A

Dissertation

On

An Efficient Mixed Pixel Resolution methodology using Bio-Inspired Heuristics

Submitted in partial Fulfilment of the requirement

For the award of Degree of

Master of Technology
In
Computer Science and Engineering

Submitted By

TANU VARSHNEY

College Roll No. 2K11/CSE/18

Under the esteemed guidance of

Dr. Daya Gupta

Department of Computer Engineering

DELHI TECHNOLOGICAL UNIVERSITY

2011-2013

DELHI TECHNOLOGICAL UNIVERSITY

DELHI - 110042

CERTIFICATE

This is to certify that the work contained in this dissertation entitled "**An Efficient Mixed Pixel Resolution methodology using Bio-Inspired Heuristics**" submitted in the partial fulfilment, for the award for the degree of Master of Technology in Computer Science and Engineering at **DELHI TECHNOLOGICAL UNIVERSITY** by **TANU VARSHNEY, Roll No. 2K11/CSE/18,** is carried out by her under my supervision. This matter embodied in this project work has not been submitted earlier for the award of any degree or diploma in any university/institution to the best of our knowledge and belief.

Date: ___________

(Dr. DAYA GUPTA)

HOD & Project Guide

Department of Computer Engineering

Delhi Technological University

ACKNOWLEDGEMENT

I would like to take this opportunity to thank everyone who directly or indirectly contributed to this work.

I am extremely grateful to my mentor Dr.Daya Gupta, HOD, Department of Computer Engineering, Delhi Technological University, for her expert guidance and constant encouragement to excel. Her in depth knowledge and constructive comments throughout the research work have been extremely helpful in successful completion of this work.

I sincerely thank Dr. Vinod Kumar Panchal, Associate Director, DTRL (Defence Terrain Research laboratory), Defence Research Development Organization (DRDO) for his valuable suggestions for further improvement of this work. I am thankful to Ms. Lavika Goel (Research scholar, Department of Computer Engineering, Delhi Technological University) for her valuable inputs and support throughout the research work.

I am grateful to my family and friends for being supportive, encouraging and inspiring to achieve my objectives.

Tanu Varshney

University Roll no: 2K11/CSE/18

M.Tech (Computer Science & Engineering)

Department of Computer Engineering

Delhi Technological University

Delhi - 110042

Abstract

Remote sensing has been used for extracting information about regions or objects for a long time. The process of gathering information without coming in direct contact with the object, is advantageous in identifying even those areas which seem to be nearly inaccessible to humans for example in monitoring earthquake hit areas. These remote sensing systems record brightness values emitted from the object at different wavelengths that commonly include not only portions of the visible light spectrum, but also infrared and, in some cases, middle infrared bands. The brightness values for each of these bands are typically stored in a separate grayscale image (raster). However, certain regions have brightness value which is composed of brightness from more than one region. Pixels depicting brightness for such regions in an image are known as the Mixed Pixels.

In the process of image classification, the mixed pixels pose a major problem. It becomes difficult to assign a accurate single class to these pixels. If these pixels are left as such then complete and accurate information may not be obtained from the remotely sensed data. Our work intends to solve the problem of mixed pixels in a remotely sensed image using recently (December, 2008) introduced bio inspired algorithm known as Biogeography based optimization. In last few years, bio-inspired methods have rapidly gained importance in computing due to the need for flexible and adaptable ways of solving engineering problems. They are a class of algorithms that imitate specific phenomena from nature. Bio-inspired algorithms are based on the structure and functioning of complex natural systems and tend to solve problems in an adaptable and distributed fashion. Several systems such as the ant-colony system, bee foraging, bird flocking etc. have been used as the basis for developing models and algorithms to solve various issues.

We have utilised the migration behaviour of a species residing in a habitat. Species immigrate to a new habitat based on several factors such as vegetation, rainfall, temperature etc. These

factors decide the fitness of a habitat for supporting the species residing in it. Based on this approach we find the fittest habitat for the mixed pixel and assign a class to it accordingly. The species are migrated taking into account the effect of external factors, in the form of sinusoidal migration pattern. A good candidate solution has relatively high emigration rate and low immigration rate, while the converse is true for a poor candidate solution.

The results produced have been compared with the previous work to show the efficiency of the proposed algorithm. In previous work, linear migration has been applied for the mixed pixel problem. We have designed our algorithm based on sinusoidal migration. The sinusoidal migration curve has been modelled in two approaches. The first approach incorporates all the bands in a remotely sensed image. The second approach finds the non redundant bands and then algorithm is applied on the selected bands.

TABLE OF CONTENTS

List of Figures	

List of Tables

Chapter 1

Introduction

In remote sensing, the main objective is to extract maximum information from the received satellite image. Thus, image classification forms an important phase in extracting required details from the satellite captured image. Remote sensing is an approach for extracting information from the object concerned without coming in direct contact with the object. It finds major application in various fields such as agriculture, forestry, geology, meteorology etc [34]. Natural surfaces are formed from a variety of features where an area is seldom made of a single uniform category [1].

In order to identify the features/objects in an image, the image is preprocessed to remove any noise and then it can be classified into desired regions/objects. However, a region/feature may not be recognized into a single class. There may be more than one interpretation for some regions due to presence of the mixed pixels. The mixed pixels are those pixels which contain spectral reflectance information from more than one region. In such a case, the pixel may not be classified to a single class leading to ambiguity. This affects the classification results. The information deduced from these results will in turn be ambiguous. Thus to obtain accurate results these mixed pixels need to be resolved.

Mixed pixels are erroneous measurement occurring near the boundary of objects in the captured image. Scattering and interference from multiple light sources are one of the sources of this error. The presence of mixed pixels in an image reduces the accuracy of data and lowers the quality of the image. The Mixed Pixel Resolution techniques are required with focus on improving the quality of the image by reducing the errors and distortions caused by multi-path interference [2].

In this work, the biogeography based technique is applied on the data of mixed pixels to optimally assign them a unique feature class.

1.1 Motivation

In the past years, researchers have been working in improving image processing techniques. Features in an image are identified by tagging a pixel to the class whose spectral signature it resembles the most. There are certain pixels which classify a region as belonging to more than one land cover type. These are called as Mixed Pixels. For a long time, problem of mixed pixels has been a hurdle in obtaining accurate results. Mixed pixels are normally left out of analysis as undesirable and problematic.

The mixed pixels often arise in low resolution image because a pixel does not cover a single feature category. Low resolution data is generally preferred for analysis of remotely sensed areas because the storage and processing costs associated with high resolution satellite data are large.

Recently nature inspired technique of Biogeography based optimization (BBO) has been applied in several domains such as Land cover feature extraction [5], Intelligent battle field planning [38], and to solve the Economic load dispatch problem [22]. Very recently BBO algorithm, modelled on linear migration concept, has been applied in resolution of mixed pixels. However, linear curve cannot represent certain natural phenomenon such as predator/prey relationships, species mobility etc effectively.

These factors motivated us to adapt BBO algorithm, which could implement these constraints using non linear migration concept.

1.2 Related work

Extracting a feature class that correctly represents the region denoted by mixed pixels is a challenging task. Research community has been continuously searching for better mixed pixel resolution techniques. The problem of mixed pixels has traditionally been solved using the linear mixing approach [23] [26] which involves two steps. First step is to find the spectral information of end member i.e. pure pixels. Second step is to find the spectral percentage of each pure feature in the mixed pixel [29]. However, their drawback is that they can unmix only as many components as there are number of bands [32]. Some researchers have extended linear mixing models [37] to take into account correlation between pixels. Geometry [33] of the targets have been included to detect effect of shadows in pixels.

Fuzzy based approaches have been proposed [23] [27] by researchers to determine constituents of mixed pixel. The fuzzy unmixing approach [36] has been used to estimate the membership of a mixed pixel in various possible mixing proportions. However, it becomes less accurate when the number of component classes is more than the number of bands. Artificial neural networks (ANN) [29] have also been proposed by the research community for decomposition of mixed pixels. Their architecture is based on characteristics of the data. ANNs have their own limitations. There are lot of parameters to adjust as well as training the network can be a long and tedious procedure. Architecture and weight selection determine the accuracy of the network.

BBO takes inspirations from geographical distribution of organisms. Migration of species is an important concept in BBO. It is a means of sharing information between the solutions. In course of the process, good solutions tend to share their features with the poor ones by using the migration operator. The poor solutions improve themselves by accepting these new features from the good ones [24]. The migration curves depict the relationship of the number of species migrating and their migration rates.

Linear migration concept of BBO has been applied [4] for resolving mixed pixels. In linear migration, the immigration rate is directly proportional to the fitness of the candidate solution. To resolve the mixed pixels, the feature class representing maximum fitness, is assigned to the mixed pixel. However, linear migration cannot take into account other constraints such as predator/prey concepts, evolution of particular species, and population size [2]. In addition, the performance of the algorithm degrades with the increase in the number of pixels.

1.3 Problem Statement

In the forgoing sections, we have explored the research on mixed pixel and seen that linear migration concept of BBO has number of constraints. It has been observed in recent literature that sinusoidal migration curves better represent the natural migration phenomenon as compared to the existing approach of using linear curves. In order to improve information extraction and analysis of mixed pixel data, we wish to develop non linear sinusoidal migration model as a new technique for resolving mixed pixels in a satellite image. Since the mixed pixels contain more than one region and every region is recorded in several wavelengths. Different wavelengths are suited for recording spectral response of different regions depending on their physical state and chemical composition. Therefore, some bands may contain redundant(irrelevant) information of the region. If we select only non redundant bands, the algorithm can be further improved.

Thus, our problem statement can be proposed as follows-

“To develop sinusoidal biogeography based optimization algorithm incorporating non redundant bands of the satellite image for efficient resolution of the mixed pixels into a specific land cover feature class.”

1.4 Scope of the work

In this research work, we design a new algorithm based on biogeography based optimization, that can effectively assign feature class to mixed pixels. Our algorithm provides improved accuracy for mixed pixel problem. Our proposed algorithms have been applied on seven band multispectral, multisensor image of Alwar, Rajasthan. This area has been chosen because of its great diversity of land features which include urban, barren, vegetation, water and rocky regions.

We have designed two algorithms for resolving mixed pixels. The first algorithm incorporates sinusoidal migration model of BBO. We have improved this algorithm by removing the redundant bands. So, in addition to using sinusoidal migration, the second algorithm does preprocessing to remove bands with redundant data. The selected bands are incorporated to calculate the fitness of a candidate solution as per the criteria specified in the BBO algorithm.

Scope of this work can be summarized as :

i. To develop an algorithm to resolve mixed pixels using non linear (sinusoidal) migration curve of BBO because this migration better resembles those found in nature.

ii. To improve the developed algorithm by selecting the non redundant bands because all bands do not contain necessary information for a particular land feature. For example vegetation is better represented in red band [3].

iii. Apply these algorithms to Alwar image and establish the results by comparing them with the linear technique [4]. We have obtained Alwar image in which the mixed pixels have been resolved. We have justified efficiency of our nonlinear migration algorithm over previous work based on linear model by comparing it with PSO/ACO classified image as a standard.

