropriate for the case of mixed pixels, an alternative tool for land cover mapping was developed. Soft or fuzzy classification assigns to each pixel multiple and partial class memberships (or area proportions of classes in some cases) and, therefore, can produce a more accurate and realistic representation of many land covers (Foody, 1996; Bastin, 1997).

In general, there exist many approaches to the classification of mixed pixels. Two common approaches are reviewed here: *spectral mixture analysis* and *fuzzy classification.*

1.4.3.1 Spectral mixture analysis

Spectral mixture analysis involves a range of techniques wherein the digital value of mixed pixels is compared to a set of pure reference spectra. The basic assumption is that the spectral variation in an image is caused by mixtures of a limited number of surface materials. The result is a prediction of the proportions of the ground area of each pixel that are occupied by each of the reference classes. The basic mixture model may be represented as

$$\mathbf{R} = \mathbf{SP} + \mathbf{e} \tag{1.7}$$

where

R = Vector of the measured sensor responses for each pixel

$\mathbf{S}_{b\cdot c}$ = the matrix of end-member spectra (b = number of spectral bands, c = number of classes)

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

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

Subject to

$$\sum \mathbf{P} = 1 \text{ and } \mathbf{P} \geq 0 \tag{1.8}$$

The proportions $\mathbf{P}$ can be resolved if $n+1 \geq c$ (n is number of MS bands). Where $n+1 > c$, it is possible to estimate $\mathbf{P}$ by the least squares principle. The linear spectral mixture model has been used widely to generate fraction images of forest cover proportions and crop area proportions (Quarmby *et al.* 1992; Townshend *et al.* 2000). The method also has potential for application with hyperspectral imagery owing to the large number of spectral bands. However, the approach has some drawbacks such as the limit to the number of land cover classes for unmixing (Lillesand and Kiefer, 2000).

1.4.3.2 Fuzzy classification

Fuzzy-set based approaches are appropriate for resolving the mixture information contained by mixed pixels. All of these approaches employ the fuzzy-set theory's membership grade concepts for the classes' proportions. An example is the supervised fuzzy maximum likelihood classifier proposed by Wang (1990). This algorithm provides membership grades for each pixel based on the fuzzy means and fuzzy covariance matrices estimated from training sets. Another algorithm for dealing with mixed pixels is the fuzzy c-means (FCM) algorithm.

An alternative algorithm for statistical fuzzy classification is provided by artificial neural networks. The potential of neural networks for mixed pixel classification was recognised by several authors (e.g. Foody, 1996; Carpenter, 1999). Foody (1996) used a multilayer perceptron neural network (MLP) to predict class proportions. The network was trained using mixed pixels with proportions outputs. Carpenter (1999) applied the ARTMAP mixture system to provide output class proportions with great accuracy. The advantage of the ARTMAP mixture system compared to the MLP soft classification is that it is not necessary to determine the number of neurons in the hidden layers.

1.5 Image fusion

The exact meaning of remote sensing data fusion varies from one scientist to another. Pohl and Van Genderen (1998) proposed that "image fusion is the combination of two or more different images to form a new image by using a certain algorithm", which is restricted to images. The spatial fusion takes place at three levels: pixel, feature and decision (Pohl and Van Genderen, 1998). Image fusion at the pixel level is the lowest processing level for merging different sources of remote sensing data and encompasses the most commonly used techniques (Steinnocher, 1999). At this level, image fusion involves combination of two or more different types of images to form a new image with finer spatial and spectral resolution than that of the original images. Fusion at the feature level requires the extraction of objects recognised in the various data sources. Similar objects from multiple sources are assigned to each other and then fused for further assessment using statistical approaches or artificial neural networks. Decision level fusion represents a method that combines information obtained from various sources for better understanding of the observed objects.

Image fusion can be applied to various types of data sets: for example, data from a single sensor, data from multiple-sensors and multi-temporal data. While temporal image fusion is mainly applicable for change detection, non-temporal fusion of single or multi-sensor data can be used in a variety of applications such as image sharpening, geometric corrections and image classification (Pohl and Van Genderen, 1998). Thus, this thesis will focus on the fusion methods that are useful for increasing the accuracy of land cover classification of remotely sensed image for urban land cover.

1.6 Super-resolution mapping

The problem of mixed pixels can be solved using soft-classification. The outputs of soft-classification are a set of proportion images, each displaying the proportion of a certain class. This produces a more appropriate and informative representation of land cover than that of hard classification. However, while the class composition of every pixel is predicted, the spatial distribution of the classes within the pixels is still unknown (Tatem *et al.*, 2002). Super-resolution mapping refers to the process of sub-pixel mapping of the soft-classified land cover proportions based on the assumption that a mixed pixel is composed of the crisp sub-pixels only.

Several techniques have been used for super-resolution mapping. Atkinson (1997) and Verhoeye and De Wulf (2001) mapped soft-classified proportion images based on the assumption of spatial dependence within and between pixels. These approaches produce accurate land cover maps at a finer spatial resolution than the initial input images. However, the resulting sub-pixel land cover maps produced by these approaches contain linear artefacts (Tatem *at al.*, 2002). Zhan *et al.* (2002) also used a similar technique by interpolating the probability measures derived from MLC and spatial information at a pixel scale. The sub-pixel probability vectors produced by inverse distance weighted interpolation from the centres (or edges) of neighbouring and central pixels. The accuracy of the resulting land cover map produced by this approach was slightly greater than the accuracy of the hard classified data.

Another algorithm, mainly developed by Tatem *et al.* (Tatem *et al.* 2000; 2001; 2002), employs the Hopfield neural network (HNN) as an optimisation tool based on the energy minimisation and constrained by the land cover proportions obtained by soft-classification. This technique is presented as a simple and robust technique for prediction of the location of class proportions within each pixel. More details on the technique will be discussed in the next chapter of the thesis. An alternative approach for the super-resolution is proposed by Atkinson (2003). This approach uses two-point histogram matching for optimisation. The first results of this approach show potential, but further research is required.

1.7 Problems and objectives

Super-resolution mapping or sub-pixel mapping resolves the uncertainty remaining in the soft-classified land cover proportions. In the super-resolution maps, the spatial locations of soft-classified land cover classes are defined. The resulting super-resolution maps are similar to the hard classified land cover classes at the sub-pixel spatial resolution. Most of the current super-resolution mapping approaches are based on the maximisation of the spatial correlation (or spatial dependence) between the sub-pixels. These approaches, therefore, can be referred to as clustering algorithms. In general, the sub-pixel maps obtained by clustering algorithms are more accurate than those produced by the conventional hard classifiers at the original coarse spatial resolution. However, there still exists a limit to the detail and accuracy of the resulting thematic maps produced by clustering algorithms since they are based only on the

soft-classified proportion data at the pixel level and the spatial dependence assumption.

Figure 1.3 shows an example of the inaccuracy of the above-mentioned super-resolution mapping methods using clustering algorithms. From the land cover proportion of a single class in Figure 1.3(a), three possible resulting thematic maps at sub-pixel resolution can be obtained using clustering algorithms as in Figure 1.3(b), 1.3(c) and 1.3(d). All these maps satisfy the assumption of spatial dependence but the location of the sub-pixels in the central pixel is different. In the real landscape, the

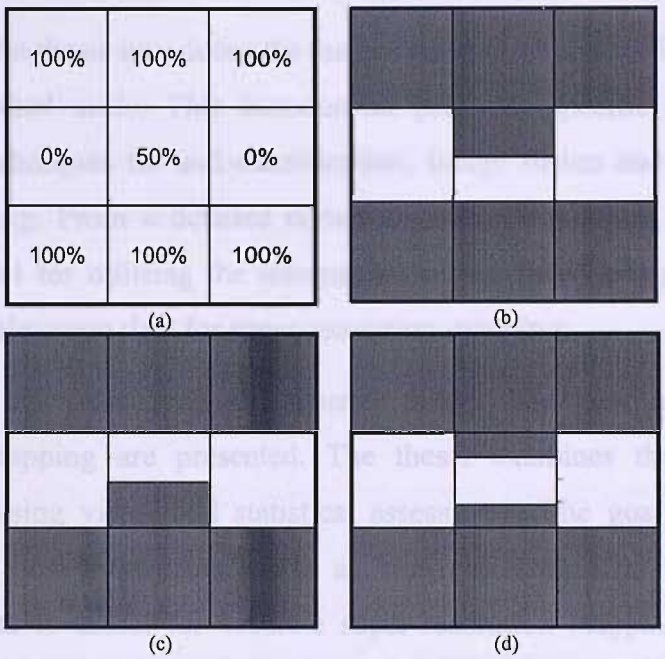


Figure 1.3 (a) Land cover proportion image and (b), (c), (d) three possible resulting super-resolution mapping images based on maximising spatial dependence.

cases which are similar to the example in Figure 1.3 are quite common. Thus, it is possible that the resulting sub-pixel maps obtained by the clustering algorithms contain some inaccuracy. This suggests that if some information can be provided at the sub-pixel or an intermediate spatial resolution (spatial resolution which finer than that of original image but coarser than that of the sub-pixel image) then more accurate sub-pixel mapping results can be achieved.

The development of this thesis emerged from the idea that the information at the sub-pixel or intermediate spatial resolution can increase the detail and accuracy of the super-resolution land cover classes. In fact, sub-pixel information obtained from vector boundary data was used previously by Aplin and Atkinson (2001). However,

this kind of accurate vector data (i.e. land parcel boundary maps) are not always available for remote sensing scenes. Nowadays, together with a MS image, commercial remote sensing systems such as Landsat, Satellite Pour l'Observation de la Terre (SPOT), IKONOS or QuickBird also provide panchromatic images with finer spatial resolution. Another source of very fine spatial resolution data is Light Detection And Ranging (LiDAR) elevation data with a high density of data points and high vertical accuracy. The main objective of the research in this thesis is utilising the above mentioned sources of fine spatial resolution data for super-resolution land cover mapping.

First of all, the thesis introduces the method used to obtain the land cover information at the sub-pixel scale. This introduction provides specific understanding of the published techniques for soft-classification, image fusion and super-resolution land cover mapping. From a detailed review on these techniques, the thesis presents a general model for utilising the information of the fused and panchromatic imagery and LiDAR elevation data for super-resolution mapping.

Developing from the proposed general model, the new approaches for super-resolution mapping are presented. The thesis examines the feasibility of these approaches using visual and statistical assessment. The goal of this process is to demonstrate that information such as from panchromatic imagery and LiDAR elevation data is useful for accurate super-resolution mapping using the HNN. In order to reach that goal, the super-resolution maps provided by the new approaches must be compared with the maps produced by hard classifiers and the super-resolution mapping without using supplementary data.

Finally, the thesis makes use of super-resolution land cover mapping to super-resolve the multispectral image. In addition, the thesis also investigates the possibility of using spatio-statistical semivariogram information at the sub-pixel level for multispectral image super-resolution and interpolation. During the investigation process, specific analysis is undertaken to provide a through understanding of the impacts of semivariograms on the super-resolution image.

Chapter 2: Methods Used to Increase the Spatial Resolution of Urban Thematic Maps

2.1 Summary

In Chapter 1, the methods for increasing the spatial resolution of the land cover classified maps were introduced briefly. This chapter considers these methods in a more detailed and specific review. This review represents the published concepts and algorithms of three useful techniques for increasing the spatial resolution of land cover mapping from remotely sensed images such as: i) soft classification, ii) super-resolution mapping and iii) data fusion. Furthermore, the review also considers the characteristics of these algorithms, their advantages and disadvantages and the possibilities of combining them for obtaining finer spatial resolution and more accurate land cover maps in comparison with the conventional hard land cover classifications.

The three techniques increase the spatial resolution of resulting land cover classified maps in different manners. Soft classification is applied for sub-pixel land cover proportion prediction given mixed pixels. In scenes where the land covers are spatially distributed with high spatial frequency (for example: urban scenes), soft classifiers should be used to achieve more accurate and informative land cover mapping. Image fusion techniques are applied to increase the spatial and spectral resolution of images from multispectral and panchromatic images. Super-resolution mapping techniques are useful to resolve the ambiguity within a pixel (i.e. mapping trees in a field) implicit in the land cover proportions predicted by soft classification. The last part of the chapter introduces a model to incorporate the information from a supplementary source of data such as fused and panchromatic images to increase the spatial resolution of the land cover map.

2.2 Soft classification

The principles of soft classification were presented briefly in Chapter 1. There are many approaches commonly used for soft classification. To choose the most accurate soft-classification, it is necessary to review the most commonly used soft-classification algorithms. Each of these algorithms is discussed in the following sections of this chapter.

2.2.1 Spectral mixture analysis

Spectral mixture analysis determines the sub-pixel proportions of ground cover based on the assumption that the energy received by the sensor is the linear sum of the energy emitted from each cover component within each pixel. Spectral mixture analysis is used widely for land cover proportion prediction. Examples are Small (2001) who used linear mixture analysis for prediction of vegetation abundance for urban land; Theseira *et al.* (1999, 2002) who used spectral mixture modelling applied to a semi-arid environment using a multiple end-member approach; Roberts *et al.* (1998) who employed spectral mixture analysis for mapping chaparral; Atkinson *et al.* (1997) who compared linear mixture modelling for sub-pixel land cover mapping with the alternative approaches such as statistical fuzzy *c*-means and artificial neural networks. Despite its wide range of applications, spectral mixture analysis has several problems that are introduced in the following sections.

2.2.1.1 Limited number of end members

A problem with spectral mixture analysis is the restriction of the number of end members by the number of available spectral features. It can be inferred that the linear mixture model is more appropriate when the number of input features is reasonably large (McGwire *et al.*, 2000). The data sources with large number of spectral bands are likely to have large number of spectral features. However, the inter-correlation between the bands of large dimensional data may pose problems for solving the set of linear equations. This problem may be solved by using orthogonal transformations such as the minimum noise fraction (Van de Meer and De Jong, 2000).

To overcome the problem of end member number restriction, Bosdogianni *et al.* (1997) proposed a method of predicting land cover proportions with higher order moments. Using this method, the proportions of sets of pixels are computed instead of land cover proportions within a single pixel. The constraint equations are computed from mean values and covariance matrices of the sets of estimated pixels. Therefore, the number of equations is augmented to estimate a larger number of end-member proportions. A search method is applied to find the solution for the composition for end members. It is obvious that the method can be used for estimation of large numbers of land cover proportions of sets of mixed pixels. However, the methods also reduce the spatial resolution of the classified data because the land cover proportions are estimated for sets of mixed pixels rather than for each single pixel.

Another approach to overcome the problem of restricted end member number is called "multiple end member spectral mixture models" (Theseira *et al.* 1999, 2002). The approach is based on the assumption that the maximum number of classes within a pixel is fewer than the number of spectral bands. The proposed technique selects the most appropriate composition of end members amongst all possible combinations of land covers within a pixel by iteration. The criterion for selection is minimum Root Mean Square Error (RMSE). However, the results are unlikely to be acceptable for some environment such as semi-arid environment (Thereisa *et al.*, 1999).

2.2.1.2 Training data

Before a mixture model is constructed, it is necessary to determine certain constants such as the values of the end member spectra. These constants may be determined from training images or from laboratory instruments such as spectrometer. Determination of these constants has to be performed carefully to avoid inaccuracy in the results of proportion estimation.

2.2.1.3 Non-linear mixture models

Another problem of the linear spectral mixture analysis is that it does not account for certain factors such as the variability of the reflectance of an end member. In this situation, a more sophisticated non-linear mixture model may be required (Borel and Gerstl, 1994). In addition, this problem may possibly be solved by non-linear classification methods such as fuzzy set theory and neural networks.

2.2.2 Fuzzy set theory and soft-classification

Fuzzy classification attempts to handle the mixed pixel problem by employing the fuzzy set concept, in which a pixel may have partial membership in more than one land cover class. In fuzzy set theory, let Ξ represent a universal of generic elements denoted by ξ. A fuzzy subset Θ of Ξ is determined by a membership function μ_Θ, which assigns a membership grade within the interval [0, 1] to each element ξ. The membership grade can be expressed as

$$\mu_\Theta : \xi \rightarrow [0,1] \tag{2.1}$$

In crisp sets, the membership grade must be 0 or 1. It is clear that, the fuzzy set is more flexible for handling the problems of vague boundaries that are common in nature. The fuzzy subset Θ can be expressed as

$$\Theta = \sum_i \mu_\Theta(\xi_i) / \xi_i \tag{2.2}$$

$$\Theta = \int \mu_{\Theta}(\xi_i)/\xi_i \tag{2.3}$$

where Equation (2.2) represents the discrete case and Equation (2.3) represents the continuous case. The symbol '/' does not represent the division operator but the link between the value of ξ and its corresponding membership grade $\mu_{\Theta}(\xi)$ in the fuzzy subset Θ. The symbols "$\sum$" and "$\int$"do not represent the sum or integral operator but they depict that the element ξ belongs to the subset Θ.

The membership grade $\mu_{\Theta}(\xi)$ can be denoted as the partial membership of a pixel for a set of land covers. The memberships then can be used as surrogates for land cover proportions $P_1, P_2, .., P_c$ within a pixel. Based on fuzzy set theory, several approaches have been developed for land cover proportion prediction such as the fuzzy *c*-means classifier, Bayesian soft classifier, fuzzy rules base classifiers and fuzzy classification by neural networks.

2.2.2.1 Fuzzy c-means clustering algorithms

Statistical fuzzy classifiers are developed from the ISODATA clustering algorithm for hard classification. The clustering criterion used in the fuzzy *c*-means algorithm is based on minimising the generalised within-groups sum of square error function J_m

$$J_m = \sum_{k=1}^{N}\sum_{i=1}^{c}(\mu_{ik})(x_k - v_i)^2 \tag{2.4}$$

where v_i is the centre of cluster i, x_k is the vector of measurements of pixel k in the feature space, μ_{ik} represents the membership values (it can represent land cover proportion P_{ik} in land cover soft classification) of the pixel k in class i, c denotes the number of classes and N denotes the number of pixels. The values μ_{ik} are the elements of the fuzzy partial matrix (or land cover proportion matrix) **P**. In equation (2.5), the square Euclidean distance $(x_k - v_i)^2$ can be replaced by the Mahalanobis distance, which requires the calculation of the fuzzy covariance matrix. The local minimum of J_m may be achieved from

$$P_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{|x_k - v_i|}{|x_k - v_j|}\right)^{2/(m-1)}} \text{ with } v_i = \frac{\sum_k (P_{ik})^m \cdot x_k}{\sum_k (P_{ik})^m} \tag{2.5}$$

where m is the membership weighting exponent.

The fuzzy *c*-means clustering is, thus, performed by iterating Equation (2.5). In the first iteration, the cluster means are calculated by the randomly initialised clustering

membership μ_{ik}. The membership values then can be derived from the calculated cluster means. In the next iteration, the cluster means are computed from the last iteration and the process continues until the difference between two consecutive iterations is small enough.

2.2.2.2 Bayesian Fuzzy classification

Another statistical approach for fuzzy classification is fuzzy maximum likelihood classification based on the Bayesian formula (Wang, 1990; Maselli *et al.*, 1995). In this approach, the fuzzy means and covariance matrices are expressed as

$$v_i = \frac{\sum_j \mu_i(x_i) x_i}{\sum_j \mu_i(x_i)}$$

and

$$F_i = \frac{\sum_j \mu_i(x_i) \cdot (x_j - v_i) \cdot (x_j - v_i)^T}{\sum_j \mu_i(x_i)} \tag{2.6}$$

The membership function values μ_i in Equation (2.6) can be calculated from the conditional probabilities as

$$\mu_k(x_i) = \frac{p_k(x_j)}{\sum_i p_i(x_j)} \tag{2.7}$$

When the membership function of class i is applied to a pixel vector, the membership grade of the pixel in class i is calculated for each pixel. In comparison with the fuzzy c-means approaches, the fuzzy maximum likelihood approach seems to be more straightforward. However, empirical evaluations by artificial data, which are implemented by Bastin (1997) and Eastman and Laney (2002) showed that the maximum likelihood sub-pixel decomposition works with important limitations. It is critical that the conditional probability distributions from the training site data overlap over the complete range of mixtures. Where the training data are pure pixels, the classification for mixed pixels is inaccurate. The classifier could perform more accurately if mixed pixels are used in the training process but the procedure may then fail to uncover the pure pixels. Häme *et al.* (2001) used conditional probabilities to compute proportions for forests in combination with an unsupervised classification. The testing procedure showed that the overall proportion of forest was underestimated.

2.2.2.3 Fuzzy rule base classification

The application of fuzzy rule base methods for soft classification is still at an early stage (Tso and Mather, 2000). Simpson and Keller (1995) used fuzzy rule base methods to segment sea ice from cloud and cloud-free oceans in polar AVHHR imagery with six fuzzy subspaces and seven rules for classification. The proportions of each information class were derived from the output of the rules based classification without implementing the defuzzification step. Bárdossy and Samaniego (1995) applied fuzzy rule based classification to produce four categories of land cover and an image depicting the degree of ambiguity of classification for each pixel.

Two issues related to the use of the fuzzy rule base for solving the mixed pixel problem should be noted. The first issue is that the membership grades for a given pixel over all classes must sum to one. The second issue concerns the number of dimensions of remotely sensed data. If the number of data dimensions is large, the number of rules may pose a problem for computation. This issue can be solved using data dimension reduction techniques.

2.2.2.4 Soft classification using neural networks

Neural network techniques have been applied in a wide variety of remote sensing applications (Atkinson and Tatnall, 1997, Tso and Mather, 2000). Several kinds of neural networks, including the multi-layer perceptron and ARTMAP system, can be used for image classification, particularly for soft classification. Evaluation of neural networks for hard classification is implemented by Paola and Schowengerdt (1995), who compared the maximum likelihood classifier and back-propagation neural network for urban land use classification. For soft classification, comparison with statistical soft classification techniques such as fuzzy *c*-means and mixture model have been undertaken by Atkinson *et al.* (1997) who achieved more accurate results with the neural network approach in terms of the accuracy of classification.

Two types of the most widely used neural network architectures for soft classification are the multi-layer perceptron and Adaptive Resonance Theory (ART) systems. In addition, application of these two architectures is introduced in the following sections.

- ***Soft classification using the multi-layer perceptron network***

The multi-layer perceptron neural network is a feed-forward network using the back-propagation learning algorithm. The architecture of the network used for image

classification can be seen in Figure 2.1. A feed-forward neural network usually has three layers. The first layer is the input layer, which contains the set of nodes that receive external inputs. The number of nodes in the input layer is equal to the dimension of the feature space. Each node in the input layer connects to all of the nodes in the hidden layer. The number of hidden layers is usually one or two depending on the application. The last layer is the output layer, which produces the output for classification. Therefore, each node in the output layer represents an information class.

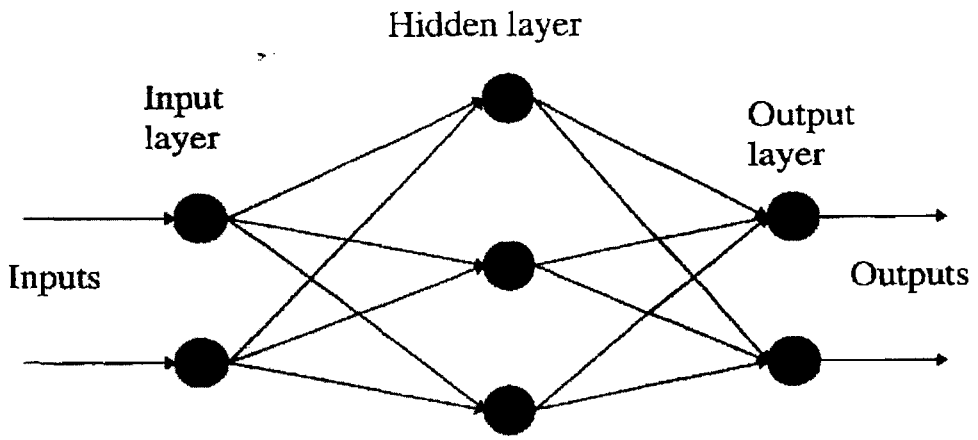


Figure 2.1. Neural network architecture for image classification

The neural network passes the signals from the input layer to the output layer via the hidden layers in a feed-forward manner. The signal received by the node of the next layer from all of its preceding nodes is as follows

$$\text{net}_j = \sum \omega_{ji} o_i \tag{2.8}$$

where ω_{ji} represents the weights between node i and node j, and o_i is the output from the node i. The output from a given node j is then computed from

$$o_i = f(\text{net}_j) \tag{2.9}$$

The function f is the activation function which is usually a sigmoid function. The procedure for neural network classification includes two stages. In the first stage called the training stage, the network is taught how to recognise the pattern of a class. Training patterns are presented to the network and the signals are fed-forwards as described above. Then, the network output is compared with the desired value and the error is computed. This error is then back-propagated through the network and the weights of connections are changed as follows

$$\Delta\omega_{ji}(n+1) = \eta(\delta_j o_i) + \alpha\Delta\omega_{ji}(n) \tag{2.10}$$

where η is the learning rate parameter, δ_j is the rate of change of the error, and α is the momentum parameter. In the classification stage, the network receives an input vector and then produces classified data in the output layer.

Multi-layer perceptron networks can be used both for hard and soft-classification. In hard classification, the training and output values are crisp. Therefore, each time a pixel spectral vector is presented, the neural network produces an output vector that is set to be 1 for the classified class and 0 for the other classes. In soft classification, the network is often trained using mixed pixels. Each input vector generates a corresponding output membership value for each of the candidate information classes. Soft classification using a feed-forward neural network approach is examined by Foody (1996), Atkinson *et al.* (1997), Foody *et al.* (1997), and Zhang and Foody (1998, 2001). All of these papers show that the class proportion of the image pixels could be predicted more accurately than that of the fuzzy *c*-means classifier and Bayesian fuzzy classification.

The weight modification process of the neural network can be achieved using alternative methods rather than the back-propagation learning algorithm. The Bayesian modification of an artificial neural network can be applied to avoid the issues related to model over-fitting (Bishop, 1995). Lampinen and Vehtari (2001) showed that the Bayesian MLP performs better than the back-propagation MLP in some classification tasks. Braswell *et al.* (2003) used an artificial neural network with the Bayesian weight modification algorithm for mapping sub-pixel land cover distributions using Multi-angle Imaging Spectroradiometer (MISR) and Moderate Resolution Imaging Spectroradiometer (MODIS) data in the Brazilian Amazon region.

Empirical studies carried out by Bernard *et al.* (1997) to test training procedures with neural networks suggest that a wide variety of mixed pixels are necessary in training sets to obtain accurate recognition of dominant and subsidiary classes. Another concerning aspect of soft classification using a MLP is the sensitivity to the absence classes in the training process (Foody 2000, 2002b). Problems of feed-forward neural network soft classification are that the output on nodes does not have to sum to 1 and the partitions of land cover classification seem to be simple or linear (Warner and Shank, 1997). Methods for solving these problems include use of a compound linear-sigmoid activation function and training using synthetic mixed pixels.

- ***Soft classification using the ARTMAP neural network***

Another kind of neural network architecture used for land cover classification is fuzzy ARTMAP (ARTMAP architecture uses two ART networks to create a map between the input and output patterns using an intermediate map field) (Carpenter *et al.*, 1997, 1999; Gopal, 1998; Mannan *et al.*, 1998). This architecture is based on the simple Adaptive Resonance Theory (ART) that was introduced by Carpenter (1987). Fuzzy ARTMAP systems can be used for both crisp and fuzzy output classification. In comparison with hard classification using maximum likelihood, the fuzzy ARTMAP system is generally more accurate, particularly for large training samples (Gopal *et al.*, 1998). For vegetation classification, fuzzy ARTMAP performed more accurately than the multi-layer perceptron neural network in terms of training time and accuracy (Carpenter *et al.*, 1997; Mannan *et al.*, 1998). However, an empirical study by Tso and Mather (2001), in contrast, shows that the training stage for fuzzy ARTMAP networks is significantly faster than other networks but the output accuracy is much lower.

For soft-classification, Carpenter *et al.* (1999) used fuzzy ARTMAP networks to estimate vegetation proportions from Landsat TM data. The soft-classification results of the ARTMAP system are more accurate compared with linear mixture modelling and Bayesian soft classification. However, it is necessary to further evaluate the performance of fuzzy ARTMAP systems versus multi-layer perceptron neural networks for land cover classification in the other landscapes.

2.2.2.5 Soft classification using support vector machines

Support Vector Machines (SVMs) are classification and regression methods based on a learning paradigm proposed by Vapnik (1998). The classification techniques are based on the optimal hyperplane that maximises the separated margin between the classes. The optimal hyperplane is used to select the data points which lie on the class boundary closest to the neighbouring classes. These data points are support vectors.

Support Vector Machines were used by Brown *et al.* (1999, 2000) for spectral unmixing. The methods were evaluated to be equivalent with the Constrained Least Squares Linear Spectral Mixture Model. Moreover, the technique selected the "pure pixel" automatically from a much larger data set. Extensions of the SVM algorithm allow the technique to be used with non-linear mixtures.

2.2.2.6 Soft classification using k-nearest neighbours

The *k*-nearest neighbour (*k*-nn) soft-classification (Lewis and Brown, 1998) estimates the land cover proportions of the unknown pixels based on the proportions of the *k* nearest neighbour training pixels or examplars. The *k*-nearest neighbour training pixels are selected using kernel-based algorithms. The proportion P_c of class c of an unknown pixel can be estimated as

$$P_c = \frac{\sum_{i=1}^{k} w_c^i \mu_c^i}{\sum_{i=1}^{k} w_c^i} \text{ for } w_c^i = 1 - \frac{D_i}{\max_{j=1}^{k}[D_i] + 0.0001} \tag{2.11}$$

where μ_c^i is land cover proportion of class c of the examplar i, w_c^i is weighting coefficient and D_i is Euclidean distance from the unknown pixel to the examplar i in the feature space.

2.2.3 Accuracy assessment for soft classification

Accuracy assessment of a conventional land cover classification is usually expressed in terms of a confusion matrix. For soft classification, these accuracy measures are not necessarily appropriate (Foody, 2002c). Alternatively, there are several other proposed measures for accuracy assessment of soft classified land cover proportions.

The most commonly used criteria for soft classification are RMSE and Mean Error (ME) (Huang and Townsend, 2003), which are calculated as

$$RMSE = \frac{1}{c}\sum_{j=1}^{c}\sqrt{\frac{\sum_{i=1}^{n}\left(P_{ij} - P_{ij}^{Known}\right)^2}{n}} \text{ and } ME = \frac{1}{c \cdot n}\sum_{j=1}^{c}\sum_{i=1}^{n}\left(P_{ij} - P_{ij}^{Known}\right)^2 \tag{2.12}$$

where c is total number of classes , N is the total number of pixels, P are the classified proportions and P_{ij}^{Known} are the known proportions.

Atkinson (1999) proposed some criteria such as correlation coefficient, standardised *RMSE*, standardised *ME* and standardised *Kernel based statistic* as alternative measures of association between estimated and target set of proportions. Distribution of errors in soft classification can be depicted by the cross-entropy (Foody, 1995), which was calculated from estimated class membership probability distribution and the ground data. Binaghi *et al.* (1999) generalised the model of the confusion matrix to fuzzy set-based error matrix, which can apply for both crisp and fuzzy classification accuracy assessment. Similarly, Lewis and Brown (2001) proposed another generalised model for confusion and error matrices that can be applied for

both hard and soft-classification. The model was demonstrated for accuracy assessment of urban land cover based on the assumption that the sub-pixel classification can be interpreted as an area estimation.

2.3 Image Fusion

2.3.1 Introduction

A variety of image fusion techniques have been introduced to date. Ranchin and Wald (2000) divided image fusion methods into three groups: (i) projection and substitution methods, (ii) relative spectral contribution methods and (iii) the methods relevant to the ARSIS (Amelioration de la Résolution Spatiale par Injection de Structures) concepts (ARSIS is a French acronym meaning spatial resolution enhancement by injection of structures). The projection and substitution group includes the Intensity-Hue-Saturation (IHS) methods (Carper *et al.*, 1990), principal component analysis (PCA) method (Chavez *et al.*, 1991), filter-based fusion (Chavez *et al.*, 1991) and standardised PCA methods. The relative spectral contributions group consists of intensity modulation (Cliché *et al.*, 1985), the PAN+XS method, and regression fusion (Price, 1999). Wavelets methods (Blanc et al., 1998, Ranchin and Wald, 2000, Ranchin *et al.*, 2003) are based on the ARSIS concept that spectral characteristics of the synthetic images must be close to the real image at the fine spatial resolution. In addition to those commonly used methods mentioned above, there are also other methods such as geostatistical digital image merging (Delgado-Garcia *et al.*, 2002), and hybrid fusion (Núñez *et al.*, 1999).

2.3.2 Projection and substitution methods

2.3.2.1 Intensity-Hue-Saturation

IHS is the most commonly used method for image fusion (Chavez *et al.*, 1991). IHS has been used for the fusion of SPOT PAN and SPOT high resolution visible (HRV) MS image data (Carper *et al.*, 1990) and Landsat TM PAN and MS imagery (Chavez *et al.*, 1991). Chen *et al.* (2003) used the IHS transformation for hyperspectral and radar data to enhance urban surface features.

IHS image fusion can be applied for merging fine spatial resolution imagery and coarser spatial resolution imagery with fine spectral resolution. Lillesand and Kiefer (2000) used IHS to fuse three bands of the lower spatial resolution data set that is

registered to the same grid as the fine spatial resolution image and then transformed these data into IHS space. The finer spatial resolution image, which is usually a PAN image, is contrast stretched so that it has variance and average approximate to the intensity component image. The stretched, finer spatial resolution image replaces the intensity component image before the images are transformed back into the original space. The main justification used for replacing the intensity component with the stretched finer spatial resolution image is that the two images are approximately equal to each other spectrally.

In the research by Carper *et al.* (1990), the intensity image was replaced by the weighted-average image, which is computed as twice the PAN plus registered resampled SPOT multispectral band 3. The weighted-average image was found to be highly correlated with the intensity derived from the multispectral data. Therefore, the resulting synthesised images were more correlated with the original MS images. The IHS data fusion methods are inappropriate for high quality transformation of MS content when increasing the spatial resolution (Ranchin and Wald, 2000). The MS content of the original images is not preserved due to the difference between the intensity component image and the PAN image.

2.3.2.2 PCA methods

There are two approaches for image fusion using the PCA transformation. Chavez *et al.* (1991) implemented a PCA-based fusion method by replacing the first component of the transformed PCA of the low spatial resolution image by the stretched higher spatial resolution image. Another approach was carried out by Yesou (1993) using PCA transformation of all multi-image data channels. In this approach, the image channels of the different sensors are combined into one image file and a PCA is applied for all the channels.

2.3.2.3 Gram-Schmidt transformation

The Gram-Schmidt transformation method was registered as US Patent by Laben and Brower (2000). A coarser spatial resolution PAN image is simulated and a Gram-Schmidt transformation (Clayton, 1974) is performed on the simulated coarser spatial resolution PAN image and the plurality of lower spatial resolution spectral band images. The simulated lower spatial resolution PAN image is employed as the first band in the Gram-Schmidt transformation. The statistics of the finer spatial resolution

panchromatic image are adjusted to match the statistics of the first transform band resulting from the Gram-Schmidt transformation and the higher spatial resolution panchromatic image (with adjusted statistics) is substituted for the first transform band resulting from the Gram-Schmidt transformation to produce a new set of transform bands. Finally, the inverse Gram-Schmidt transformation is performed on the new set of transform bands to produce the finer spatial resolution MS image.

2.3.2.4 Filter fusion (FF)

The idea of filter fusion is to transfer the high spatial frequency content of the high spatial resolution image to the MS imagery (Chavez *et al.*, 1991). Note that the filters referred to are spatial filters. The fusion is performed by combining a version of the fine spatial resolution image and a high pass filtered (HPF) version of the more highly resolved image. The method preserves a high percentage of the spectral characteristic, since the spectral information is associated with the low spatial frequencies of the MS imagery. The high-frequency information is extracted by high pass filtering the finer spatial resolution PAN band. The cut-off frequencies of the filters have to be chosen in such a way that the included data do not influence the spectra of the opposite data. The limitation of the approach is the introduction of false edges if the low spatial resolution MS band and the more highly resolved band exhibit only a weak correlation.

2.3.3 Relative spectral contribution methods

2.3.3.1 PAN+XS method (PAN+XS)

PAN+XS method is a standard merging method for SPOT HRV multispectral band 1 and band 2 and PAN image fusion. SPOT HRV provides two almost identical instruments and therefore the method can be applied to scenes that were observed by one instrument in the multispectral mode and by the other in PAN mode. Again, the fusion is based on the correlation between the bands XS_1 and XS_2 with PAN_{10}. The radiance counts of the new XSH_1 and XSH_2 bands are computed according to Equation (2.13) as

$$SX_1^H = \frac{2 \cdot PAN_{10} \cdot SX_1^L}{SX_1^L + SX_2^L} \text{ and } SX_2^H = \frac{2 \cdot PAN_{10} \cdot SX_2^L}{SX_1^L + SX_2^L} \tag{2.13}$$

2.3.3.2 Regression fusion (RF)

Due to the high correlation between the visible bands, the relation between one of the fused visible waveband images and the fine spatial resolution band can be expressed by the simple regression shown as

$$DN_{resampledi} = a_i + b_i DN_{high} \tag{2.14}$$

The bias parameter a_i and the scaling parameter b_i can be calculated by a least squares approach between the resampled digital number of the low spatial resolution multispectral band i $DN_{resampledi}$ and the high spatial resolution DN_{high}. The regression technique is not suitable for the uncorrelated bands since the global correlation is low. However, the problem can be partially solved using a local correlation approach. Instead of computing the global regression parameters, the a_i and b_i parameters are determined in a sliding window (Price, 1999). The more complicated regression method could also be implemented using both fine and coarse spatial resolution images and more parameters which are determined using geostatistical approaches (Pardo-Igúzquiza *et al.*, 2006).

2.3.3.3 Brovey transform

The Brovey transformation is based on spectral modelling (Ranchin and Wald, 2000, Liu, 2000). The algorithm is

$$DN_{fusedi} = \frac{DN_1}{DN_1 + DN_2 + .. + DN_b} DN_{highes} \tag{2.15}$$

where DN_{fusedi} means the digital number of the resulting fused image produced from the input data in b multispectral bands multiplied by the high spatial resolution image DN_{highes}. The Brovey transform assumes that the spectral range of the high spatial resolution image DN_{highes} covers the spectral range of the sum of the DN_{fusedi}. If this is not the case, there will be spectral distortion in the synthesised images.

2.3.4 Wavelets image fusion based on the ARSIS concept methods

2.3.4.1 Wavelets transform

Wavelets theory is a powerful mathematical tool developed recently to analyse non-stationary signals of finite energy for which the classical formalism based on variance and the correlation function does not hold. Remotely sensed images are such a signal (Ranchin and Wald, 1993). The wavelet transform creates a summation of elementary fusion or wavelets from arbitrary functions of finite energy. Respective weights of the

wavelets in the summation are called the wavelet coefficients. The mother wavelet or generating wavelet is a function $\Psi(t)$ i.e.

$$\psi_{a,b} = |a|^{-1/2} \psi[(x-b)/a] \tag{2.16}$$

A remotely sensed signal function $f(t)$ can then be decomposed to the approximation and detail images using wavelets as in Figure 2.2 (Mallat, 1989). The model was used to compute the wavelet coefficient for the high-pass filter and low-pass filter in the model called *multiresolution analysis*. Multiresolution analysis can be used to produce an approximation image at multi-scales represented as a pyramid (Figure 2.3). At the base of the pyramid is the original image. Each level of the pyramid is an approximation computed from the original one. The higher the level of the pyramid, the coarser is the spatial resolution of the approximation image.

A number of researchers have used wavelets for fusion of remotely sensed images. Examples are Garguet-Duport *et al.* (1996) and Yocky (1996), who replaced the approximation images of the decomposed multispectral bands of 20 m spatial resolution by an approximation of the SPOT PAN band and used the inverse wavelet transform to produce synthesised images of multispectral bands of 10 m spatial resolution; Li *et al.* (2002), used a similar technique to replace the detail image of the multispectral bands at two successive levels in the multi-scale pyramid.

2.3.4.2 ARSIS concepts for image fusion

The ARSIS concept for image fusion was first developed by Wald *et al.* (1997) for fusion of a high spatial resolution image A_h with high spectral but with low spatial resolution image B_{li}. The synthetic image of the high spectral bands is B^*_{hi}. The concept is based on three properties as

- Any synthetic image B^*_{hi}, once degraded to its original spatial resolution l, should be as similar as possible to the original image B_{li}.
- Any synthetic image B^*_{hi} should be as similar as possible to the images B_{li} that the corresponding sensor would observe with the highest spatial resolution h.

Approximation of The images at spatial resolution 1/2	"Horizontal" detail at the spatial resolution 1/2
"Vertical" detail at the spatial resolution 1/2	"Diagonal" detail at the spatial resolution 1/2

Figure 2.2. Approximation and detail images of an image using the Mallat algorithm

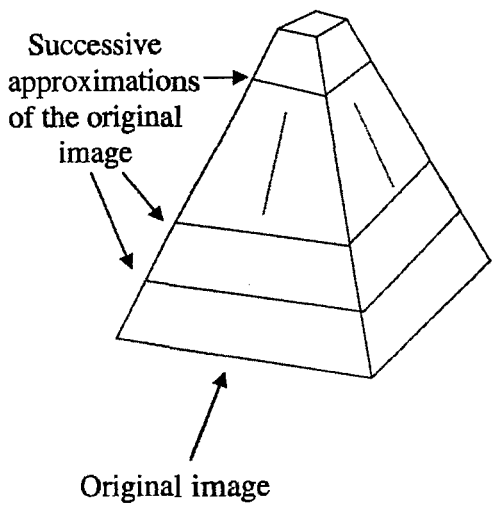


Figure 2.3 Multi-scale pyramid.

- The MS set of synthetic images B_{hi}^{*} should be as similar as possible to the MS set of images B_{li} that the corresponding sensor would observe with the highest spatial resolution h.

Ranchin and Wald (2000) and Ranchin *et al.* (2003) constructed a scheme for the application of the ARSIS concept using a multi-scale model and its inverse and used it to fuse the SPOT PAN and multispectral bands. The scheme is similar to the wavelet-based fusion presented in section 2.4.4.1 except the core procedure in the proposed scheme is the Inter-Band Structures Model (IBSM) and the quality assessment process.

Blanc *et al.* (1998) used iterated rational filter banks within the ARSIS concept to produce 10 m Landsat multispectral images using the SPOT PAN bands. The research resolved the problem of dyadic multiresolution analysis by using the rational factor of 1/3 or 2/3 and several approaches were used to synthesise the high spatial resolution TM MS images.

2.3.4 Further fusion methods

2.3.4.1 Hybrid method

Núñez et al., (1999) used the "a trous" wavelet decomposition combined with the PCA transformation and the IHS transformation to merge SPOT PAN and Landsat TM images. The "a trous" wavelet algorithm decomposes the image into wavelet planes. Given an image P decomposed to a sequence of approximations P_1, P_2,., P_r and the wavelet plane, w_B^l, defined as the difference between two successive approximations P_{l-1} and P_l. The reconstruction can be carried out as

$$P_0 = \sum_{l=1}^{n} w_l + P_r \tag{2.17}$$

The wavelet planes can be used to fuse images in two different approaches such as the substitution method and additive method. In the substitution method, the wavelet planes of the SPOT PAN image can be substituted for wavelet planes of the TM multispectral images in PCA transformation fusion. Similarly, the additive method adds the wavelet planes of the SPOT PAN image to the red, blue and green bands of the TM images or to the intensity component of the IHS fusion.

2.3.4.2 Geostatistical digital image merging

A geostatistical algorithm was used by Garcia (2002) to merge Landsat 7 ETM+ images and a 3 m spatial resolution aerial ortho-image. The method consists of a series of geostatistical techniques. The process began with a normal score transformation of the image data in order to have a unique statistical distribution for all the images. The data were then used to compute the variograms in the column and row directions. The behaviour of the obtained variograms was analysed for use in the next step. From the models it was possible to obtain the 3 m variogram model parameters applying an iterative deregularisation. The parameters were used as the input information for GSLIB SGSIM co-simulation program to simulate the data sets on 3 m-pixel size. The last operation involved conditioning the co-simulated images

with the original 30 m spatial resolution to keep the coincidence between the original digital numbers and the mean values of the 10 x 10 pixels of 3 m spatial resolution cosimulated image. The geostatistical method has many advantages over non-geostatistical methods such as IHS, PCA, high-pass filter and colour normalised from a geostatistical point of view. However, more research on quality of the fusion results should be done to investigate the potential of the method.

Pardo-Igúzquiza *et al.* (2006) proposed an approach for sharpening MS image by PAN image based on downscaling cokriging. Using linear systems theory (Papoulis, 1984), the weight parameters for random variables of pixels at fine and coarse spatial resolution were defined for the image fusion model. The comparison demonstrated that the downscaled cokriged sharpened image was more visually and quantitatively similar to the target image than the fused image from HPF fusion approach.

2.3.4.3 Spatial based technique

Steinnocher (1999) used a filter technique for image sharpening. The method is based on the assumption that an image object Z is represented by a set of neighbouring pixels z_i, whose values are Gaussian distributed. The adaptive modified sigma filters are used to assign the pixels to the objects since it matches the assumption given above.

For the fusion approach, the multispectral bands are included in the filtering process. The multispectral images are co-registered with the higher spatial resolution PAN image by nearest neighbour sampling. The pixels belonging to an object are recognised using modified sigma filter to the PAN image. The position of the object pixels are transferred to the multispectral bands, where the average of the respective sub-pixel is performed. The resulting multispectral images are sharpened by the objects defined in the higher spatial resolution.

2.4 Super resolution mapping

As briefly described in Chapter 1, super resolution mapping is the technique of increasing the spatial resolution of thematic maps in comparison with the original spatial resolution using soft-classified land cover proportion data. Super resolution mapping techniques make use of the land cover proportions information produced by soft classification and a spatial dependence assumption. This section introduces super-resolution mapping and other approaches for increasing the spatial resolution (without

using the information at sub-pixel spatial resolution) such as image sharpening, clustering techniques, neural network optimisation and two-point histogram.

2.4.1 Image sharpening

Similar to super-resolution mapping, image sharpening methods can be used to increase the spatial resolution of land cover maps. In some research, the spatial resolution of the classified data was increased using the proportion maps and finer spatial resolution data such as a PAN image. These methods are a kind of remotely sensed data fusion. An example of these methods can be cited as Gross and Schott (1998), who used synthetic PAN imagery of 4.5 m spatial resolution to sharpen the proportions image derived from the spectral mixture analysis at 9 m spatial resolution. However, the result retains a certain degree of uncertainty since the prediction was not a hard thematic map of the experimental area but the proportions image at a finer spatial resolution. Another example is Foody (1998), who used an image of 15 m spatial resolution to sharpen 30 m spatial resolution data to produce a soft classified image of a boat lake in Swansea, UK. The resulting image was refined through fitting class membership contours. A similar regression technique was used by Foody (2002b) to determine the position of ground control points (GCPs) within coarse spatial resolution image pixels of AVHRR (Advanced Very High Resolution Radiometer) data to reduce misregistration errors. The mean distance between the estimated locations of 11 GCPs in the fine versus coarse spatial resolution imagery was reduced by 37.5 percent from 655.8 m for the conventional method to 421.2 m.