1.5 Organization of the thesis

The remainder part of this thesis is organized in the following chapters:

Chapter 2: Remote Sensing in satellite image classification

This chapter describes the process of remote sensing and its related concepts that have been used in our work. The applications of remote sensing in real world have been discussed here. It also describes the basics of satellite image and its features such as resolution, spectral bands and digital numbers. This chapter also discusses the emergence of nature inspired algorithms and how they have been successfully applied in solving hard and complex computational problem. Solving these problems using conventional techniques would have been cumbersome process. These algorithms take inspiration from nature in solving the various optimization problems. Algorithms that have been discussed are ACO, PSO, Hybrid PSO/ACO.

Chapter 3: Mixed Pixels

This chapter gives an introduction to the mixed pixels and reasons for their occurrence in the satellite image. In this chapter, we have discussed the state of the art of the mixed pixel problem. It discusses the previous works on resolving mixed pixels and their drawbacks with reference to our problem.

Chapter 4: The bio-inspired technique of Biogeography based optimization

This chapter introduces the recent biogeography based optimization algorithm that mimics migration phenomenon of species in the biogeography. It further discusses the features of the algorithm such as the habitats, SIV’s, HSI, immigration and emigration rates. This chapter describes the various migration models of BBO both linear and non linear and their related concepts.

Chapter 5: Mixed Pixel Resolution

In this chapter, we introduce our proposed resolution algorithm based on BBO. We have developed two algorithms. The first algorithm applies nonlinear concept of BBO whereas second algorithm is designed to improve the first algorithm by incorporating preprocessing steps of selecting necessary bands. We present the framework of our algorithms as well as parameters and flowchart.

Chapter 6: Experiments and Results

The experimental datasets and the multispectral seven band satellite image of Alwar, Rajasthan are discussed here. We describe the training dataset of Alwar region and the Mixed pixel dataset in this section. We examine the resolved mixed pixels, first using the sinusoidal algorithm that includes all the bands. Then we consider the second improved sinusoidal algorithm that includes only non redundant bands for comparison with the algorithm using linear migration. We evaluate and judge the performance of our experimental results.

Chapter 7: Conclusion and Future Scope

In this chapter, the conclusion of the thesis work and the future scope of the work are presented.

Chapter 8: Publication from Thesis

This section gives the details of publication from the thesis and the conference details.

References: This section gives the reference details of the thesis.

Appendix A: Abbreviations used

Appendix B: Introduction to MATLAB

Appendix C: Introduction to open source tool for PSO/ACO

Appendix D: Introduction to ERDAS

Chapter 2.

Remote sensing in satellite image classification

In this chapter we have introduced the remote sensing procedure in context of satellite image. We have explained the features associated with a satellite image to ease our understanding of their problem of mixed pixel. We then discuss various recent bio-inspired techniques to denote their benefits over the traditional problem solving techniques.

2.1 Introduction to Remote Sensing

Remote sensing generally refers to detecting and classifying objects on earth. Remote sensing is a method of acquiring information about an object or feature without coming in its actual physical contact. In other words, remote sensing is the sensing of an object or a phenomenon from remote distance [18]. Therefore, the electromagnetic radiations from the sunlight form the source of gathering the required information about the object. The electromagnetic spectrum ranges from shorter cosmic and gamma rays to longer radio waves. The electromagnetic spectrum is known to be consisting of seven different regions. The regions are - gamma rays, X-rays, ultraviolet, visible light, infrared, microwaves and radio waves. The various portions of electromagnetic spectrum are useful in the field of remote sensing [14].

Essentially remote sensing has three components :

- The signal (from an object)
- The sensor (from a platform) and
- The sensing (acquiring information from the sensor).

One scene is made up of a number of images. Each image represents one spectral band. Thousands of pixels make up an image. Each pixel has associated with it a brightness value (0-255). Image can display three bands at a time - Red, green and blue and use the information in multiple bands to separate land cover classes.

The device to detect EM radiation reflected or emitted from an object is called the remote 'sensor'. A vehicle to carry sensor is called platform. Aircrafts or satellites are used as platforms. Each object has unique and different characteristics of reflection or emission depending on the chemical type of object and its environmental conditions. Thus, remote sensing is a technology to identify and understand the object or environmental conditions through its uniqueness of reflection or emission.

2.2 Process of remote sensing

The process of Remote Sensing has been depicted in figure 2.2. The steps in the process of remote sensing are as follows -

A. The source of electromagnetic radiation in this case is the sun. The EM energy from the sun is passes through the earth's atmosphere and reaches the target.

B. The target reflects the EM radiation according to its characteristic physical and chemical component of the target. A part of this energy gets scattered and absorbed. The reflected radiations travel through the atmosphere to reach the satellite sensor.

C. Sensors mounted on the satellite detect and record the radiation reflected from the target. Some sensors have additional onboard data storage devices and can temporarily store large data volumes.

D. The data which is detected and recorded by the sensor is received by ground receiving station.

E. The data processing facility converts the analog signal received from satellite sensor to digital signal.
F. The data needs to be further preprocessed to extract meaningful information from it
G. The data is now ready for analysis
H. User is provided with the required information such as thematic map, forest cover and disaster management.

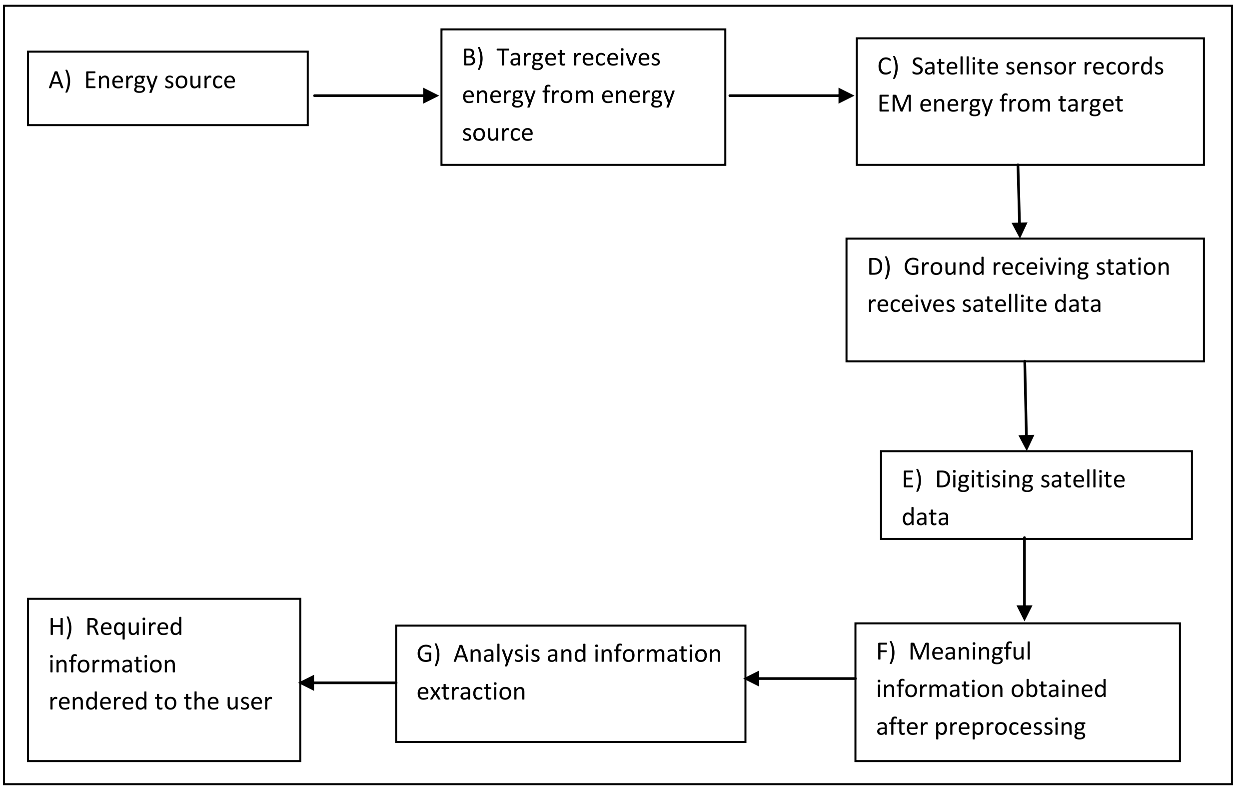

Figure 2.2 Process of remote sensing

2.2.1 Passive remote sensing

Based on the type of source of electromagnetic radiation, remote sensing can be categorised into two types – passive and active. In Passive remote sensing [12] the source of energy is natural. The sun is the most powerful and the most commonly used source of energy for passive remote sensing. The radiation measured by passive sensors is the reflected sunlight. A very basic example of passive remote sensors is the photographic film.

When passive remote sensing is carried out without the natural source of energy i.e. the sun, it is known as passive emitted remote sensing. In this case, the source of energy is the target material itself and the sensor records the emitted radiation. When a part of the source's radiation is not reflected back to the sensor, it is absorbed by the target. So, the temperature of the material rises and the radiation absorbed by the material is later emitted at a different wavelength. Remote sensing in the thermal infrared portion of the electromagnetic spectrum occurs through passive emitted remote sensing.

2.2.2 Active remote sensing

In active remote sensing, the source of energy is artificial. Mostly the satellite acts as the source of energy by sending a pulse of radiation towards the target. The nature of the emitted radiation such as wavelength, duration, power etc can be controlled by us. So, active remote sensing can be carried out in all weather conditions. Radar is an example of active remote sensing where the time difference between energy emission and return is measured to find the location, speed and direction of the target [12].

2.2.3 Applications of remote sensing

The development and deployment of manned and unmanned satellites has improved the collection of remotely sensed data. It has given us an inexpensive way to obtain information over large areas on the surface of the earth. The ability of remote sensing process to identify and monitor land surfaces and environmental conditions has enhanced to a large extent over the past few years and remotely sensed data has become an essential tool of gathering information for various domains. Satellite and aircraft platforms gather an extremely large amount of data in a short period of time. This remotely sensed data is often archived in order to track changes in the earth's surface over long periods of time.

An important application of remote sensing is identification of land cover on the earth's surface. The thematic maps are prepared which can then be used in estimation of crop yields, soil type, topography and in assessing how far a field is, from roads or non-agricultural crops. Information from these remotely sensed data could be important in detecting pest infestation or in planning chemical application in the fields [21].

Remote sensing finds application in forestry. It is used in the identification and mapping of the species residing in the forest trees, estimating the geographic extent of forests and monitoring the deforestation activities. Through the process of remote sensing, important information relevant to the geosciences such as in mineral and petroleum exploration, mapping geomorphology, and volcano monitoring can be provided.

In many of the applications specified above, remotely sensed data has been used with a range of other Earth science data. This data can further be beneficial in providing information about the natural environment. This analysis of the data on earth science, from a range of sources is generally made in a geographic information system (GIS).

2.3 Satellite image features

A digital image is represented in the form of a two dimensional matrix of picture elements. These elements are known as pixels and are arranged as elements in rows and columns of a matrix. Every pixel in an image is representative of an area on the surface of the Earth. The intensity value of a pixel is the physical quantity for example the solar radiance reflected from the ground in a given wavelength band, which can be measured by the sensor. It can be emitted infrared radiation or backscattered radar intensity. Pixel intensity is basically the average of the complete ground area covered by the pixel. The intensity value is converted into digital form and then recorded for an image. This is known as a digital number or DN value.

2.3.1 Image resolution

The intrinsic resolution of an imaging system is also determined by the instantaneous field of view (IFOV) of the sensor. The required information is collected on a pathway which is called the field of view (FOV). The measure of the area on the earth's surface as seen by the detector in a given instant of time is called as the IFOV. The IFOV is thus analogous to the pixel size of the sensor. There are various factors that lower intrinsic resolution and cause degradation of the image. Such degradation includes blurring of the image, like improper focusing, atmospheric scattering and target motion. As the satellite orbits the Earth, the basic work of the satellite sensors is to gather information about the reflected radiation from the object of study. The different resolutions specified for an image are as follows -

- **Spectral resolution**

  The spectral resolution denotes the width of the wavelength interval at which the electromagnetic radiation is recorded. Therefore, it is the capability of sensor to detect small differences in wavelength [16].

- **Spatial resolution**

  Spatial resolution is specified in terms of size of the object to be viewed. It is defined as the size of the smallest object that can be resolved by the sensor. In a digital image, the smallest area is covered by the pixel. Therefore, this resolution is limited by the pixel size. This means that the smallest resolvable object cannot be smaller than the size of the pixel. The image with a small resolution size is called as the high resolution image [21]. Finer details of the land cover objects are visible in a high resolution image. On the other hand, a low resolution image is the one with a large resolution size, which means that only coarse features can be seen in the image.

  The figure 2.3.1 given below clearly shows the difference between these two resolutions.

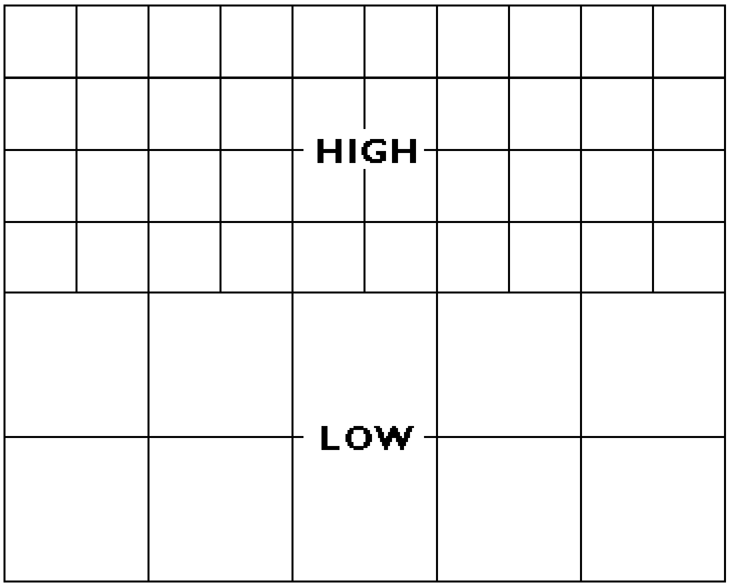

Figure 2.3.1 Spatial resolution (High Vs Low) [21]

- ## Radiometric resolution

It is defined as the smallest change in the intensity level that can be detected by the sensing system. The intrinsic radiometric resolution of a sensing system is affected by the signal to noise ratio of the detector. Quantization digitizes continuous intensity values. So, the radiometric resolution is limited by the number of discrete quantization states.

To make judicious use of the finite storage capacity, the digital number is stored in the memory using a finite number of bits. The number of bits per pixel is known as the bit depth. The bit depth determines the radiometric resolution of the image. The greater the bit depth, greater the radiometric resolution. For example, an 8-bit digital number can take on values from 0 to 255 (2^8 - 1), while an 11-bit digital number can take values from 0 to 2047.

- ## Temporal resolution

The temporal resolution applies to time series of images and describes how long the interval between two successive recordings of the same scene is. In case the scanner is carried by a satellite, the temporal resolution is determined by the satellite's orbit.

2.3.2 Data format

Information about a scene can be extracted from the EM radiation reflected or emitted by the objects. The radiation is measured using imaging scanners aboard aircraft or satellite. The resulting images are corrected and transformed to a digital image comprising of pixels. Satellite images are collected in raster format, which is a matrix of thousands of individual picture elements called pixels. The ground area covered by each pixel determines the resolution of the pixel. For instance, if the image resolution is 30 m, each picture element is restricted to an area on the ground covering 30 m, about 100 ft, square [21].

2.3.3 Spectral bands

The amount of reflected/emitted radiation from a remotely sensed object is measured in spectral bands**.** A discrete interval or range of the electromagnetic spectrum is known as the spectral band. For example the wavelength range of 0.4μm to 0.5μm is one spectral band. The different patterns of reflectance and absorption across different wavelengths are known as the spectral signature of the object such as water, tree, building. The spectral signature of a material aids in distinguishing one material from another. The reflected radiation from a single surface material, such as a rock, is not the same across the electromagnetic spectrum. It varies over the range of wavelengths in the spectrum. All Earth surface features, including minerals, vegetation, soil, water, rocks and snow, have unique spectral reflectance signatures.

Satellites collect data in groups of spectral bands. The three primary colours i.e. red, green and blue are used to display a colour composite image. To generate a colour composite image, a spectral band (not necessarily a visible band) is associated to a separate primary colour. In a natural colour image, the spectral bands are fused together such that the displayed image resembles the actual colours of the object as seen in natural light. The three bands of data are mapped to red, green, and blue colour and are also known as true colour image of the object.

To ease the detection of features that are not readily distinguishable otherwise, false colour composite image are generated. The false-colour encoding maps an RGB image in the visible spectrum, to different colour channels. Typically, the mapping is done as "NRG→RGB" with "N" being the near-infrared spectral band and the blue spectral band being unused. This mapping results in the typical false-colour images where vegetation is in red. False colour image is used for satellite and space images like remote sensing satellite images.

2.4 Bio inspired techniques

It is said that nature is the best teacher and researchers are taking inspiration from it to solve hard and complex problems. Nature provides some of the efficient ways to solve problems. Algorithms imitating processes in nature are called Bio Inspired Algorithms. Problem with traditional techniques have led to increased interest of researchers in this field. There are different methods for different types of problems.

- Constraint handling is problematic. For e.g. using penalty method is sensitive to penalty parameters.
- They often get stuck in local optima and lack global perspective.
- They usually need knowledge of first/second order derivatives of objective functions and constraints.

Natural processes tend move towards the optimal solution while maintaining perfect balance among its components. It is these dynamic, robust designs and capabilities that the researchers are trying to mimic. These bio inspired techniques are being extensively used for solving problems in various areas including computer networks, security, robotics, bio medical engineering, control systems ,parallel processing ,data mining, power systems, production engineering and many more. Swarm Intelligence (SI) is a paradigm in bio inspired which is based on collective social behaviour of organisms. SI techniques implement the collective intelligence of groups of simple agents known as swarm. The multiple agents in a swarm work together without any central control. Few of the bio inspired algorithms have been discussed next.

2.4.1 Ant Colony Optimization

In recent years, several collective decision techniques inspired by social insects have been developed to solve complex problems. One of the techniques introduced in this class, which is inspired by behaviour of ants, is known as Ant colony optimization (ACO). This technique was first introduced by Dorigo [10] and since then it has been successfully applied to several optimization problems. This algorithm uses the mechanism of ant colony foraging. The colony of ants communicates via pheromones to find the optimal path between food and nest of the colony. Inspired by the Ant behaviour, there are various modified algorithms that are developed for solving different problems.

According to ACO algorithm, the first ant wanders randomly until it finds the food. If food is found, it returns to the nest laying down a pheromone trail. If pheromone is found, with some increased probability then other ants follow the pheromone trail, also laying pheromone trail. Once it is back at the nest, it again goes out again in search of food. However, pheromones evaporate over time, such that unless they are reinforced by more ants, the pheromones will disappear. Since the ants on the shortest path lay pheromone trails faster, this path gets reinforced with more pheromone, making it more appealing to future ants. The ants become increasingly likely to follow the shortest path since it is constantly reinforced with a larger amount of pheromones.

The pheromone trails of the longer paths evaporate. Thus, the paradigm for optimization problems is to express them as short path finding problem in a graph. The number of classes is fixed a priori, therefore the size of ant is known. Ants of the population moving in the search space using an objective function measuring the fitness of each ant. The Euclidean distance between learning samples and representatives of classes has been used as objective function in various researches. In ACO, pheromone is used as a chemical messenger. The movement of an ant is controlled by pheromone which will evaporate over time. With proper pheromone evaporation, they usually behave very well.

Without such time-dependent evaporation, the algorithms will lead to premature convergence to the solutions that are generally wrong. So, there are two important issues to be considered - the probability of choosing a route, and the evaporation rate of pheromone. The pheromone concentration is the indicator of quality of solutions to a problem of interest. As the solution is associated with the pheromone concentration, the search algorithms generally produce routes and paths marked by the higher pheromone concentrations. Therefore, it is particularly suitable for discrete optimization problems.

2.4.2 Particle Swarm Optimization

It is another technique inspired by social behaviour of as fish and bird schooling in nature. PSO was first introduced in 1995 by Kennedy and Eberhart [11]. PSO has been applied to almost every area in optimization, computational intelligence, and design/scheduling applications due to its unique searching mechanism, simple concept, computational efficiency, and easy implementation. A swarm of individuals called particles, fly through the multidimensional search space with a velocity, which is constantly updated by the particle's own experience and the experience of the particle's neighbours or the experience of the whole swarm.

The term particles refer to population members which are mass-less and volume-less or with an arbitrarily small mass or volume. They are subject to velocities and accelerations towards a better mode of behaviour. Each particle in the swarm represents a solution in a high-dimensional space with four vectors, its current position, best position found so far, the best position found by its neighbourhood so far and its velocity. The movement of a swarming particle consists of two major components - a social component and a cognitive component. It adjusts its position in the search space based on the best position reached by it and on the best position reached by its neighbourhood during the search process, while at the same time it has a tendency to move randomly.