The spatial resolution can be increased directly (without using the soft-classification) from the original MS image. Pinilla Ruiz and Ariza Lopez (2002) used the deconvolution filters derived from the point spread function to restore the original signal of SPOT HRV data. The method increased the spatial resolution of the restored image, especially for finding structures such as concrete dams, bridges, highway, etc.

2.4.2 Clustering techniques

This section focuses on super resolution mapping techniques using clustering algorithms. The clustering algorithms are based on the assumption of spatial dependence, which refers to the tendency for pairs of pixels of a given property to be more alike when close together than when further apart. This assumption reflects the nature of spatial objects in remotely sensed imagery where pixels in the boundaries

between objects may cover several land cover classes and the sub-pixels of the same class are closer.

2.4.2.1 Per-field classification

The spatial resolution of the resulting land cover maps can be increased using sub-pixel mapping by per-field classification. Aplin *et al.* (2000, 2001) implemented per-field classification of the data both from hard and soft classification of *CASI* data at 4 m spatial resolution. In the first stage, the polygon vector data were integrated with raster data at a spatial resolution of 0.4 m. Following the integration of image data, the image was classified per-pixel using both conventional maximum likelihood and Bayesian fuzzy classification. In the next stage, a per-field land cover class was produced from the hard and soft classification in different ways. For the hard classified images, the number of pixels was calculated for each land cover class and the land cover label of each polygon was assigned according to the class with the largest number. For soft classification, the class assignment of each scaled-up pixel was based on the area of the original pixel covered by polygons and the proportions of the land cover classes. The results show high accuracy for super resolution mapping by per-field classification. However, the disadvantage of super resolution by per-field classification is the limited availability of accurate vector data sets in most cases.

2.4.2.2 Spatial clustering and image segmentation

Schneider and Steinwendner (1999) used spatial sub-pixel analysis combined with segmentation techniques to produce vector maps at sub-pixel accuracy. The first step in this technique is region growing segmentation started at seed pixels. The seed pixels must be pure pixel (pixel locates entirely inside an object). The region growing process is implemented with strict homogeneity requirements. Pixels in the boundary of segments are examined for candidates for sub-pixel analysis. The spatial sub-pixel analysis derives sub-pixel information from the spatial pattern of pixels in a certain neighbourhood (e.g. 3x3 pixels) of a given pixel. Different models for the scene pattern within the neighbourhood can be assumed. This sub-pixel parameter information can be used for segmentation directly or for resampling with a smaller pixel size to produce finer spatial resolution images. The technique represents a successful, automated and simplistic pre-processing step for increasing the spatial resolution of satellite sensor imagery. However, application of this technique is limited to the images containing large objects with straight boundaries at a certain

spatial resolution. Moreover, the models used still have problems with pixels containing more than two classes.

2.4.2.3 Spatial clustering techniques

Atkinson (1997) proposed the concept of sub-pixel mapping or super-resolution mapping and introduced a method for locating the land cover within a pixel using soft classified images. The sub-pixel locating process was based on the spatial dependence maximisation of the sub-pixels. A sub-pixel is assigned to a given land cover class based on its highest spatial dependence scores for each land cover class. Based on the concept of super-resolution mapping, several algorithms for allocating classes to all the sub-pixels by maximising their spatial dependence within an original pixel were proposed. These methods differ in the way to calculate and maximise the spatial dependence. Atkinson (1997) calculated the attractiveness based on a local window. The scores were calculated by the inverse squared distances for each class. The method was then extended to a sub-pixel swapping technique which have been proposed by Atkinson (2001, 2006). At first, a number of sub-pixels for each land cover class were allocated based on the increase of the spatial resolution and the soft-classification output. Subsequently, the attractiveness score for each sub-pixel was calculated and the swapping of sub-pixels was implemented to increase the spatial correlation. The algorithm ran for a specified number of iterations or until the spatial correlation was maximised. The approach was used by Thornton *et al.* (2006) to map the rural land cover objects using QuickBird image in Christchurch, UK.

Verhoeye and De Wulf (2001) also extended the work of Atkinson (1997) by establishing a linear optimisation model. The location of sub-pixels is determined by maximizing spatial dependence

$$SPD = \sum_{i=1}^{c} \sum_{j=1}^{NP} x_{kj} * PLC_{kj} \tag{2.18}$$

where the variable x_{ij} = 1 if a sub-pixel j is assigned to land cover class k and otherwise, $x_{kj} = 0$, c is the number of land cover classes, NP is number of sub-pixels in a pixel, and PLC_{kj} is spatial dependence of sub-pixel j belonging to land cover class k. The proportion image is used as a constraint for optimisation procedures. The spatial dependence parameter PLC_{kj} can be calculated from the land cover class proportions of each pixel and distance from the sub-pixel to the neighbourhood pixels.

Merterns *et al.* (2003) used a genetic algorithm to allocate the sub-pixels in a pixel based on the neighbouring value which represents the spatial dependence of the sub-pixels. The neighbouring value of each sub-pixel can be calculated as

$$\eta_i = \sum_{k=1}^{N} \frac{\delta_{ij}}{N} \tag{2.19}$$

with $\delta_{ij} = 1$ if neighbour j belongs to the same class as sub-pixel i and $\delta_{ij} = 0$ otherwise, N is number of neighbouring sub-pixels. The allocation of the sub-pixels should maximise the sum of the neighbouring values of all sub-pixels within an original pixel. The sum of the neighbouring values of all sub-pixels within an original pixel is

$$\eta = \sum_{i=1}^{NP} \eta_i \tag{2.20}$$

where NP is number of sub-pixels within an original pixel.

Zhan *et al.* (2002) implemented sub-pixel mapping based on land cover class probabilities. The sub-pixel classification procedure consisted of two stages. The first stage was standard classification using the maximum likelihood algorithm. The classified image was then used to determine the class probabilities of a sub-pixel for each class by interpolating using the land cover classes of eight neighbouring pixels and the pixel itself. Subsequently, the sub-pixel hard classification image was derived from the interpolated probabilities at the sub-pixel scale. Despite considerable increase in accuracy in comparison with the hard classification, in general, the approach often softens the land cover image rather than predicting accurately the land cover location within a pixel.

2.4.4 Two-point histogram

An alternative technique for super-resolution mapping of land cover is the two-point histogram (Atkinson, 2003). The technique is based on geostatistical post-processing with simulated annealing in the program ANNEAL (Deutsch and Journel, 1998). The spatial optimisation algorithm aims to minimise the objective function as

$$O = \sum_{h} \left(\sum_{k=1}^{K} \sum_{k'=1}^{K} \left[p_{k,k'}^{training}(h) - p_{k,k'}^{realization}(h) \right]^2 \right) \tag{2.21}$$

where K is number of outcomes or, in this case, is land cover categories, the bivariate transition probability $p_{k,k'}^{training}(h)$ and $p_{k,k'}^{realization}(h)$ are the target and corresponding

frequencies of the realization image. The two-point histogram $p_{k,k'}(h)$ at distance h is the set of bivariate transition probabilities defined as follows:

$$p_{k,k'}(h) = \Pr\left\{ \begin{array}{l} Z(i) \in category \quad k \\ Z(i+h) \in category \quad k' \end{array} \right\} \tag{2.22}$$

where Z is a random variable at position i that can take one of $k = 1, \ldots, c$ outcomes. Applied to the super-resolution problem, the realization image is initiated and contains hard classified sub-pixels, with the number of sub-pixels per class in each original pixel determined according to land cover class proportions from the soft classification. For example: in a pixel of 10 by 10 sub-pixels with proportion of 30% grassland and 70% built-land, there will be 30 grassland sub-pixels and 70 built-land sub-pixels. The two-point histogram of training and an initial image is calculated for optimisation. In each pixel, the sub-pixels are swapped in order to gain a smaller value for the objective function. The swapping is accepted only if the new value of the objective function $O_{new} < O_{initial}$ (Goovaerts, 1997). If the swapping is accepted, the value of the objective function will be updated to the new value. This process is repeated until the objective function value is minimised.

Foody *et al.* (2003, 2005) used the above described method for super-resolution mapping of shoreline in Kampung Sebarang Takir, Terenggaru, Malaysia. The prediction result shows a higher accuracy in terms of RMSE and percentage of prediction within 2 m of actual shoreline in comparison with hard classification, contouring soft classification and wavelet interpolation.

2.4.3 Super-resolution mapping using neural networks

2.4.3.1 Super-resolution mapping using MLP neural network

MLP neural network was used for super-resolution mapping by Mertens *et al.* (2004). For the super-resolution mapping task, three networks were trained to produce the horizontal, vertical and diagonal wavelet coefficients from a land cover proportion image. The input layer of the network included nine neurons. Each neuron represented a proportion value of the central original pixel or its eight surrounding pixels. In the output layer of the neural network, there was only a single neuron which represented the horizontal, vertical or diagonal wavelet coefficient value. From the predicted wavelet coefficient images and the land cover proportion image, the proportion image at the sub-pixel level was produced based on multiresolution analysis. The proportion images were then used to produce the crisp land cover images based on a hardening

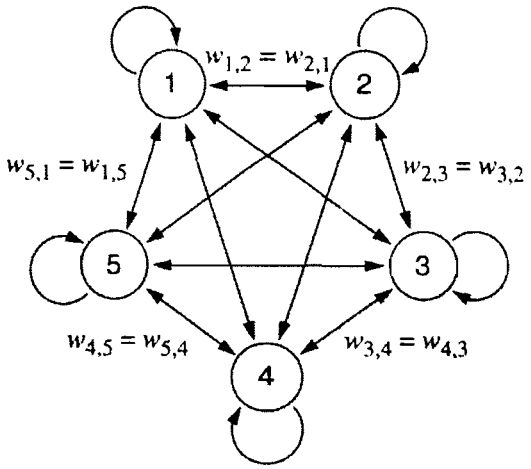


Figure 2.4. Hopfield neural network of five nodes

process. Both artificial and real proportion images were used to test the method. Visual comparison of the results showed that the predicted maps at the sub-pixel level were more detailed and contiguous than those obtained by hard classification.

2.4.3.2 Super-resolution mapping using the Hopfield neural network

- ***Hopfield neural network***

The Hopfield neural network is a fully connected recurrent network that is mostly used for auto-association and optimisation (Mehrotra *et al.*, 1997). A simple network of five neurons can be seen as in Figure 2.2.The stable states of the network are obtained when the energy function is minimised. Energy functions E can be simplified such that

$$E = -\frac{1}{2}\sum_{i}\sum_{j} T_{ij} v_i v_j - \sum_{i=1} v_i I_i \tag{2.23}$$

where T_{ij} is the weight of connection from neuron j to i, and v_i and v_j are the outputs of neurons i and j. I_i is an external bias on neuron i.

For optimisation purposes, the energy function is calculated based on the goal and constraint functions, which can be defined as

$$Energy = Goal + Constraints \tag{2.24}$$

For the task of super resolution land cover mapping, the Hopfield neural network is used as an optimisation tool by Tatem *et al.* (2001a, 2001b, 2002a, 2002b, 2002c). The class proportion images output from a soft classification are used as constraints for minimizing the energy function of the network.

- ***Target identification by Hopfield neural network***

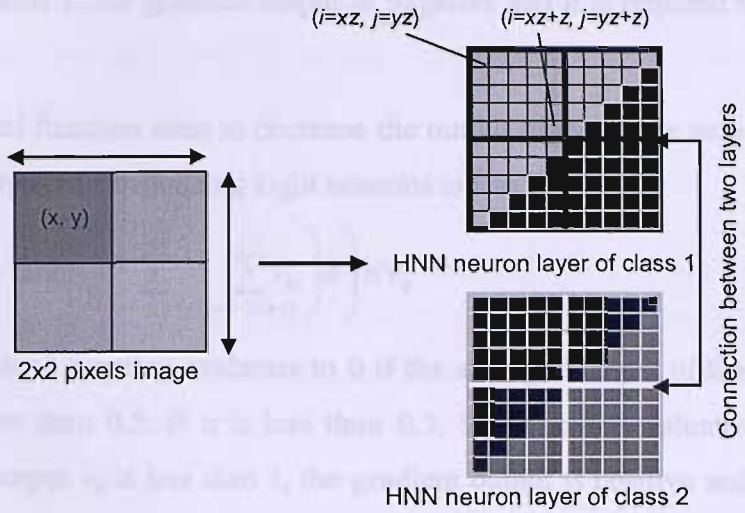


Figure 2.5 Hopfield neural network for super-resolution mapping

Tatem *et al.* (2001b) used the Hopfield neural network for target identification from remotely sensed imagery. The depiction of a Hopfield neural network designed for super-resolution mapping can be seen in Figure 2.5. In the research, each sub-pixel (i,j) in the image coordinates was represented by a neuron in the Hopfield neural network. Each pixel (x,y) in the satellite sensor image was represented by a matrix of $z \times z$ sub-pixels centred at $[xz + int(z/2), yz + int(z/2)]$, where *int* is an integer value (z is the zoom factor). The energy function E was defined as

$$E = -\sum_i \sum_j (k_1 G1_{ij} + k_2 G2_{ij} + k_3 P_{ij}) \tag{2.25}$$

where k_1, k_2 and k_3 are weighted constants, $G1_{ij}$ and $G2_{ij}$ are output values for neuron (i,j) contributed by the two goal functions, and P_{ij} is the output value contributed by the constraint. The first goal function $G1_{ij}$ aims to increase the output of the centre neuron v_{ij} to 1 if the average output of the surrounding eight neurons is greater than 0.5.

$$\frac{dG1_{ij}}{dv_{ij}} = \frac{1}{2}\left(1 + \tanh\left(\frac{1}{8}\sum_{k=i-1(k\neq j)}^{i+1}\sum_{l=j-1(l\neq j)}^{j+1} v_{kl} - 0.5\right)\lambda\right) \times \left(v_{ij} - 1\right) \tag{2.26}$$

where λ is a gain which controls the steepness of the *tanh* function. The *tanh* function controls the effect of the neighbouring neurons. If the averaged output of the neighbouring neurons is less than 0.5, then the function has no effect on the energy function. If the average of the surrounding function is greater than 1 and the output of

neural v_{ij} less than 1, the gradient output is negative and it is required to increase the neuron output.

The second goal function aims to decrease the output of the centre neuron to $v_{ij} = 0$ if the average output of surrounding eight neurons is less than 0.5

$$\frac{dG2_{ij}}{dv_{ij}} = \frac{1}{2}\left(1 - \tanh\left(\frac{1}{8}\sum_{k=i-1(k\neq j)}^{i+1}\sum_{l=j-1(l\neq j)}^{j+1} v_{kl}\right)\lambda\right) \times v_{ij} \tag{2.27}$$

This time, the *tanh* function evaluates to 0 if the averaged output of the neighbouring neurons is more than 0.5. If it is less than 0.5, the function evaluates to 1 and the centre neural output v_{ij} is less than 1, the gradient output is positive and it is required to decrease neuron output. The values of the neuron outputs are stabilised if only one of these conditions is satisfied and $G1_{ij} + G2_{ij} = 0$.

The two goal functions of the Hopfield neural network play the role of the spatial dependence maximisation in super-resolution mapping. By adjusting the values of the two goal functions, the energy function of the HNN is minimised or the adjacent sub-pixels are more likely to be in the same land cover class.

The goal functions provide the enforcement of spatial order. However, the sole application of these functions would result in all neuron outputs taking the values 1 or 0. Therefore, a method of constraining the effect of those functions to the correct image is required. The constraint P_{ij} aims to retain the pixel class proportions of the soft classified images. The constraint can be described as follows:

$$\frac{dP_{ij}}{dv_{ij}} = \frac{1}{2z^2}\left(\sum_{k=xz}^{xz+z}\sum_{l=yz}^{yz+z}(1+\tanh(v_{kl}-0.55)\lambda)\right) - a_{xy} \tag{2.28}$$

where 0.55 is threshold to define a sub-pixel belong to a class. If the area proportion of the estimate for the original pixel (x,y) is lower or greater than the target area, the output values of neurons are increased or decreased to solve the problem. Only when the area proportion estimate is identical to the target area proportion for each pixel, a zero gradient occurs, corresponding to $P_{ij} = 0$ in the energy function. The weighted constants k_1, k_2 and k_3 are used to rank the importance of the element functions in Equation (2.25). Tatem *et al.* (2001b) suggested that the values of $k_1 = k_2 = 150$ and $k_3 = 130$ are appropriate for land cover target identification.

- ***Land cover pattern prediction by Hopfield neural network***

The above mentioned energy function is applied to identify large land cover targets. For identifying the targets of sub-pixel size, Tatem *et al.* (2002c) use another energy function as follows:

$$E = -\sum_i \sum_j \left(k_1 S1_{ij} + k_2 S2_{ij} + k_3 S3_{ij} + k_4 S4_{ij} + k_5 S5_{ij} + k_6 S6_{ij} + k_7 S7_{ij} + k_8 P_{ij}\right) \quad (2.29)$$

where k_1 to k_8 are constants weighting the energy parameters, P_{ij} represents the output value for the area proportion constraint, and $S1$ to $S7$ represent the output values for neuron (i,j) of seven semivariance functions. These seven functions are defined as

$$\frac{dSk_{ij}}{dv_{ij}} = v_{ij} - v(c)_{ij} \quad (2.30)$$

where

$$v(c)_{ij} = \frac{1}{2N(h)}\left[-\sum_{i=1,j=1}^{N(h)} v_{l\pm h,j\pm h} \pm \sqrt{\left(\sum_{i=1,j=1}^{N(h)} v_{l\pm h,j\pm h}\right)^2 - 4\cdot 2N(h)\left(\sum_{i=1,j=1}^{N(h)} \left(v_{i\pm h,j\pm h}\right)^2 - 2N(h)\gamma(h)\right)}\right] \quad (2.31)$$

In Equation (2.31), $N(h)$ is the number of pixels at lag h from the centre pixel, $\gamma(h)$ is the semivariance at lag h (Atkinson *et al.*, 2000). The semivariances $\gamma(h)$ of the objects are derived from aerial photographs of the experimental area.

- ***Land cover mapping at the sub-pixel scale using the Hopfield neural network***

Super resolution mapping (or land cover mapping at a sub-pixel scale) techniques are developed from the above described target identification techniques. Tatem *et al.* (2001a) introduced the multi-layer Hopfield neural network for super resolution land cover mapping. In this network, each land cover class is represented by a neural layer in the network. The neurons in the network connect to all neurons in the same layer and other layers. From the energy function (Equation (2.25)), a multi-class constraint is added as follows

$$E = -\sum_h \sum_i \sum_j \left(k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij}\right) \quad (2.32)$$

where the goal and within-pixel area constraint functions are similar to those in Equations (2.26), (2.27) and (2.28). The constraint function M_{hij} is defined as

$$\frac{dM_{hij}}{dv_{hij}} = -\left(\sum_{k=0}^{N} v_{hij}\right) - 1 \quad (2.33)$$

Constraint (2.28) can be expressed as the sum of outputs of each set of neurons representing pixel (i,j) in the final image must be equal to 1.

Where the scene consists of features that are both larger and smaller than a pixel, Tatem *et al.* (2002a, 2002b) combined the energy functions (Equation (2.29) and Equation (2.32)) to form a comprehensive energy function

$$E = -\sum_{h}\sum_{i}\sum_{j}\left(\sum_{n=1}^{7} k_n Sn_{hij}\right) + k_8 G1_{hij} + k_9 G2_{hij} + k_{10} P_{hij} + k_{11} M_{hij}) \qquad (2.34)$$

where the first seven constraint functions are similar to those in Equation (2.29), and the other parameter functions are like those in Equation (2.32).

2.5 General model for using supplementary data for super-resolution mapping

2.5.1 The Hopfield neural network and added constraint and goal function

Comparing with the Hopfield neural network techniques, the regression techniques in section 2.3.2 are less accurate since they are unable to create a suitable computation model for the spatial dependence between sub pixels. The spatial dependence, which is often calculated by interpolation of land cover class proportions from the centre of the original pixel, is complicated and not suitable for clustering of the sub pixels in some cases such as mapping the object smaller than a pixel. Conversely, the optimisation by the HNN simply clusters the sub pixels into regions and constrains them by the proportion and multi-class functions. Therefore, it is less complicated and more effective to maximise the spatial dependence.

Another advantage of super resolution mapping by the HNN is its potential to incorporate other ancillary information for land cover mapping by using the constraints and the goals for the energy function. This ancillary information may be derived from other sources of data such as remotely sensed imagery obtained by other sensors or with different spatial resolutions such as PAN images, aerial photographs or other information on the objects. Particularly, with a flexible zoom factor z, the Hopfield neural network super resolution mapping technique can be used as a tool for data fusion, which is useful for land cover mapping.

The energy function in Equation (2.32) can be used to incorporate the supplementary information for super-resolution land cover mapping, assuming that the supplementary data can provide the information about the land cover classes for the sub-pixels or the neurons of the HNN. Similar to the semivariance functions in Equation (2.34), the information about the land cover classes can be represented by a

function F. For every sub-pixel (i,j) or neuron (h,i,j) the value of the function F (F_{hij}) can be added to the energy function as

$$E = -\sum_{h} \sum_{i} \sum_{j} (k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij} + k_F F_{hij}) \tag{2.35}$$

The weighting coefficient k_F controls the effect of the supplementary information on the energy function and the class labeling process for sub-pixel (i,j). If the value of k_F is great, the resulting land cover is more likely to be indicated by the supplementary information. Hence, when the information revealed by the supplementary data can define the land cover class accurately, the weighting coefficient k_F can be increased to obtain the accurate results. Otherwise, the small value of k_F still makes a contribution of the function F into the super-resolution mapping results.

2.5.2 General model for super-resolution mapping using fused or panchromatic images

Amongst the sources of data which can provide useful information for more accurate super-resolution mapping, panchromatic imagery are the most widely available at a spatial resolution finer than that of the original MS image. Panchromatic imagery can be obtained by the current commercial remote sensing sensors such as Landsat ETM+ with 15 m spatial resolution PAN image, SPOT 5 with 2.5 m PAN image, IKONOS with 1 m PAN image and QuickBird with PAN image of around 0.6 m spatial resolution. These panchromatic images can be fused with the MS image at a coarser spatial resolution or used directly to provide useful information for more accurate super-resolution mapping.

The general model for using information from the panchromatic imagery for super-resolution mapping is presented in Figure 2.6. The model is implemented at three different spatial resolutions: original spatial resolution of the MS image, intermediate spatial resolution PAN image, and sub-pixel resolution of the resulting land cover classes. The proposed model employs Hopfield neural network optimisation as the core algorithm since it is easy to incorporate the additional information from PAN imagery by adding new constraint or goal functions to the energy function of the HNN. Similar to the HNN used by Tatem *et al.* (2002), the HNN is constrained by the soft-classified land cover proportions. In the proposed model, these new function values can be defined based on forward and inverse models. According to the

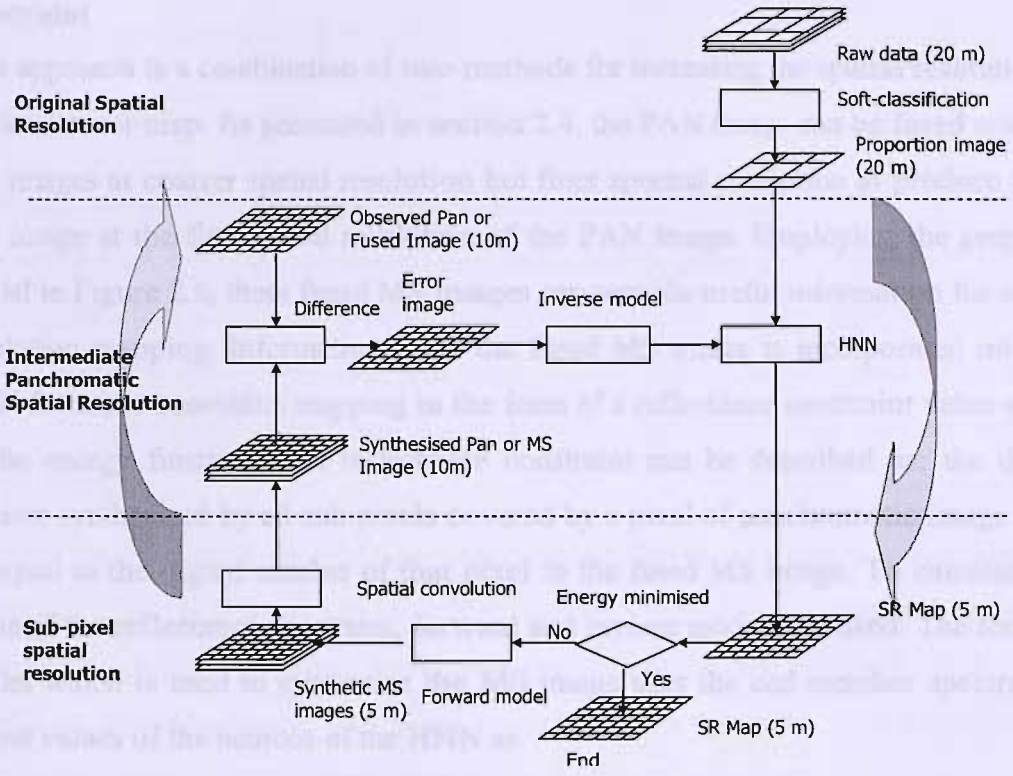


Figure 2.6. General model for super-resolution mapping using panchromatic or MS image

proposed model, the super-resolution mapping with information from PAN imagery uses the land cover proportions at the spatial resolution of the MS image as input to initialise and constrain the HNN. After the first iteration, the sub-pixel crisp land cover map is predicted. The forward model is then used to predict the synthetic MS image at sub-pixel spatial resolution using end-member spectra. With spatial and spectral convolutions, synthetic PAN or MS images at the intermediate spatial resolution of the PAN image can be predicted. The synthetic PAN or MS images then are compared with the observed PAN or fused MS images to produce an adjusted value for the HNN using an inverse model. The adjusted or *reflectance constraint value* together with the goal functions and proportion constraint of the super-resolution mapping by *Tatem et al.* (2002) are used as inputs for the HNN to calculate the output values of the neurons and the value of the energy function in the next iteration. The process continues until the value of the energy function is minimised and the HNN converges to a stable state. At this state, the outputs of the HNN neurons are binary values which represent the sub-pixel super-resolved land cover maps.

2.5.2.1 Super-resolution mapping using information from the fused image as a constraint

This approach is a combination of two methods for increasing the spatial resolution of the land cover map. As presented in section 2.4, the PAN image can be fused with the MS images at coarser spatial resolution but finer spectral resolution to produce fused MS image at the fine spatial resolution of the PAN image. Employing the proposed model in Figure 2.6, these fused MS images can provide useful information for super-resolution mapping. Information from the fused MS image is incorporated into the HNN for super-resolution mapping in the form of a reflectance constraint value added to the energy function. The reflectance constraint can be described as: the digital number synthesised by all sub-pixels covered by a pixel of panchromatic image must be equal to the digital number of that pixel in the fused MS image. To calculate the value of the reflectance constraint, forward and inverse models are used. The forward model which is used to synthesise the MS image uses the end-member spectra and output values of the neurons of the HNN as

$$R^{o}_{hmn} = P_1 S_{C1} + P_2 S_{C2} + ... + P_c S_{Cc} \tag{2.36}$$

where R^{o}_{hmn} is the estimated reflectance of pixel (m,n) in a fused image band, c is the number of land cover class, $P_1, P_2,.., P_c$ are the proportions of a given land cover class covered by a pixel of the fused images or the number of sub-pixels in a given land cover class ($\sum P=1$), and $S_1, S_2,..,S_c$ are end-member spectra of the land cover class 1, 2,.., c in a multispectral fused band. The inverse model which produces the reflectance constraint value is a linear mixture model (Settle and Drake, 1993).

2.5.2.2 Super-resolution mapping using the panchromatic image as a constraint

The panchromatic imagery can be used directly for super-resolution mapping based on the core model as in Figure 2.6. Incorporating information from the PAN image directly in the super-resolution mapping process, it is possible to skip the image fusion process which is included in the approach using the fused imagery for super-resolution mapping proposed above. To obtain the synthetic PAN image, a spectral convolution process is employed to convolve the MS band synthesised by the forward model using end-member spectra. Similar to the super-resolution mapping with the fused image, the additional *panchromatic reflectance constraint* can be described as the digital number of every pixel of the synthetic panchromatic image

derived from the sub-pixels of the super-resolved maps must be equal to the digital number of the observed PAN image.

2.5.3 Multispectral image super-resolution

Using the forward model with end-member spectra in Figure 2.6, it is possible to produce the sub-pixel spatial resolution synthetic MS image from the sub-pixel land cover maps. In this application of the proposed model, the end product of the process is MS images at sub-pixel spatial resolution instead of the sub-pixel land cover maps. These super-resolved MS images can then be used for other remote sensing applications.

Another application of the HNN for MS super-resolution mapping is semivariogram matching. The MS image at a coarse spatial resolution is matched with the semivariogram at the sub-pixel scale to produce the MS image with sub-pixel spatial variation. The method can be applied to generate the sub-pixel pattern in the MS image. The details of the method are given in Chapter 5 of this thesis.

2.5.4 General model for super-resolution mapping using the LiDAR elevation data

LiDAR is technology for determining the elevation of the ground surface plus natural and human-made features. From the LiDAR data, it is possible to acquire a digital surface model (DSM) (an elevation model contains non-terrain objects such as trees, buildings, power lines and other human-made objects). In practice, the DSM is frequently used for extraction of trees and buildings. Different from the multispectral and panchromatic imagery, the LiDAR elevation data provide useful information for prediction of just a few land cover classes. Correspondingly, the model for integrating the LiDAR for super-resolution land cover mapping must be modified from the model based on the constraint in Figure 2.6.

Figure 2.7 presents the diagram for the LiDAR elevation data integrating model. The approach assume that the information obtained from the LiDAR elevation data is useful for a a land class (for example: building class). A class probability image is produced for the building class using elevation data. The class probability value represents the probability of the sub-pixel belonging to the building class. During the

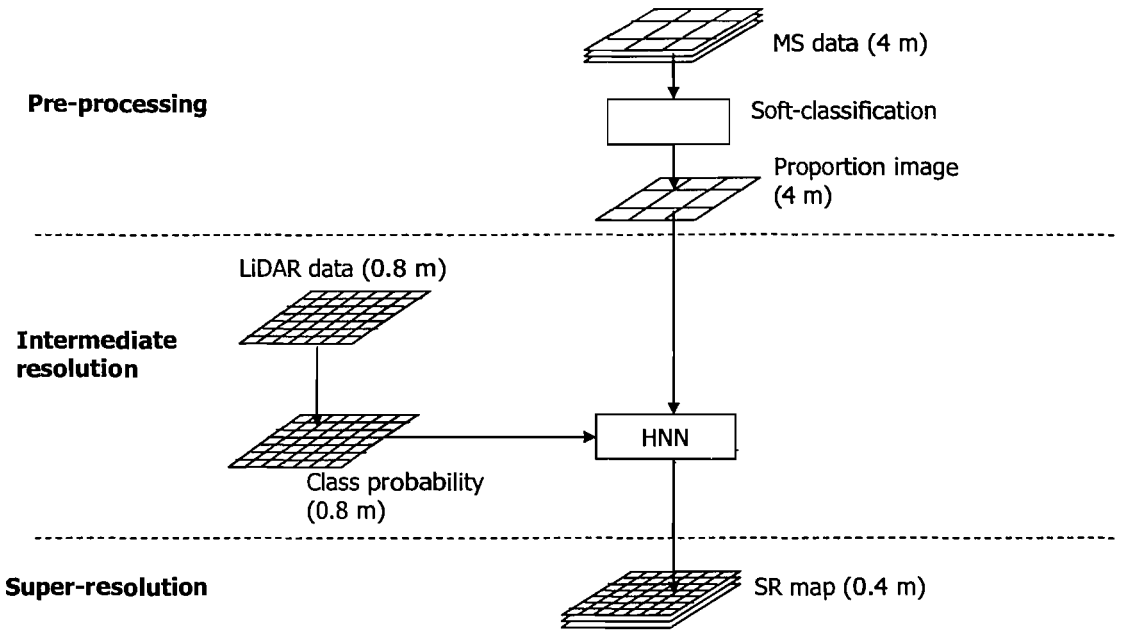


Figure 2.7 General model to incorporate LiDAR elevation data for super-resolution mapping

super-resolution mapping optimisation process, the output values of the neurons of the building class are increased according to the values in the class probability image. The increasing values for the neurons in the building class are named *height function values*. For the neurons of the other classes, the elevation data have no effect. Accordingly, the height function values for these neurons are zeros. Using this model, the LiDAR information is integrated into the super-resolution land cover mapping.

Chapter 3: Super-resolution mapping using a Hopfield neural network and a fused image[1]

3.1 Introduction

In this Chapter, a proposed model to incorporate super-resolution mapping and image fusion for increasing the spatial resolution is presented. The model is based on the Hopfield neural network. In this model, the information from the fused images is exploited in the form of an additional reflectance constraint of the energy function of the HNN. The value of the reflectance constraint is defined using forward and inverse models. The forward model is used to produce the synthetic MS images from the super-resolved land cover map at sub-pixel spatial resolution based on the end-member spectra. Comparing the synthetic MS images at the intermediate spatial resolution of the fused images with the observed fused image, the value of the reflectance constraint is determined using an inverse model based on a linear mixture model.

The proposed model was evaluated using the simulated and real remotely sensed imagery. The accuracy assessment of the resulting land cover map showed that the combination of the fusion and super-resolution mapping could increase the accuracy visually and statistically.

3.2 Hopfield neural network structure for using a fused image for super-resolution mapping

3.2.1 Fused imagery for land cover classification

There are two approaches for using a fused MS image for land cover classification. The fusion-then-classification approach uses the fused MS image directly for land cover classification. An example of the fusion-then-classification approach is given by Shackleford and Davis (2003) who used a 1 m fused MS image for urban mapping. The classification-then-sharpening approach uses the fused MS image to sharpen a land cover proportion image obtained from the original MS image. The classification-then-sharpening approach was used by Foody (1998) to sharpen a fuzzy classification output and by Gross and Schott (1998) to sharpen a proportion image obtained by spectral mixture analysis. According to the results of both approaches, the accuracy of

[1] This chapter is based on Nguyen *et al.* (2006)

the thematic map using the fused MS image increased slightly in comparison with that of the original MS image. Evaluation of the two approaches by Robinson *et al.* (2000) based on a linear mixing model indicated that the classification-then-sharpening approach was preferable in terms of accuracy. The fusion-then-classification approach produced a thematic map with lower accuracy due to the spectral distortion in the fused MS image.

3.2.2 The Hopfield neural network design[2]

This section presents a model for using the fused image as additional information for super-resolution mapping. Theoretically, the fused MS image can provide useful information at an intermediate spatial resolution for predicting land cover at a finer spatial resolution. However, the spectral distortion of the fused images may cause some errors. To reduce the effect of spectral distortion, the classification-then-sharpening approach was incorporated into the HNN for super-resolution mapping. Figure 3.1 is a graphical depiction of the proposed method to use a fused image for super-resolution mapping by the HNN. From the MS images at the original spatial resolution land cover area proportion images are produced by a soft-classification procedure. The area proportion images are then used to constrain the HNN to produce the super-resolution land cover map in the first iteration of the optimization process.

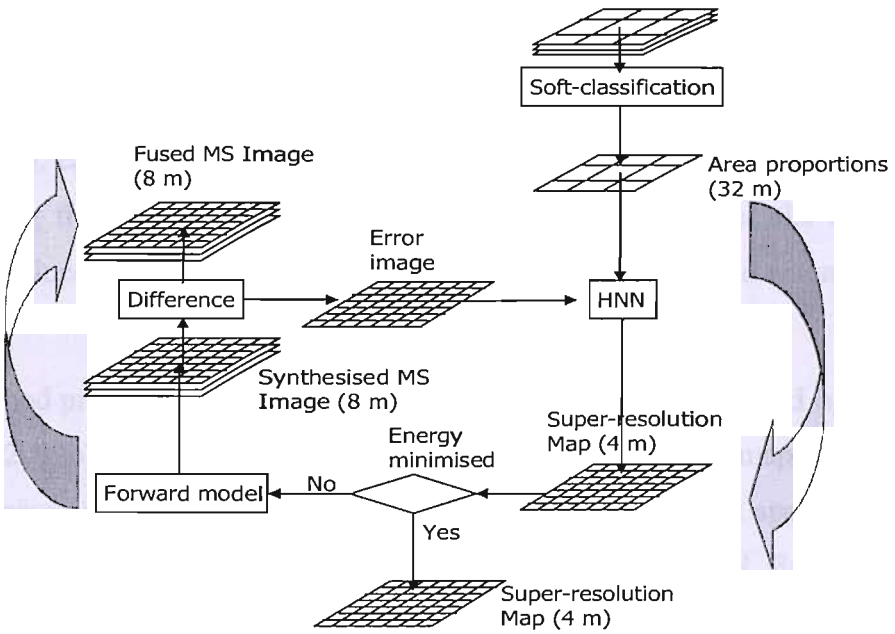


Figure 3.1 HNN super-resolution mapping using the fused multispectral mages

[2] In this section, there will be some repeatation of the model presented earlier due to use of paper in the thesis, and so that reader can read this chapter independently.

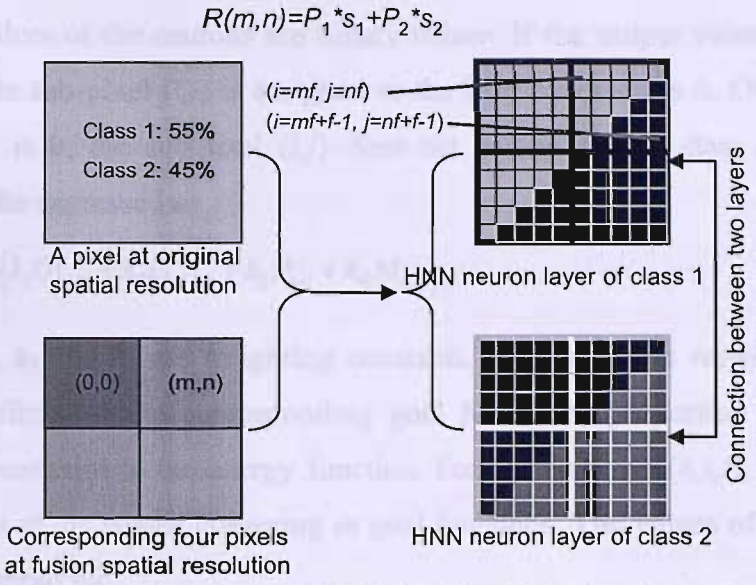


Figure 3.2 Reflectance constraint for sub-pixels covered by pixel (m,n) at the fused level.
A pixel at the original level contains four pixels at the fusion level. m, n are the coordinates of the fused pixel. $R(m,n)$ is the reflectance value of the fused pixel (m, n) and s_1 and s_2 are the end-member spectra of class 1 and class 2. P_1 and P_2 are average output values of the neurons of class 1 and 2 that are covered by the fused pixel (m, n). f is the fusion factor.

From the super-resolution map at the first iteration, an estimated MS image is then produced using a forward model. The estimated MS image is then compared with the fused image and a reflectance error image is determined. For all neurons covered by the same pixel in the fused image, a value based on the reflectance error image is produced to adjust the estimated MS image. Thus, the HNN is constrained by the reflectance values of the fused image. The adjustment value, or reflectance constraint, along with the goal and constraint values in the HNN structure proposed by Tatem *et. al.* (2001a), can be used in the optimisation process for super-resolution mapping by minimising the energy function. After the optimisation process, the estimated MS image produced by sub-pixels in the super-resolution map should resemble the fused image.

The method presented is based on the structure of the HNN proposed by Tatem *et al.* (2002a, 2002b). The structure of the HNN for super-resolution mapping of two land cover classes can be seen in Figure 3.2. A pixel at the original spatial resolution is divided into two inter-connected matrices of neurons in the HNN. Each neuron (h,i,j) represents a sub-pixel at position (i,j) in the land cover class h and each matrix of neurons represents a land cover class. The HNN is a recurrent neural network and it reaches a stable state when the energy function is minimised. For super-resolution mapping, the HNN is initialised using the soft-classified land cover proportions and

runs until it converges to a monotonic stable state (Tatem *et. al.*, 2001a). At this state, the output values of the neurons are binary values. If the output value of the neuron (h,i,j) is 1, the sub-pixel (i,j) is assigned to the land cover class h. Otherwise, if the output value is 0, the sub-pixel (i,j) does not belong to the class h. The energy function can be expressed as

$$E = -\sum_h \sum_i \sum_j \left(k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij} \right), \tag{3.1}$$

where k_1, k_2, k_3 and k_4 are weighting constants. Values of the weighting constants define the effects of the conresponding goal functions, proportion constraint and multi-class constraint to the energy function. For each neuron (h,i,j), $G1_{hij}$ and $G2_{hij}$ are the values of the spatial clustering or goal functions. The values of $G1_{hij}$ and $G2_{hij}$ can be determined by

$$\frac{dG1_{hij}}{dv_{hij}} = \frac{1}{2}\left(1 + \tanh\left(\frac{1}{8}\sum_{\substack{d=i-1 \\ d\neq i}}^{i+1}\sum_{\substack{e=j-1 \\ e\neq j}}^{j+1} v_{hde} - 0.5\right)\lambda\right)\left(v_{hij} - 1\right) \tag{3.2}$$

and

$$\frac{dG2_{hij}}{dv_{hij}} = \frac{1}{2}\left(1 + \left(-\tanh\left(\frac{1}{8}\sum_{\substack{d=i-1 \\ d\neq i}}^{i+1}\sum_{\substack{e=i-1 \\ e\neq j}}^{j+1} v_{hde}\right)\lambda\right)\right)\left(v_{hij}\right), \tag{3.3}$$

where λ is the gain or the steepness of the tanh function (usually assigned a value 100 [14]). $1/8 \equiv 1/N$, where N is the number of pixels in the neighbourhood used in the goal function, 0.5 is the threshold, and v_{hij} is the ouput value of the neuron (h,i,j). The first goal function (Equation (3.2)) is used to increase the output value v_{hij} of the neuron if the average output value of the eight surrounding neurons is greater than 0.5. In contrast, the second goal function (Equation (3.3)) decreases v_{hij} if the average output value of the eight surrounding neurons is less than 0.5.

The value P_{hij} in Equation (3.1) is the proportion constraint. This value retains the land cover proportion for each original pixel and is defined as

$$\frac{dP_{hij}}{dv_{hij}} = \frac{1}{2z^2}\sum_{d=xz}^{xz+z-1}\sum_{e=yz}^{yz+z-1}\left(1 + \tanh\left(v_{hde} - 0.5\right)\lambda\right) - a_{hxy}, \tag{3.4}$$

where $1/2z^2 \sum_{d=xz}^{xz+z-1}\sum_{e=yz}^{yz+z-1}\left(1 + \tanh\left(v_{hde} - 0.5\right)\lambda\right)$ is the estimated proportion and a_{hxy} is the input proportion of the land cover h of the pixel (x,y) which is obtained by soft classification. The pixel (x,y) is the corresponding pixel at the original spatial

resolution to which the sub-pixel or neuron (h,i,j) belongs. z is the zoom factor, which determines the increase in spatial resolution from the original image to the super-resolution mapping image. The proportion constraint function contributes a positive value if the estimated proportion of the class h is greater than the input proportion. As a result, the network reduces the output values of neurons within the pixel (x,y) in the class layer h. Conversely, if the estimated proportion is less than the input proportion, the proportion constraint produces a negative value to increase the output values of the neurons in the class h.

The multi-class value M_{hij} is used to reduce the output of the neurons if the sum of outputs of c classes at the position (i,j) is greater than 1. If the sum of outputs of c classes is less than 1, the function increases the output of the neurons at the position (i,j). The value of the multi-class constraint is calculated as

$$\frac{dM_{hij}}{dv_{hij}} = \left(\sum_{k=0}^{c} v_{kij} \right) - 1, \tag{3.5}$$

To use the fused image for super-resolution mapping by the HNN, the energy function in Equation (3.1) is modified by adding a reflectance constraint. In this experiment, a function based on the reflectance of the fused image is added to the goal functions and proportion constraint that comprise the energy function. The new energy function for each band of the fused image can be expressed in an equation as follows

$$E = -\sum_{h}\sum_{i}\sum_{j}\left(k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij} + k_5 R_{hij} \right), \tag{3.6}$$

where R_{hij} is the reflectance constraint value for each neuron (h,i,j).

The structure of the modified HNN can be seen in Figure 3.2. Each neuron in the HNN represents a sub-pixel point in the original spatial resolution image. The fusion factor f determines the increase in spatial resolution of the new super-resolution image in comparison with the fused image. Apart from the proportion constraint for each original pixel, $f \times f$ sub-pixels covered by pixel (m,n) of the fused image are constrained by a reflectance constraint. The reflectance constraint is based on the principle that the average predicted reflectance from all sub-pixels located within a pixel of the fused image should be equal to the observed reflectance (or target reflectance) of that pixel.

For each band of the fused image, there is an additional constraint for the energy function. The energy function is minimised if the derivatives of variables in Equation (3.6) converge to zero for each neuron (h,i,j)

$$\frac{dE_{hij}}{dv_{hij}} = k_1 \frac{dG1_{hij}}{dv_{hij}} + k_2 \frac{dG2_{hij}}{dv_{hij}} + k_3 \frac{dP_{hij}}{dv_{hij}} + k_4 \frac{dM_{hij}}{dv_{hij}} + k_5 \frac{dR_{hij}}{dv_{hij}}. \tag{3.7}$$

The derivative values of $G1$, $G2$, P, and M with respect to v_{hij} are computed using Equations (3.2), (3.3), (3.4), (3.5), respectively.

To derive the value dR_{hij}/dv_{hij}, the estimated reflectance R'_{smn} of the neurons representing the fused pixel (m,n) can be defined by a forward model using a linear mixture model (Settle and Drake,1993) as

$$R^t_{smn} = V_1 S_{s,1} + V_2 S_{s,2} + \ldots + V_c S_{s,c}, \tag{3.8}$$

where the estimated proportion value $V_e = 1/f^2(\sum_{p=mf}^{(mf+f-1)} \sum_{q=nf}^{(nf+f-1)} v_{epq})$ and $S_{s,e}$ is the end-member spectra of the land cover class e for a spectral band s.

Similarly, the observed reflectance for pixel (m,n) can be expressed using the same forward model as

$$R^o_{smn} = P_1 S_{s,1} + P_2 S_{s,2} + \ldots + P_c S_{s,c}, \tag{3.9}$$

where R^o_{smn} is the observed reflectance of pixel (m,n) in a fused image band, c is the total number of land cover classes, P_1, P_2,.., P_c are the proportions of a given land cover class covered by a pixel of the fused images ($\sum P=1$), and $S_{s,1}$, $S_{s,2}$,.., $S_{s,c}$ are end-member spectra of the land cover classes 1, 2,.., c in a MS fused band s.