The basic PSO algorithm is population-based approach in which the group of candidate solutions moves towards a convenient solution or group of solutions for the given problem. As the name implies, it is an optimization method, with the objective of finding a global optima of a real-valued function also known as fitness function, defined in a given space known as the search space. A population of particles flies through the problem space. The multidimensional space in which the particles move, represents the belief space. These particles have a characteristic that they can retain part of their previous state. In other words, particles can be said to have memory. Any number of particles can share the same point in belief space yet each has its own individuality.

Drawing analogy from nature, the particles form a society. Every particle has an opinion which is a part of the search space shared by every possible particle. The particles may change their "opinion state" based on the following factors:

- The knowledge of the environment i.e. its fitness value
- The individual's previous history of states i.e. its memory
- The previous history of states of the individual's neighbourhood

The neighbourhood of an individual can be specified in various ways. It is the neighbourhood that forms the social network of the individual. Each particle's movement is made up of an initial random velocity and two randomly weighted instances:

Individuality, which is the tendency of a particle to return to its best previous position, and

Sociality, which is the tendency of a particle to move towards the neighbourhood's best previous position.

At each iteration, each individual determines its nearest neighbour and replaces its velocity with that of its neighbour. This results in synchronous movement of the flock. As the individuals in the population follow these rules of interaction, they adapt their scheme of belief to those individuals who have been more successful in their social network. With time, the situation comes when the individuals hold closely related opinions.

2.4.3 PSO/ACO

As the name suggests, this approach is the hybrid of PSO and ACO algorithms which have been discussed before. It adds positive features of these algorithms together to produce an algorithm which is better than its constituent algorithms. Ant colony optimization (ACO) and Particle swarm optimization (PSO) are the two main algorithms of swarm intelligence, and because of the self-organization, cooperation, communication and other intelligent merits, they have been greatly applied in remote sensing image processing. As it is a hybrid, it combines the strengths from both the algorithms. Thus, it can directly cope with the nominal attributes, without converting nominal values into numbers in a pre-processing phase.

The Hybrid PSO-ACO as given by Nicholas and Frietas [30] uses sequential covering approach for rule extraction. PSO is quite fast speed to approach optimal solutions, so that it can optimize the parameter of ACO. Therefore ACO and PSO can combine to a Hybrid Algorithm to solve various problems. The rule set is initially empty. In each iteration of the loop, the best rule is discovered from the training samples. The rule returned does not contain any term with continuous values. Thus, it is used as a base for discovering terms with continuous values. Each particle in the PSO/ACO2 population is considered as a group of n pheromone matrices where n is the number of nominal attributes in a data set. A particle is decoded probabilistically into a rule with a predefined consequent class.

We have discussed various features that are characteristic of remote sensing procedure. Advantages of bio-inspired techniques over traditional classification techniques have been discussed which led to the development of proposed algorithm. So, before the implementation of the proposed algorithm, the problem of mixed pixels is introduced in the next chapter.

Chapter 3.

Mixed pixels

In this chapter we explain the mixed pixel problem. We discuss the reasons for the occurrence of mixed pixels and present their state of the art. We have discussed the development and evolution of techniques that have been employed so far for resolution of mixed pixels.

3.1 Mixed pixel problem description

Thematic mapping from remotely sensed data is typically based on an image classification. In image classification, each pixel is classified into one of the many land cover types. However, there may be pixels which consist of more than one land cover type. Such pixels are known as the mixed pixels.

If the Instantaneous Field of View (IFOV) of a sensor records more than one land cover type then it results in mixed pixels as shown in figure 3.1 a). The presence of mixed pixels thus seems to be affected by spatial resolution of the system. It appears that mixed pixels may be reduced by increasing the spatial resolution but it is not always the case [6]. This is because as the spatial resolution is increased, more pure pixels are set inside the boundaries, which in turn may reveal finer details of new features leading to more and new spectral classes. For example the image of a forest region, which was recorded as uniform region at coarse resolution, may display individual trees of various species along with open spaces like grass, soil etc. at finer resolution.

Further increasing the resolution would record the leaves, birds, nests and other flora and fauna in the forest. Thus, increasing the resolution may add on fine details of the land cover which leads to more feature classes to be associated with the system. Even if the numbers of feature classes are kept the same and the number of mixed pixels is reduced, the classified data may still be error prone because the variations in the class increases as local difference in elevation, humidity and illumination become more apparent. Moreover, increase in spatial resolution comes at the cost of the spectral and radiometric resolution. As the IFOV is reduced to obtain fine spatial resolution, the energy received is also reduced. To measure the reduced energy, the spectral bands at which the reflectance is measured need to be broadened, thereby affecting spectral and radiometric resolution. Another drawback is that the number of pixels can become very large, which may add to the cost of processing.

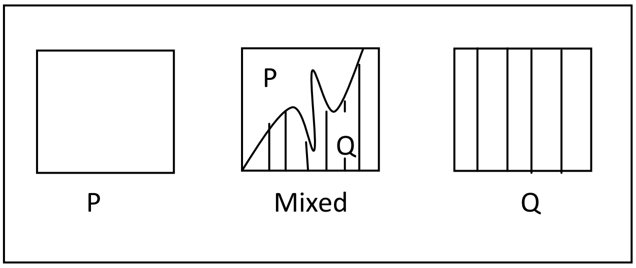

Figure 3.1 a) Mixed pixel

There is a growing interest in data of global coverage recorded on more and smaller spectral bands, new scanners may be designed to have coarser rather than fine spatial resolution. Mixed pixels can be a homogeneous combination of the materials, independent of the sensor resolution. Following are the scenarios when mixed pixels may occur as shown in figure 3.1 b) [8].

- Sub pixel objects : it occurs when there are small objects present in a pixel such as house.
- Intergrade : it is the transition between one form to another form of the same object such as transition in vegetation class from grassland to tree cover.
- Linear subpixel : it is the presence of a thin long structure such as a road which may partly occur in another pixel.

- Boundary pixel : it occurs when the boundary of two or more regions passes through the pixel.

The occurrence of mixed pixels can negatively affect the accuracy of depth data and the overall quality of the image.

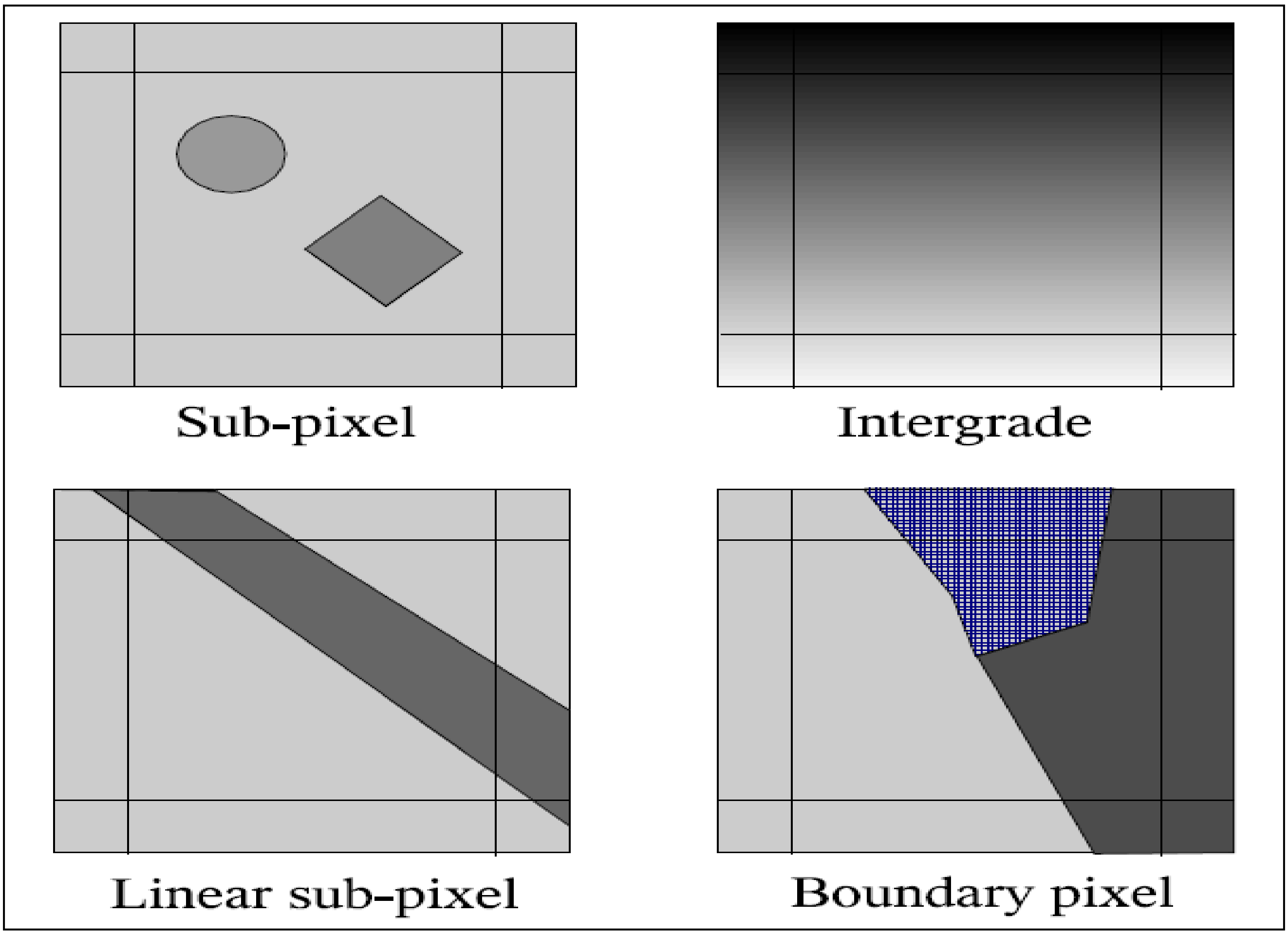

Figure 3.1.b) Occurrence of mixed pixels [8]

The usual approach of classifying all pixels is rather inaccurate because of the errors that are introduced by the classification of mixed pixels. Instead a resolution algorithm can be applied, after which a mixed pixel can be assigned to a particular category. The number of mixed pixels present in an image depends on fabric of the landscape. Since most mixed pixels lie on the boundary of two agricultural field and hence consist of only two classes. The resolving of mixed pixel is less susceptible to spectral confusion than classification so a better performance may be achieved.

3.2 State of the Art of mixed pixels

The traditional classification methods have been there for long, but their applicability to the processing of remote sensing data is limited due to the presence of mixed pixels. These mixed pixels affect the use of remotely sensed data in land cover classification and makes it difficult to detect features. Various techniques are being used to attempt to unmix the information and identify mixing proportions from the mixed pixels. There are two stages in the spectral unmixing procedure [31]. The first stage deals with the classification technique used and the second stage involves finding the fractions of components spectral response in the mixed pixel.

Several techniques have been proposed for the mixed pixel resolution such as linear unmixing, fuzzy approach as well as artificial neural network. Researches of different models and methods for unmixing the mixed pixel have been done in parallel but linear model is the widely used method. In linear model, the spectral value of the mixed pixel is considered as the linear combination of the pure pixel spectral values weighted by their corresponding proportions. Variations of linear model have been proposed which consider correlation between pixels. In some variants, geometry of the target is also considered to determine the effect of shadow. However, this model can only be used to unmix at most as many components as the number of bands and is unable to handle backscatter. Non linear mixing models are quite complex. Fuzzy and Artificial neural networks approaches use fuzzy set theory and neural networks respectively.

Overall the following methods have been applied for the mixed pixel problem.

- Linear mixture model
- Fuzzy model
- Artificial neural network

3.2.1 Linear mixture model

Linear mixture based approach introduced by Horwitz [35] has been a widely used approach for resolution of mixed pixels. This model is based on the assumption that the random variables associated with the elements are statistically independent. It assumes that mixed pixel contains elements belonging to different classes. The spectral signature of a class is taken as n dimensional Gaussian distribution with mean m^*_i and variance-covariance matrix N^*_i. The n is the number of spectral bands the scanner takes measurements at. This model describes the relationship between the spectral response associated with a pixel and the proportions and characteristics of its components. The basic linear mixture model is given as follows –

$X=Mf + e$ Equation (a)

The vector x denotes the pixel observation in the multispectral bands of size $n\times1$. The matrix M of size $n\times c$ denotes the spectra of c end members that constitute the mixed pixel. In this model, c represents the number of types of land cover. The end member spectrum is the response that is obtained from the homogeneous, unique ground object i.e. the response obtained from the pure pixel. The types of land cover are denoted by vector f of size $c\times1$. In a situation when a cell contains elements of class i only then they are represented by random variables with mean m_i as shown in figure 3.2.1 and variance-covariance matrix N_i . If the number of such elements present within the resolution cell is equal to a_i, then

$m_i = a_i\, m^*_i$

$N_i = a_i N^*_i$

provided that the variables are statistically independent.

The (nx1) error vector e, is used to model the statistical fluctuations around the mean value m(f)=Mf and is assumed to have a multivariate normal distribution with a zero mean and variance-covariance matrix N(f). The mixing equation defined by Equation (a) is accompanied by two constraints which should be satisfied explicitly when estimating f. The first constraint is the 'sum-to-one constraint'. According to this constraint, a pixel is well defined by its components, therefore its proportions should add up to unity.

$$\sum_{i=1}^{c} f_i = 1$$

The second constraint is the positivity constraint. According to this constraint, no component of a mixed pixel can make a negative contribution.

$f_i >= 0$ for i = 1,. . . c;

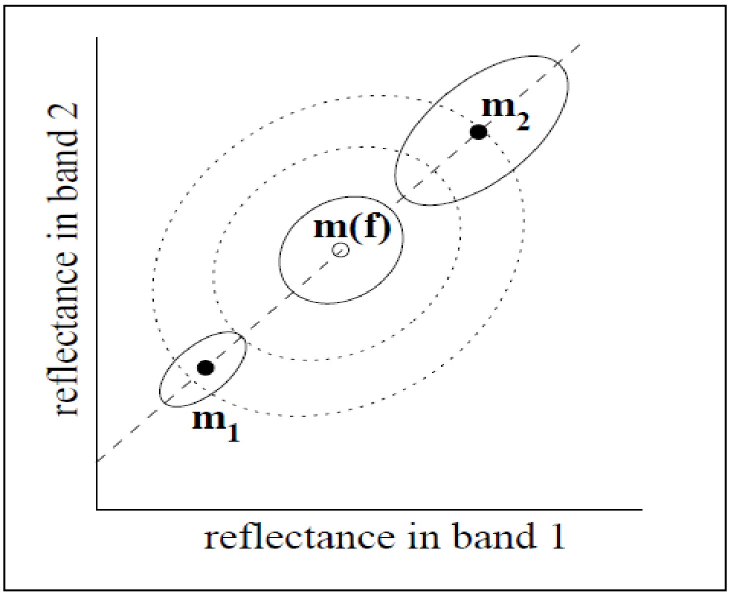

Figure 3.2.1 Spectral signature according to Linear mixture model [35]

A limitation of linear mixture model is that to ensure the uniqueness of a solution, the set of end members must be linearly independent. Another drawback of this model is that the number of end members in a scene are limited to one more than the number of spectral bands.

3.2.2 Fuzzy model

Fuzzy classification, or pixel unmixing, is the estimation of the proportion of the cover types from the composite spectrum of a mixed pixel [23]. This approach was proposed by Wang and is an extension of well known maximum likelihood classification technique. In fuzzy representation, each pixel is described by a group of membership grades which indicate the degree of similarity or strength of membership to classes considered. Just like class proportions, fuzzy membership function values must be positive and sum to one for each pixel. Basic assumption of fuzzy approach is that membership grades are informative about subpixel components. For this approach, a membership grade function is defined by the posterior probability p(k/x) of a pixel x belonging to a class k. The membership grades relate directly to the class proportions. According to the Bayesian theorem,

$$f_k = p(k/x) = \frac{p(k)p(x/k)}{\sum_{i=1}^{c} p(i)p(x/i)}$$

Where p(x/i) are class specific probabilities .

Similar to fuzzy classification, the fuzzy c-means approach is also based on fuzzy set theory. The algorithm partitions the data in the feature space into c fuzzy groups or classes. The Fuzzy C Means [27] approach outputs class membership values that may be interpreted as probabilities or degrees of sharing, which are attractive for the unmixing the class composition of pixels, and allows the fuzziness of the output to be adjusted by the analyst. The FCM has been applied successfully in a range of environments for the estimation and mapping of sub-pixel land cover composition. However, class membership is calculated with respect to all defined classes and hence the values of class membership strength are relative rather than absolute. A measure of the strength of class membership is used to indicate the class composition, with its magnitude taken as proportional coverage of the class within the area represented by the pixel. A major limitation of fuzzy approach is that it is slow and computationally expensive.

3.2.3 Artificial Neural Network

Applications of mostly artificial neural networks for remote sensing classification use a supervised, feed forward network employing a back propagation algorithm known as multilayer perceptron. It adjusts the network weights to produce convergence between the network outputs and a set of training data. These self learning structures permit the combination of multiple data types [29]. The network comprises of layers of neurons that are interconnected through weighted synapses.

All nodes in a given layer are connected to each node in the subsequent layer of the network and each internode connection has an associated weight [9]. This weight can be excitatory or inhibitory in nature. The first layer contains a separate node for each of the classification input variables and the last layer contains an output node for each of the possible output classes. Intermediate "hidden" layers provide an internal structure of neural pathways through which input data are processed to arrive at output values or conclusions. The figure 3.2.3 shown next depicts the multilayer perceptron and it input and output layer.

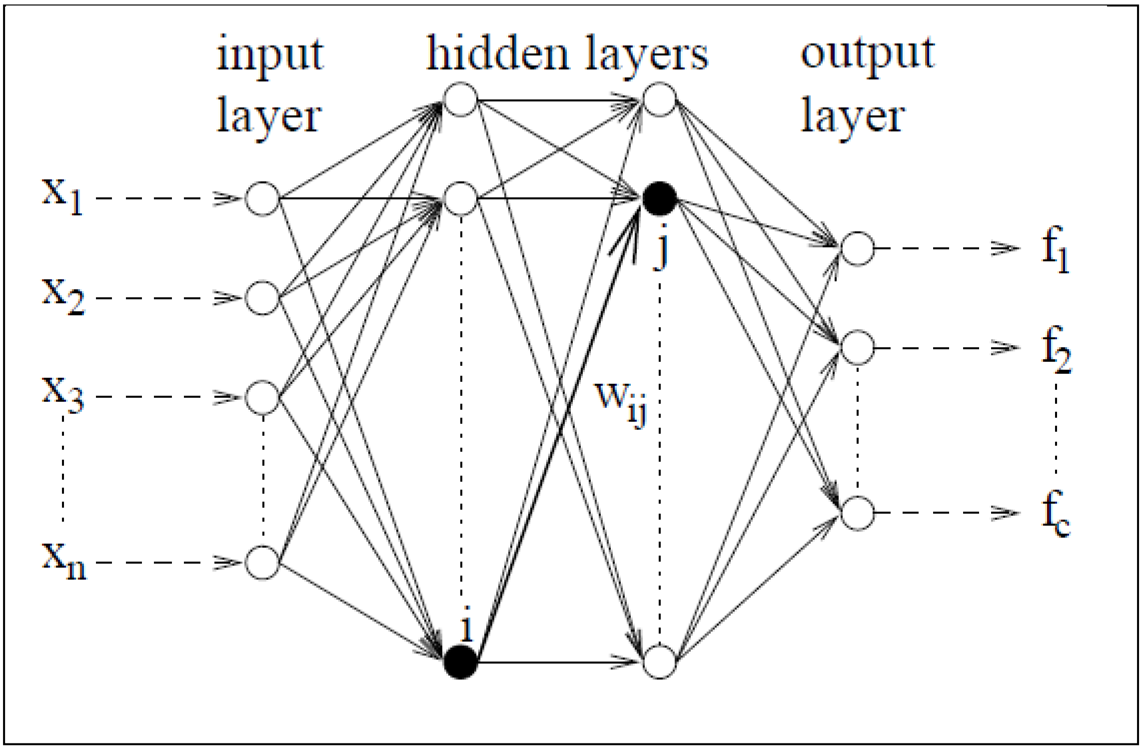

Figure 3.2.3 Mutilayer perceptron [29]

Typically, the output from the network is a hard classification, with only the code (i.e., nominal value) of the predicted class of membership indicated for each pixel. Outputs of an ANN trained with the pure pixels can be regarded as class membership grades whose strength can be used to derive class proportions. The network has thus n inputs corresponding to the spectral bands and m outputs corresponding to the pure classes. These are non-parametric techniques which have been shown to generally be capable of classifying data as or more accurately than conventional classifiers. Furthermore, although there may be problems associated with training an artificial neural network, particularly in relation to overtraining and training time, an artificial neural network, once trained, may classify data extremely rapidly as the classification process may be reduced to the solution of a large number of extremely simple calculations which may be performed in parallel.

Conventionally, a back propagation learning algorithm is used which iteratively minimizes an error function over the network outputs and a set of target outputs, taken from a training data set. A network can require large volumes of training data and take considerable computation time for the training cycle to converge. Network training and design also tend to be computationally complex and may require a large amount of trial and error to determine the appropriate number of hidden layers and the number of units within each layer. The accuracy of the network depends on the architecture and weight selection. Such disadvantages led to exploration of bio-inspired techniques for finding solution for this problem.

Chapter 4.

The bio-inspired technique of Biogeography based optimization

In this chapter we introduce the Biogeography based optimization algorithm in detail. We describe the features of BBO that we have adapted in our algorithm. We also explain the various migration models both linear and non linear. We present the previous work involving linear migration concept.