From Equations (3.8) and (3.9), the difference between the observed and estimated reflectance of the neurons representing the fused pixel (m,n) is defined by

$$R^o - R^t = \sum_{e=1}^{c} P_e S_{s,e} - \sum_{e=1}^{c} V_e S_{s,e}, \tag{3.10}$$

The value of the reflectance constraint requires that the reflectance difference in Equation (3.10) should be zero for every spectral band. Accordingly, the estimated proportion value V_e is expected to converge to the land cover proportion P_e within each pixel of the fused image. Therefore, the value of the reflectance constraint for the fused pixel (m,n) can be produced based on an inverse model using the linear mixture model equation

$$\mathbf{R}^o_{mn} = \mathbf{SP}_{mn}, \tag{3.11}$$

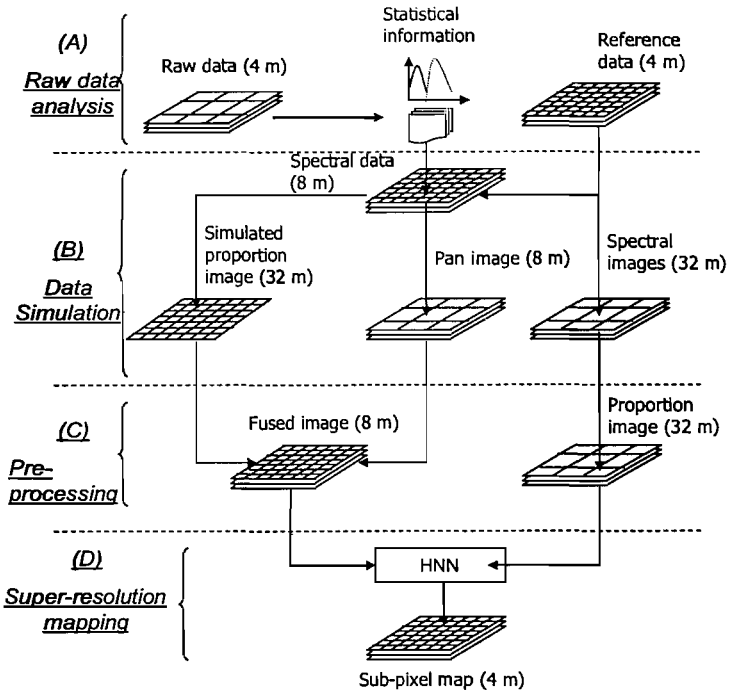


Figure 3.3 Four steps in experiment:
(A) Raw data analysis, (B)- Data simulation, (C)- Pre-processing, and (D)- Super-resolution mapping.

where $\mathbf{R}^o_{mn} = [R^o_1 \ .. \ R^o_i \ .. \ R^o_b]^T_{mn}$ is the vector of reflectance values of the fused spectral bands 1,.., s,.., b and

$$\mathbf{S} = \begin{bmatrix} S_{1,1} & \cdots & S_{1,c} \\ S_{s,1} & \cdots & S_{s,c} \\ S_{b,1} & \cdots & S_{b,c} \end{bmatrix} \tag{3.12}$$

is the matrix of end-member spectra values for b fused spectral bands and $\mathbf{P}$ is the vector of land cover proportion values.

Using the least squares linear mixture model, land cover class proportions of the fused pixel (m,n) can be predicted by

$$\mathbf{P}_{mn} = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{R}^o_{mn} = \mathbf{M}\mathbf{R}^o_{mn} \text{ with } \mathbf{M} = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T. \tag{3.13}$$

From Equation (3.13), the values for reflectance constraints for all fused spectral bands are calculated from the difference between observed and estimated reflectance of b fused multispectral bands as

$$\Delta = \begin{bmatrix} dR/dv_{C1ij} \\ \cdots \\ dR/dv_{Ccij} \end{bmatrix} = \begin{bmatrix} P_1 \\ \cdots \\ P_c \end{bmatrix} - \begin{bmatrix} V_1 \\ \cdots \\ V_c \end{bmatrix} = \mathbf{M}(\mathbf{R}^o_{mn} - \mathbf{R}^t_{mn}) = \mathbf{M}\begin{bmatrix} R^o_{1mn} - R^t_{1mn} \\ \cdots \\ R^o_{bmn} - R^t_{bmn} \end{bmatrix} \tag{3.14}$$

If the number of fused image spectral bands is less than the number of land cover classes, the value of the reflectance constraint cannot be known.

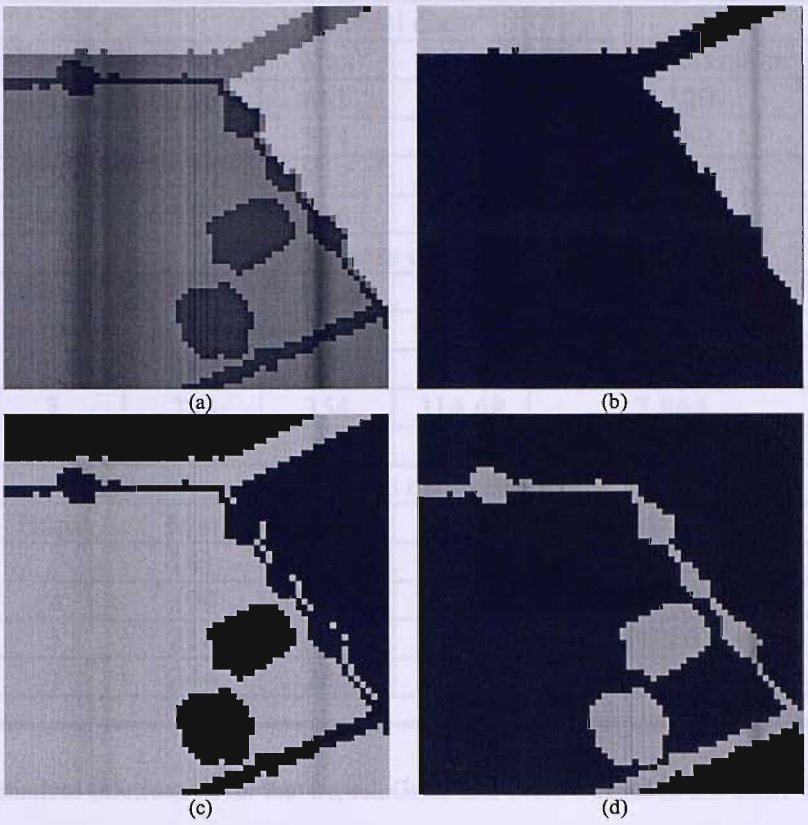


Figure 3.4 (a) Land cover map at 4 m spatial resolution used for simulating data, (b) 4 m cereal class map, (c) 4 m grass class map, (d) 4 m trees class map.

3.3 Experiment 1: Simulated IKONOS data

3.3.1 Data

In this experiment, a set of data based on IKONOS imagery was used. The set of data is a simulated 32 m MS image and an 8 m fused spectral image created by degrading a real IKONOS image. The ratio between the spatial resolution of the simulated MS and fused images is similar to the ratio between the real 4 m MS and 1 m panchromatic (PAN) IKONOS images. Thus, the algorithm should be applicable to the real imagery (i.e. 4 m MS and 1 m fused image) if it performs adequately on the simulated imagery. There are two reasons for using simulated imagery. Firstly, it was possible to evaluate the quality of the fused imagery. Secondly, the simulation ensured that there were no errors in image registration between the reference image and the land cover image obtained by super-resolution mapping. For the simulated IKONOS image, the experiment was implemented in four steps as follows: (A) Raw data analysis,

Cereal class				
Band	Min	Max	Mean	Standard deviation
1	236	461	386.99	27.520
2	336	574	491.68	27.282
3	275	432	384.64	18.565
4	257	346	322.75	9.993
Grass class				
Band	Min	Max	Mean	Standard deviation
1	220	325	259.64	12.303
2	377	623	492.91	29.851
3	283	354	314.68	7.864
4	271	309	289.37	4.618
Trees class				
Band	Min	Max	Mean	Standard deviation
1	141	238	183.98	25.127
2	194	713	421.70	105.667
3	218	311	259.71	20.061
4	244	286	264.16	8.887

Table 3.1 Statistical Information for the Cereal, Grass and Trees Classes in the IKONOS Image

(B) Data simulation, (C) Pre-processing, and (D) Super-resolution mapping (Figure 3.3).

3.3.1.1 Raw data analysis

Raw data: An IKONOS MS image was acquired over Eastleigh and Chandler Ford, Southampton, UK in 25th August 2000. The IKONOS image consisted of four 4 m MS bands in the following wavebands: Red (632-698 nm), Near-Infra Red (NIR: 757-853 nm), Green (506-595 nm) and Blue (445-516 nm) and a 1 m PAN band (450-900 nm). Based on image fusion, the MS and PAN bands can be used to produce four fused MS image bands at 1 m spatial resolution.

Reference data and statistical information: The experiment was implemented in an area of 64×64 pixels (at 4 m spatial resolution) that consisted of three land cover classes: cereal, grass, and trees (Figure 3.4(a), 3.4(b), 3.4(c) and 3.4(d) (In the experiment, several testing sites were chosen. However, the area of 64×64 pixels used above was the largest site that contained three land cover classes). These three land cover classes were produced using maximum likelihood classification of the real IKONOS image. Statistical information such as the means and standard deviations of the three land cover classes in the area was obtained (Table 3.1 and Figure 3.5). The

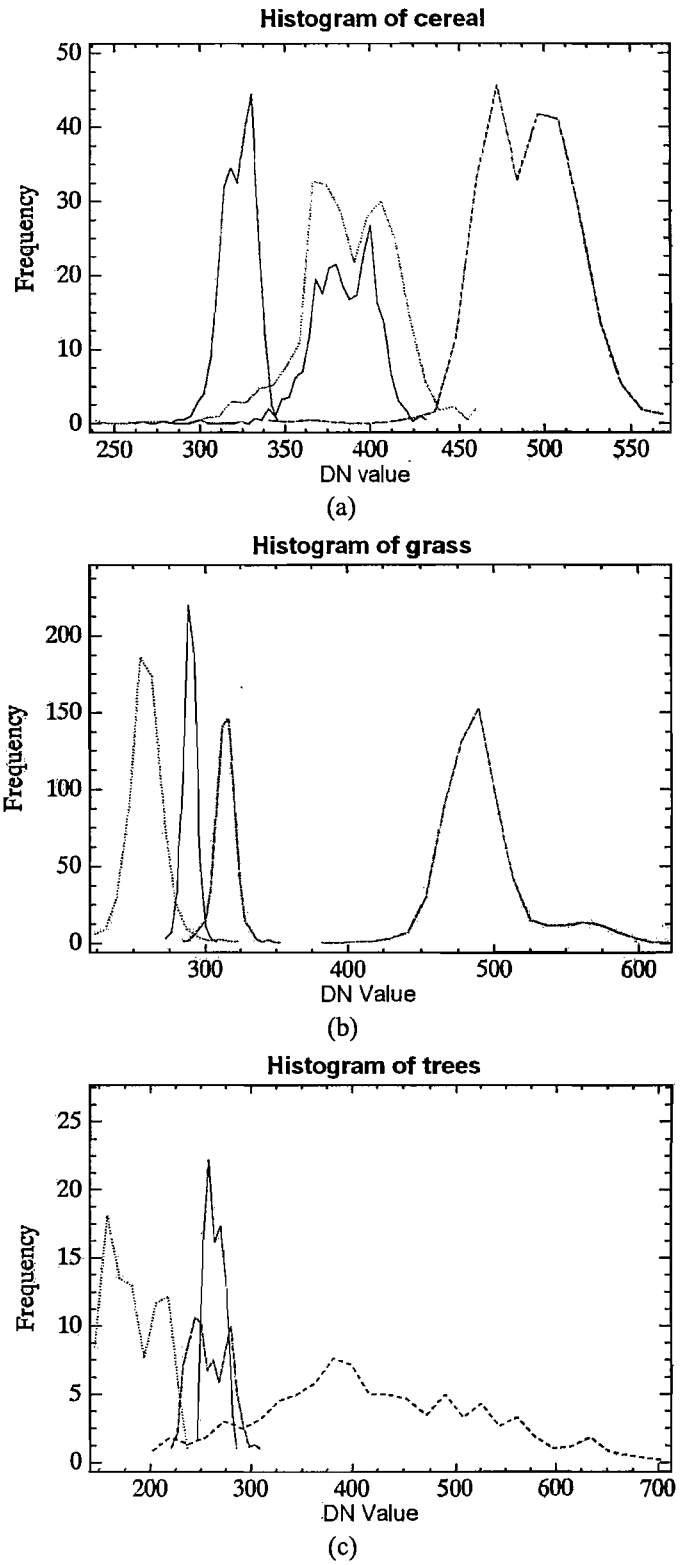


Figure 3.5 Histogram of three classes in four bands of IKONOS MS image:
(a) Cereal; (b) Grass and (c) Trees. Solid line-Blue Band; Dash line-Green Band; Dot line-Red Band; Dot and dash line-NIR Band.

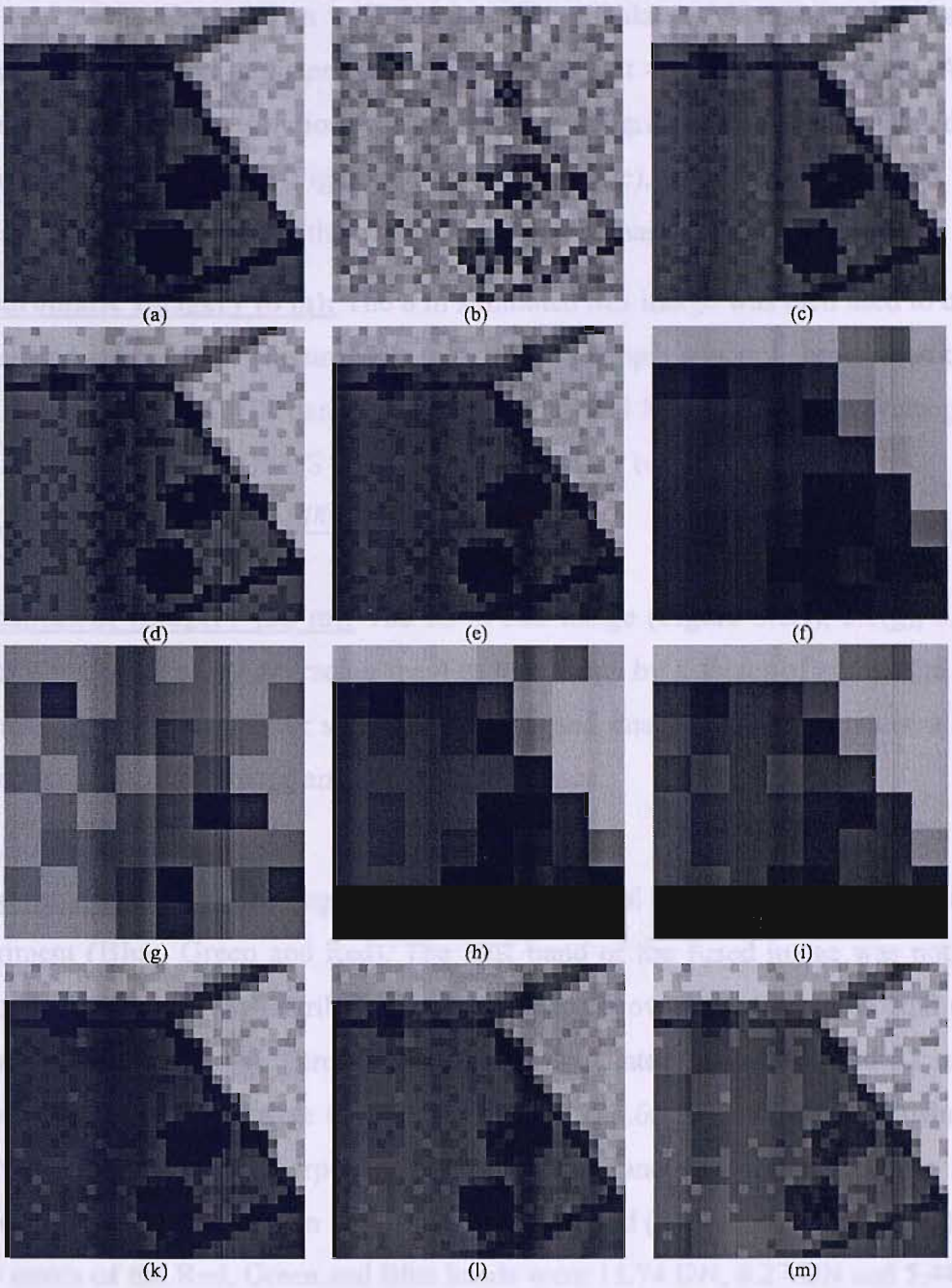


Figure 3.6 Simulated IKONOS image:
Four bands (a) Red, (b) NIR, (c) Green, and (d) Blue 8 m of simulated MS IKONOS image. (e) 8 m simulated PAN image. Four bands (f) Red, (g) NIR, (h) Green, and (i) Blue of 32 m simulated image. Three bands (k) Red, (l) Green, and (m) Blue of 8 m simulation of the fused MS image.

three land cover classes were used as a reference for the sub-pixel map obtained by the proposed algorithm.

3.3.1.2 Data simulation

Multispectral imagery (8 m): From the land cover map (Figure 3.4(b), 3.4(c), 3.4(d)) at 4 m spatial resolution, a set of multispectral images at 4 m spatial resolution was simulated based on the random normal distribution using the mean and variance of

each land cover class shown in Table 3.1. The simulated MS image, therefore, is similar spectrally to a multispectral IKONOS image at 4 m spatial resolution. A MS image at 8 m spatial resolution was generated by degrading the 4 m simulated MS image by a factor of two (Figure 3.6(a), 3.6(b), 3.6(c), 3.6(d)). These images were used as reference to evaluate the quality of the fused image produced below.

Panchromatic imagery (8 m): The 8 m simulated MS image was then used to create a simulated PAN image (Figure 3.6(e)) based on a simple spectral convolution of the Blue, Green, Red and NIR bands of the 8 m simulated MS image (the wavelength of the PAN band of the IKONOS image covers these four bands) as

$$Pan = \frac{BLUE + GREEN + RED + NIR}{4}. \tag{3.15}$$

Multispectral imagery (32 m): The 32 m MS image (Figure 3.6(f), 3.6(g), 3.6(h), 3.6(i)) was produced by degrading the 4 m MS image by a factor of eight. The 32 m MS image was then used for soft-classification and image fusion to produce a 32 m land cover proportion image and an 8 m fused image.

3.3.1.3 Pre-processing

Fused imagery (8 m): Amongst the four fused spectral bands, three were used in the experiment (Blue, Green and Red). The NIR band of the fused image was not used because of the scattered distribution and the spectral overlap of all three land cover classes over this band (Figure 3.5). From the simulated 8 m PAN and 32 m MS images, the 8 m fused image (Figure 3.6(k), 3.6(l), 3.6(m)) was predicted using the Gram-Schmidt Spectral Sharpening method (Laben and Brower, 2000). The fused image was evaluated based on the RMSE for each band (Munechika *et al.*, 1993). The RMS errors of the Red, Green and Blue bands were 15.74 DN, 8.27 DN and 5.44 DN, respectively. Comparing with the RMS errors of the fused image obtained in Munechika *et al.* (1993), the fused image produced was similar to that of real data.

Simulated land cover proportion imagery (8 m): To provide a realistic test, a set of proportion images was produced using soft-classification of the simulated 32 m MS image. The simulated MS image was used because in the simulated case the three land cover classes at the sub-pixel (4 m) level are known, facilitating direct evaluation of the technique. A *k*-nearest neighbour classifier (*k*-NN) (Lewis and Brown, 1998) was used for soft-classification with k=5. The land cover proportion image was produced

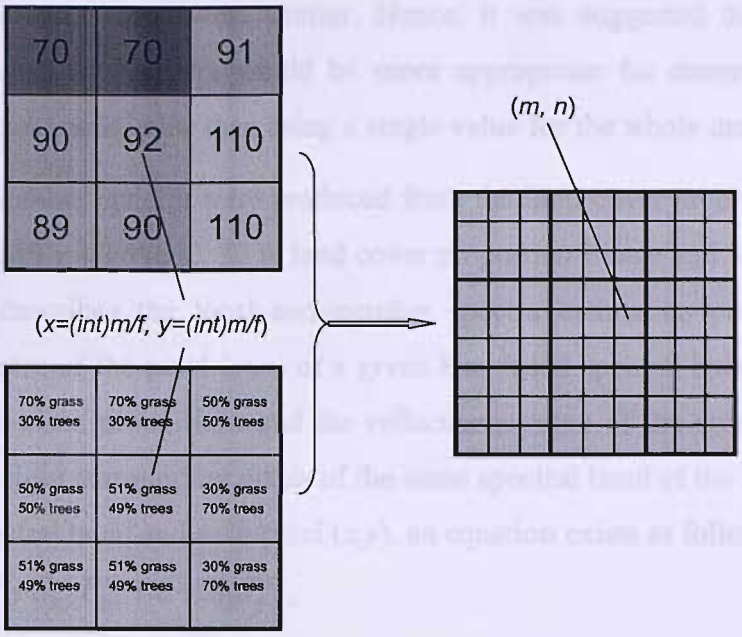


Figure 3.7 Local end-member spectra calculation:
(m,n) are coordinates of the fused image pixel and (x,y) are coordinates of the pixel in the original image that corresponds to the fused pixel (m,n). From land cover proportion and digital number of pixel (x,y) and its eight surrounding pixels, the local spectra of the pixel (m,n) can be calculated.

with overall area error proportion of 0.5552% (Lewis and Brown, 2001) and overall RMS error of 0.0838 pixels (Tatem *et al.*, 2001a). Statistics for the resulting land cover map from soft-classification show that the land cover proportion images contained an amount of error similar to that of a soft-classified real MS image. In this sense, the simulated land cover proportion image was similar to that which might be obtained from real data.

3.3.1.4 Global end-member spectra

The end-member spectra values in Equation (3.7) can be acquired from laboratory measurements or can be estimated from training data (Settle and Drake, 1993). Since the simulated MS image was created from statistics on the cereal, grass and trees classes, the end-member spectra of these classes should be the means of the spectral distributions in the Blue, Green and Red bands.

3.3.1.5 Local end-member spectra

Three land cover classes exhibited a large variance and multimodal distributions over all four spectral bands (Table 3.1 and Figure 3.5). Thus, the single set of end-member spectra values used in Equation 2 was not appropriate for every pixel in the image. Investigation of the real IKONOS image indicated that the DN values of adjacent

pixels of the same class were similar. Hence, it was suggested that using locally defined end-member spectra would be more appropriate for determining the local reflectance constraint value than using a single value for the whole image.

Local end-member spectra were produced from the land cover proportion image and the original MS image (e.g., 32 m land cover proportion image and 32 m MS image). Figure 3.7 describes the local end-member spectra estimation process. The end-member spectra of the pixel (m,n) of a given 8 m fused spectral band can be defined based on the class proportions and the reflectance value of the corresponding pixel (x,y) and its eight surrounding pixels of the same spectral band of the 32 m MS image. For each spectral band and each pixel (x,y), an equation exists as follows

$$R_s^{xy} = S_{s,1}P_1^{xy} + S_{s,2}P_2^{xy} + ... + S_{s,c}P_c^{xy}, \tag{3.16}$$

where R_s^{xy} is the DN value of pixel (x,y) in spectral band i, P_1^{xy}, P_2^{xy},..., P_c^{xy} are class proportions and $S_{s,1}$, $S_{s,2}$,.., $S_{s,c}$ are the local end-member spectra of the pixel (x, y) in spectral band i.

With eight surrounding pixels, there are eight equations which can be rewritten in matrix form as

$$\mathbf{R}_s^{xy} = \mathbf{P}\mathbf{S}_s, \tag{3.17}$$

where $R_s^{xy} = [R_s^{(x-1)(y-1)} \ \ldots \ R_s^{(x+1)(y+1)}]^T$, $S_s^{xy} = [S_{s,1} \ \ldots \ S_{s,c}]$, and

$$\mathbf{P} = \begin{bmatrix} P_1^{(x-1)(y-1)} & .. & P_c^{(x-1)(y-1)} \\ .. & .. & .. \\ P_1^{xy} & .. & P_c^{xy} \\ .. & .. & .. \\ P_1^{(x+1)(y+1)} & .. & P_c^{(x+1)(y+1)} \end{bmatrix}.$$

Using the least squares method, the local end-member spectra S_s can be resolved as

$$\mathbf{S}_s = \left(\mathbf{P}^T\mathbf{P}\right)^{-1}\mathbf{P}^T\mathbf{R}_s. \tag{3.18}$$

Amongst the pixels that are used to determine the local end-member spectra, the pixel (x,y) should be the most important since it covers the fused pixel (m,n). To emphasise the contribution of the corresponding pixel (x,y) to the end-member spectra, a weight mechanism was used such that Equation (3.19) becomes

$$\mathbf{S}_s = \left(\mathbf{P}^T\mathbf{W}\mathbf{P}\right)^{-1}\mathbf{W}\mathbf{P}^T\mathbf{R}_s, \tag{3.19}$$

where $\mathbf{W}$ was the diagonal matrix:

$$\mathbf{W}=\begin{bmatrix} w^{(x-1)(y-1)} & 0 & 0 & 0 & 0 \\ 0 & .. & 0 & 0 & 0 \\ 0 & 0 & w^{xy} & 0 & 0 \\ 0 & 0 & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & w^{(x+1)(y+1)} \end{bmatrix} \tag{3.20}$$

and $w^{(x-1)(y-1)},..,w^{xy},..,\ w^{(x+1)(y+1)}$ are weight values for each corresponding pixel. The assumption was tested using the weight value w^{xy} of 1 up to 20 and the other weight values of 1. The optimal weight value w^{xy} was determined based on the KIA of results of the super-resolution mapping using the fused image.

3.3.2 Results and discussion

3.3.2.1 Network settings

Two sources of data were used in super-resolution mapping using fused imagery. The first data source was the land cover proportion image obtained by soft-classification. The second data source was the fused image. In the experiment using the simulated land cover proportion image, both data sources contained an amount of error similar to that of real data.

From the predicted soft-classified land cover proportion image (Figure 3.8(a), 3.8(b), and 3.8(c)), the 4 m sub-pixel land cover maps were obtained using the traditional HNN (Figure 3.8(g), 3.8(h), and 3.8(i)), the HNN using the fused image with global end-member spectra (Figure (j), 3.8(k) and 3.8(l)) and the HNN using the fused image with local end-member spectra (Figure 3.8(m), 3.8(n) and 3.8(o)). The greatest accuracy land cover map was obtained with the weighting coefficients of k_1=70, k_2=70, k_3=70, k_4=70 and k_5=70 after 6000 iterations and the optimal weight value of 14 to determine the local spectra. The choice of weighting coefficients was implemented empirically after testing a number of sets of weighting coefficient values. The 32 m hard classified land cover image (Figure 3.8(d), 3.8(e) and 3.8(f)) was produced from the 32 m multispectral image (Figure 3.8(g), 3.8(h), 3.8(i), and 3.8(j)) based on the highest land cover proportions estimated by the k-NN.

3.3.2.2 Accuracy assessment

The accuracy assessment was implemented at the sub-pixel spatial resolution based on the statistics which are mostly used for hard classification such as KIA, overall accuracy, and per-class omission and commission errors. Although Foody (1992)

Statistics for the hard classified image					
	Cereal	Grass	Trees	ErrorO (%)	ErrorC (%)
Unclassified	0	0	0		
Cereal	1005	126	21	3.74	12.76 %
Grass	39	2305	280	8.42	12.16 %
Trees	0	86	234	56.26	26.88 %
KIA – κ =		0.7430	Overall accuracy =		86.52 %
Statistics for the HNN super-resolution mapping without using the fused image					
	Cereal	Grass	Trees	ErrorO (%)	ErrorC (%)
Unclassified	0	4	3		1.000
Cereal	977	89	2	4.60	8.52 %
Grass	63	2350	231	8.14	11.12 %
Trees	4	74	299	36.82	20.69 %
KIA – κ =		0.7814	Overall accuracy =		88.53 %
Statistics for the HNN super-resolution mapping using the fused image (Global spectra)					
	Cereal	Grass	Trees	ErrorO (%)	ErrorC (%)
Unclassified	11	4	2		1.0000
Cereal	931	39	0	10.82	4.02
Grass	91	2346	209	6.79	11.34
Trees	11	128	324	39.44	30.02
KIA – κ =		0.7897	Overall accuracy =		87.92 %
Statistics for the HNN super-resolution mapping using the fused image (Local spectra value of 14)					
	Cereal	Grass	Trees	ErrorO (%)	ErrorC (%)
Unclassified	6	0	0		1.000
Cereal	994	39	0	4.79	3.68
Grass	38	2412	206	4.17	9.19
Trees	6	67	329	38.50	18.16
KIA – κ =		0.8320	Overall accuracy =		91.19 %

Table 3.2 Confusion matrices and Accuracy Statistics of Simulated IKONOS Results

suggested that the KIA is overestimated and the overall accuracy is underestimated, these statistics were used because they still represented for the accuracy of the hard land cover maps super-resolution mapping results. Accuracy statistics for each class based on KIA, overall accuracy, and per-class omission and commission errors are presented in Table 3.2.

3.3.2.3 Visual evaluation

The map produced by the new HNN super-resolution technique with real proportion image data was more accurate than the hard classification and traditional HNN in retaining small and linear objects. Despite the effect caused by the error in the class proportion image, the linear features in the trees class were recreated by the new HNN

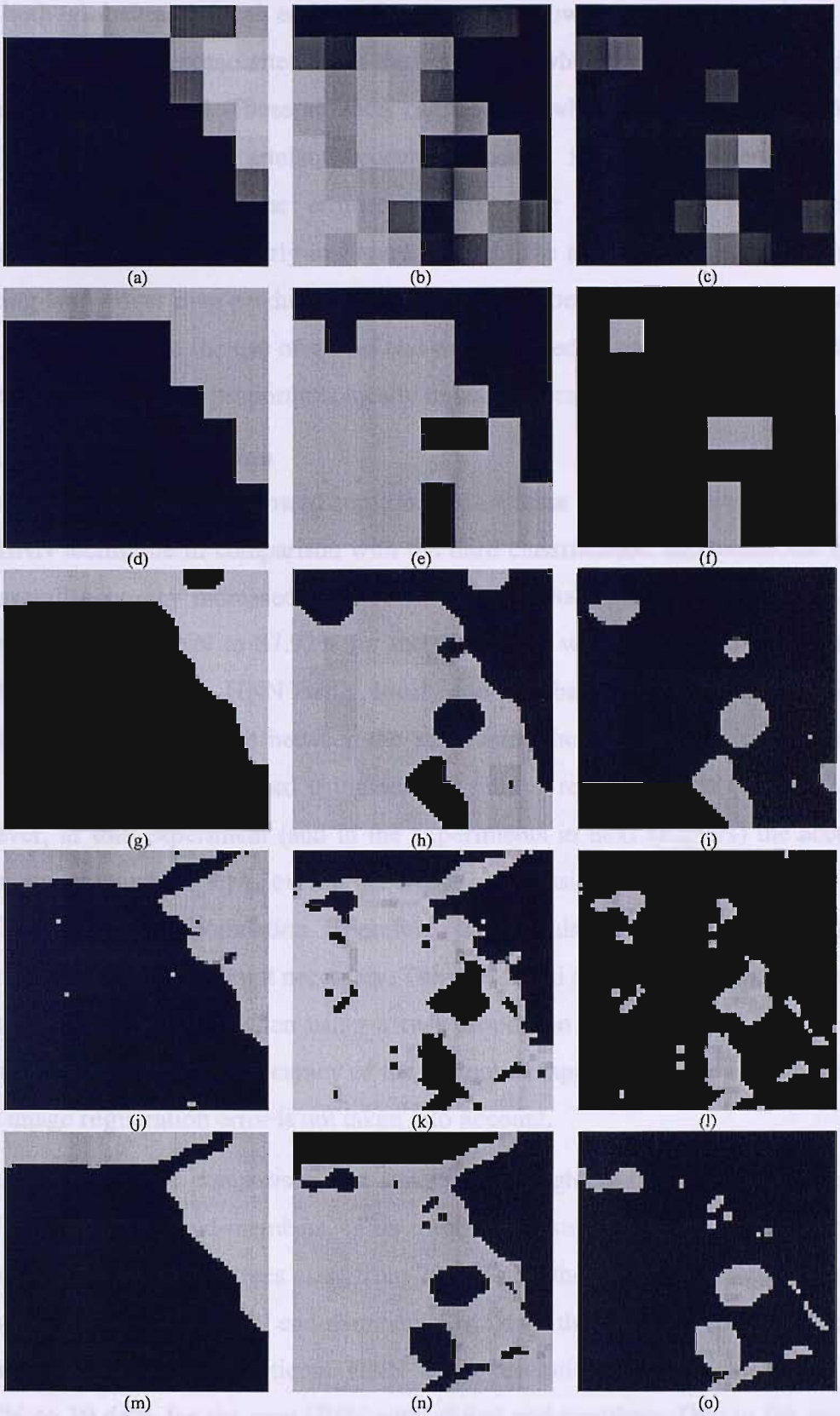


Figure 3.8 Results for simulated IKONOS image:
4 m Cereal (a), Grass (b), and Trees (c) land cover proportion image. 4 m Cereal (d), Grass (e), Trees (f) hard classified land cover image. 4 m Cereal (g), Grass (h), Trees (i) HNN super-resolution mapping image. 4 m Cereal (j), Grass (k), Trees (i) HNN super-resolution mapping using the fused image with the global end-member spectra resulting image. 4 m Cereal (m), Grass (n), Trees (o) HNN super-resolution mapping using the fused image with the local end-member spectra resulting image.

with both global and local end-member spectra. However, the errors from soft-classification caused some artefacts in the trees class when the new HNN with global end-members was used. These artefacts can be seen when comparing Figure 3.4(d) and Figure 3.8(l). The artefacts occurred mostly for pixels where the soft-classification predicted some erroneous land cover proportions (the erroneous proportions can be seen clearly in Figure 3.8(b)). The artefacts did not occur in the resulting land cover map produced by the HNN with local end-members. This can be explained as follows: the use of global end-members led to an inability to reduce the effects of the erroneous proportions locally in some cases.

3.3.2.4 Statistical evaluation

The statistics in Table 3.2 showed considerable increase in all accuracy values for the new HNN technique in comparison with the hard classification and traditional HNN. The overall accuracy increased from 86.52% for the hard classification and 88.53% for the traditional HNN to 87.92% for the new HNN with global end-members and 91.19% for the new HNN with local end-members. Usually, the statistical significance of a difference between the accuracies should be taken into account if these accuracies are calculated using sample sets of reference data (Foody, 2004). However, in this experiment (and in the experiments in next chapters) the accuracy measurements such as KIA, overall accuracy, commission and omission errors were calculated using the population. Therefore, the calculation for significance of the difference in accuracies is not necessary. The visual and statistical improvement of the resulting sub-pixel maps when using a real proportion image showed that the new algorithm can increase the accuracy of the thematic mapping with the real image data if the image registration error is not taken into account.

Similar to the visual comparison, the statistics highlighted a problem with the new HNN with global end-members. This problem resulted in an increase in the commission error for the trees class from 20.69% for the traditional HNN to 30.02% for the new HNN with global end-members. Similarly, the omission error obtained for the trees class by the traditional HNN super-resolution mapping increased from 36.82% to 39.44% for the new HNN with global end-members. Due to the problem with the erroneous proportion, the overall accuracy of the new HNN with global end-members decreased just slightly in comparison with the map obtained with the HNN super-resolution mapping, with the overall accuracy decreasing from 88.52% to

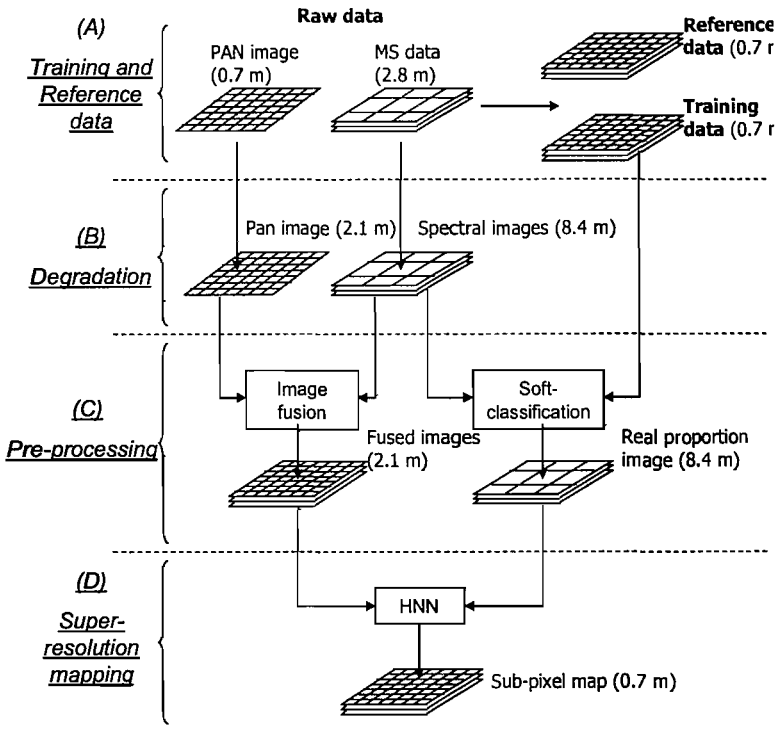


Figure 3.9 . Four steps in experiment:

(A) Training and reference data, (B)- Degradation, (C)- Pre-processing, and (D)- Super-resolution mapping.

87.92%. With local end-members, the HNN using the fused image can resolve the problem that occurred when using a single set of global end-members. The overall accuracy of the resulting map produced by the new HNN with the local end-members increased greatly by approximately 3% to 91.19%.

3.4 Experiment 2: Degraded QuickBird image

3.4.1 Data

Although simulated imagery provides greater control than real imagery for evaluating new algorithms, a common criticism is that simulated imagery may not provide a realistic test, primarily because image registration error is not included in the data. Therefore, to provide a more realistic test, and to address such concerns, a second set of proportion images was produced using a degraded QuickBird MS and PAN image. A degraded (rather than real) image was used because in the degraded case the three land cover classes at the subpixel level are known, facilitating direct evaluation of the impact of image registration error on the technique. The experiment process can be depicted in Figure 3.9 with four steps: (A) Training and reference data, (B) Degradation, (C) Pre-processing and (D) Super-resolution mapping.

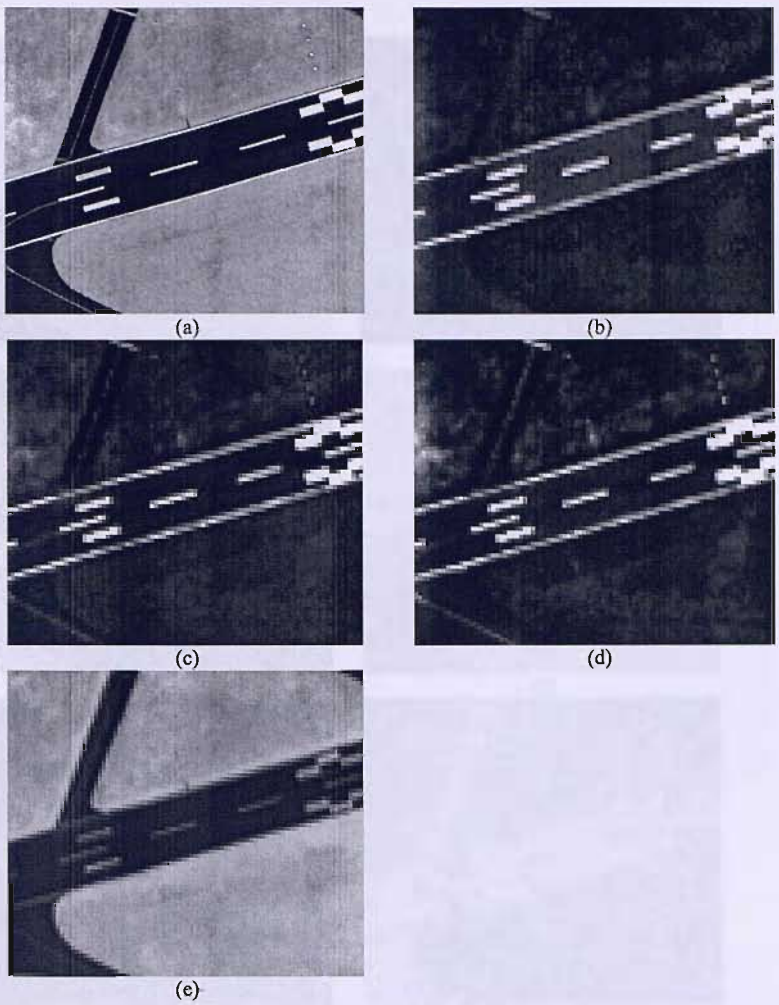


Figure 3.10 QuickBird MS and PAN image
(a) 0.7 m PAN image (360×300 pixel), (b) Blue, (c) Green, (d) Red and (e) Near IR bands of 2.8 m MS image (90×75) pixel.

3.4.1.1 Training and reference data

Raw data: A QuickBird MS image was acquired over an area of Christchurch, UK on 1st June 2002. A sub-area of PAN (360×300 pixel) (Figure 3.10(a)) and MS (90×75 pixel) (Figure 3.10(b), 3.10(c), 3.10(d), 3.10(e)) images was extracted from an area in the airport. Three land cover classes in the area were: grass, white surface and dark surface of the runway. The MS image was co-registered to the PAN image with a RMSE of 0.25 pixel.

Training and Reference data: Three land cover classes at 0.7 m spatial resolution were obtained by manual digitising from the panchromatic image (Figure 3.11(a), 3.11(b), 3.11(c)). These land cover images were used as the reference data for the results of super-resolution mapping. Thus, there was no image registration error of the PAN image against the reference data.

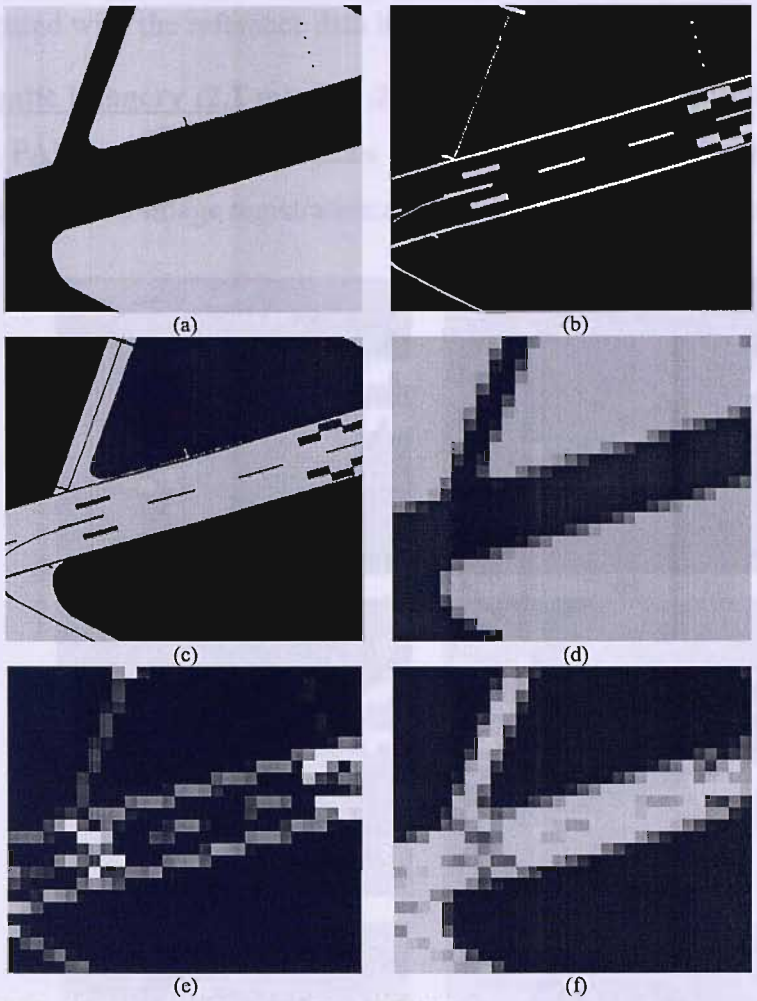


Figure 3.11 Three land cover classes image for reference:
(a) 0.7 m grass, (b) 0.7 m white surface and (c) 0.7 m dark surface. Three land cover classes image for training: (d) 8.4 m grass, (e) 8.4 m white surface and (f) 8.4 m dark surface.

Super-resolution methods use land cover proportions obtained by soft-classification as input. To implement the soft-classification, training data are required. In this research, the soft-classification was implemented at 8.4 m spatial resolution. The training data, therefore, could be produced by degrading the land cover image at 0.7 m spatial resolution by twelve times as in Figure 3.11(d), 3.11(e), and 3.11(f).

3.4.1.2 Image degradation

Multispectral imagery (8.4 m): The QuickBird MS image at 2.8 m spatial resolution was degraded by three times to produce a MS image at 8.4 m spatial resolution (Figure 3.12(a), 3.12(b), 3.12(c) and 3.12(d)). This MS image was then used to produce the land cover proportions at 8.4 m spatial resolution using soft-classification with training data in Figure 3.11. The land cover proportions were then used to

produce a 0.7 m land cover image using super-resolution mapping and the results were compared with the reference data in Figure 3.11.