4.1 Introduction to Biogeography based optimization

Biogeography based optimization (BBO) is an evolutionary algorithm (EA) proposed recently by Dan Simon [1]. Biogeography studies deal with how the biological organisms are geographically distributed [3]. It analyses the various factors that affect the birth of a species, flourishing of species or the extinction of species in a habitat. The rainfall, water, flora-fauna, temperature, existence of predators etc are various such factors affecting the population of a habitat. For example, migratory birds travel to warmer areas to escape extreme cold conditions of their native place, is one such phenomenon occurring in nature where the species migrate to a more favourable habitat [19].

Unlike various other methods of optimization, EA's do not require information about the objective function. Hence they are efficient in solving complex problems. BBO is a population based EA for global optimization [25], inspired from the migration phenomenon of the ecosystem. Migration is an adaptive process that brings about changes in the existing habitats. The process of migration is carried out through immigration and emigration.

Immigration is the migration of a species into a habitat. Emigration is the migration of a species out of the habitat [20].

The migration of species of a habitat is decided upon the basis of several factors. The decision of selecting a habitat which is optimal for the survival of the species is carried out by migratory species in their lifetime. The species tend to move towards a better habitat as it provides better chances of survival. These aspects form the basis of this algorithm [7]. Every individual of the population has its own immigration and emigration rates. The immigration and emigration rates are function of number of species in the habitat [1] [17].

As with any other EA, a solution has some probability of mutation. A habitat's suitability can change suddenly due to random events such as outbreak of disease, natural disaster like flood, drought etc. It is modelled in BBO using species count probabilities to determine mutation rates. Mutation is not an essential feature of BBO and it works well for small number of species [1]. Hence, mutation is not considered in our proposed algorithm.

4.2 Features of BBO

The figure 4.2 shows the immigration and emigration behaviour of species in BBO.

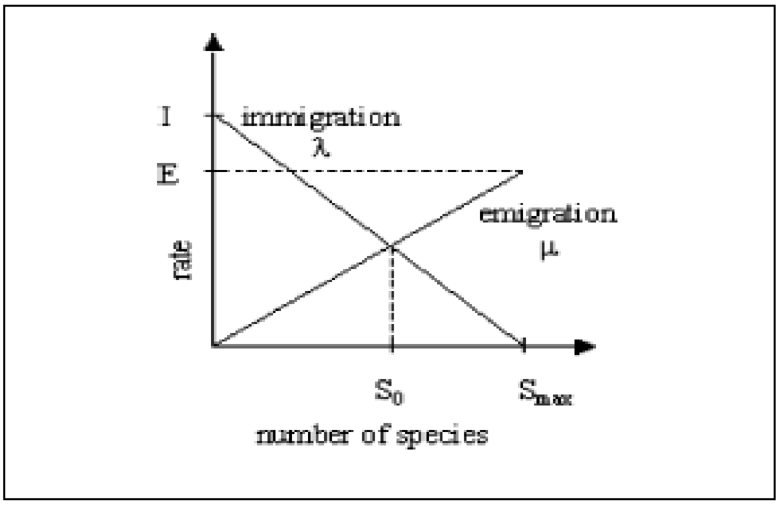

Figure 4.2 Immigration and emigration in BBO [1]

- **Habitat** : it is a region that is geographically isolated from others. Habitats comprises of diverse species depending upon how favourable the habitat turns out to be for that species. In the algorithm, these habitats denote the candidate solutions for the optimization problem.
- **Suitability index variables (SIV)** : the various factors such as rainfall, diversity of vegetation, diversity of topographic features, land area, and temperature determine how suitable the habitat is for the species. They denote the parameters affecting the solution of an optimization problem.
- **Habitat suitability index (HSI)** : it is an index based on SIV's that is used to find which habitat is more suitable for the species to flourish. It is a dependent variable and is calculated using the independent SIV variables. When applied to an optimization problem, it determines the goodness of the candidate solutions [2]. The candidate solution with high HSI is considered as good solution whereas the candidate solution with low HSI can be improved further. HSI is equivalent to the 'fitness' in other population based optimization algorithm [25].
- **Immigration rate** : it is denoted by λ. Immigration is the process of moving from one habitat into another habitat. The immigration curve is monotonically decreasing as can be seen in figure 4.2. The immigration rate is maximum when there are no species in the habitat and minimum when the number of species in the habitat is maximum. This is because the habitat can easily accept more species when it is empty as it can readily provide the necessary and required resources to the immigrating species. On the other hand, when the habitat is full, it can no longer provide the necessities and hence has low immigration rate.

- **Emigration rate** : it is denoted by μ. emigration is the process of moving out from a habitat to another habitat. The emigration curve is monotonically increasing as can be seen in figure 4.2. The emigration rate is minimum when there are no species in the habitat and maximum when the numbers of species in the habitat are maximum. As there are no species to migrate out, emigration is nearly zero. When the habitat is full, more number of species can move out to explore the other habitats. It does not mean that the entire species would move out of the habitat. A few representatives of the species may migrate to explore other avenues.

In BBO, habitats represent candidate problem solutions, and species migration represents the sharing of features between candidate solutions according to the fitness of the habitats [2]. Migration process is the means of sharing information among candidate solutions. The migration rates are a function of fitness of a habitat. This fitness is used to determine the quality of the solution.

4.3 Migration models of BBO

BBO is a global optimization algorithm based on the mathematical model of species distribution in biological systems. The information sharing in this evolutionary process is achieved by biogeography based migration operators. It can be thought of as application of biogeography to evolutionary algorithms. Biogeography is a process that brings about equilibrium in habitats.

As candidate problem solutions represent habitats, migration of the species from one feature habitat to another becomes the means of sharing features between candidate solutions according to the fitness of the habitats. The migration rates of a habitat determine the fitness of prospective solutions. For a particular habitat H_i, its immigration rate λ_i is used by the species to make the decision whether to immigrate or not. If immigration habitat is decided, then the emigrating habitat H_j is selected on the basis of emigration rate μ_j. Migration [2] is represented as follows -

$H_i\,(SIV) \leftarrow H_j\,(SIV)$

Different mathematical models of biogeography can be used to obtain different migration models. Thus, different constraints of a problem can be employed using different models. The different migration models can broadly be classified into two major categories-

- Linear
- Non linear

4.3.1 Linear migration

Linear migration phenomenon is based on linear immigration and emigration curves. The linear migration is not followed in nature but these migration curves are generally used to simplify the process. It becomes computationally easy to implement the algorithm.

a) Constant immigration and Linear emigration

The constant immigration and linear emigration curve is illustrated by figure 4.3.1 a).

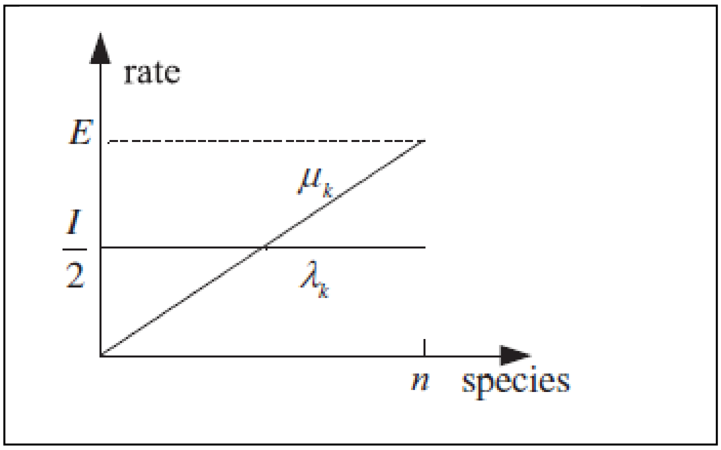

Figure 4.3.1 a) Constant immigration and linear emigration [2]

$\lambda_k = I/2$

$\mu_k = \frac{kE}{n}$

The emigration rate μ_k is linear with respect to the number of species k whereas the immigration rate λ_k is a constant. The emigration rate is zero when there are no species in the habitat. The emigration rate increases linearly with the increase in the number of species. The maximum emigration rate, E, is reached when the habitat contains the maximum number of species that it can support.

b) Linear immigration and Constant emigration

Linear immigration and constant emigration is illustrated by figure 4.3.1 b).

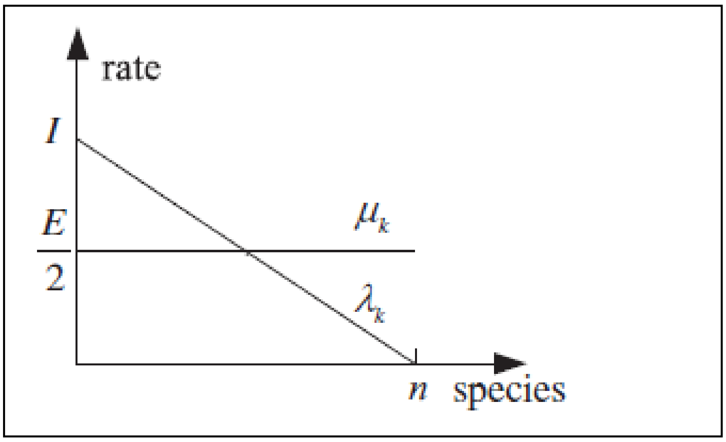

Figure 4.3.1 b) Linear immigration and constant emigration [2]

$\lambda_k = I\left(1 - \frac{k}{n}\right)$

$\mu_k = \frac{E}{2}$

The immigration rate λ_k is linear with respect to the number of species k whereas the emigration rate μ_k is a constant. The immigration rate decreases linearly with the increase in the number of species. The maximum immigration rate, I, is reached when the number of species in the habitat are zero. The immigration rate becomes zero when the habitat contains the maximum possible number of species,.

c) Linear immigration and Linear emigration

The linear immigration and linear emigration is illustrated by figure 4.3.1 c).

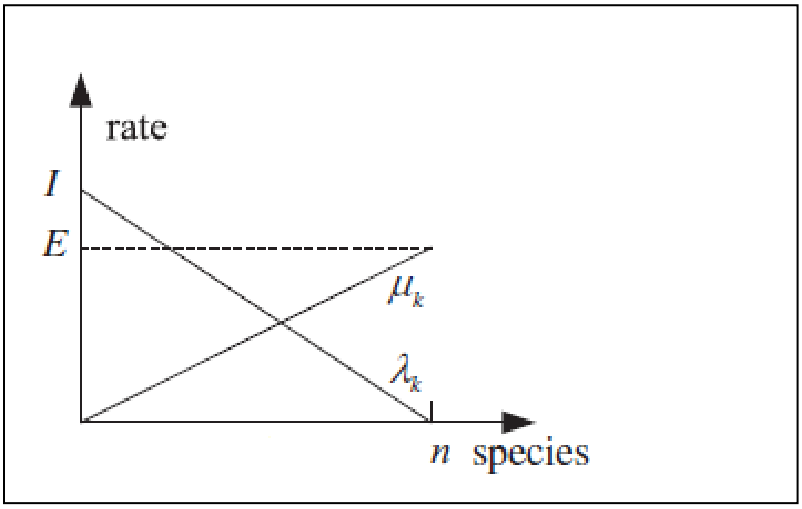

Figure 4.3.1 c) Linear immigration and linear emigration [2]

$\lambda_k = I\left(1 - \frac{k}{n}\right)$

$\mu_k = \frac{kE}{n}$

Dan Simon proposed this model in his paper. The immigration rate λ_k and the emigration rate μ_k are linear functions of the number of species k in the habitat. The immigration rate decreases linearly with the increase in the number of species. This happens because as the species enter the habitat, it gets more crowded and less number of species are able to successfully survive immigration to that habitat. On the other hand, the emigration rate increases linearly with the increase in the number of species because more number of species can leave the habitat to explore other possible options.

4.3.2 Non linear migrations

Our ecosystem is nonlinear in nature. Small changes in one area of the ecosystem leads to complex and related effects throughout the system. Therefore, the process of migration is more complicated than a linear curve. The concept of linear migration is too simple to explain the complicated phenomena such as migration in biogeography. Non linear curves closely model the natural processes occurring in a habitat such as predator/prey relationships, species mobility, evolution of particular species, and population size.

a) Trapezoidal migration

The immigration and emigration curves in trapezoidal migration are shown in figure 4.3.2 a).

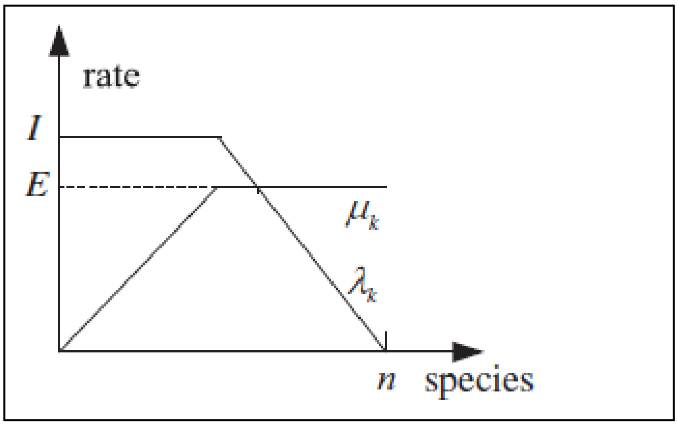

Figure 4.3.2 a) Trapezoidal migration curves [2]

$\lambda_k = I$ $k<=p$

$\lambda_k = 2I\left(1 - \frac{k}{n}\right)$ $p<k<=n$

$\mu_k = \frac{2Ek}{n}$ $k<=p$

$\mu_k = E$ $p<k<=n$

where $p = \text{ceil}\left(\frac{n+1}{2}\right)$

The immigration rate is constant and equals to the maximum immigration rate I in case when the habitat has a small number of species whereas the emigration rate increases linearly. The immigration rate decreases linearly when the species count exceeds its midpoint. On the other hand, the emigration rate is constant and equal to the maximum emigration rate E.

b) Quadratic migration

The immigration and emigration curves in quadratic migration are shown in figure 4.3.2 b).

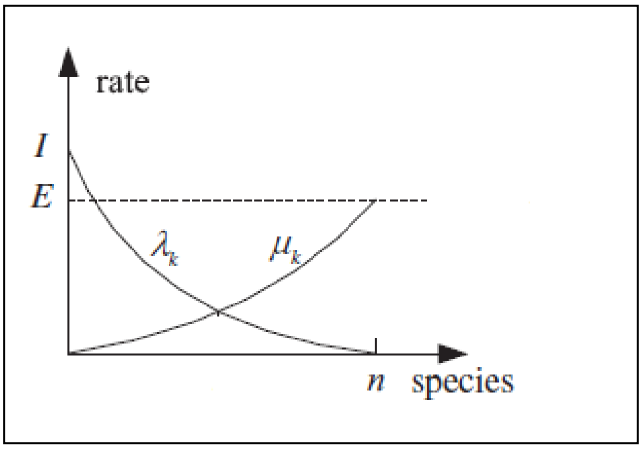

Figure 4.3.2 b) Quadratic migration curves [2]

$\lambda_k = I\left(\frac{k}{n} - 1\right)^2$

$\mu_k = E\left(\frac{k}{n}\right)^2$

The immigration rate, λ_k and the emigration rate, μ_k are quadratic functions of the number of species k. It is known [2] through an experimentally tested theory of island biogeography that migration in a single habitat follows a quadratic function of the size of the habitat and geographical proximity. The immigration rate decreases quickly from its maximum when the habitat has a small number of species whereas the emigration rate increases slowly from zero. The immigration rate gradually decreases from its maximum when the habitat is nearly saturated with species. On the other hand, the emigration rate rapidly increases from zero.

c) Sinusoidal migration

The immigration and emigration curves in sinusoidal migration are shown in figure 4.3.2 c).

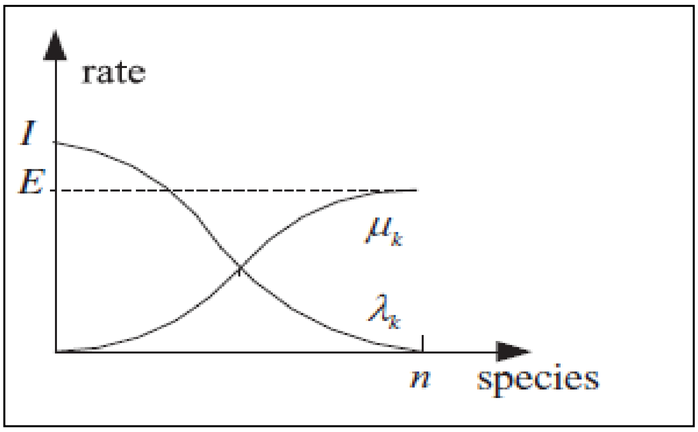

Figure 4.3.2 c) Sinusoidal migration curves [2]

$\lambda_k = \frac{I}{2}\left[\left(\cos\left(\frac{k\pi}{n}\right)\right) + 1\right]$

$\mu_k = \frac{E}{2}\left[\left(-\cos\left(\frac{k\pi}{n}\right)\right) + 1\right]$

The immigration rate, λ_k and the emigration rate, μ_k are sinusoidal functions of the number of species k. The immigration rate and the emigration rate, both change slowly from their extremes when the habitat has a small number of species or a large number of species. The migration rates change rapidly from their equilibrium values when the habitat has a medium species count. Therefore, it can be deduced that in nature, it requires a large amount of time for the species counts to reach equilibrium.

4.4 Linear BBO

Nature inspired techniques have been increasingly used for several optimization and other computationally expensive problems. BBO algorithm that models linear migration curve has been adapted to solve the problem of mixed pixels [4]. The various BBO parameters are assigned with respect to the mixed pixel problem. The set of features to which a mixed pixel may belong, forms the set of candidate habitat solutions.

The next aim of the algorithm is to find the habitat that is best for the species. The parameter HSI, known as Habitat Suitability Index, denotes the habitability of the habitat. It is used to decide how suitable the habitat is. Suitability Index Variables (SIV's) are the independent parameters of the algorithm. HSI is the function of these SIV's. Like other Evolutionary Algorithms, each candidate solution in BBO probabilistically shares information with other candidate solutions to improve candidate solution fitness. This sharing process is analogous to migration in biogeography.

Each candidate solution immigrate decision variable from other candidate solutions based on its immigration rate, and emigrate decision variable to other candidate solutions based on its emigration rate. Figure 4.3 depicts linear BBO method where S_1 represents a low HSI solution, while S_2 represents a high HSI solution. S_1 represents a habitat with only a few species, while S_2 represents a habitat with many species. The emigration and immigration rates of each solution are used to probabilistically share information between habitats.

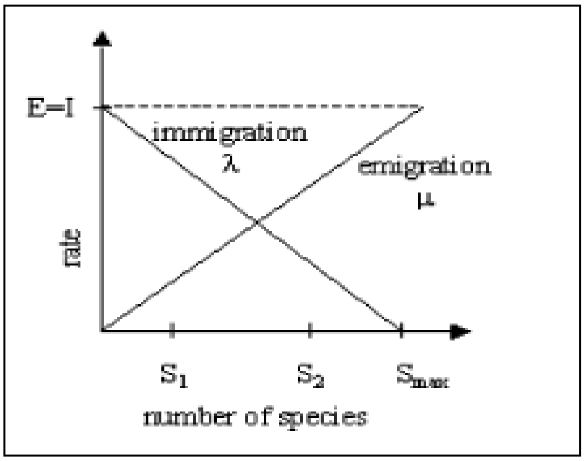

Figure 4.4 Linear BBO [1]

The linear BBO method is explained in the following steps -

1) The population of the candidate solutions is represented as vectors of integers. Each integer in the solution vector is considered to be an SIV. This is because the habitat suitability depends on SIV's.

2) High HSI solutions represent habitats with many species, and low HSI solutions represent habitats with few species. Each solution (habitat) is assumed to have an identical species curve for simplicity, but the value represented by the solution depends on its HSI.
3) The component pixel classes are the candidate solutions for the algorithm. For example – incase of a mixed pixel with components (p, q,), the candidate solutions are ‘p’ and ‘q’.
4) The HSI is calculated for pure spectral classes which is defined to be the mean of standard deviation of spectral values of the pixels in all the image bands.
5) For each pixel in the set of mixed pixel, the HSI is calculated in the same manner as specified for pure pixels.
6) Set of candidate solutions are prepared by moving the mixed pixel to each of the pure pixel set belonging to each of its component class.
7) The fitness of the candidate solution is determined by calculating the immigration rate using the HSI. As this method considers linear immigration and emigration, thus fitness is directly proportional to the HSI.
8) The candidate solution with least deviation from the fitness of pure component pixel is the class to which the mixed pixel belongs. The mixed pixel is immigrated to its class.

The above proposed models show the effective application of BBO for resolution of mixed pixels. We have implemented BBO algorithm in our approach and our work is focused on sinusoidal migration model. In the next chapter we discuss our algorithm in detail with BBO adapted to our mixed pixel problem requirements.

Chapter 5.

Mixed pixel resolution

In this chapter we will explain the model that we considered for our mixed pixel problem. We explain the two algorithms to solve the mixed pixel problem in an efficient manner. We also explain the terminology used in our algorithms and how we solve mixed pixel problem by adapting biogeography based optimization.

5.1 Model for mixed pixel resolution

In our model shown in figure 5.1 below, we have used a set of mixed pixels and pure pixels. This data is used in determining the fitness of the solution using Biogeography Based Optimization (BBO). We have considered the set of mixed pixels consisting of two different classes such as urban-barren, urban-vegetation, water-vegetation, rocky-vegetation and barren-vegetation.