Panchromatic imagery (2.1 m): The 2.1 m PAN image was produced by degrading the 0.7 m PAN image by three times (Figure 3.12(e)). The PAN image in Figure 3.12(e) contained no image registration error. To evaluate the effect of the image

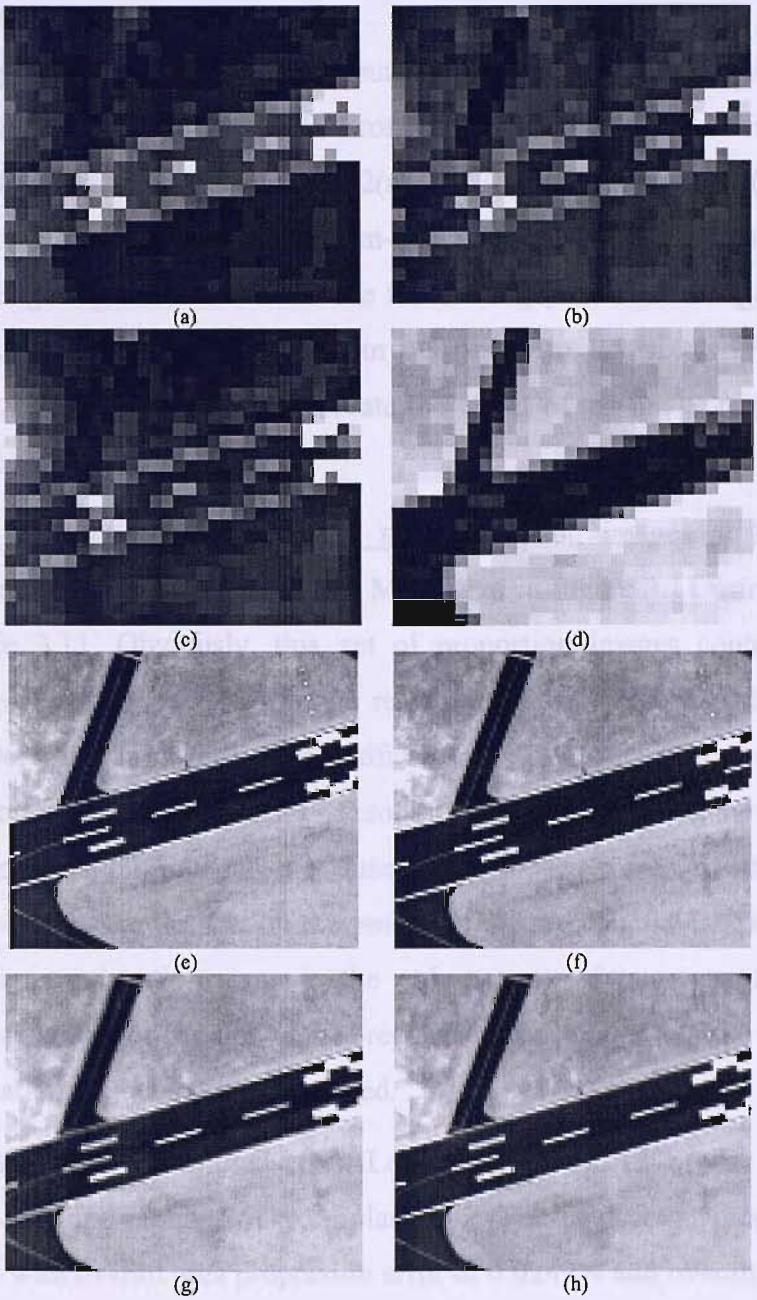


Figure 3.12 Degraded MS and PAN images
Four bands (a) Blue, (b) Green, (c) Red, and (d) NIR 8.4 m of simulated MS image. (e) 2.1 m degraded PAN image, (f) 0.5 pixel image registration error PAN image at 2.1 m spatial resolution, (g) 1 pixel image registration error PAN image at 2.1 m spatial resolution and 1.5 pixel image registration error PAN image at 2.1 m spatial resolution.

registration error on the accuracy of the resulting land cover map, the PAN image was geo-coded with RMSEs of 0.5 pixels (Figure 3.12(f)), 1 pixel (Figure 3.12(g)) and 1.5 pixels (Figure 3.12(h)). The proposed algorithm was then tested using the fused images obtained from these geocoded PAN images and the results were compared with the results obtained from the fused images without image registration error.

3.4.1.3 Pre-processing

Fused imagery (2.1 m): Amongst the four fused spectral bands, three were used in the experiment (Blue, Green and Red). From the 2.1 m degraded PAN image (Figure 3.12(e)) and 8.4 m MS image (Figure 3.12(a), 3.12(b), 3.12(c) and 3.12(d)), the 2.1 m fused image was obtained using the Gram-Schmidt Spectral Sharpening method. To include the image registration error in the fused image, the PAN images with image registration RMS error in the range from 0.5 pixels to 1.5 pixels (Figure 3.12(f), 3.12(g) and 3.12(h)) were fused to evaluate the effect of the image registration error on the algorithms.

Land cover proportion imagery (8.4 m): Proportion images at 8.4 m spatial resolution were produced from the 8.4 m MS image in Figure 3.11 using the training data in Figure 3.11. Obviously, this set of proportion images contains a certain amount of error including the MS image registration error. The predicted land cover proportions were then used for hard classification to produce a land cover map at 8.4 m spatial resolution, traditional super-resolution mapping and the new method for super-resolution mapping using the fused image. The error of the proportion predicting process has an impact on the results of all these methods. That means that the PAN image registration error is the only source of image registration error affecting the accuracy of the new super-resolution mapping when the results of the three classification methods were compared.

A *k*-nearest neighbour classifier (*k*-NN) (Lewis and Brown, 1998) was used for soft-classification with the number of exemplars k=5. The land cover proportion image was produced with overall area proportion error of 0.0242% and overall RMS error of 0.0182 pixels (Tatem *et al.*, 2001a). The proportions of three land cover classes can be seen in Figure 3.13(a), 3.13(b) and 3.13(c).

Statistics for the hard classified image					
	Grass	W surface	D surface	ErrorO (%)	ErrorC (%)
Unclassified	0	0	0		
Grass	72945	749	1474	1.47	2.96
W surface	44	598	222	85.31	30.79
D surface	1043	2725	28200	5.67	11.79
KIA – κ =		0.8685	Overall accuracy =		94.21
Statistics for the HNN super-resolution mapping without using the fused image					
	Grass	W surface	D surface	ErrorO (%)	ErrorC (%)
Unclassified	4	11	14		
Grass	73746	424	280	0.39	0.95
W surface	152	1414	786	65.28	39.88
D surface	130	2223	28816	3.61	7.55
KIA – κ =		0.9166	Overall accuracy =		96.27
Statistics for the HNN super-resolution mapping using the fused image without image registration error (Local spectra value of 14)					
	Grass	W surface	D surface	ErrorO (%)	ErrorC (%)
Unclassified	2	12	16		
Grass	73755	1006	331	0.37	1.78
W surface	2002	1861	212	54.30	18.20
D surface	73	1193	29337	1.87	4.14
KIA – κ =		0.9365	Overall accuracy =		97.18
Statistics for the HNN super-resolution mapping using the fused image with PAN image registration error of 1 pixel (Local spectra value of 14)					
	Grass	W surface	D surface	ErrorO (%)	ErrorC (%)
Unclassified	3	13	32		
Grass	73525	972	477	0.68	1.93
W surface	275	1716	391	57.86	27.96
D surface	229	1371	28996	3.01	5.23
KIA – κ =		0.9217	Overall accuracy =		96.52

Table 3.3 Confusion Matrices and Accuracy Statistics of Degraded QuickBird Results

3.4.2 Results and discussion

3.4.2.1 Network settings and accuracy assessment

In the experiment using degraded QuickBird imagery, the results produced by the three approaches were compared, as in the simulated data case. The 8.4 m hard classified land cover map was obtained from the land cover proportion image (Figure 3.13(a), 3.13(b), 3.13(c)) by assigning each 8.4 m pixel to the class of the largest proportion (Figure 3.13(d), 3.13(e), 3.13(f)). The HNN super-resolution mapping by

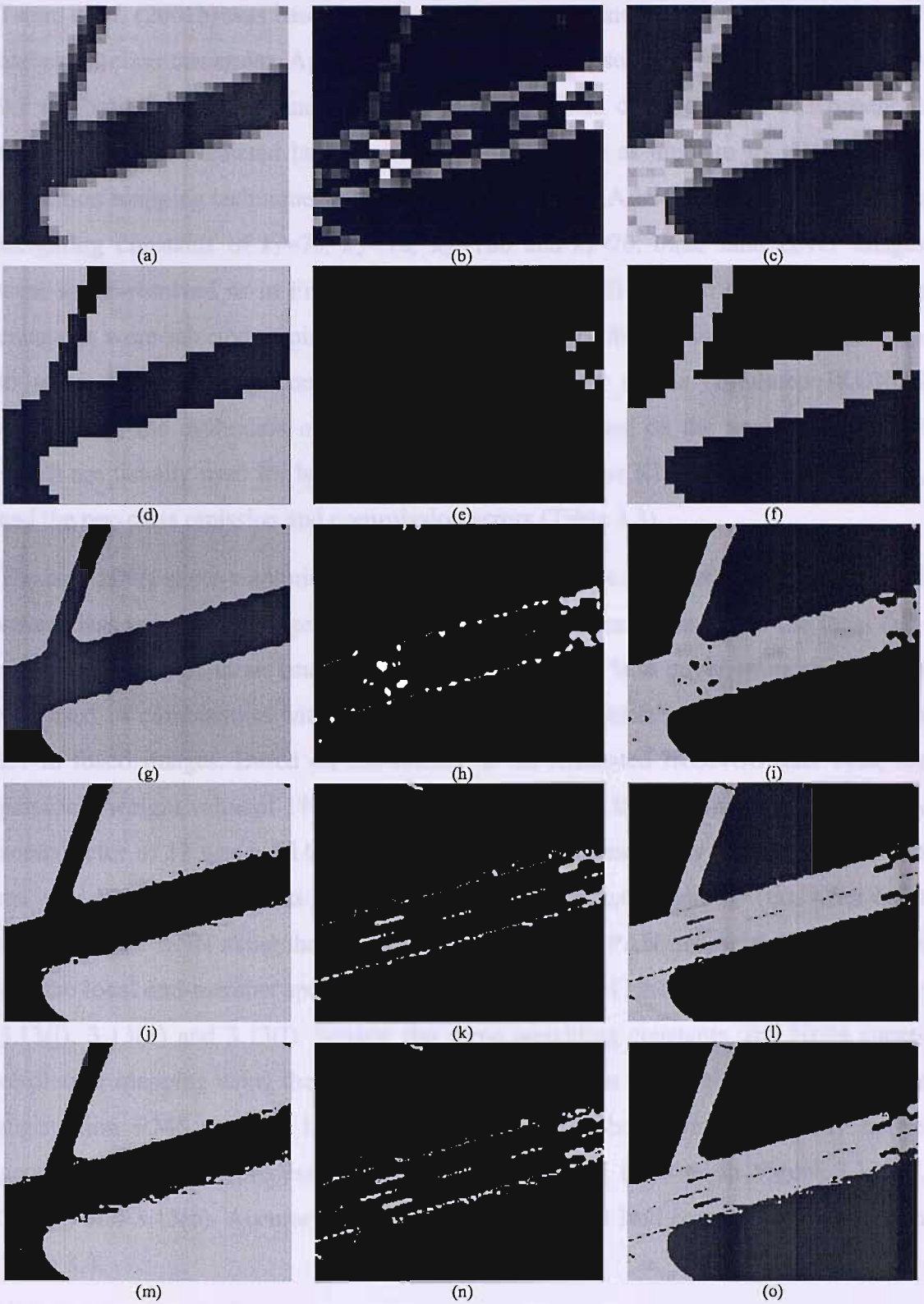


Figure 3.13 Results for the Degraded QuickBird image:
(a) Grass, (b) White surface, and (c) Dark surface 8.4 m land cover proportion image. (d) Grass, (e) White surface, and (f) Dark surface 8.4 m hard classified land cover image. (g) Grass , (h) White surface and (i) Dark surface 0.7 m HNN super-resolution mapping image. (j) Grass, (k) White surface, and (i) Dark surface 0.7 m HNN super-resolution mapping using the fused image without image registration error. (m) Grass, (n) White surface and (o) Dark surface 0.7 m HNN super-resolution mapping using the fused image with RMS image registration error of 1 pixel.

Tatem *et al.* (2001b) was based on the clustering goal functions, proportion constraint and multi-class constraint. Applying this HNN super-resolution mapping approach, a 0.7 m spatial resolution land cover map of three land cover classes was produced using the 8.4 m predicted land cover proportion image as input to the HNN super-resolution mapping technique, with a zoom factor of 12. After 6000 iterations with the weighting constants of k_1=70, k_2=70, k_3=100 and k_4=70, three land cover images were super-resolved as in Figure 3.13(g), 3.13(h), 3.13(i) (These optimal weighting constants were selected empirically. In the next parts of this thesis, the same principle to select the weighting constants was used). Similar to the simulated IKONOS experiment, the evaluation of the new method was based on the accuracy statistics which are usually used for hard classification such as the KIA, the overall accuracy, and the per-class omission and commission errors (Table 3.3).

The new HNN super-resolution mapping technique was constrained by the 8.4 m land cover class proportion image and the 2.1 m fused images. To estimate the local end-member spectra of three land cover classes, the 8.4 m land cover proportion image was used in combination with the 8.4 m MS images which were degraded from the 2.1 m fused images. Based on the results of the simulated IKONOS data case, the empirical weight value of 14 was used to determine the local end-member spectra. A zoom factor of 12 was used to produce a 0.7 m spatial resolution map. The HNN was set with weighting constants of k_1=70, k_2=70, k_3=100, k_4=70 and k_5=100. After 6000 iterations, the HNN using the fused MS image without PAN image registration error and the local end-member spectra produced the sub-pixel land cover images in Figure 3.13(j), 3.13(k) and 3.13(l). Setting the same weighting constants, the HNN super-resolution mapping using the fused image obtained from the PAN image with image registration RMS error of 1 pixel (the accuracy can be obtained normally in the geometric correction process) produced the land cover maps as in Figure 3.13(m), 3.13(n) and 3.13(o). Accuracy statistics of the predicted land cover map are given in Table 3.3.

3.4.2.2 Visual evaluation

In comparison with the hard classification and the traditional HNN super-resolution mapping, the resulting land cover map produced by the HNN using the fused image is visually more accurate. Similar to the simulated IKONOS data set, the greatest improvement can be seen in the white surface class, where almost all sub-pixels

belong to small objects. Without information from the fused image, the white surface sub-pixels of the linear objects in Figure 3.11(b) were clustered into larger objects to satisfy the HNN goal functions as in Figure 3.13(h). Although the fused image contains error due to the image registration error of the PAN image, the small and linear white surface objects can still be mapped and their shapes look similar to those in the reference image. This fact suggests that the new technique can be used for applications such as target identification.

3.4.2.3 Statistical evaluation

The accuracy statistics showed a considerable increase in accuracy with the new technique. Overall accuracy of the land cover map increased by around 3% from 94.21% for the hard classification to 97.18% for the super-resolution mapping using the fused image without image registration error. With the fused image produced by the PAN image with image registration error of 1 pixel, the accuracy of the resulting 0.7 m land cover map increased around 2% and 0.5% in comparison with the results of hard classification and the HNN super-resolution mapping technique, respectively. The KIA value increased from 0.8685 for the hard classified map and 0.9166 for the traditional HNN sub-pixel map to 0.9365 for the super-resolution mapping using fused image without PAN image registration error and 0.9166 for the super-resolution mapping using fused image with PAN image registration error of 1 pixel. Comparing with the resulting sub-pixel map produced by the HNN when the fused image was not used, the accuracy of the thematic map produced by the new technique (without PAN image registration error) increased approximately 1% in terms of overall accuracy.

Similar to the experiment involving simulated IKONOS imagery, amongst the three land cover classes, the accuracy of the white surface class increased most with the omission error reduced from 85.31% for the hard classified image and 65.28% for the traditional HNN super-resolution mapping to approximately 57.85% and 54% for the new HNN super-resolution mapping technique with and without PAN image registration error, respectively. The commission error reduced from 30.79% and 39.88% to 27.96% and 18.20%, respectively after using the fused image. The increase in accuracy of the other two classes was not as great as that of the white surface class since most sub-pixels in these classes were grouped into larger objects.

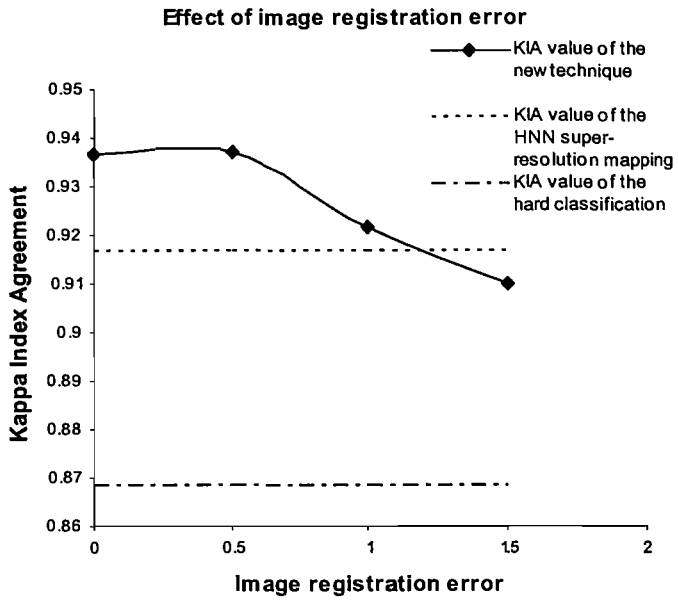


Figure 3.14 The effect of the image registration on Kappa Index Agreement value of resulted sub-pixel map using the HNN super-resolution mapping using the fused image.

3.4.2.4 Image registration error effect

To determine the effect of the image registration error on the results of the new technique, a series of PAN images with image registration error ranging from 0.5 pixels to 1.5 pixels were used to produce the fused images. Accuracy evaluation of the resulting land cover maps using these fused images based on the KIA value was implemented and presented in the plot in Figure 3.14. Obviously, the plot shows that the KIA value reduced as the RMS image registration error increased. When the RMS image registration error increased to 1.5 pixels, the resulting sub-pixel map predicted by the new technique was less accurate than the results of the HNN without using the fused image. However, with the fused image produced from a PAN image with RMS error of 1 pixel (the accuracy of image registration that can be usually obtained in geometric correction of remotely sensed images) the accuracy of the sub-pixel map produced by the HNN super-resolution mapping using the fused image was greater than that produced by the HNN super-resolution mapping technique without using the fused image. It is recommended that the new technique should be used only if the PAN image is registered with an RMS error equal to or smaller than one pixel.

3.5 Conclusions

This Chapter introduces the use of fused images for super-resolution mapping. Data from the fused images were incorporated into the HNN optimisation using forward

and inverse models in the form of the reflectance constraint. The value of the constraint was calculated based on a linear mixture model, which used both global and local end-member spectra. The effectiveness of the technique was examined using both (i) a simulated IKONOS data set and (ii) a degraded QuickBird image (with and without image registration error). In both cases, the proportions images were supplemented by a simulated fused image and original MS image. The accuracy evaluation was implemented based on the KIA, overall accuracy, and omission and commission errors.

The results demonstrated that fused images can be used as a source of supplementary information for the HNN to predict land cover accurately at sub-pixel spatial resolution from simulated and real land cover proportion images. The analysis demonstrated a considerable increase in accuracy with the new technique, particularly for land cover features at the sub-pixel scale. For larger features, the technique increased the accuracy slightly. In addition, visual inspection of the resulting image showed pleasing improvements. The analysis also suggests that the new technique can be applied only if the RMS image registration error of the PAN image is equal to or smaller than 1 pixel.

The results of the experiments suggest the potential for combining image fusion and super-resolution mapping processes for real data. Thus, future research will develop a HNN to incorporate directly real panchromatic imagery as supplementary data to increase the accuracy and detail of the predicted sub-pixel land cover map.

Chapter 4: Super-resolution mapping using a Hopfield neural network with the panchromatic image[3]

4.1 Introduction

In this Chapter, a new technique of using panchromatic imagery for super-resolution mapping is proposed. The new technique is based on the general model for using the panchromatic imagery for super-resolution mapping which is presented in Chapter 2. The panchromatic (PAN) imagery was used as an additional source of information for super-resolution mapping. Information from the PAN image was used in the form of a new panchromatic reflectance constraint for the energy function of the Hopfield neural network (HNN). The value of the new panchromatic reflectance constraint was defined based on forward and inverse models with local end-member spectra and local convolution weighting factors. Two sets of simulated and degraded data were used to test the new technique. The results indicated that panchromatic imagery can be used as a source of supplementary information to increase the detail and accuracy of sub-pixel land cover maps produced by super-resolution mapping.

4.2 Hopfield neural network structure for using the PAN image for super-resolution mapping

4.2.1 General model

The new method proposed in this paper aims to use a PAN image as additional information at the intermediate spatial resolution to produce a sub-pixel land cover map at the super-resolution. To make use of the PAN image for land cover mapping, a forward model, a spectral convolution model and an inverse model were employed in the form of a panchromatic reflectance constraint within the HNN approach developed by Tatem *et al.* (2001a). The panchromatic reflectance constraint was obtained using a minimum grey value difference method.

Figure 4.1 is a graphical depiction of the method proposed for incorporating a PAN image into the super-resolution mapping using a HNN. From the MS images at the original MS spatial resolution, land cover area proportion images are produced using a soft-classifier. From the super-resolution land cover map at the first iteration, an

[3] This chapter is partly based on Nguyen *et al.* (2005b)

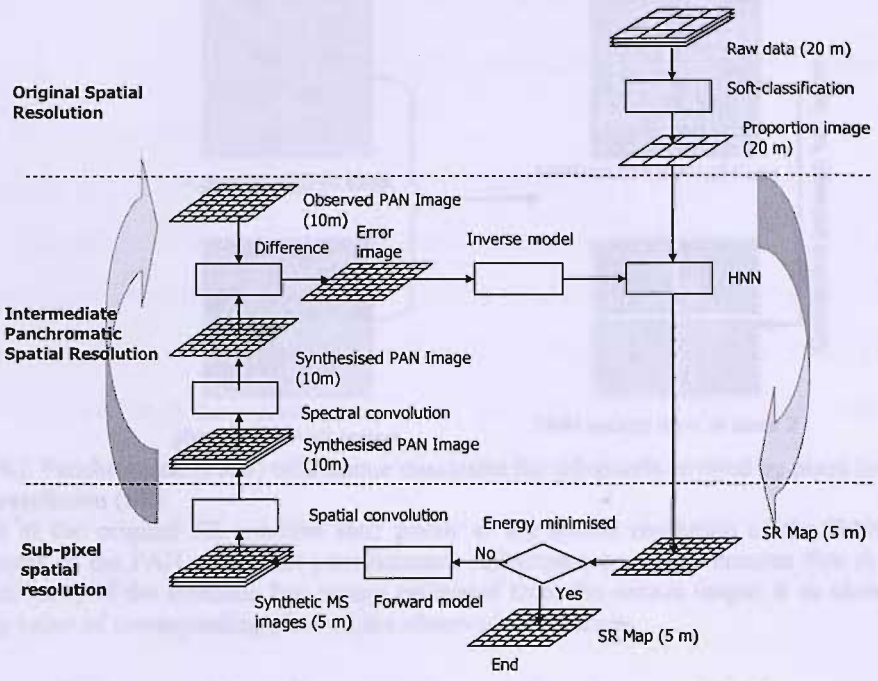


Figure 4.1 HNN super-resolution (SR) mapping using the panchromatic (PAN) mages.

estimated MS image (at the PAN image spatial resolution) is then produced using a forward model. The estimated MS image is then convolved spectrally to create a synthetic PAN image. Following a comparison of the observed and synthetic PAN images, a value is produced for all neurons covered by the same pixel in the PAN image to force the synthetic PAN image to converge to the observed PAN image. Thus, the HNN is constrained by the grey values of the PAN image. The adjustment value, or panchromatic reflectance constraint value, along with the goal and constraint values in the HNN structure proposed by Tatem *et al.* (2002a), can be used in the optimisation process for super-resolution mapping by minimising the energy function. After the optimisation process, the estimated grey values of the synthetic PAN image produced from the sub-pixel land cover map should be very similar to the observed PAN image.

4.2.2 HNN structure

The method presented is based on the structure of the HNN proposed by Tatem *et al.* (2002a). The structure of the HNN for super-resolution mapping using a PAN image of two land cover classes can be seen in Figure 4.2. A pixel at the original MS spatial resolution is divided into two inter-connected matrices of neurons in the HNN. Each neuron (h,i,j) represents a sub-pixel at position (i,j) in the land cover class h and the

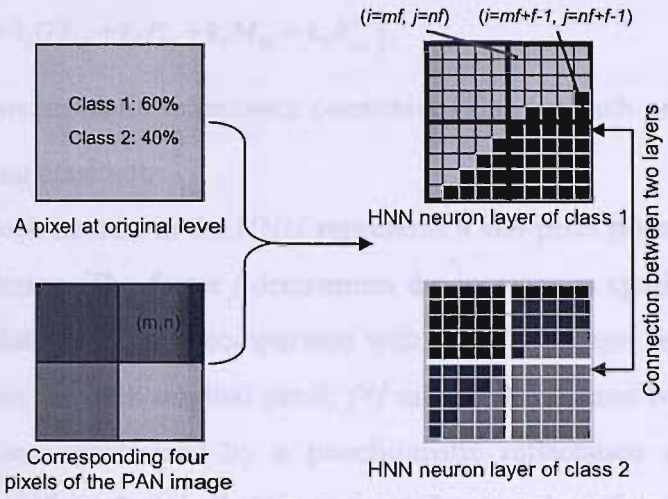


Figure 4.2 Panchromatic (PAN) reflectance constraint for sub-pixels covered by pixel (m,n) at the PAN spatial resolution (SR).
A pixel at the original SR contains four pixels at the spatial resolution of the PAN. m, n are the coordinates of the PAN pixel. The panchromatic reflectance constraint ensures that the grey value of the pixel (m,n) of the synthetic Pan image estimated from the neuron output is as close as possible to the grey value of corresponding pixel of the observed PAN image.

value allocated to each neuron represents the likelihood of belonging to that land cover class. For super-resolution mapping, the HNN is initialised using the soft-classified land cover proportions and runs until it converges monotonically to a stable state when the energy function is minimised. At the stable state, the output values of the neurons are binary values. If the output value of the neuron (h,i,j) is one, the sub-pixel (i,j) is assigned to the land cover class h. Otherwise, if the output value is zero, the sub-pixel (i,j) does not belong to the class h. The energy function can be expressed as

$$E = -\sum_{h}\sum_{i}\sum_{j}\left(k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij}\right), \tag{4.1}$$

where k_1, k_2, k_3 and k_4 are weighting constants, which determine the effects of the goal functions $G1_{hij}$ and $G2_{hij}$, proportion constraint P_{hij} and multi-class constraint M_{hij} on the energy function. The optimal values of the weighting constants are determined empirically for super-resolution mapping (Tatem *et al.* 2002a). The values of the goal functions $G1_{hij}$ and $G2_{hij}$, proportion constraint P_{hij} and multi-class constraint M_{hij} by Equation (3.2), (3.3), (3.4) and (3.5).

To use the PAN image for super-resolution mapping by the HNN, a new panchromatic reflectance constraint was added to the energy function in Equation 1. The new energy function can be expressed in an equation as follows:

$$E = -\sum_h \sum_i \sum_j \left(k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij} + k_5 R^P_{hij} \right), \quad (4.2)$$

where R^P_{hij} is the panchromatic reflectance constraint value for each neuron (h,i,j) and k_5 is a fifth weighting constant.

As in Figure 4.2, each neuron in the HNN represents a sub-pixel point in the original spatial resolution image. The factor f determines the increase in spatial resolution of the new super-resolution image in comparison with the PAN image. In addition to the proportion constraint for each original pixel, $f \times f$ sub-pixels covered by pixel (m,n) of the PAN images are constrained by a panchromatic reflectance constraint. This constraint is based on the principle that the estimated grey value of a pixel of the PAN image should be equal to the observed grey value (or target reflectance) of that pixel.

The new energy function in Equation (4.2) is minimised if the derivatives of variables in Equation (4.2) converge to zero for each neuron (h,i,j),

$$\frac{dE_{hij}}{dv_{hij}} = k_1 \frac{dG1_{hij}}{dv_{hij}} + k_2 \frac{dG2_{hij}}{dv_{hij}} + k_3 \frac{dP_{hij}}{dv_{hij}} + k_4 \frac{dM_{hij}}{dv_{hij}} + k_5 \frac{dR_{hij}}{dv_{hij}} = 0 . \quad (4.3)$$

The derivative values of $G1$, $G2$, P, and M with respect to v_{hij} are computed by Equation (3.2), (3.3), (3.4) and (3.5), respectively. The value dR_{hij}/dv_{hij} is derived based on the forward and inverse models, spatial and spectral convolutions as in Figure 4.1.

4.2.2.1 Forward model

The forward model is used to estimate a set of MS bands from a land cover image. The estimated reflectance of the neurons representing the pixel (m,n) of three MS bands at the PAN spatial resolution can be defined by a forward model as

$$\begin{aligned} R^t_1 &= s_{1,1} V_1 + \ldots + s_{1,c} V_c \\ R^t_2 &= s_{2,1} V_2 + \ldots + s_{2,c} V_c \text{ or } \mathbf{R}^t = \mathbf{S}\mathbf{V}^t, \\ R^t_3 &= s_{3,1} V_1 + \ldots + s_{3,c} V_c \end{aligned} \quad (4.4)$$

where the estimated proportion value of a class e at the iteration t is $V^t_i = 1/f^2 (\sum_{p=mf}^{mf+f-1} \sum_{q=nf}^{nf+f-1} v_{ipq})$, $s_{i,j}$ is the spectral value of the land cover class j for a spectral band i, $\mathbf{R}'_{m,n} = [R^t_1 \;\; R^t_2 \;\; .. \;\; R^t_b]^T_{m,n}$, $\mathbf{V}'_{m,n} = [V^t_1 \;\; V^t_2 \;\; .. \;\; V^t_3]^T_{m,n}$ and

$$\mathbf{S} = \begin{bmatrix} s_{1,1} & \cdots & s_{1,c} \\ s_{2,1} & \cdots & s_{2,c} \\ s_{b,1} & \cdots & s_{b,c} \end{bmatrix}. \quad (4.5)$$

4.2.2.2 Spectral convolution

The spectral convolution procedure is employed to create a synthetic PAN image from a set of MS bands. Spectral convolution can be based on the synthesising method proposed by Zhang (1999) as

$$R_{PAN} = \sum \varphi_i R_i , \tag{4.6}$$

where R_{PAN} is the digital value of the synthetic PAN image, R_i is the reflectance value of the MS band i, and φ_i is a weighting factor for the MS band i. The weighting factor φ_i can be estimated directly from the PAN and MS images using multiple regressions of the original PAN image and the original MS bands.

4.2.2.3 Error image

From the observed PAN image, it is possible to produce an error image that can be used for HNN super-resolution mapping. The error image is the difference between the observed PAN image and the synthetic PAN image and can be calculated as

$$E = R^o - R^t_{PANSyn} , \tag{4.7}$$

where R^t_{PANSyn} is the synthetic PAN image at the iteration t, which can be calculated by the spectral convolution model in Equation (4.6)

$$R^t_{PANSyn} = \sum \varphi_i R^t_i . \tag{4.8}$$

4.2.2.4 Inverse model and panchromatic reflectance constraint

From the error image, a panchromatic reflectance constraint value for each neuron in the HNN can be produced based on an inverse model. If the grey value of a synthesised PAN pixel is greater than that of the corresponding observed PAN pixel, the constraint produces a value to reduce the grey level of the synthetic image. Conversely, if the synthetic reflectance value of a PAN pixel is smaller than that of the observed PAN image, the constraint produces a value to increase the grey value of the synthetic image.

Supposing that the MS bands at a spatial resolution of a PAN image are 1, 2,.., b and the estimated land cover proportions of c land cover classes for pixel (m,n) are P_1, P_2,.., P_c. Using the forward model in Equation (4.4), the relationship between the reflectance of the MS bands and the land cover classes can be expressed as

$$\begin{aligned} R^o_1 &= s_{1,1}P_1 + \ldots + s_{1,c}P_c \\ R^o_2 &= s_{2,1}P_1 + \ldots + s_{2,c}P_c \\ R^o_3 &= s_{b,1}P_1 + \ldots + s_{b,c}P_c \end{aligned} \quad \text{or} \quad \mathbf{R}^o_{m,n} = \mathbf{S}\mathbf{P}_{m,n} , \tag{4.9}$$

where $\mathbf{R}_{m,n}=[R_1 \ .. \ R_i \ .. \ R_b]^T_{m,n}$, is the vector of reflectance values of the bands 1, 2, .., b and $\mathbf{P}_{m,n}=[P_1 \ P_2 \ .. \ P_c]^T_{m,n}$ is the vector of land cover proportion values. In sub-pixel mapping, based on the assumption that a pixel consists of a matrix of crisp sub-pixels, the land cover proportions P_1, P_2,.., P_c must be a quotient of a whole number of sub-pixels belonging to a given class divided by the number of sub-pixels within an original pixel. For example, if a pixel in the PAN image consists of 2x2 sub-pixels, the land cover proportion of a land cover class P_i should be selected from the set of the values {0/4, 1/4, 2/4, 3/4 and 4/4}. The land cover proportions of all c land cover classes of a pixel in the PAN image are then a combination of the values such as P_1=1/4, .., P_i=0/4,.., P_c=3/4 with P_1 +..+ P_i + ..+P_c = 1. The synthetic grey value of a PAN pixel would converge to the observed grey value or the error image would converge to zero if the estimated proportion value v_i^t of class c at the iteration t is equal to the land cover proportion P_i.

The land cover class proportions for each PAN pixel can be determined from a limited number of possible combinations of land cover classes. For example, the land cover proportions of a PAN pixel (m,n) which contains two land cover classes P_1 and P_2 can be one of the possible combinations $\mathbf{P}_1$ = [P_1=0/4, P_2=4/4], $\mathbf{P}_2$ = [P_1=1/4, P_2=3/4], $\mathbf{P}_3$ = [P_1=2/4, P_2=2/4], $\mathbf{P}_4$ = [P_1=3/4, P_2=1/4], and $\mathbf{P}_5$ = [P_1=4/4, P_2=0/4] (provided that a pixel in the PAN image covers 2x2 sub-pixels). In this research, the land cover proportions $\mathbf{P}$ of a PAN pixel can be determined based on the minimum grey value difference method.

The minimum grey value difference method is based on the forward model (Equation (4.4)) and spectral convolution (Equation (4.6)). To calculate the land cover proportions for pixel (m,n) in the PAN image, the difference between the synthetic grey value computed from each possible land cover proportion and the observed grey value is calculated. The land cover proportion of the pixel can be chosen based on the minimum difference between the synthetic grey value $R_p^{Synthetic}$ determined from the set of land cover combinations {$\mathbf{P}_1$, $\mathbf{P}_2$,.., $\mathbf{P}_{NCB}$} and the observed grey value $R^{Observed}$ of the PAN image (NCB is total number of the possible combinations). The process can be expressed by a rule as follows:

$$\mathbf{P}_p \text{ is } \mathbf{P} \text{ if } \left(\left|R_p^{Synthetic} - R^{Observed}\right|_{mn} = \min\right) \tag{4.10}$$

where **P** is a vector of land cover proportions for PAN pixel (m,n). When the land cover proportion image **P** is obtained, the panchromatic reflectance function value can be calculated as

$$dR / dv_{hij} = P_{hmn} - V'_{hmn} \tag{4.11}$$

where (m,n) is PAN pixel which covers the pixel (i,j), P_{hmn} is the land cover proportion of class h and V'_{hmn} is the estimated land cover proportion of class h of the pixel (m,n) at the iteration t.

4.2.2.5 Weighting mechanism for proportion and panchromatic reflectance constraint

Using the panchromatic reflectance constraint, the land cover proportions produced from all pixels of the PAN image covered by a pixel (x,y) of the original image should be close to the land cover proportions of the pixel (x,y) predicted by soft-classification. To ensure that this is the case, an empirical weighting mechanism was used to adjust the weighting constant k_5 in Equation (4.3) as

$$\text{If}\left((\mathbf{P}^{xy}_{PAN} - \mathbf{P}^{xy}_{Soft-classified}) > 0.2\right) \text{ then } (k_5 = k_5 / 10), \tag{4.12}$$

where $\mathbf{P}^{xy}_{PAN}$ is the land cover proportion produced from all pixels of the PAN image located within the pixel (x,y) and $\mathbf{P}^{xy}_{Soft-classified}$ is the land cover proportion predicted by soft-classification.

In the next sections, the newly proposed method of using PAN image for super-resolution mapping is evaluated on two data sets: simulated IKONOS and degraded QuickBird data.

4.3 Example 1: Simulated IKONOS image

4.3.1 Data

Prior to application to real data, the proposed method was tested using the simulation image in Chapter 3. In this set of simulated data, a simulated 32 m MS image and an 8 m PAN image were simulated using a real IKONOS image. The ratio between the spatial resolutions of the simulated MS and PAN image is similar to the ratio between the real 4 m MS and 1 m PAN IKONOS image. Thus, the algorithm is likely to be applicable to real imagery if it performs accurately on the simulated image. Although the size of features or spatial frequency of variation should be taken into account, super-resolution mapping tasks are considerably similar for the degraded and the real imagery. For example, mapping trees or roads using a 32 m image is similar to

mapping bushes and alleyways using a 4 m image. The simulation ensured that there were no errors in image registration between the reference image and the land cover image obtained by super-resolution mapping. Therefore, the focus of the evaluation is on the performance of the algorithm independent of registration issues.

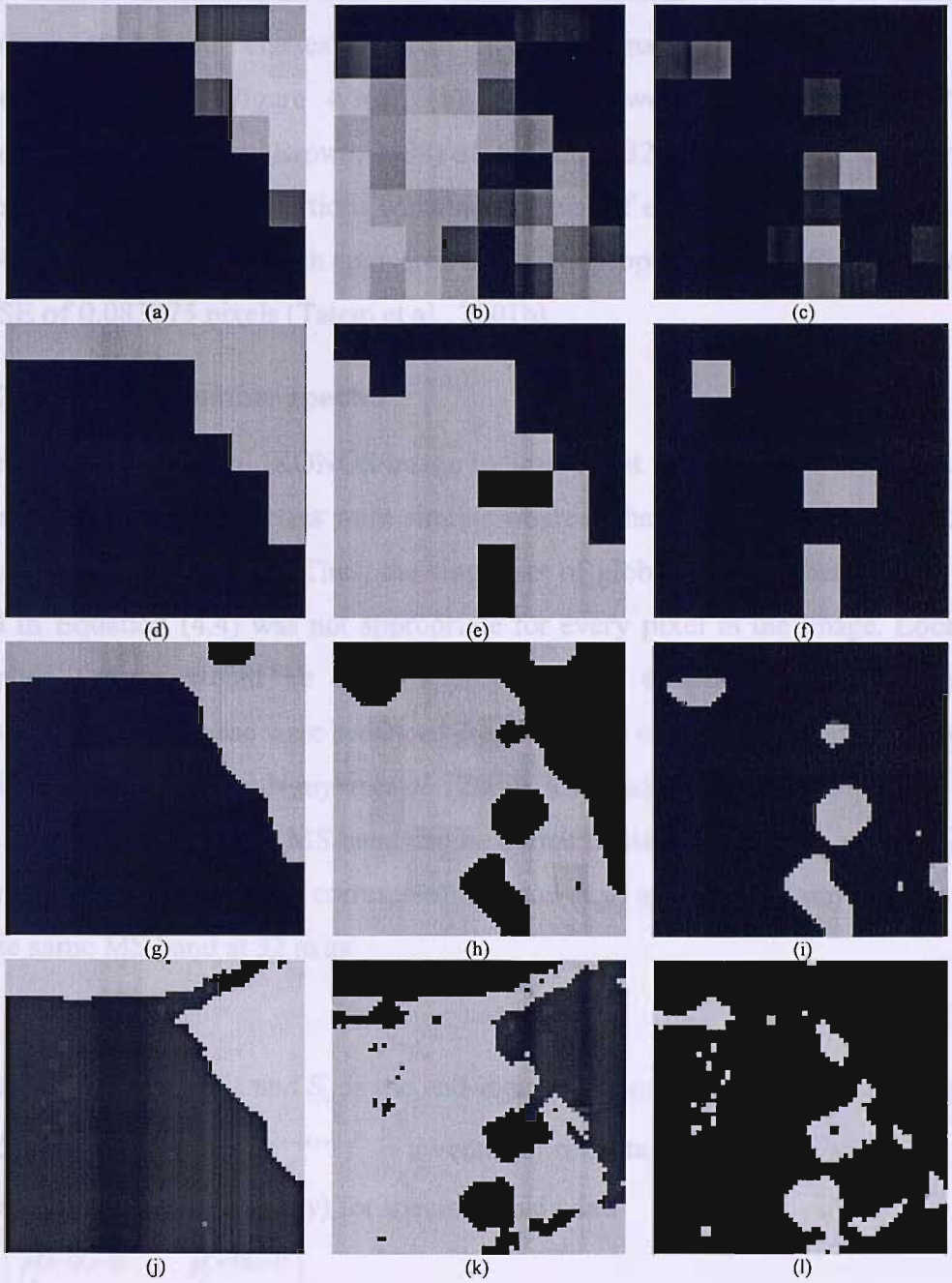


Figure 4.3 Results of the simulated data
(a) 4 m Cereal, (b) Grass and (c) Trees land cover proportion image. (d) 4 m Cereal, (e) Grass, (f) Trees hard classified land cover image. (g) 4 m Cereal, (h) Grass, (i) Trees HNN super-resolution mapping image. (j) 4 m Cereal, (k) Grass, (i) Trees HNN super-resolution mapping using the PAN with the local end-member spectra resulting image.

The simulation data set included 32 m MS (Figure 3.6(f), (g), (h), (i)) and 8 m (Figure 3.6(e)) PAN images which were generated from 4 m reference cereal, grass and trees classes (Figure 3.4(a), (b), (c), (d)) using a random normal distribution based on the mean and variance of each class in the MS bands. The means and variances of the three land cover classes in each MS band were obtained based on an analysis of reference land cover classes and real IKONOS imagery. The 32 m land cover proportion images (Figure 4.3(a), (b), and (c)) were produced by *k*-NN soft classification (Lewis and Brown, 1998) of simulated 32 m MS imagery with k=5. The estimated land cover proportions contained a level of error similar to that which can be observed for real data with an overall area error proportion of 0.56% and an overall RMSE of 0.083775 pixels (Tatem et al., 2001b).

4.3.2 Local end-member spectra

Investigation of the real IKONOS image indicated that the end-member spectra of the same class in adjacent pixels were similar whereas the end-member spectra of more distant pixels varied more. Thus, the single set of global end-member spectra values used in Equation (4.4) was not appropriate for every pixel in the image. Local end-member spectra would be more appropriate for determining local reflectance constraint values. These were produced from the land cover proportion image and the original MS image as in Nguyen *et al.* (2006). The end-member spectra values of the pixel (m,n) of a given 8 m MS band can be defined based on the class proportions and the reflectance values of the corresponding pixel (x,y) and its eight surrounding pixels of the same MS band at 32 m as

$$\mathbf{S}_i = \left(\mathbf{P}^T\mathbf{P}\right)^{-1}\mathbf{P}^T\mathbf{R}_i . \tag{4.13}$$

where $\mathbf{S}_i^{xy} = [S_{i,1} \ \ldots \ S_{i,c}]$ and $S_{i,j}$ is the end-member spectra value of class c_j on spectral band i, $\mathbf{R}_i = [R_i^{(x-1)(y-1)} \ \ldots \ R_i^{(x+1)(y+1)}]^T$ is a vector of reflectance values in which R_i^{xy} is the reflectance value of pixel (x,y) for spectral band i and

$$\mathbf{P} = \begin{bmatrix} P_1^{(x-1)(y-1)} & .. & P_c^{(x-1)(y-1)} \\ .. & .. & .. \\ P_1^{xy} & .. & P_c^{xy} \\ .. & .. & .. \\ P_1^{(x+1)(y+1)} & .. & P_c^{(x+1)(y+1)} \end{bmatrix}. \tag{4.14}$$

where P_j^{xy} is the proportion value of pixel (x,y) for class j.

Amongst the pixels that are used to determine the local end-member spectra, the pixel (x,y) should be the most important since it covers the PAN pixel (m,n). To emphasise the contribution of the corresponding pixel (x,y) to the end-member spectra, a weighting mechanism was used, such that Equation 16 becomes

$$\mathbf{S}_i = \left(\mathbf{P}^T\mathbf{W}\mathbf{P}\right)^{-1}\mathbf{W}\mathbf{P}^T\mathbf{R}_i, \tag{4.15}$$

where $\mathbf{W}$ is the diagonal matrix:

$$\mathbf{W} = \begin{bmatrix} w^{(x-1)(y-1)} & 0 & 0 & 0 & 0 \\ 0 & .. & 0 & 0 & 0 \\ 0 & 0 & w^{xy} & 0 & 0 \\ 0 & 0 & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & w^{(x+1)(y+1)} \end{bmatrix}, \tag{4.16}$$

and $w^{(x-1)(y-1)},..,w^{xy},..,\ w^{(x+1)(y+1)}$ are weight values for each corresponding pixel. The optimal weight value w^{xy} was tested using the weight values w^{xy} in the range [1, 20] and the other weight values of one.

4.3.3 Local spectral convolution weighting factor

As for the end-member spectra, the spectral convolution weighting factors can be determined locally using the least squares method. The spectral convolution weighting factors of a pixel (m,n) of the PAN image were calculated based on the digital value of the MS and PAN pixels that were covered by its corresponding original pixel (x,y) and eight surrounding pixels at the MS image spatial resolution.

$$\mathbf{\Phi}^{m,n} = \left(\mathbf{R}_{MS}^T\mathbf{R}_{MS}\right)^{-1}\mathbf{R}_{MS}^T\mathbf{R}_{PAN} \tag{4.17}$$

where $\mathbf{\Phi}^{xy} = \begin{bmatrix} \varphi_1 \\ .. \\ \varphi_b \end{bmatrix}$, $\mathbf{R}_{MS} = \begin{bmatrix} R_1^{(x-1)(y-1)} & .. & R_b^{(x-1)(y-1)} \\ .. & .. & .. \\ R_1^{(x+1)(y+1)} & .. & R_b^{(x+1)(y+1)} \end{bmatrix}$, $\mathbf{R}_{PAN} = \begin{bmatrix} R_{PAN}^{(x-1)(y-1)} \\ .. \\ R_{PAN}^{(x+1)(y+1)} \end{bmatrix}$ where R_i^{xy} is the digital value of the pixel (x,y) of the MS band i and R_{PAN}^{xy} is the digital value of the pixel (x,y) of the PAN band.

4.3.4 Results and discussions

4.3.4.1 Network settings

Using the 32 m soft-classified land cover proportion image (Figure 4.3(a), 4.3(b), and 4.3(c)), the 4 m sub-pixel land cover maps were obtained using the traditional HNN (Figure 4.3(g), 4.3(h), and 4.3(i)), and the HNN using the PAN image with local end-member spectra (Fig. 4.3(j), 4.3(k) and 4.3(l)). The greatest accuracy land cover map

Statistics for the hard classified image

	Cereal	Grass	Trees	ErrorO (%)	ErrorC (%)
Unclassified	0	0	0		
Cereal	1005	126	21	3.74	12.76
Grass	39	2305	280	8.42	12.16
Trees	0	86	234	56.26	26.88
KIA – κ =		0.7430	Overall accuracy =	86.52%	

Statistics for the HNN super-resolution mapping without using the PAN image

	Cereal	Grass	Trees	ErrorO (%)	ErrorC (%)
Unclassified	0	4	3		1.000
Cereal	977	89	2	6.42	8.52
Grass	63	2350	231	8.14	11.12
Trees	4	74	299	36.82	20.69
KIA – κ =		0.7814	Overall accuracy	88.53%	

Statistics for the HNN super-resolution mapping using the PAN image

	Cereal	Grass	Tree	ErrorO (%)	ErrorC (%)
Unclassified	10	70	31		1.000
Cereal	983	30	2	5.84	3.15
Grass	34	2349	106	6.67	5.62
Trees	17	68	396	25.98	17.67
KIA – κ =		0.8361	Overall accuracy	91.02%	

Table 4.1 Accuracy statistics of simulated results.

was obtained with the weighting coefficients of k_1=100, k_2=100, k_3=100, k_4=150 and k_5=100 after 1000 iterations. The 32 m hard classified land cover image (Figure 4.3(d), 4.3(e) and 4.3(f)) was produced from the 32 m soft-classified proportion image (Figure 4.3(g), 4.3(h), 4.3(i), and 4.3(k)) by assigning a land cover class to each pixel based on its largest proportion. To evaluate the resulting hard land cover maps at sub-pixel spatial resolution of 4 m, accuracy statistics as traditionally applied for hard classification based on the error matrix were used. These included the KIA value – κ, overall accuracy, and per-class omission and commission errors (Table 4.1).

4.3.4.2 Visual evaluation

Comparing the results of the two prediction techniques visually, it is clear that the super-resolution mapping using the PAN image is preferable to hard classification and the traditional HNN super-resolution mapping technique (Figure 4.3). The increase in accuracy is most obviously seen in the trees class, where most of the trees objects are

smaller than the 32 m simulated image pixel (Table 4.1). Without information from the PAN image, the trees class sub-pixels contained in the linear features in the right-bottom and in the top of Figure 3.4(c) were clustered into larger objects to satisfy the goal functions as in Figure 4.3(i). The panchromatic reflectance constraint produced a value to retain reflectance for the neurons in those linear objects. Therefore, objects smaller than an original pixel were mapped more accurately (Figure 4.3(l)).

4.3.4.3 Statistical evaluation

The accuracy statistics (Table 4.1) showed a considerable increase in accuracy for the new technique. Overall accuracy of the land cover map increased by around 5% from 86.52% for the hard classification to 91.02% for the super-resolution mapping using the PAN image with local spectra. The KIA value - κ increased from 0.7533 for the hard classified map to 0.8361 for the new HNN technique. In comparison with the HNN super-resolution mapping without using the PAN image, the accuracy of the thematic map produced by the new HNN technique increased 3.5% in terms of overall accuracy.

Amongst the three land cover classes, the accuracy of the trees land cover class increased most with the omission error reduced from 56.26% for the hard classified image and 36.82% for the traditional HNN super-resolution mapping to 25.98% for the new HNN super-resolution mapping technique. Similarly, the commission error reduced from 26.25% and 20.69% to 17.67% after using the PAN image. In the other two classes, the increase in accuracy was not as great as that of the trees class since most sub-pixels in these two classes were grouped into objects larger than a 32 m pixel. However, the increases in accuracy are relatively large if the small initial omission and commission errors are taken into account. This can be explained by the fact that the clustering goal functions are suitable for large features, yet the detailed information at the PAN spatial resolution is essential for more accurate super-resolution mapping of the boundary pixels.

The number of unclassified pixels (Table 4.1) was slightly increased in the super-resolved land cover image using the PAN image in comparison with the results of the traditional HNN super-resolution mapping. These unclassified sub-pixels occurred as the goal functions, proportion constraint and multi-class constraint could not be satisfied simultaneously. However, the number of unclassified pixels did not reduce

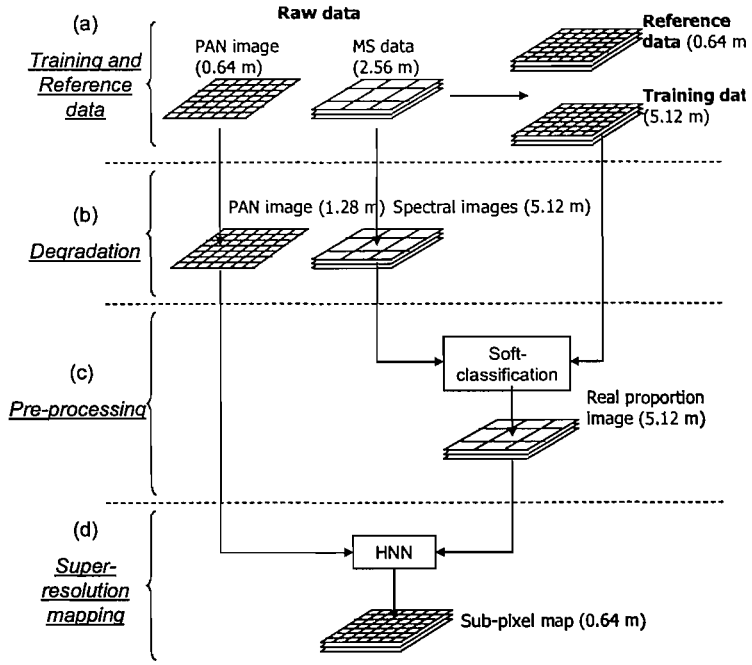


Figure 4.4 Four steps in experiment:
(a) Raw data analysis, (b) Data simulation, (c) Pre-processing, and (d) Super-resolution mapping.

the accuracy of the new technique.

4.4 Example 2: Degraded QuickBird image

4.4.1 Data

To provide a more realistic test for the new technique, a second set of proportion images was produced using a degraded QuickBird MS and PAN image. The degraded QuickBird MS and PAN images are similar spectrally to the real QuickBird MS and PAN images. Furthermore, with the degraded QuickBird MS and PAN images, it is possible to evaluate directly the impact of the image registration error on the new algorithms because the land cover classes at the sub-pixel level are known. For the degraded QuickBird image, the experiment was implemented in four steps as follows: (a) Training and reference data, (b) Degradation, (c) Pre-processing, and (d) Super-resolution mapping (Figure 4.4).

4.4.1.1 Training and reference data

Raw data: QuickBird imagery with a spatial resolution of 2.56 m (MS) and 0.64 m (PAN) was acquired over an area of Christchurch, UK on 1st June 2002. A sub-area of 328×296 PAN pixels (Figure 4.5(a)) and 82×74 MS pixels (Figure 4.5(b), 4.5(c), 4.5(d), 4.5(e)) was extracted from a rural area. Four land cover classes in the area

were identified as green cereal, grass, asphalt, and ripe cereal. The MS image was co-registered to the PAN image with a root mean square error of 0.25 pixels.

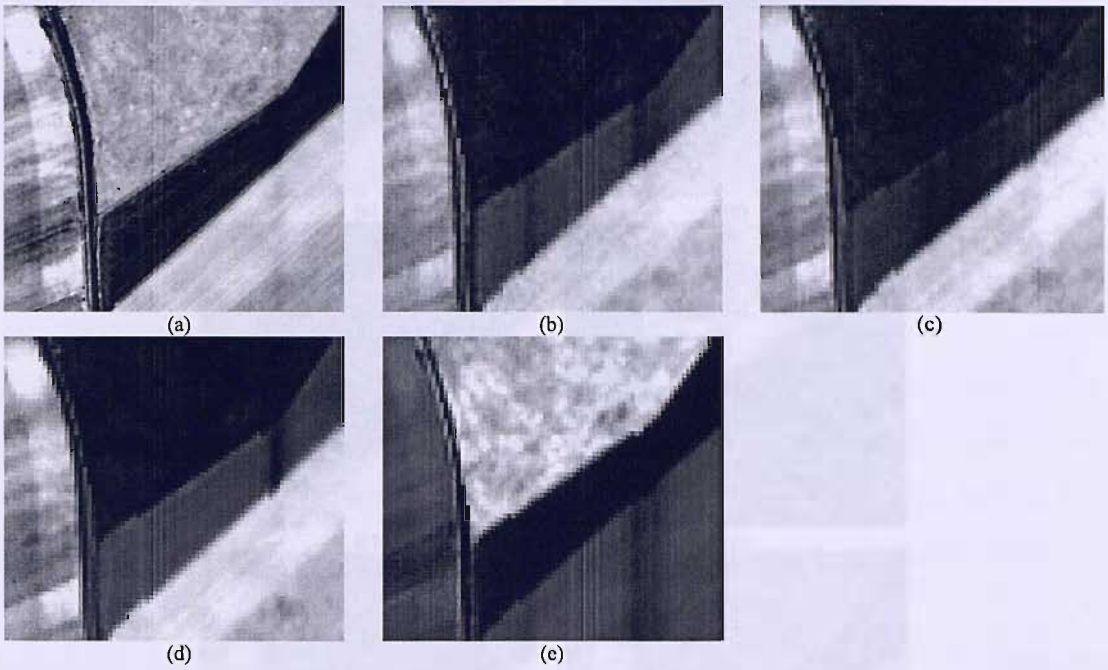


Figure 4.5 Original PAN and MS images:
(a) 0.64 m PAN image (360×300 pixels), (b) Blue, (c) Green, (d) Red and (e) NIR bands of 2.56 m MS image (90×75 pixels).

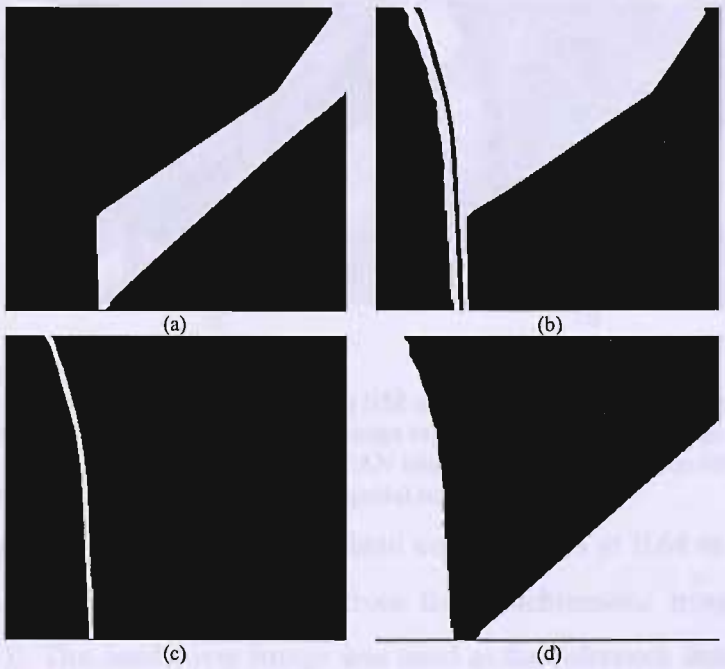


Figure 4.6 Four land cover class images for reference:
(a) 0.64 m green cereal, (b) 0.64 m grass, (c) 0.64 m asphalt and (c) 0.64 m ripe cereal.

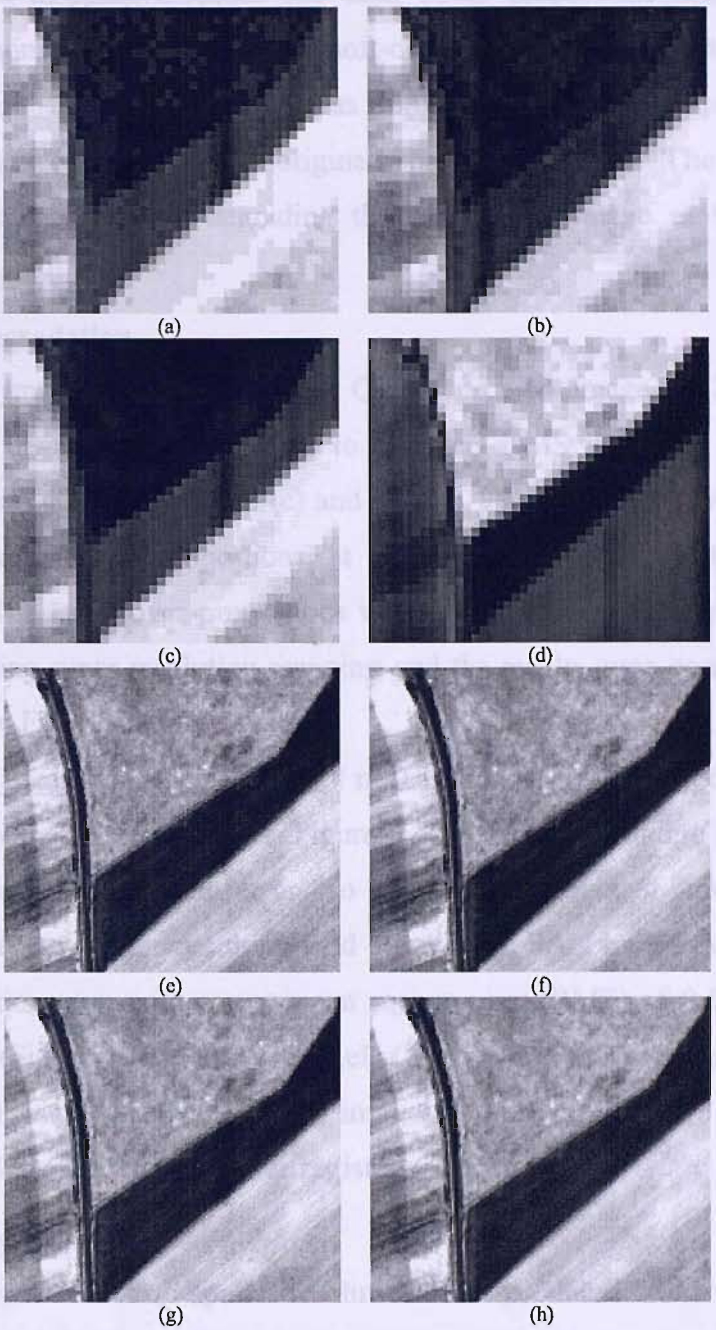


Figure 4.7 Degraded PAN and MS images:
Four bands (a) Blue, (b) Green, (c) Red and (d) NIR degraded MS image at 5.12 m spatial resolution. (e) 1.28 m degraded PAN image, (f) 0.5 pixel image registration error PAN image at 1.28 m spatial resolution, (g) 1 pixel image registration error PAN image at 1.28 m spatial resolution and 1.5 pixel image registration error PAN image at 1.28 m spatial resolution.

Training and Reference data: Four land cover classes at 0.64 m spatial resolution were obtained by manual digitising from the panchromatic image (Figure 4.6(a), 4.6(b), 4.6(c)). This land cover image was used as the reference data for the results of super-resolution mapping. Thus, there was no image registration error between the PAN image and the reference data.

Super-resolution mapping methods use land cover proportions obtained by soft-classification as input. To implement the soft-classification, training data are required. In this research, the soft-classification was implemented at 5.12 m spatial resolution (i.e. eight times coarser than the original spatial resolution). The training data, therefore, were produced by degrading the land cover image at 0.64 m spatial resolution by eight times.

4.4.1.2 Data degradation

Multispectral imagery (5.12 m): The QuickBird MS image at 2.56 m spatial resolution was degraded by two times to produce a MS image at 5.12 m spatial resolution (Figure 4.7(a), 4.7(b), 4.7(c) and 4.7(d)). This MS image was then used to produce the land cover proportions at 5.12 m spatial resolution using soft-classification. The land cover proportions were then used to produce a 0.64 m land cover image using super-resolution mapping and the results were compared with the reference data in Figure 4.6.

Panchromatic imagery (1.28 m): A 1.28 m PAN image was produced by degrading the 0.64 m PAN image by two times (Figure 4.7(e)). The PAN image in Figure 4.7(e) contained no image registration error. To evaluate the effect of image registration error on the accuracy of the resulting land cover map, PAN images were created by adding registration error with a root mean square error (RMS) of 0.5 pixels (Figure 4.7(f)), 1 pixel (Figure 4.7(g)) and 1.5 pixels (Figure 4.7(h)). The proposed algorithm was then tested using these altered PAN images and the results were compared with that of the PAN images without image registration error.

4.4.1.3 Pre-processing

Proportion images at 5.12 m spatial resolution were produced from the 5.12 m MS image in Figure 4.7. Obviously, this set of proportion images contained a certain amount of error including the MS image registration error. The predicted land cover proportions were then used for hard classification, traditional super-resolution mapping and super-resolution mapping using the PAN image. That means the results of all these methods were affected by the same amount of error contained in the land cover proportions. The only difference between the new method of super-resolution mapping and the other methods was that the results of this method were also affected by the registration error of the PAN image. Therefore, it was necessary to evaluate the impact of the PAN image registration error on the methods.

A k-nearest neighbour (k-NN) classifier was used for soft-classification with k=5. The land cover proportion image was produced with overall area error proportion of P = 0.0321% and overall RMS error of 0.0332 pixels. The proportions of the four land cover classes can be seen in Figure 4.8(a-d).

4.4.2 Results and discussions

4.4.2.1 Network settings

To evaluate the performance of the new algorithm with degraded QuickBird imagery, the results produced by the three approaches were compared visually and statistically, as for the simulated example. The 5.12 m hard classified land cover map was obtained from the land cover proportion image by assigning each 5.12 m pixel to the class of the largest proportion (Figure 4.8(e-h)). Applying the HNN super-resolution mapping approach of Tatem *et al.* (2001a), a 0.64 m spatial resolution land cover map was produced (Figure 4.8(i), 4.8(j), 4.8(k) and 4.8(l)) from the 5.12 m proportion images with a zoom factor of 8, weighting constants of k_1=100, k_2=100, k_3=150 and k_4=100, and 1000 iterations.

The new HNN super-resolution mapping technique was tested with the 5.12 m land cover class proportion image and the 1.28 m PAN images. With 1000 iterations, zoom factor of 8, and weighting constants of k_1=100, k_2=100, k_3=150, k_4=100 and k_5=100, the HNN network using the PAN image without registration error produced the land cover images in Figure 4.8(m-p). Using the same weighting constants with the PAN image and an image registration RMS error of 1 pixel (this accuracy can often be obtained in the geometric correction process) the HNN produced the land cover maps in Figure 4.8(q-t). To evaluate the new technique statistically, the error matrix and accuracy statistics such as the KIA value, omission, commission and overall errors for the resulting maps of the three methods are given in Table 4.2.

4.4.2.2 Visual evaluation

Figure 4.8 shows the advantage of the new technique in comparison with the hard classification and the traditional HNN super-resolution mapping technique. Similar to the simulated IKONOS data set, the greatest improvement can be seen in the asphalt class (Figure 4.8(o)), where most sub-pixels belong to small and linear features. Without information from the PAN image, the road sub-pixels in Figure 4.8(c) were split into discrete objects to satisfy the HNN goal functions (Figure 4.8(g)). Even

when the PAN image was registered with an RMS error of 1 pixel, the small and linear road object remained fairly contiguous (Figure 4.8(s)). This fact suggests that the new technique is applicable to map features smaller than a pixel.

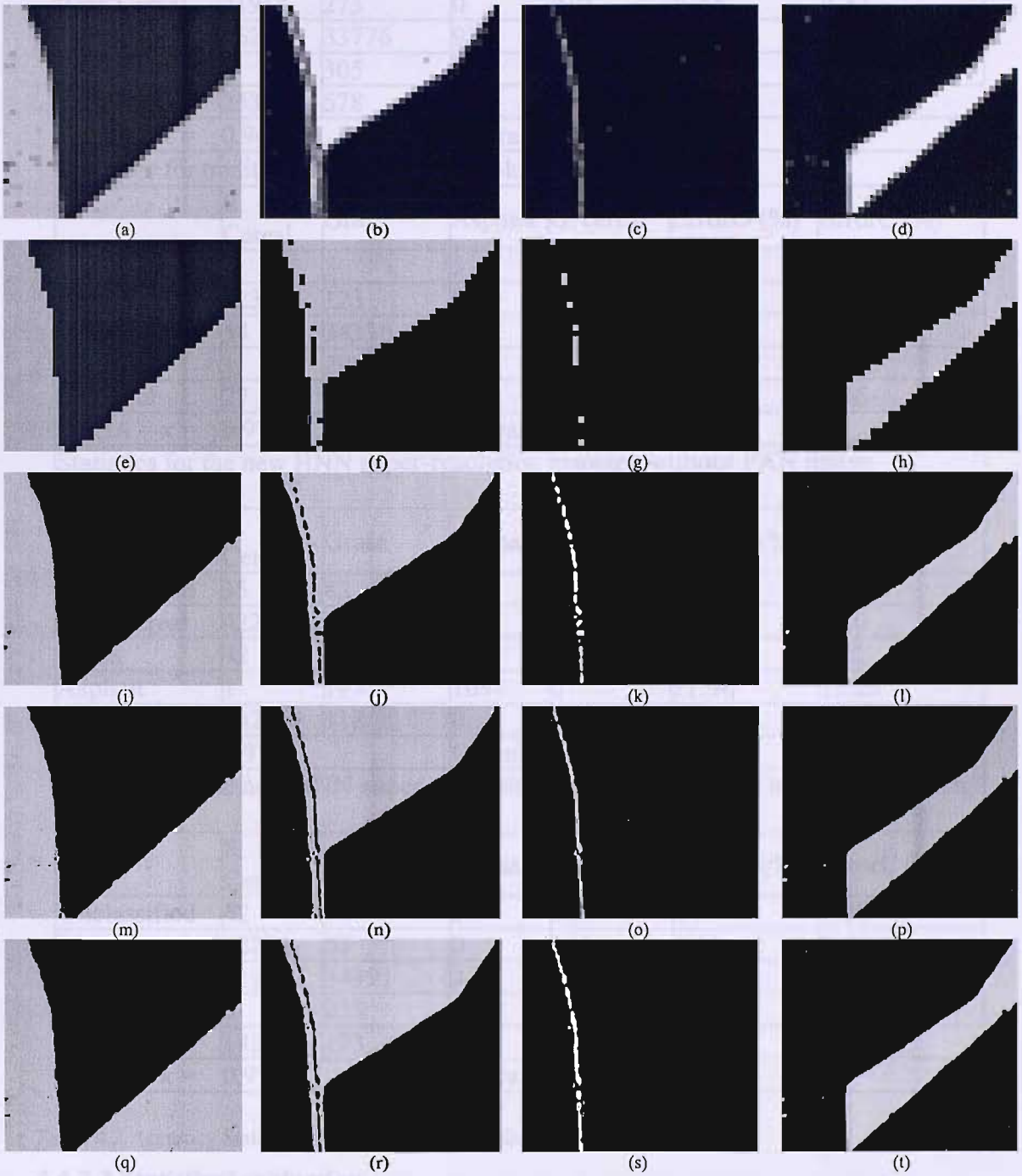


Figure 4.8 Results for the degraded QuickBird images:
5.12 m land cover proportion image for (a) Ripe cereal, (b) Grass, (c) Asphalt and (d) Green cereal. 5.12 m hard classified land cover image for (e) Ripe cereal, (f) Grass, (g) Asphalt and (h) Green cereal. 0.64 m HNN super-resolution mapping image for (i) Ripe cereal, (j) Grass, (k) Asphalt and (l) Green cereal. (m) 0.64 m Ripe cereal, (n) Grass, (o) Asphalt and (p) Green cereal land cover image of the HNN super-resolution mapping using the PAN without image registration error. (q) 0.64 m Ripe cereal, (r) Grass, (s) Asphalt and (t) Green cereal land cover image of the HNN super-resolution mapping using the PAN image with RMS image registration error of 1 pixel.

Statistics for hard classification						
	R. Cereal	Grass	Asphalt	G. cereal	ErrorO (%)	ErrorC (%)
Unclassified	0	0	0	0	0	0
Ripe Cereal	41914	273	0	501	1.45	1.81
Grass	362	33776	935	191	3.31	4.22
Asphalt	0	305	463	0	39.71	66.88
G Cereal	253	578	0	17537	4.52	3.80
KIA – κ =	0.9454		Overall accuracy =		96.50	
Statistics for traditional HNN super-resolution mapping image						
	R. Cereal	Grass	Asphalt	G. cereal	ErrorO (%)	ErrorC (%)
Unclassified	0	5	5	1	0	1.00
Ripe Cereal	42397	123	0	295	0.31	0.98
Grass	54	34336	612	64	1.71	2.08
Asphalt	1	298	781	0	27.69	44.13
G Cereal	77	170	0	17869	1.97	1.36
KIA – κ =	0.9726		Overall accuracy =		98.24	
Statistics for the new HNN super-resolution mapping without PAN image registration error						
	R. Cereal	Grass	Asphalt	G. cereal	ErrorO (%)	ErrorC (%)
Unclassified	35	82	55	30	0	1.00
Ripe Cereal	42279	129	0	126	0.59	0.60
Grass	72	34408	252	30	1.50	1.02
Asphalt	1	195	1091	0	21.96	15.23
G Cereal	142	118	0	18043	1.02	1.42
KIA – κ =	0.9797		Overall accuracy =		98.69	
Statistics for the new HNN super-resolution mapping with PAN image registration error						
	R. Cereal	Grass	Asphalt	G. cereal	ErrorO (%)	ErrorC (%)
Unclassified	42	92	57	32	0	1.00
Ripe Cereal	42234	53	0	131	0.69	0.44
Grass	111	34391	343	24	1.55	1.37
Asphalt	0	223	998	0	28.61	18.26
G Cereal	142	173	0	18041	1.03	1.72
KIA – κ =	0.9772		Overall accuracy =		98.53	

Table 4.2 Accuracy Statistics of Degraded QuickBird Results

4.4.2.3 Statistical evaluation

The advantage of the new technique was also demonstrated by the accuracy statistics. Overall accuracy of the land cover map increased by around 2% from 96.50% for the hard classification to 98.69% for the super-resolution mapping using the PAN image without image registration error. For the PAN image with an image registration error of 1 pixel, the accuracy of the resulting 0.64 m land cover map increased around 2%

and 0.25% in comparison with the results of hard classification and the HNN super-resolution mapping technique, respectively. The KIA value increased from 0.9454 for the hard classified map and 0.9726 for the traditional HNN sub-pixel map to 0.9797 for the super-resolution mapping using the PAN image without registration error and 0.9772 for the super-resolution mapping using the PAN image with registration error of 1 pixel. Although the increase in KIA and overall accuracy was not great, this still showed a more accurate results for the new method.

The accuracy statistics also demonstrated that amongst the four land cover classes, the accuracy of the asphalt class increased the most, with the omission error decreasing from 39.71% for the hard classified image and 27.69% for the traditional HNN super-resolution mapping to 21.96% for the new HNN super-resolution mapping technique using the PAN image without registration error. Although the omission error for the new method increased slightly from 27.69% for the traditional HNN super-resolution mapping to 28.61% due to the impact of the PAN image registration error of 1 pixel, the commission error reduced from 66.88% for the hard classification and 44.13% for the traditional HNN technique to 18.26%. For the other three classes, where most sub-pixels were grouped into larger objects, the increase in accuracy was not as great as that of the asphalt class with the commission and omission errors decreasing by around 0.5%. However, it is still a considerable improvement over hard classification and the HNN super-resolution mapping without PAN image if the already very high accuracy of these classes is taken into account.

4.4.2.4 Image registration error effect

The impact of the PAN image registration error was determined based on the variation of the KIA value statistics. The variation of the KIA values obtained by applying the new method to the PAN images with RMS image registration errors ranging from 0.5 pixels to 1.5 pixels is presented in Figure 4.9. It is logical that the KIA value decreases as the RMS image registration error increases. However, with a RMS error of 1 pixel (the accuracy of image registration that can commonly be obtained in the geometric correction of remotely sensed images) the KIA value for the sub-pixel map produced by the new method was still greater than that produced by the traditional HNN super-resolution mapping technique. When the RMS image registration error increased to 1.5 pixels, the resulting sub-pixel map predicted by the new technique was less accurate than the results of the HNN without using the PAN image. It is

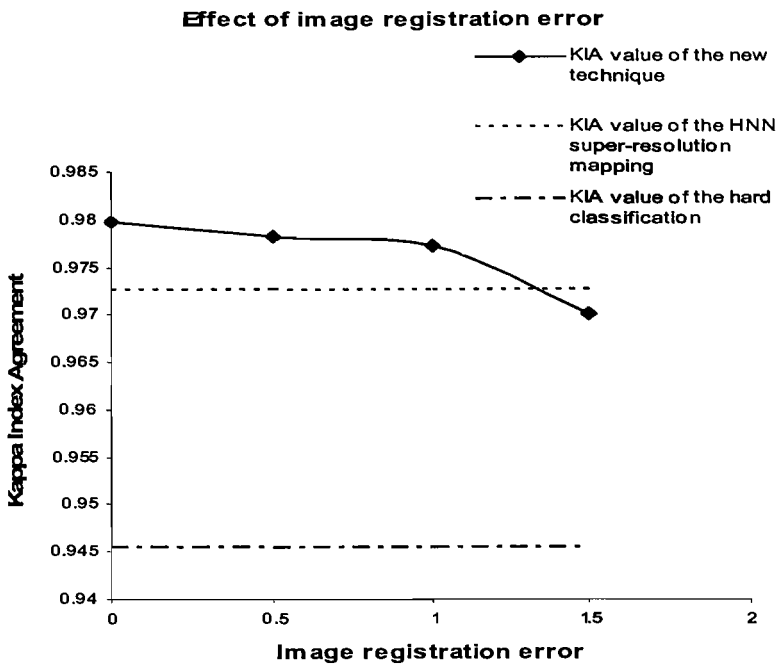


Figure 4.9 The effect of image registration error on Kappa Index of Agreement value of the sub-pixel map produced using the HNN super-resolution mapping using the PAN image.

therefore recommended that the new technique may be applicable even with PAN images that contain image registration error. Further research will be necessary to explore the range of situations in which the technique might be operationally applicable.

4.5 Conclusions

This chapter introduces the use of panchromatic imagery as supplementary data for super-resolution mapping. Information from the PAN images was incorporated into the HNN optimisation in the form of a panchromatic reflectance constraint, which was added to the energy function. The value of the panchromatic reflectance constraint was estimated based on forward and inverse models using local end-member spectra and a local spectral convolution weighting factor. The technique was examined using both simulated IKONOS data and a degraded QuickBird image (with and without image registration error). The accuracy of the results was evaluated based on the error matrix and accuracy statistics such as the KIA value, overall accuracy, and omission and commission errors.

The results from both simulated and real data sets demonstrated that PAN images can be used as a source of supplementary information for the HNN to predict land cover at a sub-pixel spatial resolution. Both the qualitative and quantitative analysis showed that the new technique can increase the accuracy of all land cover classes, but most significantly for land cover features at the sub-pixel scale. The results also showed that the algorithms can produce a sub-pixel map accurately with PAN imagery containing an image registration error of around one pixel.

Chapter 5: Super-resolution of multispectral imagery based on the HNN

5.1 Introduction

Chapters 3 and 4 introduced two new methods for using information from fused and panchromatic imagery to increase the detail and accuracy of a sub-pixel super-resolved land cover map. Both of the two methods used synthetic multispectral (MS) images, which were produced from a sub-pixel land cover map using a forward model, as an intermediate step for calculating the reflectance constraint value. This suggested that a synthetic super-resolved MS image can be predicted from the sub-pixel land cover map using the same forward model. This synthetic sub-pixel image prediction process is referred to as MS image super-resolution.

The task of image super-resolution is to increase the spatial resolution of the imagery. In fact, image super-resolution commonly refers to the process of combining a set of coarse spatial resolution images of the same scene to obtain a single fine resolution image. There have been a large number of studies on super-resolution. Examples can be cited such as Elad and Feuer (1999), Freeman *et al.* (2002), and Tipping and Bishop (2003). Although widely applied in image processing, these approaches were hardly applicable for super-resolution of remotely sensed MS imagery because of the lack of a sequence of images in a scene at the same time. Amongst the remotely sensed data sources, the super-resolution approaches using image sequences are more applicable to hyperspectral imagery (Akgun *et al.*, 2005). For common multispectral remotely sensed imagery, several methods for increasing the spatial resolution have been proposed such as a Point Spread Function-derived convolution filter (Pinilla Ruiz and Ariza Lopez, 2002) and a segmentation technique (Schneider and Steinwendner, 1999).

In this chapter, three methods for increasing the spatial resolution of the MS image are introduced. Firstly, a new model for predicting the MS image at the sub-pixel spatial resolution is presented. The new model is based on the super-resolution mapping technique, soft-classification and a forward model using the local end-member spectra. The soft-classification is based on unsupervised classification to reduce the land cover proportion prediction error. Secondly, a method of semivariogram matching is presented. The method is based on the HNN and prior

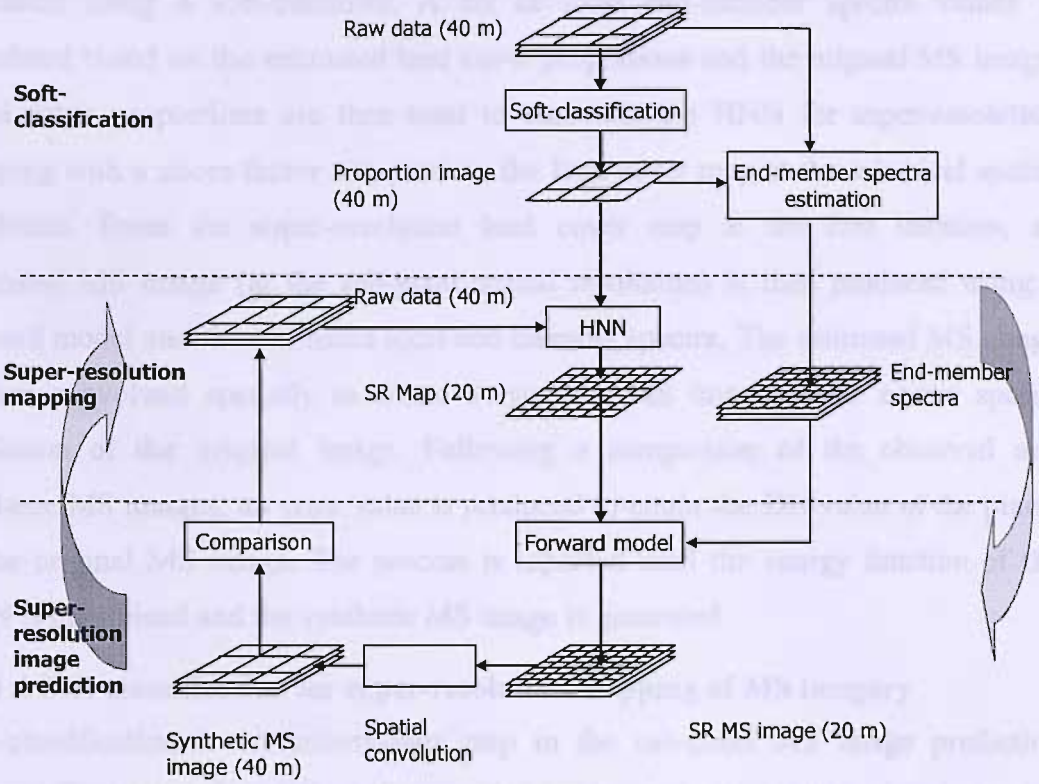


Figure 5.1 General model for super-resolution of MS imagery.

semivariance values to generate a sub-pixel image with the spectral features of the coarse spatial resolution original image and typical spatial character at the sub-pixel scale. Finally, a new method for image interpolation, or smoothing based on the HNN and zero semivariance values, is introduced. The method utilises the smoothing effect of minimising the values between neighbours process to produce a smooth sub-pixel MS image with smaller RMSE.

5.2 Super-resolution mapping and forward model for super-resolution of MS imagery

5.2.1 General model

The proposed model is an extension of super-resolution mapping using HNN optimisation. The prediction of a synthetic MS image at the sub-pixel spatial resolution is based on the forward model with local spectra which was used in Chapters 3 and 4. In addition to the goal functions and the proportion constraint of the HNN for super-resolution mapping, a reflectance constraint is used to retain the digital number (DN) values of the original MS image.

Figure 5.1 presents the HNN sub-pixel MS image prediction algorithm. From the MS images at the original MS spatial resolution, land cover area proportion images are

estimated using a soft-classifier. A set of local end-member spectra values is calculated based on the estimated land cover proportions and the original MS image. Land cover proportions are then used to constrain the HNN for super-resolution mapping with a zoom factor z to produce the land cover map at the sub-pixel spatial resolution. From the super-resolution land cover map at the first iteration, an estimated MS image (at the sub-pixel spatial resolution) is then produced using a forward model and the estimated local end-member spectra. The estimated MS image is then convolved spatially to create a synthetic MS image at the coarse spatial resolution of the original image. Following a comparison of the observed and synthetic MS images, an error value is produced to retain the DN value of the pixels of the original MS image. The process is repeated until the energy function of the HNN is minimised and the synthetic MS image is generated.

5.2.1.1 Soft-classification for super-resolution mapping of MS imagery

Soft-classification is an intermediate step in the sub-pixel MS image prediction process. The estimation of the MS image is based on super-resolution mapping using proportions which are obtained from soft-classification. Conventionally, there must be a set of training data for some soft-classifiers. Accordingly, it is necessary to have some prior information about the spectral distribution of land cover classes in the MS bands but training data are not always available for the scene. However, the land cover map is only an intermediate step here, and is not the target of the proposed algorithm. Therefore, soft-classification based on unsupervised classification should be suitable in this case.

The soft-classification that is applicable for the sub-pixel MS image prediction is the unsupervised-based fuzzy c-means classifier (Benzek *et al.*, 1999). Alternative supervised-classification-based soft-classifiers could also be used such as Bayesian, neural network or k-NN classifiers. However, the training data for these soft-classification techniques should be obtained from the unsupervised classifications. In the experiment in this chapter, the MS image was at first clustered into spectral classes using the unsupervised Iterative Self-Organizing Data (ISODATA) classifier. The training data based on these spectral classes were then used for the soft-classifiers to estimate the land cover proportions. By categorising the classes based on statistical clustering, the within-class spectral variation is minimised. As a result, the error in the predicted proportions is reduced.

5.2.1.2 Forward model and end-member spectra

The forward model is used to produce a sub-pixel MS image from the sub-pixel land cover classes. The DN value of a sub-pixel (m,n) of a spectral band s can be predicted as

$$R_{smn} = V_1 S_{s,1} + V_2 S_{s,2} + \ldots + V_c S_{s,c} \tag{5.1}$$

where V_e is the output neuron of the class e and $S_{s,e}$ is the end-member spectra of the land cover class e for a spectral band s. As presented in Chapter 3 and Chapter 4, the end-member spectra vector $\mathbf{S}_s$ ($\mathbf{S}_s = [S_{s,1},..,S_{s,e},..,S_{s,c}]$) of the original pixel (x,y) of the spectral band s can be estimated locally using the predicted land cover class proportions and the MS image at the original coarse spatial resolution (Equation (3.19) and (3.20)) as

$$\mathbf{S}_s = \left(\mathbf{P}^T\mathbf{W}\mathbf{P}\right)^{-1}\mathbf{W}\mathbf{P}^T\mathbf{R}_s ,$$

where $\mathbf{P}$ is a matrix of land cover proportions

$$\mathbf{P} = \begin{bmatrix} P_1^{(x-1)(y-1)} & .. & P_c^{(x-1)(y-1)} \\ .. & .. & .. \\ P_1^{xy} & .. & P_c^{xy} \\ .. & .. & .. \\ P_1^{(x+1)(y+1)} & .. & P_c^{(x+1)(y+1)} \end{bmatrix}$$

and $\mathbf{W}$ is matrix of weights

$$\mathbf{W} = \begin{bmatrix} w^{(x-1)(y-1)} & 0 & 0 & 0 & 0 \\ 0 & .. & 0 & 0 & 0 \\ 0 & 0 & w^{xy} & 0 & 0 \\ 0 & 0 & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & w^{(x+1)(y+1)} \end{bmatrix}.$$

5.2.2 Evaluation of the super-resolved image

The target of the new method is to produce a sub-pixel spatial resolution MS image which is as close to the real MS image at the sub-pixel spatial resolution as possible. The resulting predicted MS image, therefore, must be evaluated based on statistics which quantify the closeness to the real MS image at the sub-pixel spatial resolution. In this case, the RMSE is appropriate for evaluation of the new method. The RMSE of a spectral band s can be calculated as

$$RMSE_s = \sqrt{\frac{\sum_{i=1}^{n}\left(DN_{s,i}^{Real} - DN_{s,i}^{Predicted}\right)^2}{n}} \tag{5.2}$$

where n is number of pixels used to calculate the RMSE, DN_i^{Real} is DN value of pixel i of the real spectral band s and $DN_i^{Predicted}$ is DN value of the pixel i of the predicted spectral band s. For all b spectral bands, the RMSE can be defined as

$$RMSE = \frac{1}{b}\sum_{s=1}^{b} RMSE_s \tag{5.3}$$

5.3 Experiment

5.3.1 Data

Similar to the evaluation process for the other image fusion techniques, a degraded image produced from a remotely sensed image was used in this experiment. Since the reference fine spatial resolution image being known, it was possible to evaluate the feasibility of the newly proposed approach. The experiment was accomplished on a SPOT MS image of Southampton and the Isle of Wight, UK.

5.3.1.1 Reference data

SPOT MS imagery with a spatial resolution of 20 m was acquired over an area of Southampton and the Isle of Wight, UK on 28th June 1994. A sub-area of 512×512 pixels (Figure 5.2) was extracted from a rural area. The SPOT MS image consists of three spectral bands: Green (0.50-0.59 μm), Red (0.61-0.68 μm) and NIR (0.78-0.89μm) (Figure 5.2(a), (b) and (c)).

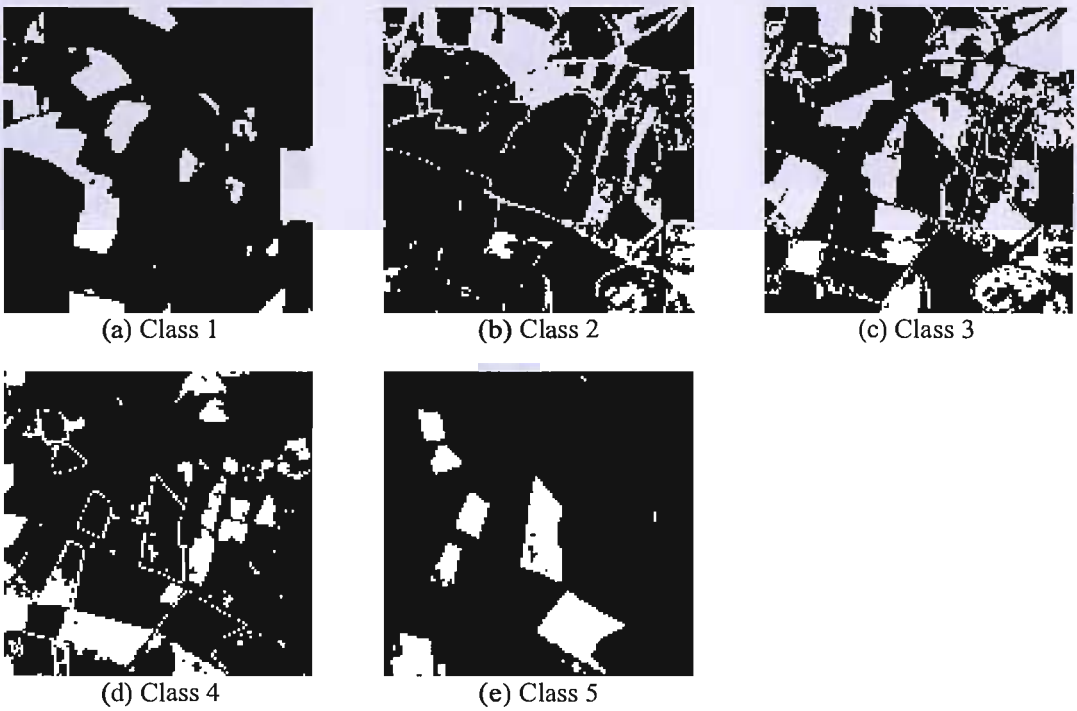
5.3.1.2 Degraded images and soft-classification

Training and Reference data: The land cover proportions can be estimated using training imagery obtained by the unsupervised classification. To assess the accuracy of the proportion estimation, the reference land cover classes are required. These reference land cover classes were produced based on the unsupervised ISODATA classification of the SPOT image. Five clusters were obtained as shown in Figure 5.3. These classes were then used for soft-classification to estimate the land cover proportions.

Degraded images: The HNN super-resolution algorithm was tested using degraded MS images obtained from the reference MS images in Figure 5.2. To evaluate the effect of the super-resolution zoom factor on the RMSEs of the resulting super-resolved images, the reference MS images were degraded by a factor of two to 40 m spatial resolution (Figure 5.4(a), 5.4(b) and 5.4(c)), a factor of three to 60 m spatial

(a) Green band (b) Red band (c) NIR band

Figure 5.5.2 Reference 20 m MS image



(a) Class 1 (b) Class 2 (c) Class 3

(d) Class 4 (e) Class 5

Figure 5.3 Unsupervised classes

resolution (Figure 5.4(d), 5.4(e) and 5.4(f)) and a factor of four to 80 m spatial resolution (Figure 5.4(g), 5.4(h) and 5.4(i)) .

Soft-classification: The soft-classification was used to estimate land cover proportions from the 40 m, 60 m and 80 m degraded MS images. The estimated proportions were then used to predict sub-pixel land cover map and sub-pixel MS image at the 20 m spatial resolution (the spatial resolution of the reference MS

image). In this experiment, the k-NN soft-classifier with $k = 5$ was used to produce the proportions from the degraded images in Figure 5.4. Since the training classes were

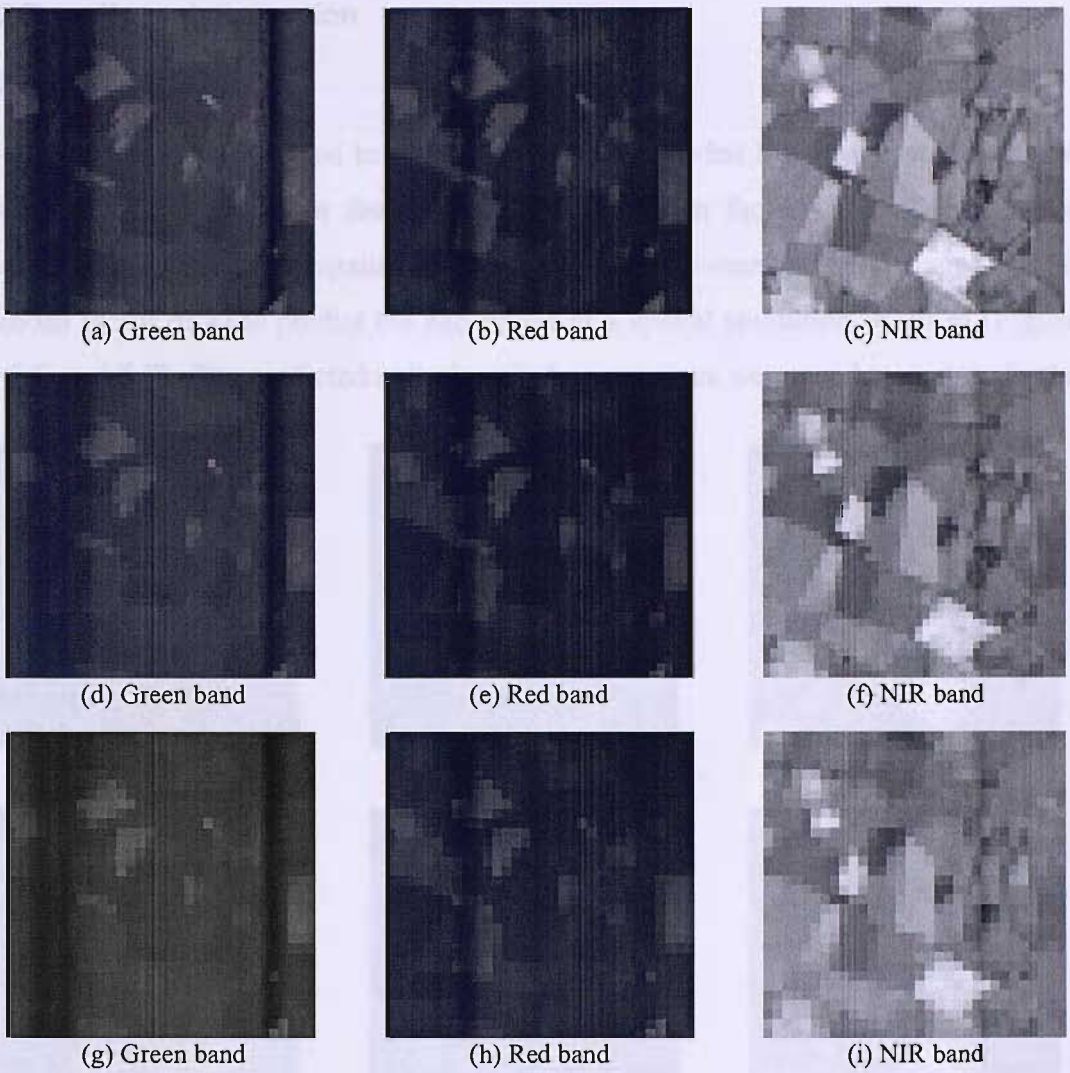


Figure 5.4 Degraded MS images:
(a), (b) and (c) - The 40 m degraded MS image. (d), (e) and (f) - The 60 m degraded MS image. (g), (h) and (i) - The 80 m degraded MS image.

MS image (Resolution)	Overall area proportion area (%)	Overall RMS error
40 m	0.0327	0.0482
60 m	0.0538	0.0487
80 m	0.0805	0.0533

Table 5.1 Accuracy statistics of the soft-classification.

clustered using an unsupervised method based on the spectral distribution, the proportions were estimated accurately (Table 5.1) and the effect of proportion estimation error on the sub-pixel super-resolution image was reduced as a result.

5.3.2 Results and discussion

5.3.2.1 Results

The new method was applied to super-resolve the degraded MS image at the spatial resolutions of 40 m (zoom factor of 2), 60 m (zoom factor of 3. Zoom factor represents the increasing in spatial resolution of the super-resolved MS image) and 80 m (zoom factor of 4) to predict the MS image at a spatial resolution of 20 m (Figure 5.5, 5.6 and 5.7). The predicted soft-classified proportions were used to constrain the

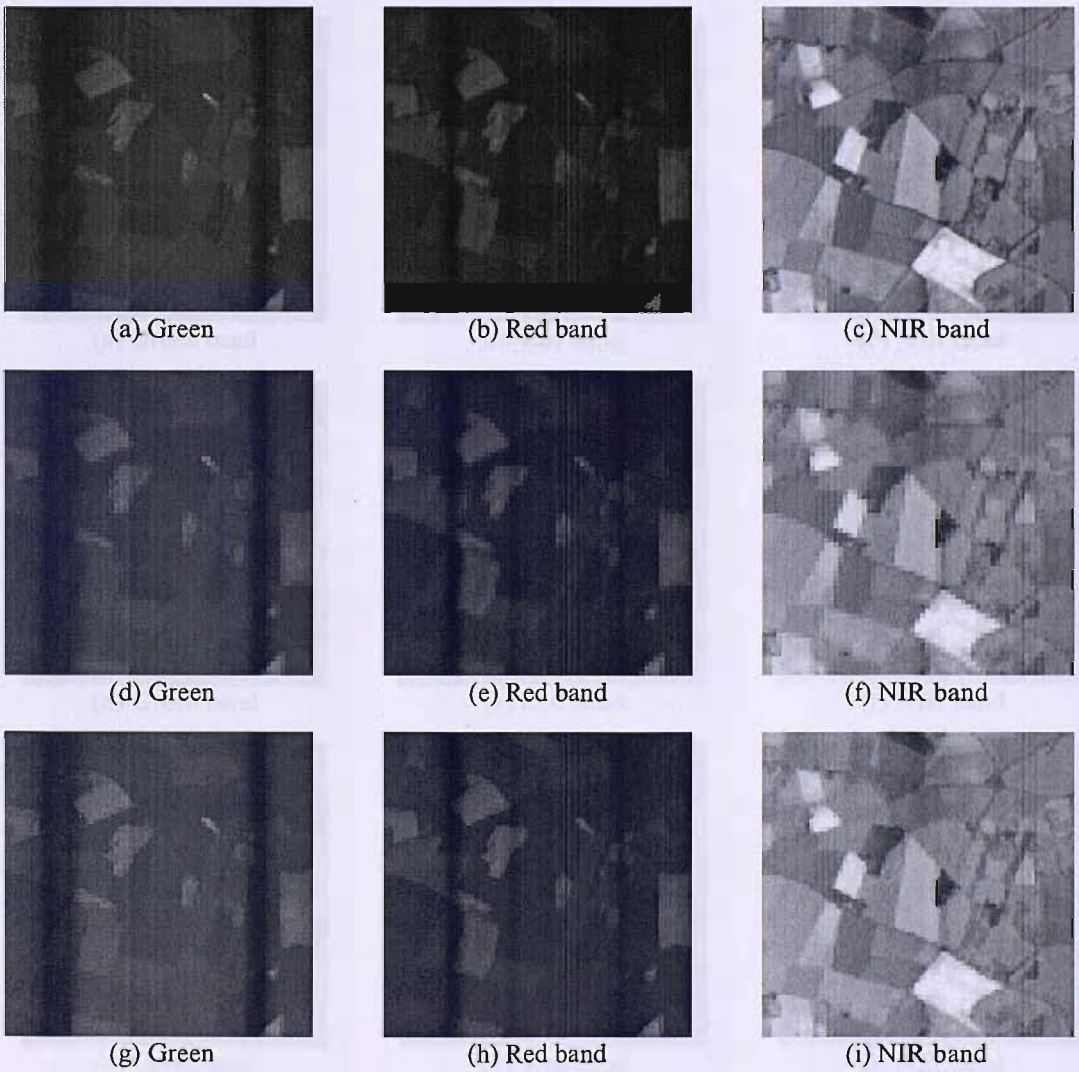
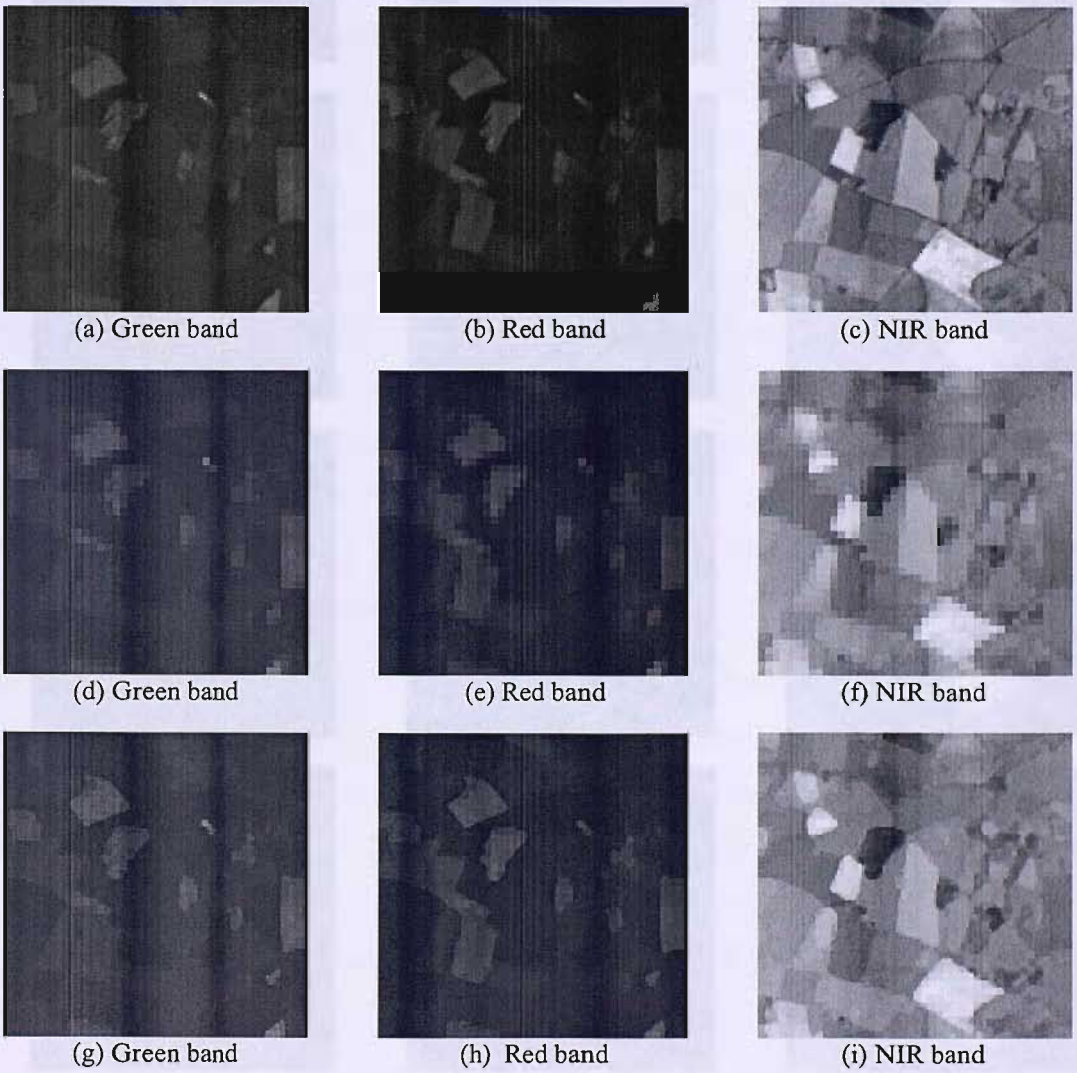


Figure 5.5 Super resolution of the 40 m MS image:
(a), (b) and (c) - The 20 m reference MS image. (d), (e) and (f)- The 40 m degraded image. (g), (h) and (i) - The 20 m super-resolved image.

HNN with weighting factors of k_1 = 100, k_2 =100, k_3 = 100 and k_4 = 100 to predict the sub-pixel land cover and then the 20 m MS image. The resulting 20 m MS images were then compared with the reference 20 m MS image and the RMSE for each spectral band was calculated (see Tables 5.2, 5.3 and 5.4).

Band	RMSE	
	Degraded image	Super-resolved image
Green	3.1276	2.6933
Red	3.2049	2.7655
NIR	7.3703	5.9094
Σ	13.7028	11.3682

Table 5.2 RMSE of the spectral bands of the 40 m degraded MS and the resulting 20 m super-resolved images.



(a) Green band (b) Red band (c) NIR band

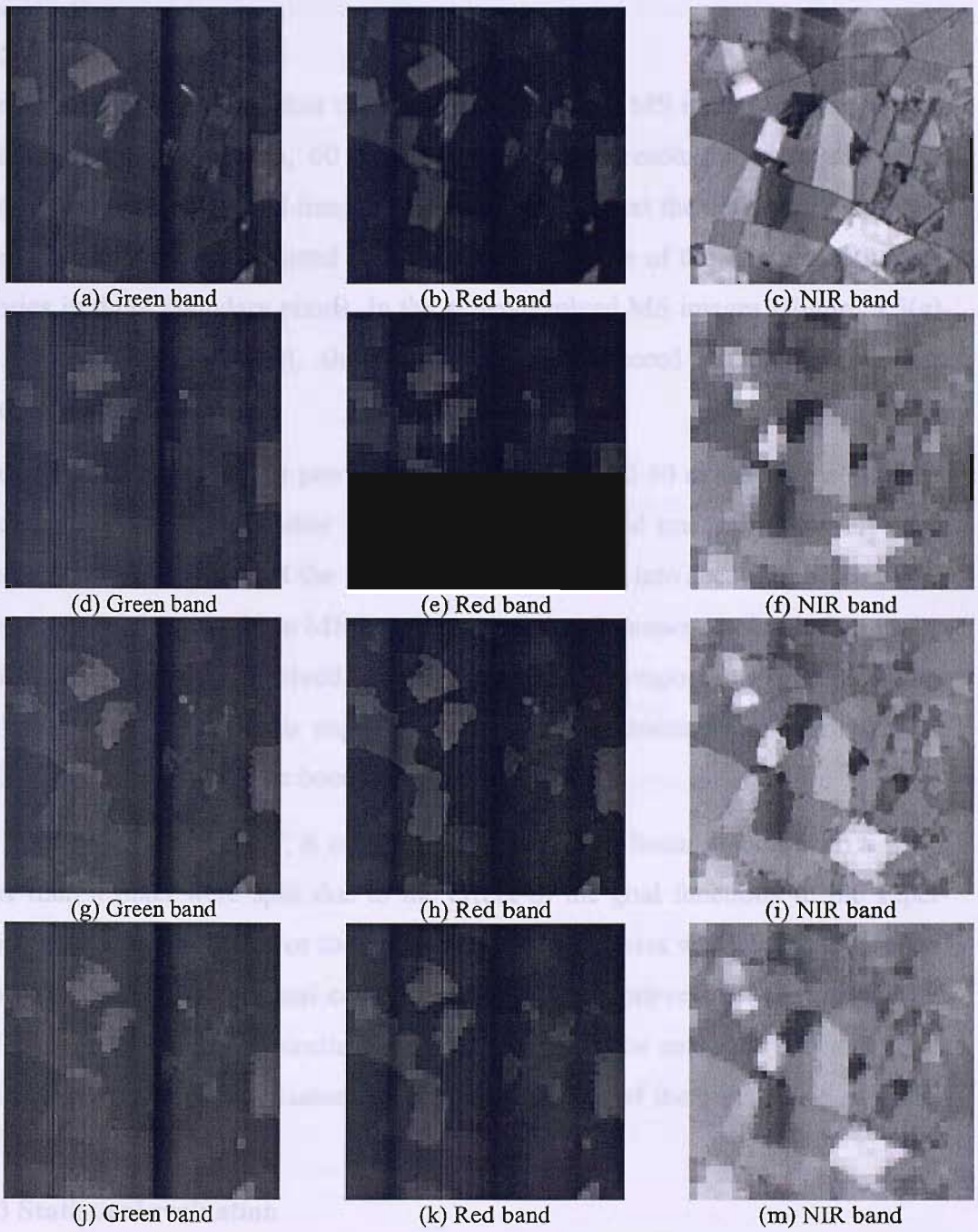
(d) Green band (e) Red band (f) NIR band

(g) Green band (h) Red band (i) NIR band

Figure 5.6 Super resolution of the 60 m MS image
(a), (b) and (c) - The 20 m reference MS image. (d), (e) and (f) - The 60 m degraded image. (g), (h) and (i) - The super-resolved 20 m image.

Band	RMSE	
	Degraded image	Super-resolved image
Green	4.2126	3.6575
Red	4.5125	3.8873
NIR	10.1797	7.1811
Σ	18.9047	14.7258

Table 5.3 RMSE of the spectral bands of the 60 m degraded MS and the resulting 20 m super-resolved images.



(a) Green band (b) Red band (c) NIR band

(d) Green band (e) Red band (f) NIR band

(g) Green band (h) Red band (i) NIR band

(j) Green band (k) Red band (m) NIR band

Figure 5.7 Super resolution of the 80 m MS image
(a), (b) and (c) - The 20 m reference MS image. (d), (e) and (f) - The 80 m degraded image. (g), (h) and (i) - The super-resolved 20 m image. (j), (k) and (m) - The 40 m super-resolved image.

Band	RMSE		
	Degraded image	Super-resolved image (20 m with $z = 4$)	Super-resolved image (40 m with $z = 2$)
Green	5.4117	4.6645	4.7067
Red	5.9209	4.9374	4.9379
NIR	11.9074	10.2393	10.6188
Σ	23.2400	19.8412	20.2633

Table 5.4 RMSE of the spectral bands of the 80 m degraded MS image and the resulting 20 m and 40 m super-resolved images.

5.3.2.2 Visual evaluation

Figure 5.5, 5.6 and 5.7 show that the super-resolved 20 m MS images produced from the degraded image at 40 m, 60 m, and 80 m spatial resolution were sharper in comparison with the degraded images. Figure 5.4 shows that the boundary features in the degraded images were blurred and fragmented because of the mixing of the land categories in these boundary pixels. In the super-resolved MS images (Figure 5.5(g)-(i), 5.6(g)-(i), and 5.7(g)-(i)), these features were restored to make the image apparently sharper.

The resulting 20 m MS image predicted from the degraded 40 m spatial resolution is apparently sharper than the other two 20 m super-resolved images. However, when the coarse spatial resolution of the original images is taken into account, the 20 m MS image generated from the 80 m MS image is most visually super-resolved. Comparing the three resulting super-resolved images with their corresponding original coarse spatial resolution images, it is suggested that when the zoom factor increases, the resulting super-resolved image becomes sharper.

From Figures 5.5, 5.6 and 5.7, it can also be seen that the linear objects with a width smaller than a pixel were split due to the effect of the goal functions in the super-resolution mapping process. For the pure pixels, the sharpness was the same in the 20 m super-resolved as the original coarse images, as the super-resolution mapping has no effect on pure pixels. Theoretically, this problem can be solved if the land cover class is segmented into sub-classes, hence defining some of the pure pixels as mixed pixels.

5.3.2.3 Statistical evaluation

Table 5.2, 5.3 and 5.4 illustrates the reduction in the RMSE of the super-resolved images in comparison with their corresponding coarse degraded image for every

spectral band. As with the visual evaluation, the statistics show the correlation between an increase of the zoom factor and the decrease in the RMSEs of the resulting super-resolved MS images. In comparison with the RMSE of the original 40 m degraded MS image, the RMSE of the 20 m super-resolved image (zoom factor of two) reduced by 2.3346 DN from 13.7028 DN to 11.3682 DN. When the zoom factor increased to four (down-scaling from 80 m spatial resolution to 20 m spatial resolution), the decrease for the RMSE rose to 3.3988 DN. For the same original 80 m degraded coarse image, the RMSE of the 20 m super-resolved image (19.8412 DN) was smaller than that of the 40 m super-resolved image (20.2639 DN).

5.4 Super-resolution of MS imagery using semivariogram matching

5.4.1 General model

In section 5.3, a model for super-resolution mapping of MS imagery based on the HNN and forward model was introduced. Super-resolution mapping using the HNN and forward model can predict the MS image at sub-pixel spatial resolution from the

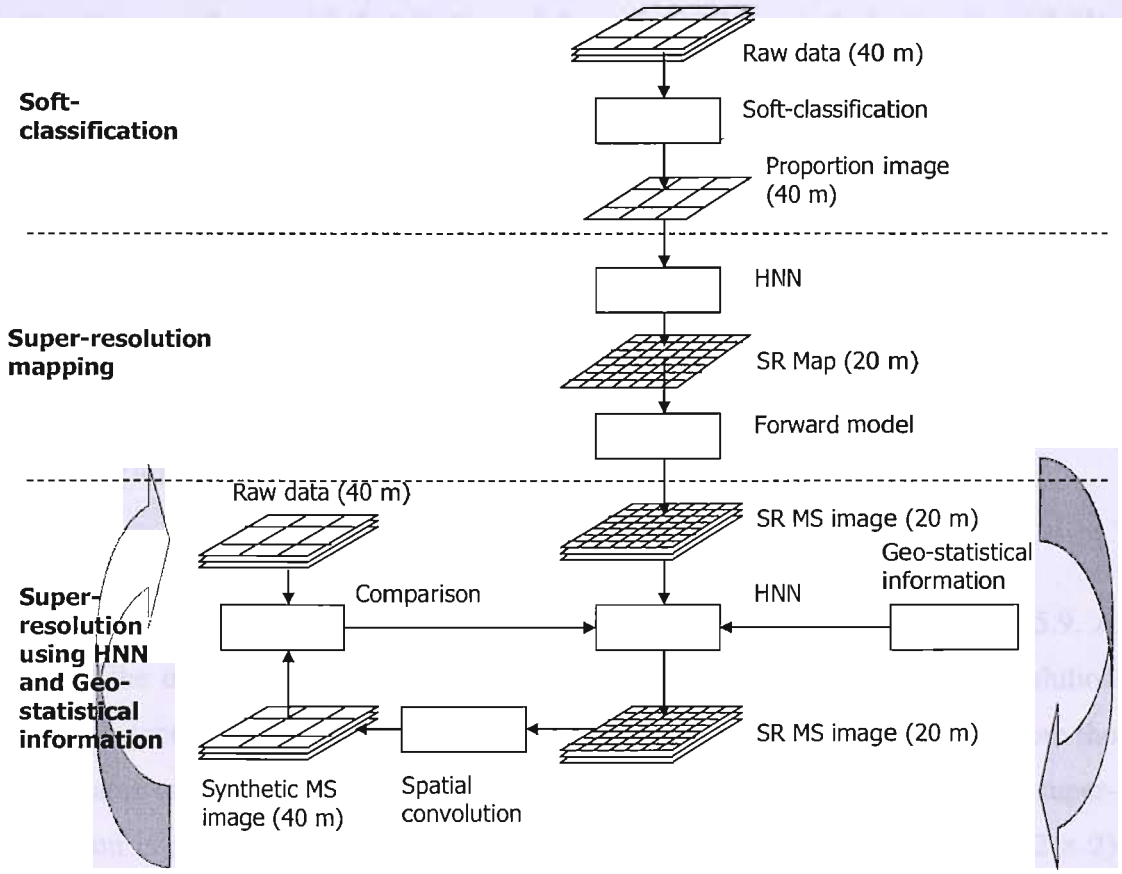


Figure 5.8 Super-resolution and semivariogram matching

degraded MS image with smaller RMSE and sharper visual appearance for mixed pixels. However, for pixels in the same class, the method does not increase the spatial resolution because the classification defines these pixels as pure pixels. To obtain further super-resolution for some of these pixels, the same algorithm can be applied when a land cover class is re-categorised into several sub-classes and a number of pure pixels will be re-defined as mixed pixels as a result. Obviously, this is not a comprehensive solution because pure pixels still exist in these sub-classes. Thus, the process of super-resolution using a HNN and the forward model is similar to image segmentation.

This section introduces a geostatistics-based method to obtain sub-pixel spatial resolution for pure pixels that cannot be super-resolved by super-resolution mapping. The method employs the HNN and prior geostatistical information in the form of semivariograms to produce a sub-pixel image with desirable spatial statistics. Figure 5.8 shows process of the method. The method can utilise the output image of the HNN super-resolution for MS image or coarse spatial resolution images as input data. In that input image, the land cover classes are already defined. The assumption is that information on the spatial distribution of these defined spectral classes is available. For example: if class A in the image is defined as cereal, then semivariograms of cereal can be extracted from other sources of available data such as air photo or field surveying. The prior semivariograms and the input MS image are then used in the HNN to produce sub-pixel images with the desirable spatial character. In this HNN model, the input MS image is used to constrain the network in the form of a reflectance constraint. *Semivariance functions* are used to change the output values of neurons to match the available semivariograms. The method, therefore, can be named as *semivariogram matching* or *pattern matching.*

5.4.2 HNN structure

The structure of the HNN for semivariogram matching is depicted in Figure 5.9. A pixel of the original image is divided into ($z \times z$) sub-pixels in the super-resolution map (z is zoom factor). The coordinates of a sub-pixel (m,n) are determined from the coordinates of the original pixel (x,y) as in Figure 5.9. Each sub-pixel in the super-resolution image is represented by a neuron in the HNN. Thus, an image of (2×2)

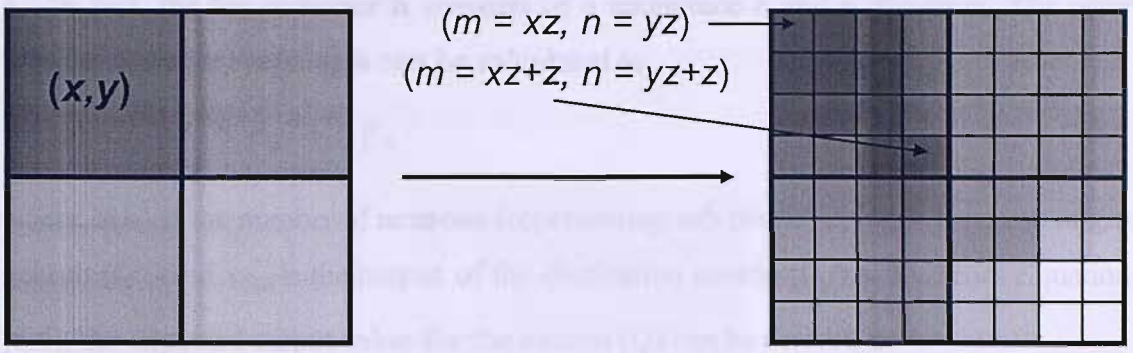


Figure 5.9 Structure of the HNN for semivariogram matching
(x,y) is coordinates of pixels in the original image, (m,n) is coordinates of sub-pixel in the super-resolution image, z is zoom factor.

pixels in the original image can be represented by a matrix of (8 × 8) interconnected neurons in the HNN with the zoom factor $z = 4$.

The HNN for semivariogram matching comprises a reflectance constraint and a number of semivariance functions. The HNN is initialised using the DN values of the original image and it runs to a stable state in which the energy function is minimised and the output values of the nodes (sub-pixels) represent the DN values of the semivariogram matching image. The energy function of the HNN is defined as

$$E = -\sum_i \sum_j \left(k_r R_{ij} + \sum_{m=1}^{M} k_m S_{m,ij} \right), \tag{5.4}$$

where R_{ij} is the reflectance constraint of the neuron (i,j), $S_{m,ij}$ is the semivariance function for the semivariance m of the neuron (i,j), M is the number of prior semivariances (in this thesis $M = 32$), and k_r and k_m are the weighting constants.

The reflectance constraint retains the DN values of the original image. This is based on the assumption that the reflectance of a pixel in the original image is an average of the reflectance of the corresponding sub-pixels (represented by the neurons in the HNN). The reflectance constraint of a neuron (i,j) can be determined as

$$\frac{dR_{ij}}{dv_{ij}} = \frac{1}{z^2} \sum_{d=xz}^{xz+z-1} \sum_{e=yz}^{yz+z-1} (v_{de}) - DN_{xy}, \tag{5.5}$$

where v_{de} is the output of the neuron (d,e) at time t, z is the zoom factor, and DN_{xy} is the DN value of pixel (x,y).

The semivariance function for neuron (i,j) describes the semivariance calculated from the output of neuron (i,j) for a lag **h** and must be matched with the prior semivariance

γ_m. In fact, the lag or vector **h** consists of a magnitude h and a direction. The prior semivariance for every lag **h** can be calculated as

$$\gamma_m(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{1}^{N(\mathbf{h})} [v_{ij} - v_{ij+\mathbf{h}}]^2 , \tag{5.6}$$

where $N(\mathbf{h})$ is the number of neurons (representing sub-pixels) at lag **h** from the origin neuron (i,j), and $v_{ij+\mathbf{h}}$ is the output of the destination neuron $((i,j) + \mathbf{h})$. From Equation (5.6), the expected output value for the neuron (i,j) can be derived as

$$v_{ij}^{expected} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \tag{5.7}$$

where $a = N(\mathbf{h})$, $b = -2 \sum_{1}^{N(\mathbf{h})} v_{ij+\mathbf{h}}$ and $c = \sum_{1}^{N(\mathbf{h})} v_{ij+\mathbf{h}}^2 - 2\gamma_m N(\mathbf{h})$.

The value returned by the semivariance function can be calculated as

$$\frac{dS_{m,ij}}{dv_{ij}} = v_{ij} - v_{ij}^{expected} . \tag{5.8}$$

5.4.3 Semivariogram matching for degraded MS image

5.4.3.1 Data

A degraded QuickBird image was used to evaluate the feasibility of the proposed method of semivariogram matching. The QuickBird image, which was acquired over an area of Christchurch, UK on 1st June 2002, consisted of four spectral bands at 2.5 m spatial resolution. The experiment was accomplished on a 36 × 40 pixels reference image of the red band of the QuickBird image. From the reference MS image (Figure 5.10 (a)), a number of training, or prior, semivariance values were extracted (Figure 5.10(d)). The testing experiment for the new method utilised eight semivariograms based on directions: North, North-East, East, South-East, South, South-West and West (Figure 5.11). The values of semivariance for the opposite directions were the same. In this experiment, semivariance values based on four lags were used to determine these eight semivariograms. Thus, 32 semivariance functions as Equation (5.8) were included in the HNN energy function (Equation (5.4)) to implement the semivariogram matching.

The degraded MS image (Figure 5.10(b)) was created by degrading the reference image (Figure 5.10(a)) by four times to produce the 10 m MS image. This degraded image was then used by the HNN to predict the MS image at a 2.5 m spatial resolution using the prior semivariance values in Figure 5.10(d). The semivariance

values of the degraded image were smaller than those of the original image since the spatial variability of neighbouring data points were decreased by the image degrading process. The semivariograms for the degraded image can be seen in Figure 5.10(e).

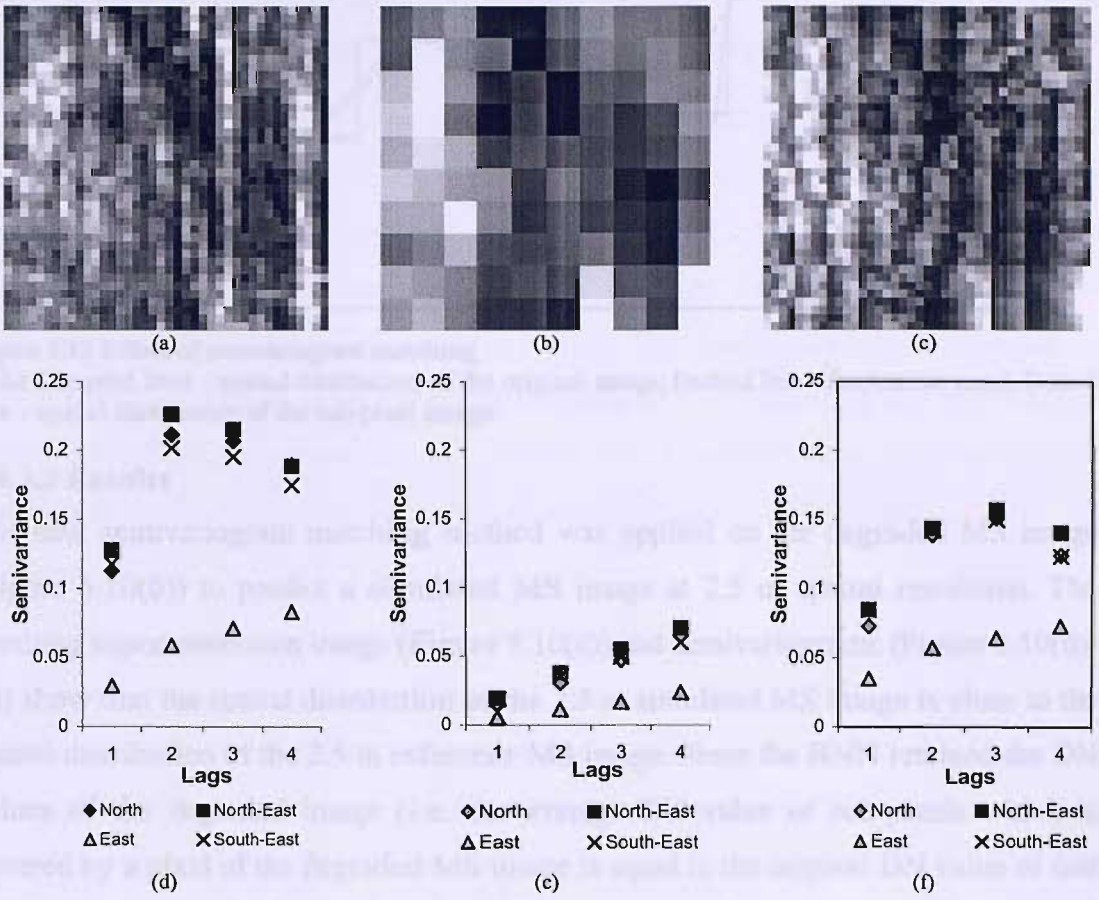


Figure 5.10 Sub-pixel semivariogram matching of a degraded MS image
(a) Reference 2.5 m QuickBird MS image, (b) Degraded image 10 m QuickBird image, and (c) 2.5 m simulated MS image using semivariogram matching. (d) Training semivariograms, (e) semivariograms of the degraded image, and (f) Semivariograms of the resulting 2.5 m simulated MS image.

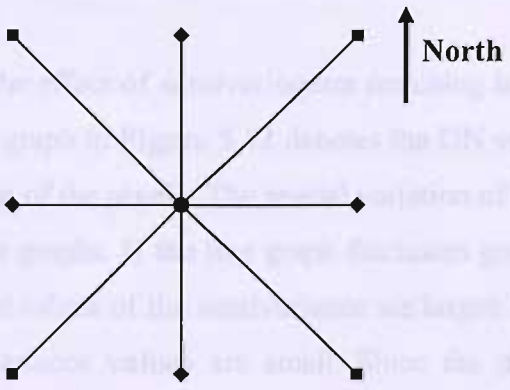


Figure 5.11 Calculation of semivariance values based on the supporting pixels in eight directions

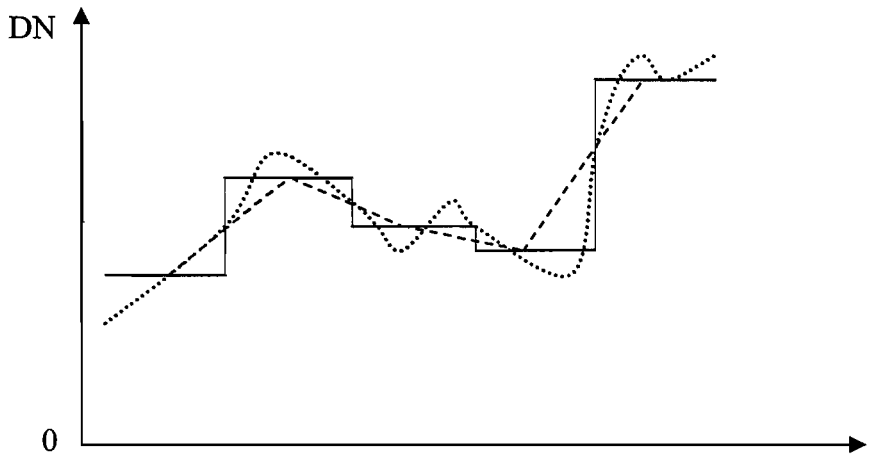


Figure 5.12 Effect of semivariogram matching
Solid disrupted lines – spatial distribution of the original image; Dashed line – fluctuation trend; Dotted line – spatial distribution of the sub-pixel image.

5.4.3.2 Results

The new semivariogram matching method was applied on the degraded MS image (Figure 5.10(b)) to predict a simulated MS image at 2.5 m spatial resolution. The resulting super-resolution image (Figure 5.10(c)) and semivariograms (Figure 5.10(d)-(f)) show that the spatial distribution of the 2.5 m simulated MS image is close to the spatial distribution of the 2.5 m reference MS image. Since the HNN retained the DN values of the degraded image (i.e. the average DN value of sub-pixels which is covered by a pixel of the degraded MS image is equal to the original DN value of that pixel), the resulting 2.5 m simulated MS image contains both the spectral features of the original degraded MS image and the prior spatial variation. With the patterns generated by the prior geostatistical information, the simulated sub-pixel image in Figure 5.10(c) was apparently more similar to the reference image than the degraded image (Figure 5.10(b)).

Figure 5.12 illustrates the effect of semivariogram matching in a single direction. The vertical axis in the line graph in Figure 5.12 denotes the DN values and the horizontal axis denotes the position of the pixels. The spatial variation of an image is depicted by the steepness of the line graphs. If the line graph fluctuates greatly, the image is more spatially variable (or the values of the semivariance are large). Otherwise, the image is smooth and the semivariance values are small. Since the pixel size of the coarse spatial resolution degraded image was large, the spatial distribution of the image was disrupted, as indicated by the solid lines in the plot in Figure 5.12. In the sub-pixel image, the increase of spatial resolution by semivariogram matching resulted in a

more continuous form of spatial distribution, which is depicted by the dotted curve in Figure 5.12. By matching with the prior semivariance values, the DN values of the sub-pixels were altered to smooth the degraded image where spatial variation is greater than that of the prior data. Conversely, the semivariogram matching increased the spatial variation in the smooth area of the degraded image.

Under the effect of the semivariance functions, the DN values of the sub-pixels were adjusted to match the prior semivariances. If only the semivariance functions were used, the DN values of the sub-pixels might be changed completely in some areas of the degraded image where the original spatial variation was too high or too low in comparsion with the prior semivariance values. This might make the dotted line in Figure 5.12 move away from the disrupted lines of the original spatial distribution. With the reflectance constraint, the dotted line is always kept attached to the fluctuation trend (the dashed line in the Figure 5.12) of the spatial distribution of the degraded image.

The weighting constants k_r and k_m control the effects of the reflectance constraint and semivariance functions on the optimisation process. If the weighting constant for the reflectance k_r was greater than the semivariance values k_m, the spatial distribution of the resulting sub-pixel image would be more similar to the spatial distribution of the degraded image. Otherwise, the network could more freely adjust the sub-pixel DN values to conform with the prior semivariograms. In this experiment, the optimal values for the weighting constants were selected based on the empirical assumption of balanced influences between the reflectance constraint and 32 semivariance functions. Supposing that the influences from the semivariance functions to the energy function are equal, the weighting constants were chosen as

$$k_r = \sum_{m=1}^{32} k_m \,. \tag{5.9}$$

If the weighting constant for the reflectance constraint k_r = 100, the weighting constant for the semivariance $k_m = 100/32 \cong 3.0$.

5.5 Combination of the HNN super-resolution using the forward model and semivariogram matching

The two methods for generating sub-pixel images from the coarse MS image in sections 5.3 and 5.4 can be combined. This combination is achieved by using images produced by the HNN super-resolution based on the forward model as input for the

semivariogram matching. In the experiment in this section, the super-resolved images from the HNN with forward model (Figure 5.5, 5.6 and 5.7) were matched with the prior geostatistical information extracted from the sub-pixel reference images in Figure 5.3. These input data included the 20 m super-resolution images produced from the degraded SPOT images at 40 m, 60 m and 80 m spatial resolution. Since the pure pixels were not super-resolved in the HNN super-resolution process, the DN values of the sub-pixels located within these pixels were the same as the original DN values. Thus, these DN values were utilised to constrain the corresponding neurons of the HNN as Equation (5.5). The semivariogram matching process, using 32 semivariance values of eight directions and 4 lags (as described in section 5.5.3), generated the simulated images as in Figure 5.13.

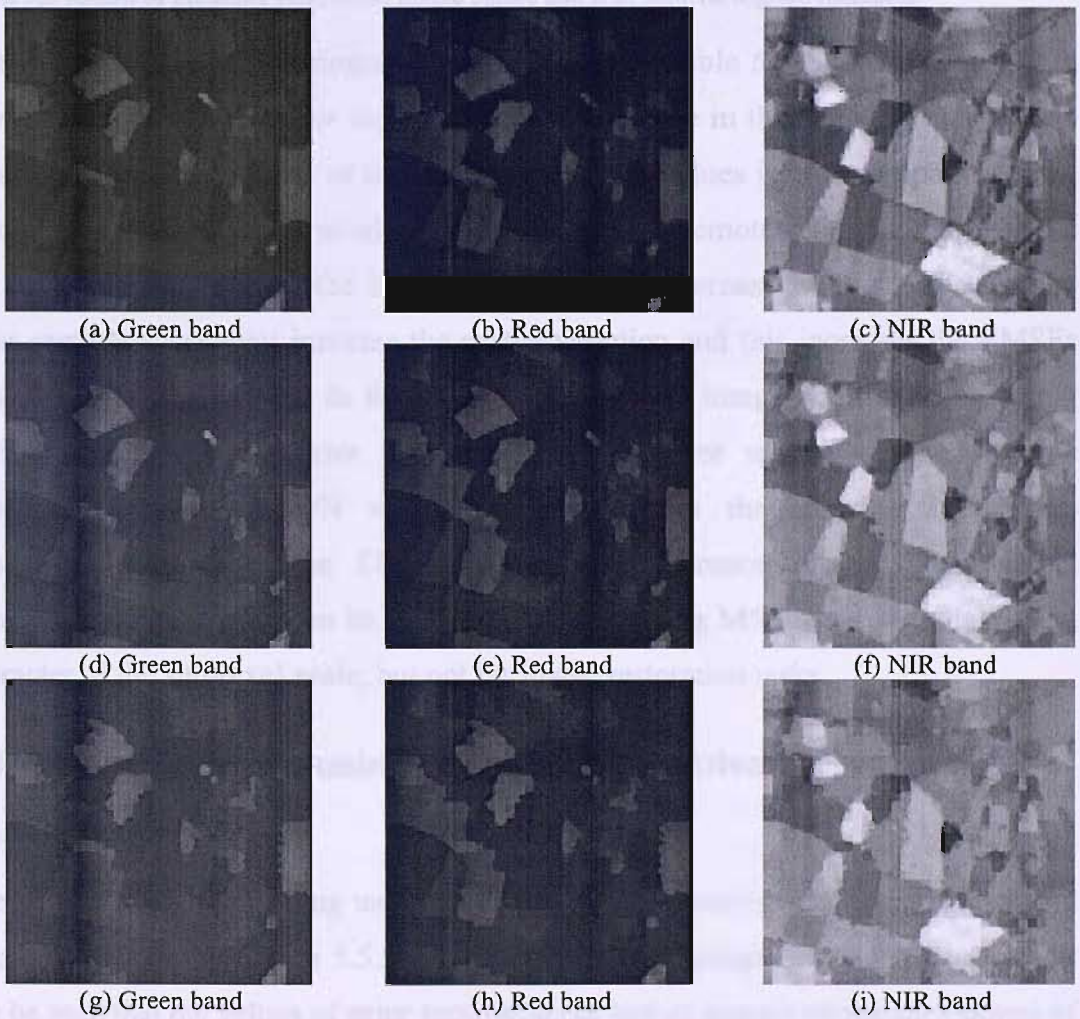


Figure 5.13 Sub-pixel images simulated from semivariogram matching
(a), (b) and (c) - The 20 m MS image simulated from the HNN super-resolved images of 40 m. (d), (e) and (f) - The 20 m MS image simulated from the HNN super-resolved image of 60 m. (g), (h) and (i) - The 20 m MS image simulated from HNN super-resolved image of 80 m.

Image		RMSE			
		Band 1	**Band 2**	**Band 3**	$\sum$
20 m MS image predicted from 40 m MS image	Before semivariogram matching	2.6934	2.7655	5.9094	11.3682
	After semivariogram matching	2.7614	2.8030	5.8933	11.4577
20 m MS image predicted from 40 m MS image	Before semivariogram matching	3.6575	3.8873	7.1811	15.6850
	After semivariogram matching	3.6200	4.2132	7.2153	15.0486
20 m MS image predicted from 40 m MS image	Before semivariogram matching	4.6645	4.9374	10.2393	19.8412
	After semivariogram matching	4.7067	4.9378	10.6188	20.2633

Table 5.5 RMSE of the super-resolution image before and after semivariogram matching.

The RMSEs of the semivariogram matching images (Table 5.5) increased slightly in comparison with those of the input images. The increase in the RMSEs is due to two reasons. Firstly, a single set of the prior semivariance values is not appropriate for the whole image because the spatial variation of a class in remotely sensed images is not the same in every area of the image. In some smooth areas, the high values of the prior semivariances may increase the spatial variation and this increases the RMSEs as a result. Secondly, even in the area of the reference image with the same spatial variation (presented by prior semivariances), the prior spatial variation can be generated whereas the DN values of sub-pixels in the resulting images are unnecessarily close to the DN values of the reference image. Thus, global semivariogram matching can be used only for generating MS images with the spatial character of the sub-pixel scale, but not for image restoration tasks.

5.6 Image smoothing using the HNN with semivariogram matching

5.6.1 General model

The idea of image smoothing using the HNN with semivariogram matching emerged from the analysis in section 5.5.3. Observing the semivariogram matching process, it can be seen that the values of prior semivariances had an impact on the smoothness of the simulated image. If the values of the prior semivariances are smaller than those of the original image, then semivariogram matching generates a smoother simulated

image. Accordingly, a smoother simulated sub-pixel image can be obtained with semivariances of zero (the mimimum value of semivariances).

The mechanism and effect of the image smoothing method using zero semivariogram matching for a single direction are demonstrated in Figure 5.14 and Figure 5.15. Since the values of prior semivariances are zero, i.e. the expected DN values of the neighbouring sub-pixels should be the same, a semivariance function for a separation h (Equation (5.8)) will produce a positive value to increase the output of the neuron

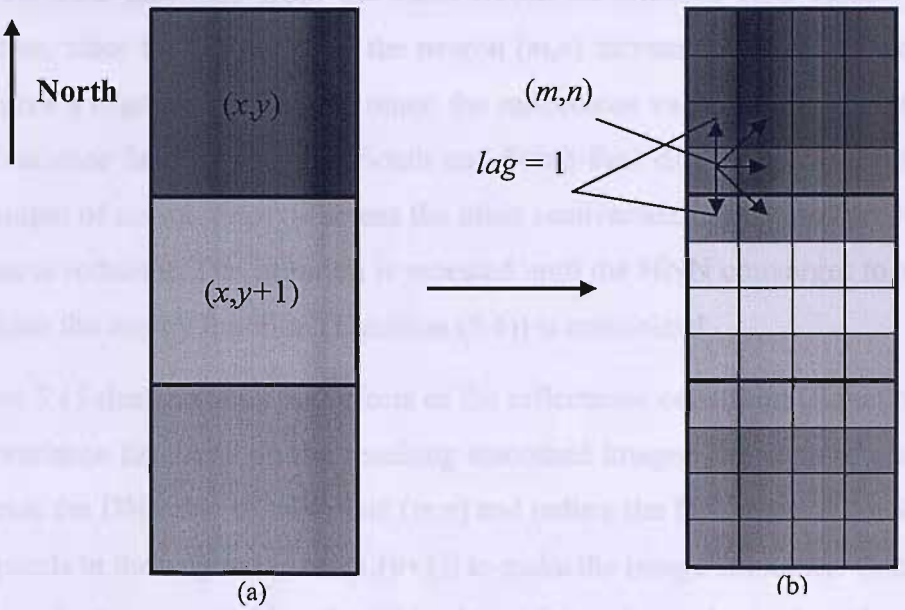


Figure 5.14 Image smoothening by zero semivariance
(a) Original image of 3 pixels, (b) Smoothened image.

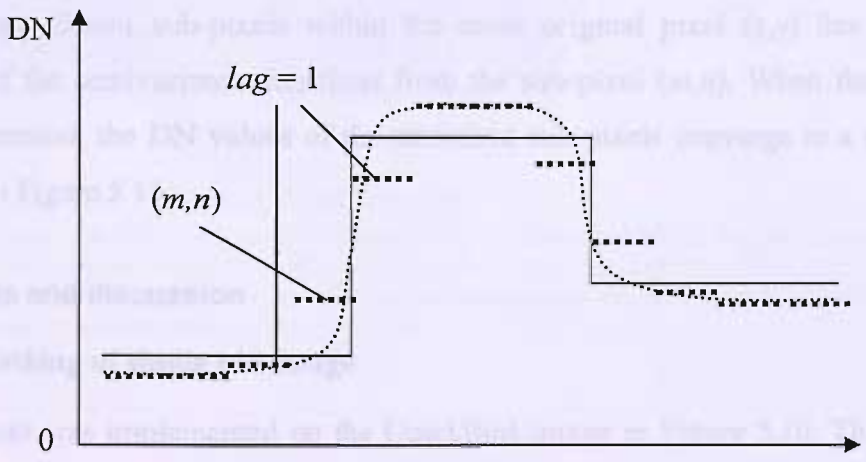


Figure 5.15 Effect of the reflectance constraint and the semivariance function in the South direction. Solid line - digital number of the original degraded image; Dashed line – digital number of the smoothened image; and Dotted line – the digital number of the smoothened image when the zoom factor increases.

(m,n) if the corresponding DN value at separation h is greater than that of the neuron (m,n). In case the DN value at separation h is smaller, the semivariance function value will reduce the output of the centre neuron (m,n). The semivariance function value is zero if the DN values of the two neurons (represents for the sub-pixels) are the same.

Considering the HNN neuron (m,n) in Figure 5.14, at the first iteration of the HNN optimisation process, the semivariance functions for separation $h = 1$ from the South and South-East directions increase the DN value for neuron (m,n) while the semivariance functions from the other directions produce zero values. In the next iteration, since the DN value of the neuron (m,n) increases, the reflectance constraint produces a negative gradient to retain the reflectance value of the original pixel. The semivariance functions for the South and South-East directions continue to increase the output of neuron (m,n) whereas the other semivariance functions produce gradient values to reduce it. The iteration is repeated until the HNN converges to a stable state in which the energy function (Equation (5.4)) is minimised.

Figure 5.15 demonstrates the effects of the reflectance constraint (Equation (5.5)) and semivariance functions on the resulting smoothed image. The semivariance functions increase the DN value of sub-pixel (m,n) and reduce the DN value of the neighbouring sub-pixels in the original pixel $(x,(y+1))$ to make the image smoother. Under the effect of the reflectance constraint, the DN values of the other sub-pixels within the original pixel (x,y) decrease to retain the DN value for (x,y). The DN values of the sub-pixels within the original pixel (x,y), which are adjacent to sub-pixel (m,n), are greater than those of more distant sub-pixels within the same original pixel (x,y) due to the influences of the semivariance functions from the sub-pixel (m,n). When the zoom factor is increased, the DN values of the smoothed sub-pixels converge to a smooth dotted line in Figure 5.15.

5.6.2 Results and discussion

5.6.2.1 Smoothing of single MS image

An experiment was implemented on the QuickBird image in Figure 5.10. The input images were generated by degrading the reference image (Figure 5.16(a)) by a factor of two to 5.2 m spatial resolution (Figure 5.16(b)) and a factor of four to 10.04 m

spatial resolution (Figure 5.16(e)). The smoothed MS images are shown in Figure 5.16(c) and Figure 5.16(f). The bilinear interpolation images (Figure 5.16(d) and

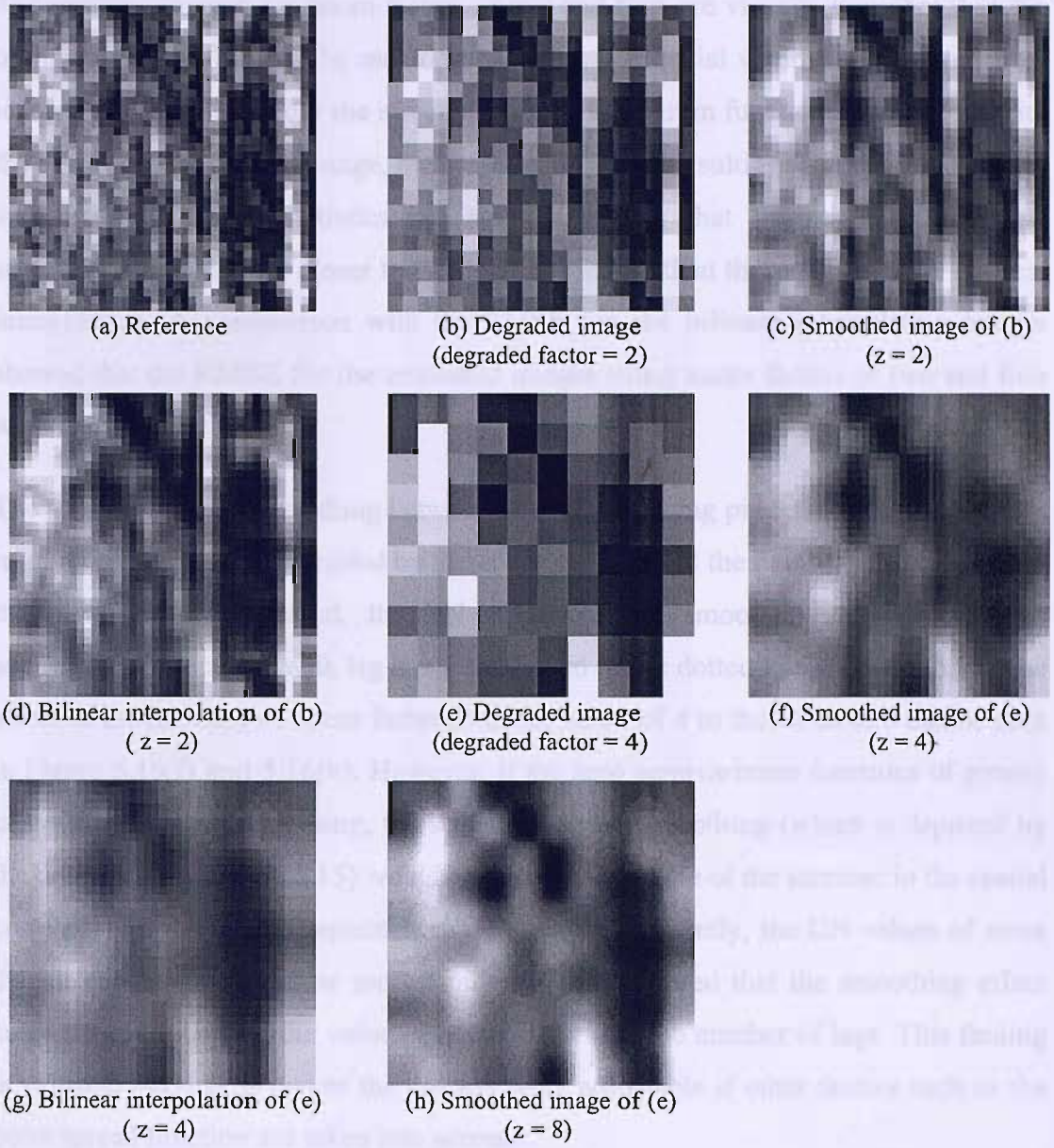


Figure 5.16 Smoothing of the degraded QuickBird image
(a) Reference image; (b) Degraded image by a factor of 2; (c) Smoothed image from the image (b), z = 2; (d) Bilinear interpolation image from the image (b) z = 2; (e) Degraded image by a factor of 4; (f) Smoothed image from the image (e), z = 4; (g) Bilinear interpolation image from the image (e), z = 4; (h) Smoothed image from the image (e), z = 8.

Degrading factor	RMSE		
	Degraded image	Smoothed image	Bilinear
2	3.8437	3.5795	5.0285
4	5.9147	6.1084	6.4994

Table 5.6 RMSE of the degraded, smoothed and bilinear interpolation images

5.16(g)) produced from the degraded images were used to compare the new image smoothing method with conventional interpolation methods.

The resulting images for zoom factors of two and four are visually smoother than the original degraded images. In each original pixel, the spatial variation of the smoothed image was increased under the effect of the semivariogram functions. Comparing with the bilinear interpolation image, the smoothness of the resulting (smoothed) images is visually similar. The statistics in Table 5.6 shows that the results of the new smoothing method were closer to the reference image than those produced by bilinear interpolation. A comparison with the RMSE for the bilinear interpolation images showed that the RMSE for the smoothed images using zoom factors of two and four were both smaller.

The steepness of the smoothing between two neighbouring pixels of the coarse spatial resolution image was controlled by the zoom factor z and the number of lags $\mathbf{h}$. When the zoom factor increased, the DN values of the smoothed image using zero semivariance functions with lag $\mathbf{h} = 1$ converged to the dotted line in Figure 5.15. The effect of the increase of zoom factor from the value of 4 to the value of 8 can be seen in Figure 5.16(f) and 5.16(h). However, if the zero semivariance functions of greater lags were used for smoothing, the steepness of the smoothing (which is depicted by the dotted line in Figure 5.15) would be reduced by virtue of the increase in the spatial correlation between two separated sub-pixels. Consequently, the DN values of more distant sub-pixels would be more similar. This suggested that the smoothing effect could be controlled by the value of zoom factor and the number of lags. This finding is valuable because it makes the process more adjustable if other factors such as the point spread function are taken into account.

5.6.2.2 Smoothing of SPOT MS image

In this experiment, the super-resolved images produced by the HNN and forward model in section 5.3 were used as input data. The resulting smoothed images are reproduced in Figure 5.17. The RMSE values of each band of the input super-resolved and smoothed image are in Table 5.6. The smoothing process was applied only for the pure pixels which were not super-resolved by the HNN super-resolution using unsupervised soft-classified proportions and a forward model.

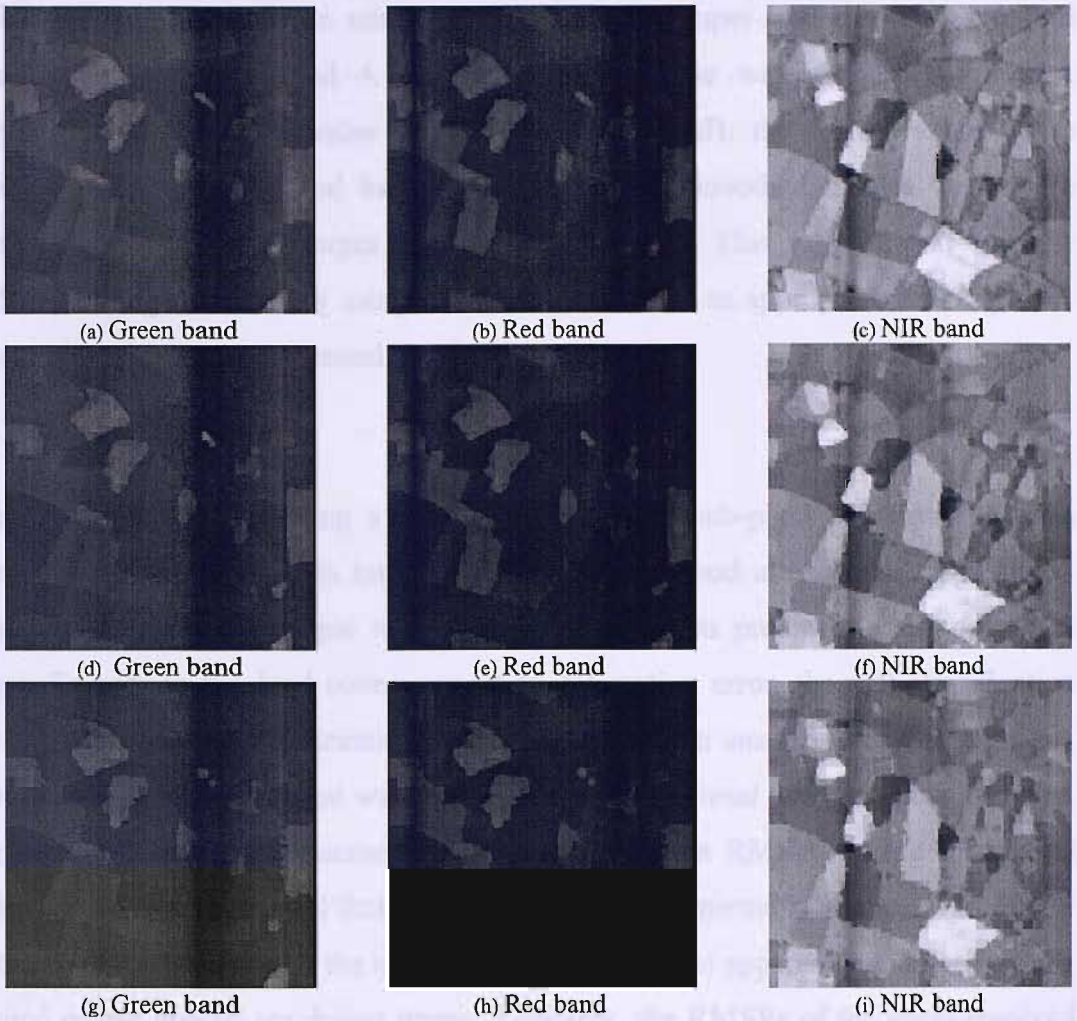


Figure 5.17 Smoothing of the SPOT the HNNsuper-resolved images using forward model
(a), (b) and (c) - 20 m smoothed MS image with zoom factor of 2. (d), (e) and (f) - 20 m smoothed image with zoom factor of 3. (g), (h) and (i) - 20 m smoothed with zoom factor of 4.

Degraded factor	Band	RMSE	
		Super-resolved image	Smoothed image
2	Green	2.6933	2.6888
	Red	2.7655	2.7090
	NIR	5.9094	5.7971
	Σ	11.3682	11.1949
3	Green	3.6575	3.6488
	Red	3.8873	3.8435
	NIR	8.1403	8.0607
	Σ	15.6850	15.5530
4	Green	4.6645	4.6706
	Red	4.9374	4.9256
	NIR	10.2393	10.1999
	Σ	19.8412	19.7961

Table 5.7 RMSE of the super-resolved image using the HNN and smoothed SPOT images

Statistical information in Table 5.6 shows that the RMSEs of the smoothed images were smaller in comparison with those of the input super-resolved images for all zoom factors of 2, 3, and 4. Although the decrease was small in each case (approximately 0.01 DN value for each spectral band), the RMSE values were decreased for every spectral band. That means the smoothed images were more similar to the reference images than the input images. This suggests that the new method of image smoothing using zero semivariance is an appropriate technique for downscaling the remotely sensed imagery.

5.7 Conclusions

Three methods of generating a fine super-resolution sub-pixel MS image from a coarse MS image have been introduced. The first method utilised the HNN super-resolution mapping technique with a forward model to predict the sub-pixel MS image. To reduce the land cover proportion estimation error, the soft-classification was implemented based on training data obtained from an unsupervised classification. The feasibility of the method was evaluated based on visual and statistical analyses. The key statistic used for accuracy assessment was the RMSE. Both a visual and statistical evaluation showed that the new method can generate MS image with a finer spatial resolution. Visually, the super-resolved image was apparently sharper than the original coarse spatial resolution image. Similarly, the RMSEs of the super-resolved image compared to the reference image were smaller than those of the original degraded image due to the super-resolution of the mixed pixels. In addition, the statistics also demonstrated that when the zoom factor increased, the resulting sub-pixel images were closer to the references.

The second method required prior information on the spatial distribution of the land cover patterns at sub-pixel spatial resolution. This information was utilised in the form of discrete semivariance functions to combine with the reflectance constraint in the HNN model. The method can be used to create a sub-pixel MS image with the spectral features of the coarse resolution image and spatial character at the sub-pixel spatial resolution. The analysis showed that the semivariogram matching process smoothed the coarse image if the prior spatial variation at the sub-pixel resolution had a small variance, and increased the spatial variation in the smooth area.

The semivariogram matching was then used for image smoothing of MS imagery. Using semivariance values of zero at a lag of one pixel, the HNN semivariogram matching model generated smoothed sub-pixel images with smaller RMSEs than those of the original degraded and bilinear interpolation imagery. This method can be used to smooth the resulting sub-pixel image from the HNN super-resolution with forward model.

Chapter 6: Super-resolution Mapping Using a Hopfield Neural Network with LiDAR Data[4]

6.1 Introduction

Together with the panchromatic and fusion imagery, Light Detection And Ranging (LiDAR) data is a source of very fine spatial resolution data on elevation. Comparing with the multispectral imagery, the LiDAR elevation data are less informative and hardly used for land cover classification. However, the elevation information provided by LiDAR data is useful for prediction of some special classes such as buildings or trees. In other words, supplementary information at the sub-pixel level of the LiDAR elevation data can be used to produce more detailed and accurate land cover maps

This Chapter proposes a new method to combine multispectral and LiDAR elevation data for land cover classification at the sub-pixel level for an urban area. From the multispectral data, land cover proportion images were produced using soft-classification. The proportion images were then used for super-resolution mapping with a HNN. To make use of the LiDAR elevation data for the super-resolution mapping, a height function was added to the energy function of the HNN. The height function increased the output value of certain classes such as "building" based on probability theory.

6.2 Literature reviews

LiDAR is an active remote sensing system that uses pulses of laser light to illuminate the terrain (Lillesand and Kiefer, 1996). The most common application of LiDAR is to produce very high accuracy elevation data. The most accurate LiDAR, which uses an airborne laser scanning technique, can provide elevation data at vertical accuracies of 15-20 cm (Baltsavias, 1999). Due to its very high accuracy and fine spatial resolution, LiDAR data have been used for building and tree extraction and 3D model construction at very fine spatial resolution (Luzum et al., 2004). In addition, LiDAR data can be combined with multispectral imagery as a complementary data source to increase the accuracy of the land cover classification (Haala and Brenner, 1999, Sohn, and Dowman, 2003).

[4] This chapter is based on Nguyen *et al.* (2005a).

Elevation data can be used as a supplementary source of information for land cover classification. Slope and aspect data produced from a coarse spatial resolution digital elevation model (DEM) can be integrated with multispectral data for land cover analysis (Giannetti *et al.*, 2001). At a very fine spatial resolution, elevation data can be used to extract buildings and trees based on DSM and digital terrain model (DTM) data (Haala and Brenner, 1999). A DSM is different from a DTM in that it not only represents the terrain but also contains non-terrain objects. The difference image between a DSM and DTM is a useful information source to extract objects such as buildings and trees.

6.3 Method

6.3.1 Hopfield neural network

In this section, a modified HNN model which makes use of the LiDAR elevation data is introduced. Figure 6.1 is a graphical depiction of the original HNN approach for super-resolution mapping. A pixel at the original spatial resolution is divided into two inter-connected matrices of (5 x 5) neurons in the HNN. Each neuron (h,i,j) represents a sub-pixel at position (i,j) in the land cover class h and each matrix of neurons represents a land cover class. For super-resolution mapping, the HNN is constrained using the soft-classified land cover proportions and runs until it converges to a stable state. At the stable state, the spatial correlation of the sub-pixels is maximised. Tatem *et al.* (2001) indicated that the convergence of the energy value of the HNN is monotonic and that convergence is faster if the network is initialised using the proportion images. If the output value of the neuron (h,i,j) is 1, the sub-pixel (i,j) is assigned to the land cover class h. Otherwise, if the output value is 0, the sub-pixel (i,j) does not belong to the class h. The energy function can be expressed as

$$E = -\sum_{h}\sum_{i}\sum_{j}\left(k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij}\right), \tag{6.1}$$

where k_1, k_2, k_3 and k_4 are weighting coefficients. Values of the weighting coefficients determine the effects of the goal functions, proportion constraint and multi-class constraint to the enegy function. The values of these weighting coefficients are defined empirically for best performance of the HNN. The values of the goal functions $G1_{hij}$ and $G2_{hij}$, proportion constraint P_{hij} and multi-class constraint M_{hij} by Equation (3.2), (3.3), (3.4) and (3.5).

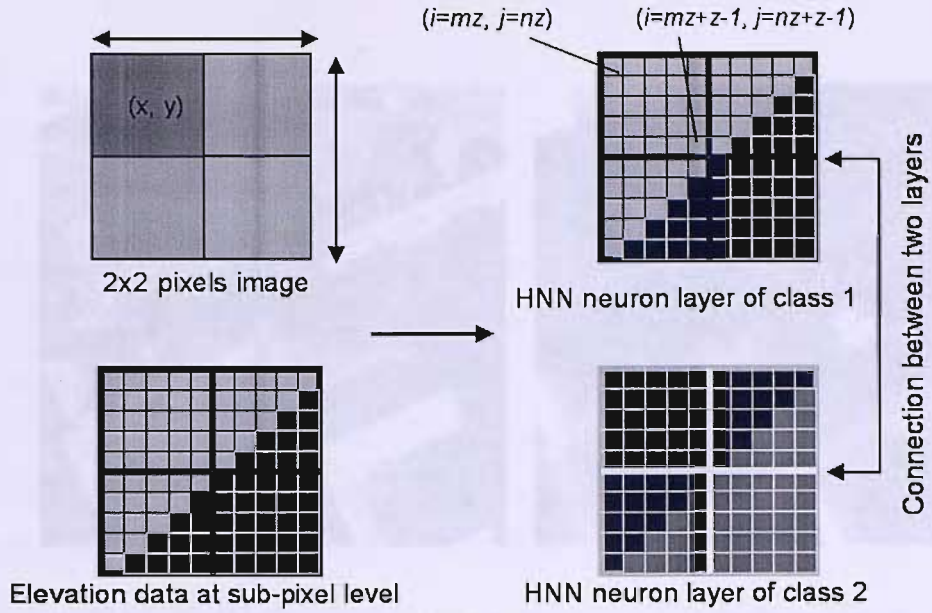


Figure 6.1 Super-resolution mapping using the HNN.
From 2x2 pixels proportion image, the HNN is used to produce 10x10 sub-pixels image of two land cover classes. (x,y) are the coordinates at original level. (i,j) are coordinates at sub-pixel level. z is increasing factor in spatial resolution.

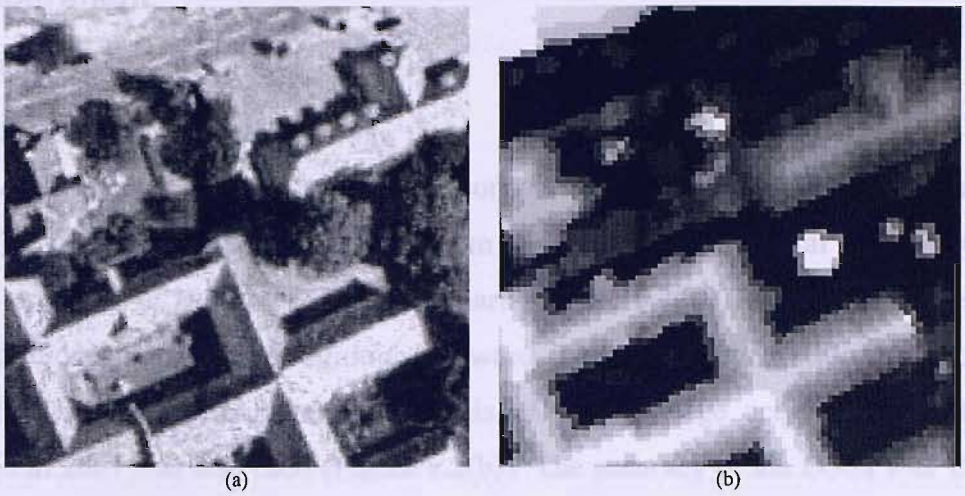
6.3.2 Height function

To use the LiDAR elevation data for super-resolution mapping by the HNN, the energy function in Equation (6.1) was modified by adding a height function. The value of the height function is defined based on the probability of a sub-pixel belonging to a class using LiDAR elevation data. The new energy function is

$$E = -\sum_h \sum_i \sum_j \left(k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij} + k_5 H_{hij} \right), \tag{6.2}$$

where H_{hij} is the height function value for each neuron (h,i,j) and k_5 is a weighting coefficient for the height function.

The structure of the HNN in Figure 3.1 can be modified to make use of the LiDAR data. Elevation data are processed from the DSM (Sohn, and Dowman, 2003) to produce an image of the *normalised height* of non-terrain objects. Based on these height data, it is possible to predict the location of sub-pixels in a certain land cover class. For example, a sub-pixel with a height of 10 m is likely to be a building or tree sub-pixel. In this research, the prediction of land cover class based on the normalised height can be calculated for each neuron as



(a) (b)

Figure 6.2 Optical and LiDAR elevation data used for land cover map and elevation data simulation. (a) 200x200 pixels digital orthophoto at 40 cm. (b) 200x200 pixels LiDAR elevation data image.

$$\frac{dH_{hij}}{dv_{hij}} = P(C_h \mid N_{ij}), \tag{6.3}$$

where $P(C_h|N_{ij})$ is the conditional probability of a sub-pixel (i,j) with normalised height N_{ij} belonging to class C_h. The probability $P(C_h|N_{ij})$ can be defined based on the Gaussian distribution as

$$P(C_h \mid N_{ij}) = \frac{1}{\delta_{C_h}\sqrt{2\pi}} e^{-(N_{ij}-\mu_{C_h})^2/(2\delta_{C_h})}, \tag{6.4}$$

where μ_{Ch} and σ_{Ch} are the mean and standard deviation of class C_h in the normalised height image (Richards, 1993).

As in Equation (6.4), the height function would increase the output value of a neuron (i,j,h) in correlation with the conditional probability of a sub-pixel (i,j) belonging to a class h. If the height function produces the same value for the neurons in position (i,j) of different classes, the output value produced by the goal functions, proportion and multi-class constraints will determine the land cover class of the sub-pixel. In this case, the algorithm based on normalised height alone might fail to separate land cover classes if the statistics of these classes based on the normalised height data are similar.

6.4 Experiment

6.4.1 Data[5]

6.4.1.1 Optical and Elevation Image

In this experiment, a simulation of the proportion image at 8 m spatial resolution was used. The simulation was based on a 40 cm spatial resolution digital orthophoto of Odense provided in 1999 by COWI, Denmark (Figure 6.2(a)). The image is used as demonstration data for the eCognition software. The image registration accuracy of the orthophoto is approximately 2-3 pixels. A DSM image acquired by LiDAR TopoSys in 2001 was also used (Figure 6.2(b)). The accuracy of the DSM was 15 cm (vertical) and 50 cm (planimetric).

6.4.1.2 Land class proportion image

A proportion image was simulated from the 200x200 pixels digital orthophoto by manual digitising. In addition, a 40 cm spatial resolution map of three classes was extracted manually from the orthophoto. Three land classes were identified: Building, Tree, and Background (Background is the class that consists of no non-terrain objects) (Figure 6.3(a), 6.3(b), 6.3(c)). The map was then used as reference data. From the map, a simulated 8 m proportion image (Figure 6.3(d), 6.3(e), 6.3(f)) was created by degrading the 40 cm map by a factor of ten. The simulated proportion image was then used as input for super-resolution mapping by the HNN super-resolution mapping to produce a 40 cm thematic map. Evaluation of the method was implemented by comparison of the 40 cm reference map with the 40 cm map predicted by the HNN super-resolution mapping.

6.4.1.3 Statistical analysis of LiDAR elevation data

The normalised height data were produced using a moving window filter. A part of the land cover images in Figure 6.3(a), 6.3(b), and 6.3(c) were used for training to calculate the mean and standard deviation of each land class on the normalised height (Figure 6.3). Investigation showed that the mean normalised height of the Building class was 21.59 m and its standard deviation was 2.627 m (Background class normalised height was assigned a value of 12 m). The normalised height of the Tree class was similar to that of the Background class except for the tree trunks since the LiDAR data were produced by the second return of the laser pulse. Therefore,

[5] The data in this chapter is different from the data used in Nguyen *et al.* (2005)

amongst the three classes in the experiment, the LiDAR data were the most informative for discriminating the Building class.

6.4.2 Results and discussions

6.4.2.1 Network settings and hard classification

To evaluate the performance of the technique, the resulting maps using the LiDAR data were compared with a 8 m hard classified image (Figure 6.4(g), 6.4(h), 6.4(i)), 0.8 m hard minimum-distance-to-means classification using LiDAR elevation data (Figure 6.4(j), 6.4(k), 6.4(l)), and 0.4 m HNN super-resolution mapping obtained without the LiDAR data (Figure 6.4(m), 6.4(n), 6.4(o)). The 8 m hard classified image was obtained from the proportion image by assigning each 8 m pixel to the class of

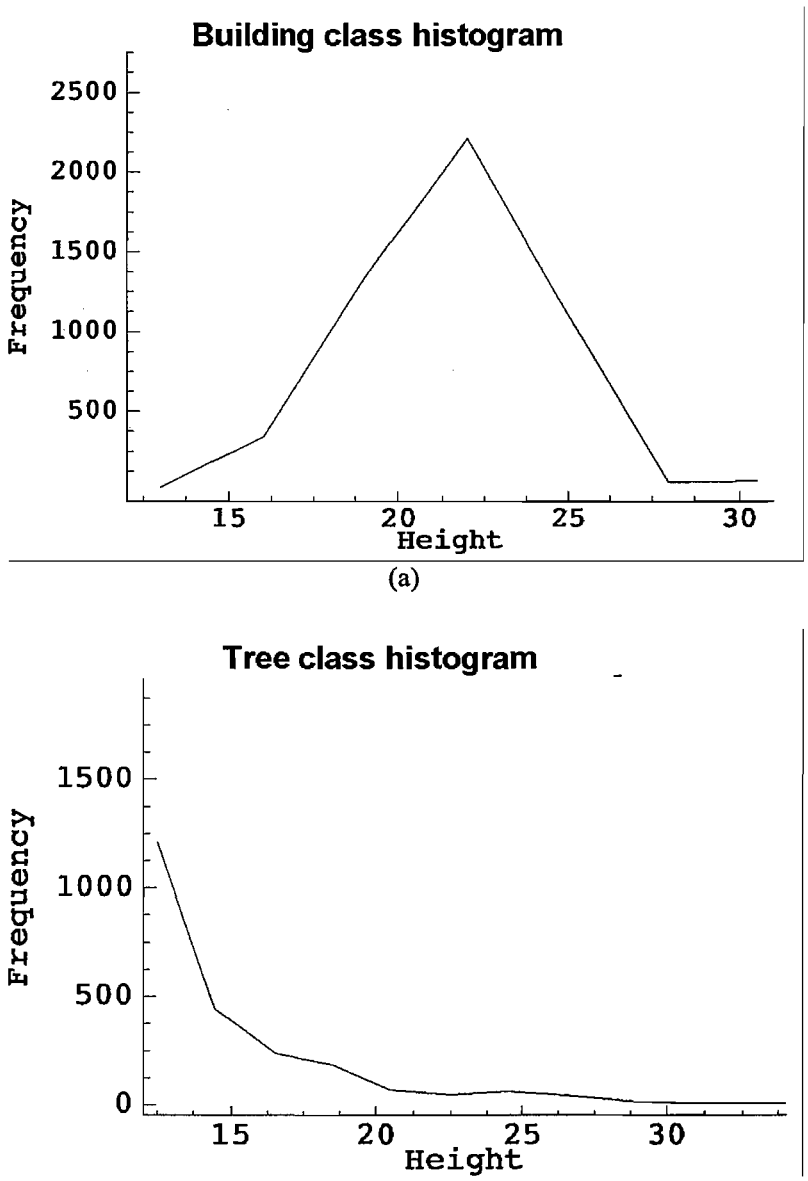


Figure 6.3 Histograms of trees and building classes in LiDAR elevation data

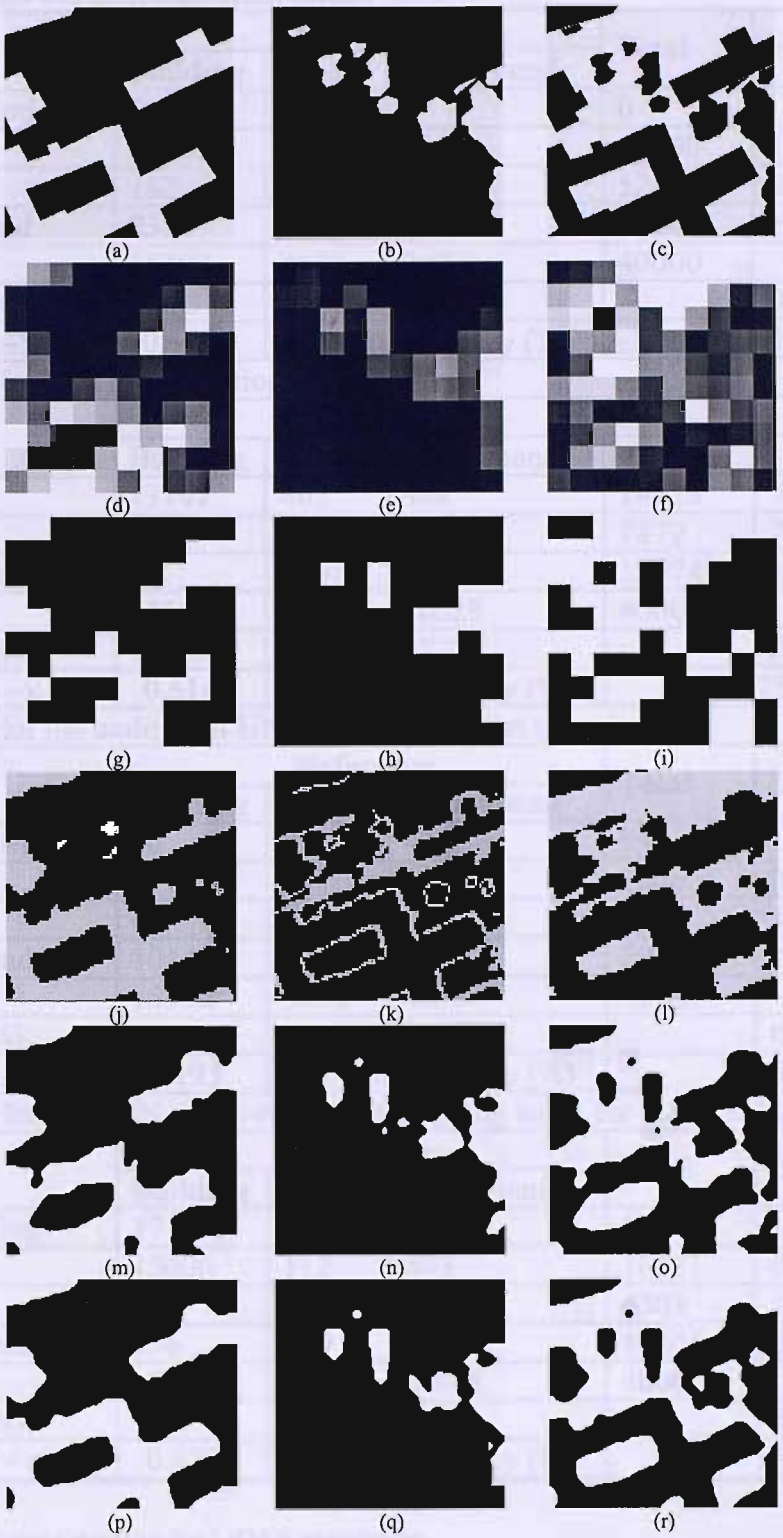


Figure 6.4 Reference and resulting images.
Reference images (0.4 m) of (a) Building, (b) Tree and (c) Background; proportion images (8 m) of (d) Building, (e) Tree and (f) Background; hard classified images (8 m) of (g) Building, (h) Tree and (i) Background; hard classified LiDAR images (0.8 m) of (j) Building, (k) Tree and (l) Background; HNN super-resolution mapping (0.4 m) without using LiDAR data for (m) Building, (n) Tree and (o) Background; HNN super-resolution mapping (0.4 m) using the 0.8 m LiDAR Data for (p) Building, (q) Tree and (r) Background.

Statistics for the hard classified image					
	Reference			Total	ErrorC (%)
	Building	Tree	Background		
Unclassified	0	0	0	0	
Building	12605	152	2043	14800	14.83
Tree	182	3458	1560	5200	33.50
Background	2515	1060	16425	20000	17.88
Total	15302	4670	20028	40000	
ErrorO (%)	17.63	25.95	17.99		0.1878
KIA – κ	0.6833	Overall accuracy (%)			81.22
Statistics for hard classification of LiDAR					
	Reference				
	Building	Tree	Background		
Building	13142	602	484	14228	7.63
Tree	2141	1666	4065	7872	78.84
Background	19	2402	15479	17974	13.71
Total	15302	4670	20028	40000	
ErrorO (%)	14.12	64.33	22.71		0.2428
KIA – κ	0.6145	Overall accuracy (%)			75.52
Statistics for the traditional HNN super-resolution mapping					
	Reference			Total	ErrorC (%)
	Building	Tree	Background		
Unclassified	30	23	42	95	1.000
Building	14183	79	1203	15465	8.29
Tree	45	3546	761	4352	18.52
Background	1044	1022	18022	20088	10.28
Total	15302	4670	18367	40000	
ErrorO (%)	7.31	24.07	10.02		0.1062
KIA – κ	0.8193	Overall accuracy (%)			89.38
Statistics for the HNN super-resolution mapping using the 0.8 m LiDAR data					
	Reference			Total	ErrorC (%)
	Building	Tree	Background		
Unclassified	17	10	56	83	1.000
Building	15006	112	893	16011	6.28
Tree	23	3569	709	4301	17.02
Background	256	979	18370	19605	6.30
Total	15302	4670	20028	40000	
ErrorO (%)	1.93	23.58	8.28		0.0764
KIA – κ	0.8703	Overall accuracy (%)			92.36

Table 6.1 Accuracy statistics for LiDAR experiment

the largest proportion. The normal HNN super-resolution mapping was implemented based on the HNN with the goal functions, proportion and multi-class constraints as in Tatem *et al.* (2001). In this process, the 8 m simulation proportion image was used as input for the HNN super-resolution mapping with a zoom factor of twenty to produce the land class map at 0.4 m. Through an empirical process, the optimal weighting

coefficients of k_1=150, k_2=150, k_3=200 and k_4=150 were determined. Using these values of weighting coefficients, the HNN super-resolution mapping produced the most accurate results after 6000 iterations as in Figure 6.4(j), 6.4(k), 6.4(l).

The newly proposed HNN super-resolution mapping using LiDAR elevation data was accomplished using the 8 m proportion image and 0.8 m spatial resolution LiDAR elevation data. The HNN was set with a zoom factor of twenty and weighting coefficients of k_1=150, k_2=150, k_3=200, k_4=150 and k_5=150. After 6000 iterations, the HNN network using the 0.8 m LiDAR elevation data produced a 0.4 m spatial resolution land class map in Figure 6.4(p), 6.4(q), 6.4(r). Since the resulting sub-pixel land class maps were hard-classified, statistics which are usually applied for accuracy assessment of hard classified maps were used. These statistics were Kappa Index of Agreement (KIA), overall accuracy, and per-class omission and commission errors (Table 6.1).

6.4.2.2 Visual evaluation

Visual comparison of the results of the two prediction techniques shows that the super-resolution mapping using the LiDAR data is preferable to the hard classification and the traditional HNN super-resolution mapping, especially for the Building class. The sub-pixels in the edges of buildings in Figure 6.4(a) were not preserved as straight lines in Figure 6.4(g) and 6.4(j). These sub-pixels were assigned to the Tree class in a hard minimum-distance-to-means classification of LiDAR elevation data (Figure 6.4(k)). Using the height function (Equation 6.4), a positive value was produced for the neurons in the Building class. Combining with the goal functions and proportion constraint, this height function value retained the linear edges of the building objects (Figure 6.4(p)).

6.4.2.3 Statistical evaluation

The accuracy statistics (Table 6.1) showed a considerable increase in accuracy for the new technique. Overall accuracy increased from 89.99% for the 8 m hard classification and 76.52% for the 0.8 m hard classification using only LiDAR elevation data to 96.22% for the HNN super-resolution mapping using 0.8 m LiDAR elevation data. In comparison with the traditional HNN super-resolution mapping, the accuracy of the thematic map produced by the HNN super-resolution mapping using the LiDAR elevation data increased approximately 5% in terms of overall accuracy.

Amongst the three classes, the accuracy of the Building class increased significantly. The omission error of the Building class decreased from the value of 8.23% for 8 m hard classification, 14.12% for the hard classification of 0.8 m LiDAR data and 6.14% for the traditional HNN super-resolution mapping to 1.82% for the new HNN super-resolution mapping technique using 0.8 m LiDAR elevation data. However, the commission error increased slightly from 1.53% for the traditional HNN super-resolution mapping to 3.18% after using the LiDAR elevation data. The results showed that the elevation data were the most informative for the Building class.

6.5 Conclusions

This Chapter introduces a technique for combining the optical imagery and LiDAR elevation data for super-resolution mapping. Information provided by the LiDAR data was incorporated into the HNN optimisation using a height value function. The value of the new function was calculated based on statistical theory. The effectiveness of the technique was examined on a simulated 0.8 m DSM and an 8 m proportion image. The accuracy evaluation was implemented based on the KIA, overall accuracy, omission and commission errors.

The results demonstrated that LiDAR elevation data can be fused with optical data for the HNN to predict accurately land cover at a sub-pixel spatial resolution. The results showed a considerable increase in all accuracy statistics for the new technique, particularly for building objects. For the other classes, the technique results in a slight increase in accuracy. In addition, visual inspection of the resulting images also demonstrated the improvement in super-resolution mapping using the elevation data.

The new technique is generic for the combination of different sources of elevation and optical data. The same principle can be applied to other sets of data at various spatial resolutions such as 20 m SPOT MS data and 1 m LiDAR or 4 m MS data and 0.5 m and 1 m LiDAR. Thus, future research will exploit other sources of information, which can be extracted from LiDAR data such as LiDAR reflectance or slope and aspect, for fusion with the optical imagery to predict land cover at the sub-pixel spatial resolution.

Chapter 7: Discussion and Future Research

7.1 Introduction

This chapter presents an overview on the principles, limitations and potential of super-resolution mapping. Firstly, spatial dependence and issues related to super-resolution mapping, such as spatial scale and spatial information content are discussed. Secondly, several problems which may influence the application of the approaches presented in the previous chapters are considered. Finally, future research and extension for the developed approaches are proposed.

7.2 Spatial dependence and super-resolution mapping

7.2.1 Spatial dependence at sub-pixel spatial resolution

7.2.1.1 Spatial dependence and small objects

Most of the present super-resolution mapping approaches are based on the spatial dependence assumption, which refers to the tendency of adjacent pixels to be more alike than those far apart. The principle of this assumption is the fact that all the geographic variables, including land cover, are spatially dependent at least at some scale (Atkinson, 1997). The super-resolution mapping process allocates sub-pixels within pixels to maximise the spatial dependence; the sub-pixels are clustered as a result.

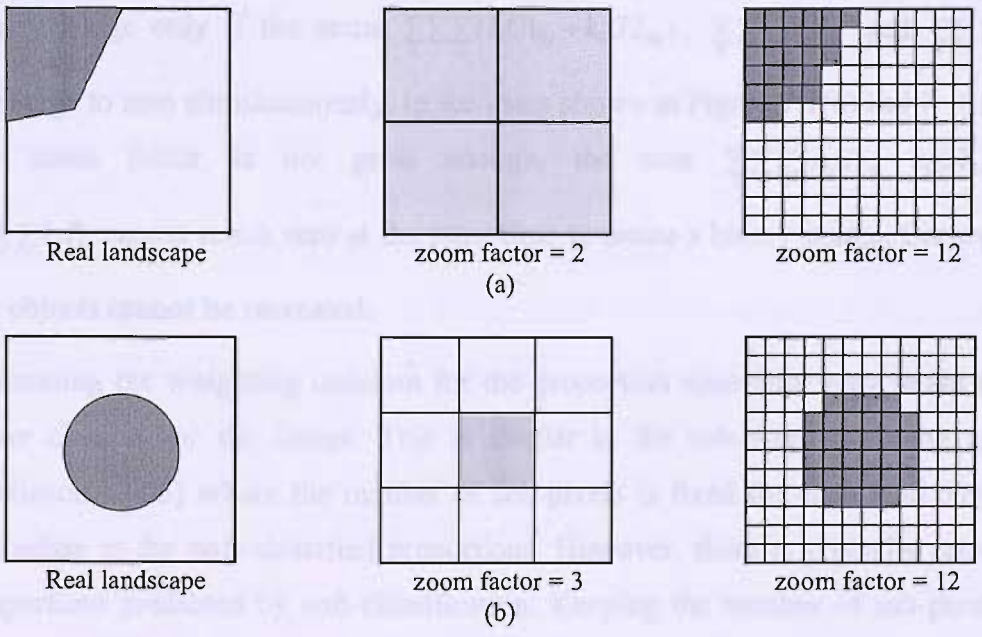


Figure 7.1 Zoom factor and spatial dependence

In super-resolution mapping by the HNN, there always exists a relationship between the spatial dependence and the spatial resolution of the resulting sub-pixel image. Increasing the spatial resolution may reinforce of the goal functions if one of the land cover proportions represents small or separated objects, such as lone trees or small roads in Landsat image. This argument can be demonstrated by Figure 7.1(a) and 7.1(b). The real landscape in Figure 7.1(a) is represented by a single pixel at the coarse spatial resolution. Supposing that the proportion of a land cover class A (depicted by the darker colour in the top-left corner of the pixel) is approximately 25%. If a zoom factor of two is used, the goal functions will produce the values to change the land cover class of the top-left sub-pixel to land cover B (represented by the white colour in the pixel) because the surrounding sub-pixels belong to the land cover B. Even though the proportion constraint retains the proportions of the land cover classes A and B, the value produced by this constraint is not great enough when compared to the goal spatial dependence function values. Similarly, if a sub-pixels is surrounded by sub-pixels of the other land cover class as in Figure 7.1(b), the goal functions will convert the centre sub-pixel of land cover class A to land cover class B.

Super-resolution mapping using the HNN is an optimisation process based on the minimisation of an energy function. For each pixel, the energy function is the sum of the goal function and proportion constraint values for all the sub-pixels within that pixel. The energy function $E = -\sum_h \sum_i \sum_j (k_1 G1_{hij} + k_2 G2_{hij} + k_3 P_{hij} + k_4 M_{hij})$ will converge to a binary image only if the sums $\sum_h \sum_i \sum_j (k_1 G1_{hij} + k_2 G2_{hij})$, $\sum_h \sum_i \sum_j k_3 P_{hij}$ and $\sum_h \sum_i \sum_j k_4 M_{hij}$ converge to zero simultaneously. In the cases shown in Figure 7.1(a) and 7.1(b), when the zoom factor is not great enough, the sum $\sum_h \sum_i \sum_j (k_1 G1_{hij} + k_2 G2_{hij})$ and $\sum_h \sum_i \sum_j k_3 P_{hij}$ cannot reach zero at the same time to create a binary image. Consequently, the objects cannot be recreated.

Increasing the weighting constant for the proportion constraint may retain the land cover class A for the image. This is similar to the sub-pixel swapping approach (Atkinson, 2005) where the number of sub-pixels is fixed for each land cover class according to the soft-classified proportions. However, there is typically error in the proportions predicted by soft-classification. Keeping the number of sub-pixels fixed

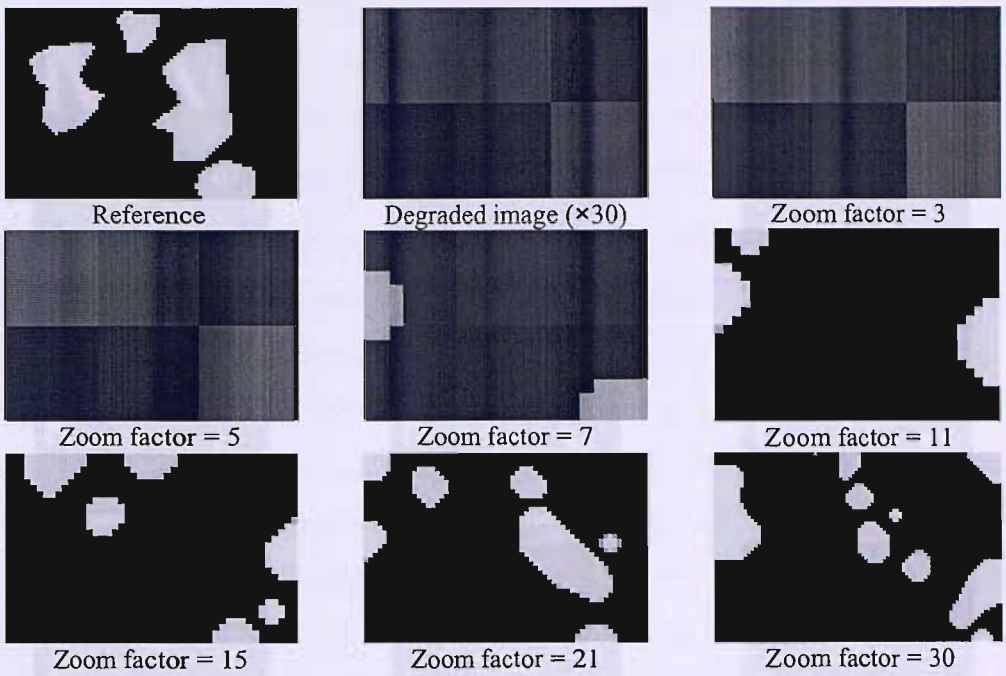


Figure 7.2 Spatial dependence effects on the small objects.

for the land cover classes means that this error is contained in the results. Another solution to increase the zoom factor. When the zoom factor is increased to twelve as in Figure 7.1(a) and 7.1(b), the goal functions produce values that increase spatial dependence between the sub-pixels of class A. Therefore, the land cover class A object is recreated.

Figure 7.2 presents an example of the effect of zoom factor on recreating small objects by HNN super-resolution mapping. An image of 90 × 60 pixels which contains several small objects was degraded by a factor of 30 to generate a 3 × 2 proportion image. The HNN super-resolution mapping was implemented with zoom factors of 3, 5, 7, 11, 15, 21, and 30, the weighting constants $k_1 = k_2 = 150$ for the goal functions, $k_3 = 200$ for the proportion constraint and 1000 iterations to evaluate the effect of the zoom factor on the ability of the HNN to recreate small objects. With small zoom factor values of 3 and 5, the HNN could not converge to a binary stable state. When the zoom factor increased to greater values (7 and 11), the small objects started to be recreated. With the largest zoom factors (15, 21 and 30), the HNN converged fully to produce binary images. In the experiment process, it was observed that a small object can be mapped if it comprises more than nine sub-pixels. Thus, it is suggested that the size of the object to be mapped should be taken into account for choosing the zoom factor for super-resolution mapping and that high zoom factors

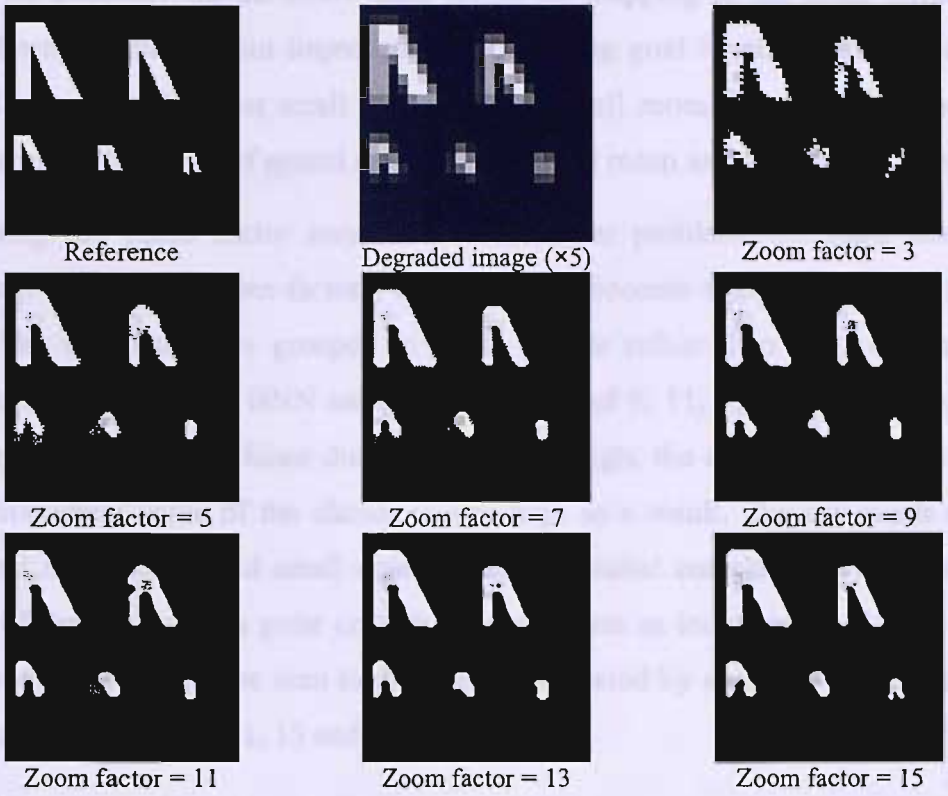


Figure 7.3 Super-resolution for objects with different sizes

are essential for mapping small objects.

7.2.1.2 Spatial dependence and accuracy of the sub-pixel maps

Tatem *et al.* (2001a, 2002b) insisted that the accuracy of the resulting super-resolution maps increases with increasing zoom factor (also does the spatial resolution of resulting the sub-pixel land cover map). The method achieves this by smoothing the linear borders of large objects. This can be seen by the experiment in Figure 7.3, where the HNN was used with weighting constants $k_1 = k_2 = 150$ for the goal functions, $k_3 = 200$ for the proportion constraint and 2000 iterations. With small zoom factor values (3 and 5), the linear borders in the two largest objects in Figure 7.3 are tortuous. When the zoom factor increases, these borders become smoother under the effect of the goal functions. High values of the zoom factor can also increase the accuracy by regenerating the small objects (as discussed above). Comparing the images produced by the HNN using zoom factors of 3 and 5 with the those obtained by using higher zoom factor values, the improvement in the recreation of small objects is clearly seen.

Although high values of the zoom factor increase the overall accuracy of the resulting sub-pixel image, a number of limitations remain. Despite increasing the zoom factor

value, the accuracy for the HNN super-resolution mapping of the sharp corners and small linear features is not improved. The clustering goal functions always produce curved corners and cluster small linear objects for all zoom factors. In this case, the increasing applicability of spatial resolution does not mean an increase in accuracy.

Increasing the zoom factor may cause some other problems for super-resolution mapping. With high zoom factors, the HNN can become trapped in local minima, where the sub-pixels are grouped in small objects rather than large objects. The images produced by the HNN using zoom factors of 9, 11, 13 and 15 in Figure 7.3 illustrates this problem. Since the zoom factor is high, the number of sub-pixels for each proportion value of the classes is also high as a result. The sub-pixels can be clustered into unexpected small objects and the spatial correlation within the sub-pixels of these objects is great enough to retain them as independent objects. These unexpected objects can be seen in the images generated by super-resolution mapping with zoom factors of 9, 11, 13 and 15.

7.2.2 Using supplementary data for super-resolution

In this thesis, the problems for super-resolution mapping mentioned in section 7.2.1 can be resolved based on information from supplementary data sources such as fused and PAN imagery and LiDAR elevation data. It should be expected that these sources of supplementary data, which are usually at a finer spatial resolution than the multispectral image, can provide useful information about land cover at sub-pixel spatial resolution. However, this information is relatively limited due to the low spectral resolution (for PAN imagery and LiDAR elevation data) and the spectral distortion (for fused imagery). The proposed approaches described in previous chapters, which can be considered as data fusion mechanisms, attempted to utilise the land cover proportions and useful information from these supplementary data to resolve the problems with sharp corners and small objects in section 7.2.1.

7.2.2.1 Super-resolution mapping by the HNN with fused imagery

Fused imagery is produced by combining a high spectral MS image with a finer spatial resolution PAN image. Although fusion between other sources of imagery such as multispectral and Synthetic Aperture Radar (SAR) data are widely used for land cover classification (Solberg *et al.*, 1994, Solaiman *et al.*, 1999), the use of fused imagery is still limited due to the distortion in the fused spectral bands. Because of the

spectral distortion in the fused image, the classification using the fused image is usually based on an object-oriented approach (Zhang and Maxwell, 2006).

The approach for HNN super-resolution mapping using the fused imagery proposed in this thesis achieved two goals: making use of the spectrally distorted fused image and providing useful information for accurate mapping of sharp corners and small linear features. The forward and inverse models based on linear mixture modelling ensured that only useful land cover information was extracted from the fused image. The number of land covers used for the linear mixture model to produce the reflectance constraint values for the HNN in Equation (3.14) is indicated by the soft-classified land cover proportions. For example, if the pixel (x,y) (in the original image) contains three land covers: trees, grass and asphalt, then the reflectance constraint values for all fused pixels within pixel (x,y) will be estimated based on the linear mixture modelling for these three land cover classes only. This prevents the occurrence of unexpected land covers due to the spectral distortion of the fused image. Furthermore, the retention of the land cover proportions at the coarse spatial resolution by the proportion constraint also controls the effect of the reflectance constraints. In essence, the HNN super-resolution mapping using the fused imagery is a mechanism for information fusion of the coarse spatial resolution MS image and the fine spatial resolution PAN image to produce land cover maps at sub-pixel spatial resolution.

In practice, the approach was fairly successful in resolving sharp corners and small linear objects problems (Figure 3.13) due to the information provided by the fused image. Analyses in Chapter 3 showed that the increase in accuracy of the HNN super-resolution mapping using the fused image was mainly due to the improvement in mapping sharp corners and small linear objects. Obviously, the capability of the reflectance constraint to map objects with sharp corners and linear small objects depends on the spatial resolution of the fused image and the size of the objects. If the spatial resolution is finer than the size of the objects, the sharp corners can be recreated.

7.2.2.2 Super-resolution mapping by the HNN with PAN imagery

The approach for utilising PAN image aims to circumvent imagery fusion as an intermediate process. Similar to the approach for the fused image, the HNN super-resolution mapping using PAN imagery is a mechanism for data fusion and resolves

the same problems. Visual and statistical analyses in Chapter 4 showed that the proposed approach was capable of increasing the spatial resolution and accuracy of the land covers. Using the PAN image directly for super-resolution mapping with new forward and inverse models, the approach also resolved the traditional problems with the linear mixture model, which can only estimate land cover proportions if the number of land covers does not exceed the number of spectral bands and assumes that the proportions sum to unity.

7.2.2.3 Super-resolution mapping by the HNN with LiDAR elevation data

The HNN super-resolution mapping using LiDAR elevation data is different from the two previous approaches in the conceptual design of the general model. The forward and inverse models which constrain the neural network iteratively is replaced by a height function. This height function is a goal function used for maximising the probability of the sub-pixels belonging to a particular class (In this research, the probability of the sub-pixels belonging to the building class is used). This probability is defined based on a Gaussian distribution. Hence, this approach is an optimisation process that maximises the spatial correlation between the sub-pixels and the probability of sub-pixels belonging to the building class.

LiDAR elevation data provide useful information for only a few classes. However, the objects in these classes are relatively small with sharp corners or comprise linear shapes such as buildings and power lines. This makes the LiDAR elevation data very useful in increasing the accuracy of the sub-pixel maps. The results in Chapter 6 demonstrated that the super-resolution based on probability maximising optimisation can produce accurate maps of the buildings and that the accuracy of the other classes also increases as a result.

7.2.3 Spatial dependence for image smoothing using variogram matching

A spatial dependence maximisation process was employed in the image smoothing approach (in Chapter 5) in the form of zero semivariance functions with lags $\mathbf{h} = 1$. These zero semivariance functions adjust the output of a neuron (representing a sub-pixel) according to the output values of surrounding neurons to maximise the spatial correlation in the image. Under the influence of a zero semivariance function for a lag $\mathbf{h} = 1$, the output value of the neuron (i,j) is changed to make the DN value of the sub-pixel (i,j) as similar to the DN value of the destination sub-pixel $((i,j)+\mathbf{h})$ as possible.

The sum of eight directional semivariance functions alters the DN value of the centre sub-pixels (i,j) to maximise the spatial correlation of the super-resolved sub-pixel image. Thus, image smoothing using zero semivariance functions is an optimisation process with a spatial correlation maximising goal and a DN constraint at the original coarse spatial resolution.

Although using the same optimisation mechanism, the image smoothing using the HNN with zero semivariance functions is different from the super-resolution mapping and the class-based super-resolution for MS imagery (a modification of the super-resolution mapping algorithm) in the manner of formulating the goal functions for spatial correlation maximisation. For the super-resolution land cover mapping, the spatial correlation (spatial dependence) is maximised by clustering the sub-pixels belonging to the same class (for the super-resolution of MS imagery, the correlation is maximised within the spectral class) whereas the image smoothing algorithm increases the spatial dependence for the entire image. The super-resolution of MS imagery using the forward model algorithm segments the image into several categories and then uses spatial correlation maximisation to separate the sub-pixels of different categories. This sharpens the sub-pixel image at the borders between two different categories. In contrast, spatial correlation maximisation is employed by the image smoothing algorithm to reduce the contrast between the two neighbouring pixels. This makes the sub-pixel image apparently smoother.

7.3 Problems for the new approaches

7.3.1 Computation problem

Computational cost is a drawback of these approaches. Although the computation capacity of computers has increased, the great number of neurons and layers used in the HNN super-resolution mapping process make the computation expensive. The problem is made more severe when the constraint and goal functions are added into the energy function to utilise supplementary data. For the super-resolution mapping of an image of 360 × 300 sub-pixels in Figure 3.13, a computer with an Intel Pentium IV 2.6 MHz processor and 256 MB memory needs 1 hour and 18 minutes to produce the sub-pixel map. The example shows that the expensive computation problem is an obstacle for the practical application of the new HNN. For the HNN pattern matching

algorithm, the computation is less problematic as the optimisation process is implemented in a single image.

Another drawback of the HNN super-resolution mapping is the subjective choice of parameters. The choice of the parameter values influences the optimisation outputs and the convergence of the HNN. Inappropriate choices for parameter values may bring about slow convergence or unexpected convergence states for the HNN. When a new goal or constraint function is added to the optimisation process, the weighting parameters have to be changed to obtain the best convergence. In this research, the weighting parameter values have been chosen empirically based on the accuracy statistics of the results. However, this method of choosing the parameter values is not convenient for a real application of the HNN super-resolution mapping.

7.3.2 The limit of number of land cover classes

For the HNN super-resolution mapping using fused imagery, the number of land cover classes is limited by the number of spectral features obtained from a fused image because the inverse model is based on the linear mixture model which derives the land cover proportions using a system of linear equations. This limitation is a problem for the practical use of the algorithm because the number of land cover classes is not always limited to a few land cover classes in reality. To solve this problem, the linear mixture model can be replaced by other models such as Bayesian probability or a feed-forward neural network system. However, this makes the approach more complicated.

7.3.3 Image registration error

The usefulness of the information from supplementary data for super-resolution mapping depends upon the accuracy of the image registration. The misregistration between the PAN and MS images may make the fusion image appear blurred or produce edge phenomena (Hong and Zhang, 2005). Analyses in Chapter 3 and 4 show that the fused and PAN image with RMS registration error of 1 pixel can increase the accuracy of the sub-pixel land cover maps. Although this accuracy for image registration is often obtainable, more accurate image registration is needed to ensure a considerable increase in accuracy of the resulting sub-pixel map. In addition, the misregistration between remotely sensed images and ground data is an obstacle for the

training and accuracy assessment of the input soft-classified proportions and the resulting sub-pixel maps.

7.4 Future research

7.4.1 Downscaling without classification step

Most super-resolution mapping methods use soft-classified land cover proportions as input for the optimisation model. However, the sub-pixel land cover map can be produced directly from the raw MS image. Kasetkasem *et al.* (2003, 2005) used the MS image directly as input for an optimisation process based on Markov random fields and a simulated annealing algorithm to produce the sub-pixel land cover image. The HNN can also implement the super-resolution mapping using MS image as input instead of the soft-classified land cover proportions. Using the forward and inverse models mechanism, the MS image can constrain the HNN in the super-resolution mapping process. This circumvents the soft-classification as an intermediate step for super-resolution mapping.

7.4.2 The HNN model for image fusion

The use of a forward model and super-resolution mapping to predict the sub-pixel image shows a potential of using this mechanism for image fusion. The PAN image can also constrain the HNN to produce a MS image at sub-pixel resolution. The HNN model for image fusion will be a combination of the HNN super-resolution for MS imagery and the HNN super-resolution using PAN imagery. Future research should focus on testing alternative inverse models that can be used for fusing different sources of remotely sensed data.

7.4.3 Improvement and alternatives for the proposed approaches

Future work should also focus on the improvement and alternatives of the proposed approaches according to the points below:

- An algorithm for choosing the optimal parameters for the HNN automatically.
- A mechanism to obtain fast convergence for the HNN such as a flexible set of parameters.
- Alternative inverse models to increase the accuracy of the sub-pixel land cover maps.

- Alternative optimisation algorithms to obtain the spatial dependence maximisation.

Chapter 8: Conclusions

8.1 Introduction

In this chapter, the works on developing, applying and evaluating the various models for increasing the accuracy of the sub-pixel land cover maps in previous chapters are summarised. From this summarisation, the principal conclusions are provided.

8.2 Summary

8.2.1 Background

Land cover mapping from remotely sensed imagery provides useful information for many authorities at different spatial scales and with different thematic contents. Thank to the emergence of remote sensing systems with increasing spatial and spectral resolution, land cover information can be obtained with finer spatial resolution and accuracy. Although the spatial resolution of remotely sensed imagery has been increased, the spatial resolution of the resulting land cover using conventional classifiers is still limited by the size of a pixel. This problem can be resolved partly by current super-resolution mapping techniques which are based on spatial dependence maximisation. However, limitations still exist for these super-resolution mapping techniques since they are based solely on the spatial dependence.

8.2.2 Objectives

The overall aim of this work was to develop new approaches to make use of the available fused and panchromatic imagery and LiDAR elevation data to increase the spatial resolution and accuracy of thematic maps. The new approaches aimed to surpass the limit of existing approaches in terms of the accuracy and spatial resolution increase. In addition, the thesis also develops new methods for super-resolving multispectral images.

8.2.3 Development and analysis

The development of the approach for using fused imagery to increase the accuracy and spatial resolution of the thematic land cover mapping was described in Chapter 3. The fused image, which is obtained by combining the coarse spatial MS image and the fine spatial resolution PAN, were employed as supplementary data for super-resolution mapping using forward and inverse models. The forward and inverse

models were developed based on a linear mixture model and local end-member spectra values. The forward and inverse models were incorporated into the super-resolution mapping process through a HNN model in the form of a single reflectance function for each spectral band of the fused image.

The analyses in Chapter 3 show the potential of the proposed approach. The land cover classes generated by the new method were visually and statistically more accurate after using the fused image. For both the simulated and degraded data, the overall accuracy of the super-resolution maps increased by approximately 3% with local end-member spectra values. In particular, the increased accuracy was mostly seen in the land cover classes which contained small objects. Small linear objects, which were split by the clustering function of the super-resolution mapping without the fused image, were re-mapped owing to the information at the finer spatial resolution of the fused image.

The approach developed in Chapter 4 modified the super-resolution mapping using the fused image approach. Here, the information from the PAN image was utilised directly for super-resolution mapping to circumvent the need for image fusion as an intermediate step. The reflectance functions for spectral bands in Chapter 3 were replaced by a single panchromatic reflectance function. This function value was derived from forward and inverse models based on the local end-member spectra and spectral convolution.

Visual and statistical analyses both demonstrated the more accurate results of the new approach in comparison with hard classification and super-resolution without using the PAN image. Similar to the results obtained with the fused image, the information from the simulated PAN image increased the overall accuracy by approximately 3% for the same set of simulated data. For the degraded data, the overall accuracy was increased by around 2%. Amongst the land cover classes, the accuracy is increased most for classes which consist of sub-pixel objects such as lone trees, hedgerows or roads. This argument was confirmed by the reduction of commission and omission errors for both simulated and degraded data.

Misregistration between PAN and MS images has an effect on both of the two new super-resolution mapping approaches. The investigation in Chapter 3 and Chapter 4

for the effect of misregistration on the KIA values showed that the two approaches are applicable for RMS image registration errors of up to 1 pixels.

In Chapter 5, three approaches of super-resolving fine sub-pixel MS images from coarse MS imagery were introduced. The first approach utilised the HNN super-resolution mapping technique with the forward model to predict the sub-pixel MS image from spectral class proportions which were estimated based on unsupervised classification. The feasibility of the method was evaluated based on the visual evaluation and RMSE. Both the visual inspection and RMSE showed that the new method can generate the MS image with a finer spatial resolution. Visually, the super-resolved image was sharper than the original coarse resolution image. Similarly, the RMSEs of the super-resolved images compared to the reference images were smaller than those of the original degraded images, due to the super-resolution of the mixed pixels. The statistics also demonstrated that when the zoom factor increased, the resulting sub-pixel images were closer to the references.

The second approach required prior information in the form of discrete semivariance functions to combine with the reflectance constraint in the HNN model. The approach was used to create a sub-pixel MS image with the spectral features of the coarse resolution image and spatial variation at sub-pixel resolution. The analyses revealed that the semivariogram matching process smoothed the coarse image if the prior spatial variation at the sub-pixel resolution has a small variance and increased the variance in the smooth area.

Semivariogram matching was subsequently used for image smoothing of the MS image. Using semivariance values of zero, the HNN semivariogram matching model generated smooth sub-pixel images with smaller RMSEs than those produced by bilinear interpolation. The method can be used to smooth the resulting sub-pixel image from the HNN super-resolution with forward model.

Chapter 6 presented work on utilising LiDAR elevation data for super-resolution mapping. In contrast with the fused and PAN image, the LiDAR elevation data revealed sub-pixel information on only a few class objects such as trees, buildings, power lines and other human-made objects. The general model of the method, therefore, was different from the model used in Chapter 3 and 4. In this model, a new height function was added to the HNN energy function based on the class probability.

The model was implemented for identifying the buildings out of the trees and ground classes. The results demonstrated that LiDAR elevation data can be fused with optical data for the HNN to predict accurately the land cover at a sub-pixel resolution. The results showed a considerable increase in all accuracy statistics of the new technique, particularly for building objects.

8.3 Conclusions

- Accurate mapping of land cover types, which is essential to various users, can be obtained with the HNN super-resolution mapping using supplementary data.
- Fused and panchromatic imagery can be used to increase the accuracy and details of sub-pixel maps predicted by the HNN for super-resolution mapping.
- The reflectance value of land cover classes varies geographically and local-end member spectra values are more suitable than a single set of global end-member spectra values for incorporating the information obtained from the intermediate spatial resolution fused and panchromatic imagery.
- Misregistration errors have an effect on the new super-resolution mapping approaches. However, with the panchromatic image registered at the RMSE of 1 pixel, the new approaches still produce more visually and statistically accurate results.
- LiDAR elevation data can provide information to increase the accuracy and detail for a certain number of land cover classes such as trees, buildings and other human-made objects.
- The HNN super-resolution mapping technique with the forward model to predict the sub-pixel MS image from spectral class proportions which were estimated based on unsupervised classifications.
- The HNN can be used to create the sub-pixel MS image with spectral features of the coarse resolution image and spatial variation at sub-pixel resolution based on prior information in form of discrete semivariance functions. The semivariogram matching process smooths the coarse image if the prior spatial variation at sub-pixel resolution is low and increases the spatial variation in the smooth area.
- The semivariogram matching can be used to smooth the MS image. Using semivariance values of zero, the HNN semivariogram matching model can

generate smooth sub-pixel images with smaller RMSEs than those produced by bilinear interpolation.

References

Akgun, T. Altunbasak, Y. Mersereau, R.M. , 2005, Super-resolution reconstruction of hyperspectral images , *IEEE Transactions on Image Processing,* Vol. 14, 1860-1875.

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

Atkinson, P. M., Cutler, M. E. J., and Lewis, H., 1997, Mapping sub-pixel proportional land cover with AVHRR imagery, *International Journal of Remote Sensing*, Vol. 18, 817-935.

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

Atkinson, P. M., 1997, Mapping sub-pixel boundaries from remotely sensed images, *Innovation in GIS 4*, 166-180.

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

Atkinson, P. M., 1999, Assessing accuracy in fuzzy land cover maps, *Proceedings of the 25th Annual Technical Conference of the Remote Sensing Society, Earth Observation: from Data to Information,* Cardiff University, Wales, UK.

Atkinson, P. M. and Lewis, P., 2000, Geostatistical classification for remote sensing: an introduction, *Computers & Geosciences,* Vol. 26, 361-371.

Atkinson, P. M. and Tate, C., 2000, Spatial Scale Problems and Geostatistical Solutions: A review, *The Professional Geographer*, Vol. 52, 607-623.

Atkinson, P. M., 2001, Super-resolution target mapping from soft-classified remotely sensed imagery, *Fifth International Conference on GeoComputation*, University of Leeds, CD-ROM.

Atkinson, P. M., 2003, Super-resolution land cover classification using geostatistical optimization, Edited by X. Sanchez-Villa, *GeoENV IV: Geostatistics for Environmental Applications*, Kluwer-Dordrecht.

Atkinson, P. M. , 2006, Super-resolution target mapping from soft-classified remotely sensed imagery, *Photogrammetric Engineering and Remote Sensing*, **in press**.

Baltsavias, E. P., 1999, A comparison between photogrammetry and laser scanning, *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 54, no. 2-3, 68-82.

Bárdossy, A. and Samaniego, L., 2002 Fuzzy rule-based classification of remotely sensed imagery, *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 40, 362-374.

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

Bezdek, J. C., Keller, J. M., Krishnapuram, R., and Pal, N. R. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing.* Boston, 1999.

Bernard, A. C., Winkinson, G. G., and Kanellopoulos, I., 1997, Training strategies for neural network soft classification of remotely sensed imagery, *International Journal of Remote Sensing*, Vol. 18, 1851-1856.

Binaghi, E, Brivio, P. A., Ghezzi, P. and Rampini, A., 1999, A fuzzy set-based accuracy assessment of soft classification, *Pattern Recognition Letters,* Vol. 20, 935-948.

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

Blanc, P., Blu, T., Ranchin, T., Wald, L., and Aloisi, R., 1998, Using iterated rational filter banks within the ARSIS concept for producing 10m Landsat multispectral images, *International Journal of Remote Sensing*, Vol. 19, 2331-2343.

Borel, C. C. and Gerstl, S. A. W., 1994, Nonlinear spectral mixing models for vegetative and soil surfaces, *Remote Sensing of Environment,* Vol. 47, 403-416.

Bosdogianni, P, Petrou, M. and Kittler J., 1997, Mixture Models with higher order moments, *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 35, 341-353.

Braswell, B., Hagen, S., Frolking S. and Salas, W., A multivariable approach for mapping sub-pixel land cover distributions using MISR and MODIS: Application in the Brazilian Amazon region, *Remote Sensing of Environment,* In press.

Brown, M, Lewis, H. and Gunn, S., 1999 Support Vector Machines for optimal classification and spectral unmixing, *Ecological Modelling,* Vol. 120, 167-179.

Brown, M, Lewis, H. and Gunn, S., 2000, Linear spectral mixture models and Support Vector Machines for Remote Sensing, *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 38, 2346-2360.

Campbell, J. B., 1996, *Introduction to Remote Sensing*, Second edition, Taylor & Francis, London.

Carpenter, G. A, and Grossberg, 1987, A massive parallel architecture for a self-organizing neural pattern recognition machine, *Computer Vision Graphics Image process*, Vol. 37, 54-115.

Carpenter, G. A, Gjaja, M. N., Gopal, S., and Woodcock, C. E., 1997, ART neural networks for remote sensing: Vegetation classification from Landsat TM and terrain data, *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 35, 308-325.

Carpenter, G. A, Gopal, S., Macomber, S., Martens, S., and Woodcock, C. E., 1999, A neural network method for mixture estimation for vegetation mapping, *Remote Sensing of Environment,* Vol. 70, 138-152.

Carper, J. W., Lillesand, T. M., and Kiefer R. W., 1990, The use of Intensity-Hue-Saturation transformations for merging SPOT panchromatic and multispectral image data, *Photogrammetric Engineering and Remote Sensing*, Vol. 56, 459-467.

Chavez, P. S., Sides, S. C. Jr., and Anderson, J. A., 1991, Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic, *Photogrammetric Engineering and Remote Sensing*, Vol. 57, 459-467.

Chen, C. M., Hepner, G. F., and Forster, R. R., 2003, Fusion of hyperspectral and radar data using the IHS transformation to enhance urban surface features, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 58, 19-30.

Clayton, D. G., 1974, Gram-Schmidt Orthogonalization, *Applied Statistics*, Vol. 20, 335-338.

Cliché, G. and Bonn, F, 1985, Integration of SPOT panchromatic channel into multispectral mode for images sharpness enhancement, *Photogrammetric Engineering and Remote Sensing*, Vol. 51, 811-816.

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

Davidson, D. A., Watson, A. I., 1995, Spatial variability in soil moisture as predicted from airborne thematic mapper (ATM) data, *Earth Surface Processes and Landforms* Vol. 20, 219-230.

de Bruin, S., 2000, Predicting the Areal Extent of Land-Cover Types Using Classified Imagery and Geostatistics, *Remote Sensing of Environment*, Vol. 74, 387-396.

Deutsch, C. V. and Journel, A. G. (1998), *GSLIB: Geostatistical Software and User's Guide*, Second Edition, Oxford University Press: Oxford.

Dungan, J. L., 1998, Spatial prediction of vegetation quantities using ground and image data, *International Journal of Remote Sensing*, Vol. 19, 267-285.

Eastman, J. R. and Laney, R. M., 2002, Bayesian soft classification for sub-pixel analysis: A critical evaluation, *Photogrammetric Engineering and Remote Sensing*, Vol. 68, 1149-1154.

Elad, M. Feuer, A., 1999, Super-resolution reconstruction of image sequences, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, 817-834.

Foody, G. M., 1992, On compensation for the chance agreement in image classification accuracy assessment, *Photogrammetric Engineering and Remote Sensing*, Vol. 58, 1459-60.

Foody, G. M., 1995, Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 50, 1-12.

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

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*, Vol. 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*, Vol. 19, 2593-2599.

Foody, G. M., 2000, Estimation of sub-pixel land cover composition in the presence of untrained classes, *Computer & Geosciences*, Vol. 26, 469-478.

Foody, G. M., 2002a, Hard and soft classifications by a neural network with a non-exhaustively defined set of classes, *International Journal of Remote Sensing*, Vol. 23, 3853-3864.

Foody, G. M., 2002b, The role of soft classification techniques in the refinement of estimates of Ground Control Point location, *Photogrammetric Engineering and Remote Sensing*, Vol. 68, 897-903.

Foody, G. M., 2002c, Status of land cover classification accuracy assessment, *Remote Sensing of Environment*, Vol. 80, 185-201.

Foody, G. M., Muslim, A. M., and Atkinson, P. M., 2003, Super-resolution mapping of the shoreline through soft classification analyses, *Proceedings of IEEE Geoscience and Remote Sensing International Symposium, IGARSS 2003*, Vol. 6, Toulouse, France.

Foody, G. M., 2004, Thematic map comparison: evaluating the statistical significance of differences in classification accuracy, *Photogrammetric Engineering and Remote Sensing*, Vol. 70, 627-633.

Foody, G. M., Muslim, A. M., and Atkinson, P. M., 2005, Super- resolution mapping of the waterline from remotely sensed data, *International Journal of Remote Sensing*, Vol. 26, 5381 - 5392.

Freeman, W. T., Jones, T. R., Pasztor, E. C., 2002, Example-based super-resolution, *IEEE Computer Graphics and Applications*, Vol. 22, 56-65.

Delgado-Garcia, J., Chica-Omo, M. and Arbarca-Hernandez, F, 2002, Geostatistical digital image merging, *Proceedings of geoENV 2002*, Barcelona, Spain.

Goovaerts, P., 1997, *Geostatistics for Natural Resource Evaluation*, Oxford University Press, Oxford.

Garguet-Duport, B., Girel, J., Chassery, J.-M., and Pautou, G., 1996, The use of multiresolution analysis and wavelets transform for merging SPOT panchromatic and multispectral image data, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, 1057-1066.

Giannetti, F., Montanarella, L. and Salandin, R., 2001, Intergreted use of satellite images, DEMs, soil and substrate data in studying the mountainous land, *International Journal of Applied Earth Observation and Geoinformation,* Vol. 3, 25-29.

Gopal, S., Woodcock, C. E., and Strahler, A. H., 1998, Fuzzy neural network classification of global land cover from a 1° AVHRR data set, *Remote Sensing of Environment,* Vol. 67, 230-243.

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

Goodchild, M. F., 1987, Spatial autocorrelation, *Concepts and Techniques in Modern Geography Catmog* 47, Norwich: Geo-Abstracts.

Haala, N. and Brenner, C., 1999, Extraction of building and trees in urban environments, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 54, no. 2-3, 130-137.

Häme, T., Stenberg, P., Anderson, K., Rauste, Y, Kennedy, P., Folving, S. and Sarkeala, J., 2001, AVHRR-based forest proportion map of the Pan-European area, *Remote Sensing of Environment,* Vol. 77, 76-91.

Hong, G. and Zhang, Y., 2005, The image registration technique of high resolution remote sensing image in hilly area, *Proceedings of the ISPRS joint conference 3rd URBAN 2005 and URS 2005*, Tempe, AZ, USA.

Huang, C. and J. R. G. Townshend, 2003, A stepwise regression tree for nonlinear approximation: applications to estimating sub-pixel land cover, *International Journal of Remote Sensing*, Vol. 24, 75-90.

Kasetkasem, T. Arora, M.K. and Varshney, P.K., 2003, Sub-pixel land cover mapping based on Markov random field models, *Geoscience and Remote Sensing Symposium, 2003, IGARSS '03 Proceedings*, 3456- 3458.

Kasetkasem T., Arora, M. K. and Varshney, P. K., 2005, Super-resolution land cover mapping using a Markov random field based approach, *Remote Sensing of Environment,* Vol. 96, 302-314.

Laben, C. A. and Brower, B. V., 2000, Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening, *United States Patent*, 6,011,875.

Lampinen, J. and Vehtari, A., 2001, Bayesian Approach for Neural Networks – Review and Case Studies, *Neural Networks*, Vol. 14 3, 7-24.

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., 2001, A generalised confusion matrix for assessing area estimates from remotely sensed data, *International Journal of Remote Sensing*, Vol. 22, 3223-3235.

Li, S., Kwok, J. T., and Wang, Y., 2002, Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images, *Information Fusion*, Vol. 3, 17-23.

Lillesand, T. M. and Kiefer, R. W., 2000, *Remote Sensing and Image Interpretation,* Forth edition, John Wiley & Sons, Inc, USA.

Liu, J. G., 2000, Evaluation of Landsat-7 ETM + panchromatic band for image fusion with multispectral bands, Natural Resources Research, Vol. 9, 269-276.

Luzum, B. J., Slatton, K. C. and Shrestha, R. L., 2004, Identification and analysis of airborne laser swath mapping data in a novel feature space, *IEEE Geoscience and Remote Sensing Letters*, Vol. 1, 268-271.

Mannan, B., Roy, J. and Ray, A. K., 1998, Fuzzy ARTMAP supervised classification of multi-spectral remotely sensed images, *International Journal of Remote Sensing*, Vol. 19, 767-774.

McGwire, K., Minor, T. and Fenstermaker, L., 2000, Hyperspectral Mixture Modelling for Quantifying Sparse Vegetation Cover in Arid Environments, *Remote Sensing of Environment*, Vol. 72, 360-374.

Marceau, D. J., Howarth, P. J., Gratton, D. J., 1994, Remote sensing and the measurement of geographical entities in a forested environment. 1. The scale and spatial aggregation problem, *Remote Sensing of Environment*. Vol. 49, 93-104.

Maselli, F., Conese, C., De Filippis, T., and Norcini, S., 1995, Estimation of forest parameters through fuzzy classification of TM data, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 33, 77-84.

Mehrotra, K., Mohan, K. C. and Ranka S., 1997, *Elements of Artificial Neural Networks*, The MIT press, Cambridge, Massachusetts.

Merterns, K. C., Verbeke, L., Ducheyne, E. I., and De Wulf, R., Using genetic algorithms in sub-pixel mapping, *International Journal of Remote Sensing*, Vol. 24, 4241-4247.

Merterns, K. C., Verbeke, L., Westra, T., and De Wulf, R., 2004, Sub-pixel mapping and sub-pixel sharpening using neural network predicted wavelet coefficients, *Remote Sensing of Environment*, Vol. 91, No. 2, 225–236.

Miranda, F. P., Fonseca, L. E. N., and Carr, J. R., 1998, Semivariogram textural classification of JERS-1 (Fuyo-1) SAR data obtained over a flooded area of the Amazon rainforest, *International Journal of Remote Sensing*, Vol. 19, 937-953.

Munechika, C. K., Warnick, J. S., Salvaggio, C. and Schott, J. R., 1993, Resolution enhancement of multi-spectral image data to improve classification accuracy, *Photogrammetric Engineering and Remote Sensing*, Vol. 59, No. 1, 67-72.

Nguyen, Q. M., Atkinson, P. M. & Lewis, H. G., 2005a, Super-resolution mapping using Hopfield neural netwok with LIDAR data. *IEEE Geoscience and Remote Sensing Letters*, Vol. 2, 366-370.

Nguyen, Q. M., Atkinson, P. M. & Lewis, H. G, 2005b, Super-resolution mapping using Hopfield neural netwok with panchromatic image. *Asian Conference on Remote Sesing ACRS*, Hanoi, Vietnam, CD-ROM.

Nguyen, Q. M., Atkinson, P. M. & Lewis, H. G, 2006, Super-resolution mapping using Hopfield neural netwok with fused images. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 44, No. 3, 736-749.

Núñez, J., Otazu, X., Fors, O., Prades, A., Pala, V. and Arbiol, R., 1999, Multiresolution-based image fusion with additive wavelet decomposition, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, 1204-1211.

Papoulis, A., 1984, *Probability, random variables and stochastic processes*, Second edition, Singapore: McGraw-Hill International Editions.

Pardo-Igúzquiza, E., Chica-Olmo, M., and Atkinson, P. M., 2006, Downscaling cokriging for image sharpening, *Remote Sensing of Environment*, Vol. 102, 86-98.

Pellemans, A. H. J. M., Jordans, R. W. L., and Allewijn, R., 1993, Merging multispectral and panchromatic SPOT images with respect to the radiometric properties of the sensor, *Photogrammetric Engineering and Remote Sensing*, Vol. 59, 81-86.

Pohl, C. and Van Genderen, J. L., 1998, Multisensor image fusion in remote sensing: concepts, methods and applications, *International Journal of Remote Sensing*, Vol.19, 823 – 854.

Price, J., 1999, Combining multispectral data of differing spatial resolution, *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 37, 1199-1203.

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*, Vol.13, 415 – 425.

Ranchin, T. and Wald, L., 1993, The wavelet transform for the analysis of remotely sensed images, *International Journal of Remote Sensing*, Vol.14, 625-619.

Ranchin, T. and Wald, L., 2000, Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation, *Photogrammetric Engineering and Remote Sensing*, Vol. 66, 49-61.

Ranchin, T., Aiazzi, B, Alparone, L., Baronti, S. and Wald, L., Image fusion- the ARSIS concept and some successful schemes, 2003, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 58, 4-18.

Richards, J. A., 1993, *Remote Sensing Digital Image Analysis,* Second Edition, Berlin: Springer-Verlag.

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G. and Green, R. O., 1998, Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models, *Remote Sensing of Environment*, Vol. 65, 267-279.

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G. and Green, R. O., 1998, Mapping Chaparral in the Santa Monica Mountains using multiple end-member spectral mixture models, Remote Sensing of Environment, Vol. 65, 998, 267-279.

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, Vol. 71, No. 2, 272-281.

Pinilla Ruiz, C. and Ariza Popez, F., J., 2002, Restoring SPOT images using PSF-derived deconvolution filters, *International Journal of Remote Sensing*, Vol. 23, 2379-2391.

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

Simpson, J. J. and Keller, R. H., 1995, An improved fuzzy logic segmentation of sea-ice, clouds, and ocean in remotely sensed Arctic imagery, *Remote Sensing of Environment,* Vol. 54, 290-312.

Schackleford, A. and Davis, C. H., 2003, A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9, 1920-1932.

Schneider, W. and Steinwendner, J., 1999, Land cover mapping by interrelated segmentation and classification of satellite images, *International Archives of Photogrammetry and Remote Sensing,* Vol. 32, part 7-4-3.

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

Sohn, G. and Dowman, I., 2003, Building extraction using LiDAR DEMs and IKONOS images, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 33(3/W13).

Solberg, A.H.S. Jain, A.K. and Taxt, T., 994, Multisource classification of remotely sensed data: fusion of Landsat TM and SAR images, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, 768-778.

Solaiman, B., Pierce, L.E. and Ulaby, F.T., 1999, Multisensor data fusion using fuzzy concepts: application toland-cover classification using ERS-1/JERS-1 SAR composites, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, 1316-1326

Steinnocher, K., 1999, Adaptive fusion of multi-source raster data applying filter techniques, *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, part 7-4-3.

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,* Vol. 39, 781-796.

Tatem, A. J., Lewis, H. G., Atkinson, P. M. and Nixon, M. S., 2001b, Multi-class land cover mapping at the sub-pixel scale using a Hopfield neural network, *International Journal of Applied Earth Observation and Geoinformation,* Vol. 3, 184-190.

Tatem, A. J., Lewis, H. G., Atkinson, P. M. and Nixon, M. S., 2002a, Super-resolution land cover pattern prediction using a Hopfield neural network, in G. M. Foody and P. M. Atkinson (eds), *Uncertainty in Remote Sensing and GIS*, John Wiley & Sons, 59-76.

Tatem, A. J., 2002b, Super-resolution land cover mapping from remotely sensed imagery using a Hopfield neural network, *Unpublished Ph.D. Thesis,* University of Southampton.

Tatem, A. J., Lewis, H. G., Atkinson, P. M. and Nixon, M. S., 2002c, Super-resolution land cover pattern prediction using a Hopfield neural network, *Remote Sensing of Environment,* Vol. 79, 1-14.

Theseira, M. A., Thomas, G., Sannier, C. A. D., 2002, An evaluation of spectral mixture modelling applied to a semi-arid environment, *International Journal of Remote Sensing*, Vol. 23, 687-700.

Theseira, M. A., and Thomas, G., 1999, An evaluation of multiple end-member spectral mixture model applied to an agricultural landscape in England, *Proceedings of the 25th Annual Conference of The Remote Sensing Society*, Cardiff, United Kingdom.

Thornton, M. W., Atkinson, P. M. and Holland, D. A., 2006, Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping, *International Journal of Remote Sensing*, Vol. 27, 473–491.

Tipping, M. E. and Bishop, C. M., Bayesian image super-resolution, in *Advances in Neural Information Processing Systems (NIPS 2002)*, Vol. 15. The MIT Press, 2003.

Townshend, J.R.G., Huang, C., Kalluri, S. N. V., Defries, R. S., Liang, S., and Yang, K. 2000, Beware of per-pixel characterization of land cover, *International Journal of Remote Sensing*, 21, 839-843.

Tso, B. and Mather P. M., 2001, *Classification Methods for Remotely Sensed Data*, Taylor & Francis, London.

Van der Meer, F. and De Jong, S.M., 2000, Improving the results of spectral unmixing of Landsat Thematic Mapper imagery by enhancing the orthogonality of end-members, *International Journal of Remote Sensing*, Vol. 21, 2781- 2797.

Vapnik, V, 1998, *Statistical Learning Theory,* John Wiley & Sons, Inc, USA

Verhoeye, J. and De Wulf, R., 2002, Land cover mapping at sub-pixel scales using linear optimisation techniques, *Remote Sensing of Environment,* Vol. 79, 96-104.

Wald, L., Ranchin, T., and Mangolini, M., 1997, Fusion of satellite images of different spatial resolution: Assessing the quality of resolution images, *Photogrammetric Engineering and Remote Sensing*, Vol. 63, 691-699.

Wang, F., 1990, Improving remote sensing image analysis through fuzzy information representation, *Photogrammetric Engineering and Remote Sensing*, Vol. 56, 1163-1168.

Wang, F., 1990, Fuzzy supervised classification of remote sensing images, *IEEE Transactions on Geoscience and Remote Sensing*, 28, 194-201.

Yesou, H, Besnus, Y. and Rolet, J., 1993, Extraction of spectral information from Landsat TM data and merger with SPOT panchromatic imagery-a contribution to the study of geological structures, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 48, 23-36.

Yocky, D., 1996, Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data, *Photogrammetric Engineering and Remote Sensing*, Vol. 68, 1067-1073.

Zhan, Q., Molenaar, M., and Lucieer, A., 2002, Pixel unmixing at the sub-pixel scale based on land cover class probabilities: Application to urban areas, in G. M. Foody and P. M. Atkinson (eds), *Uncertainty in Remote Sensing and GIS*, John Wiley & Sons, 59-76.

Zhang, J. and Foody, G. M., 1998, A fuzzy classification of sub-urban land cover from remotely sensed imagery, *International Journal of Remote Sensing*, Vol. 19, 2721-2738.

Zhang, J. and Foody, G. M., 2001, Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: statistical and artificial neural network approaches, *International Journal of Remote Sensing*, Vol. 22, 615-628.

Zhang, Y., 1999. A new merging method and its spectral and spatial effects. *International Journal of Remote Sensing*, Vol. 20, 2003-2014.

Zhang, Y. and Maxwell, T., 2006, A fuzzy logic approach to supervised segmentation for object-oriented classification, *ASPRS 2006 Annual Conference*, Reno, Nevada, USA.