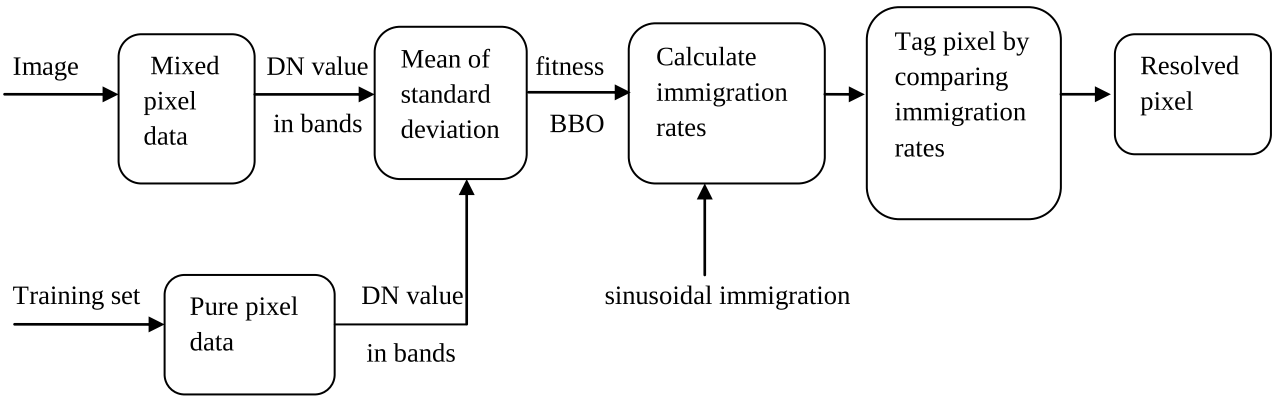

Figure 5.1 Our model for mixed pixel resolution

Our process of mixed pixel resolution is shown by the model before. The main steps of this model are as follows –

1. We fetch the mixed pixel data from the image and pure pixel data from the training set.
2. We determine fitness of the component mixed pixel class in comparison with pure pixel class. In this step we have used mean of standard deviation as fitness criteria.
3. We determine the immigration rates for every mixed pixel component using sinusoidal immigration of BBO.
4. Finally we get the resolved mixed pixels by comparing immigration rates of their component classes.

We have implemented the model in two approaches. In the first approach, the algorithm employs sinusoidal migration of BBO. In the second approach, the previous algorithm is improved by adding preprocessing steps of selection of typical feature bands. We present these two algorithms next.

5.2 Sinusoidal BBO – First approach

In our proposed approach, migration rates are obtained using sinusoidal migration concepts. The sinusoidal migration is nonlinear migration and it better represents the natural phenomenon such as predator/prey relationships, species mobility, evolution of particular species, and population size [2]. The habitat to which immigration will occur, is decided on the basis of immigration rate [5]. For our problem, the habitat from which emigration occurs, is restricted to the set of mixed pixels. Hence, only immigration rate is considered in the proposed algorithm.

5.2.1 Framework of algorithm

The set of mixed pixels with not more than two components and set of pure pixels of every feature class are extracted. Initially set of pure pixels belonging to the components of mixed pixel are retrieved from training set.

5.2.1.1 Parameters in algorithm

We have modified certain parameters to adapt the BBO in our approach as per our problem requirements.

- The set of pure pixels denoting the land features on earth such as vegetation, rocky, water, urban and barren are analogous to the feature habitat. A habitat H Є SIV^m where m is the number of bands.

- It is the SIV's that determine the fitness of the candidate solutions. SIV's are the independent parameters of the problem. Each of the multi-spectral bands of image represents one Suitability Index Variable (SIV) of the habitat. Thus, SIVЄ C is an integer and C Є [0,255].
- The bands contain brightness value for eacn pixel in that particular band. All bands are included in calculation of HSI for a feature habitat.
- HSI is calculated as the mean of standard deviations using DN values of feature habitat including least correlated bands.
- The mixed pixels are analogous to the species migrating into and out of a habitat.
- The feature classes that comprise the mixed pixel form the set of candidate solution habitats.

- As species immigrate to a habitat which is more suitable, similarly mixed pixel also migrates to a suitable feature class. In other words, the mixed pixel is resolved to belong to that particular feature class.

The framework of our sinusoidal BBO algorithm is as shown in figure 5.2.1. The figure clearly demonstrates the processes carried out during the implementation of our algorithm.

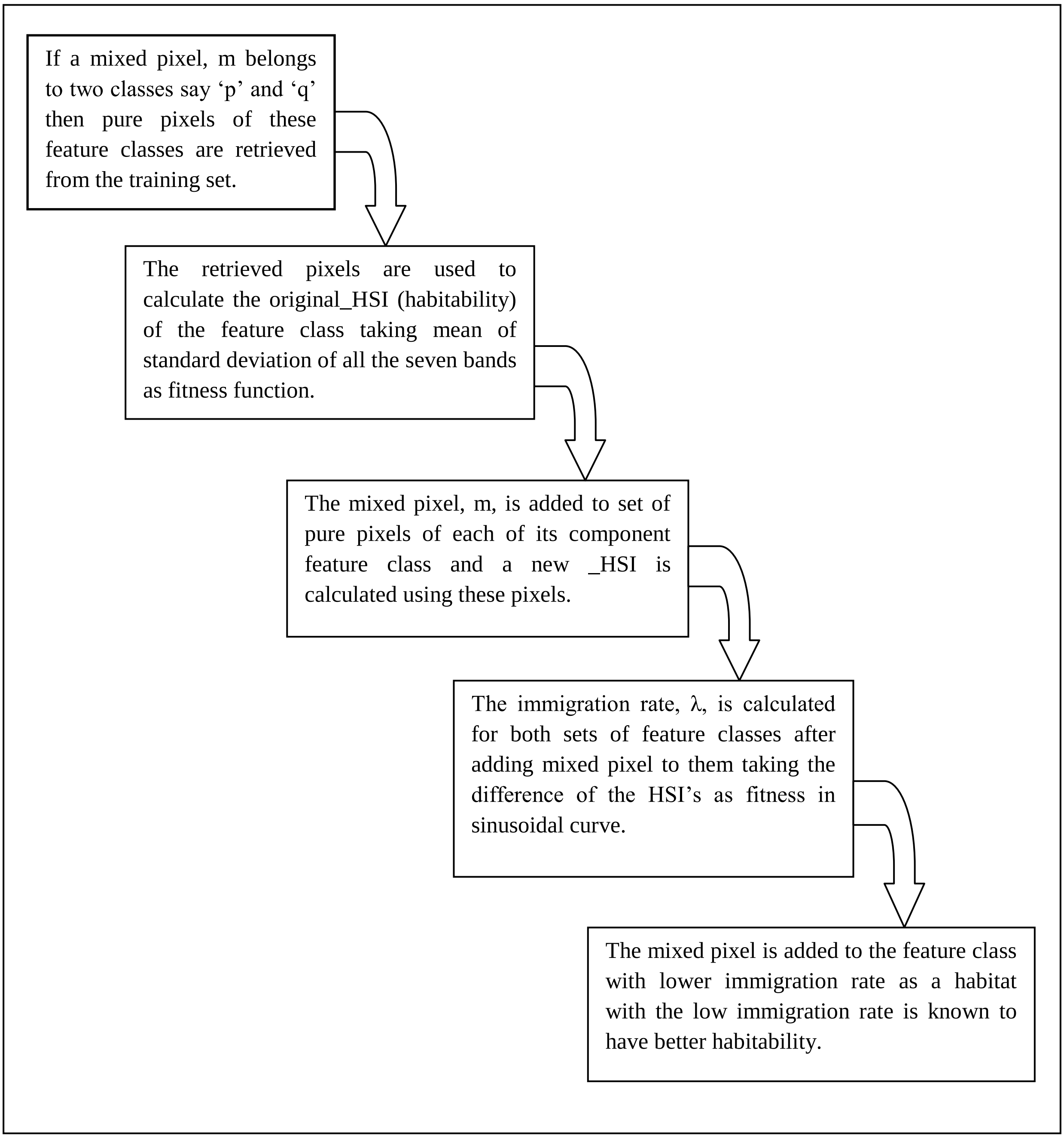

Figure 5.2.1 The framework of sinusoidal BBO algorithm

5.2.1.2 Algorithm I (Sinusoidal BBO)

The mixed pixel dataset and pure pixel dataset for all image bands are obtained and fed to the algorithm implementing sinusoidal BBO. The algorithm is as follows -

a) Calculate HSI for the pure pixels using fitness function using equations given below and term it as Original_HSI.

$$\text{HSI} = (\textstyle\sum_{k=1\ to\ b} Sd(k))/b \tag{1}$$

$$\text{Sd}=\sqrt{((\textstyle\sum_{i=1}^{m} dev(i))/m)} \tag{2}$$

$$\text{dev(i)} = (x_i - (\textstyle\sum_{j=1}^{m} x)\ /m)^2 \tag{3}$$

b= number of bands included in implementation

m = number of pixels in feature class

x= DN values in specific bands

b) For each pixel in the set of mixed pixels, the pure pixels comprising a mixed pixel are obtained.

   i. Determine HSI for both cases and term them as New_HSI for respective component class.

   ii. Calculate HSI for both the temporary sets using equation (1), (2), (3). Term it as New_HSI.

   iii. Find difference of New_HSI and Original_HSI and normalize it to calculate immigration rate using equation given below.

$$\lambda = I/2[(\cos(\text{fitness} \times \pi) + 1] \tag{4}$$

I is the maximum immigration rate

c) Assign the mixed pixel to class with lower immigration rate.

5.3 Improved Sinusoidal BBO – Second approach

The intrinsic dimensionality of the image is the number of components that hold the majority of the information in a multispectral image. Each data image may have different intrinsic dimensionality. When dealing with multispectral images, the amount of information to be treated can be very large. Moreover, the spectral information is highly correlated along a given spectral range. Therefore, instead of having an exhaustive representation of the whole spectrum, selecting some key bands can considerably reduce the amount of data without practically losing relevant information [15] [28].

5.3.1 Framework of algorithm

When there are several bands associated with an image, there may be redundancy in acquired data. Few bands may contain information that is already present in other bands. The main objective of band selection in multispectral imaging is getting rid of redundant information and reducing the amount of data to be processed.

5.3.1.1 Parameters in algorithm

The parameter K is defined that determines the number of bands to select to effectively represent the data of a defined feature area. The other parameters for calculating HSI and immigration rates are same as those used in first approach. The following figure 5.3.1 outlines the basic preprocessing steps involved in improved algorithm. As seen in figure5.3.1, the main idea is to find the subset of K bands that are as much independent as possible among them. Taking the selected bands, the sinusoidal migration process is used to find suitable feature habitats for the mixed pixel.

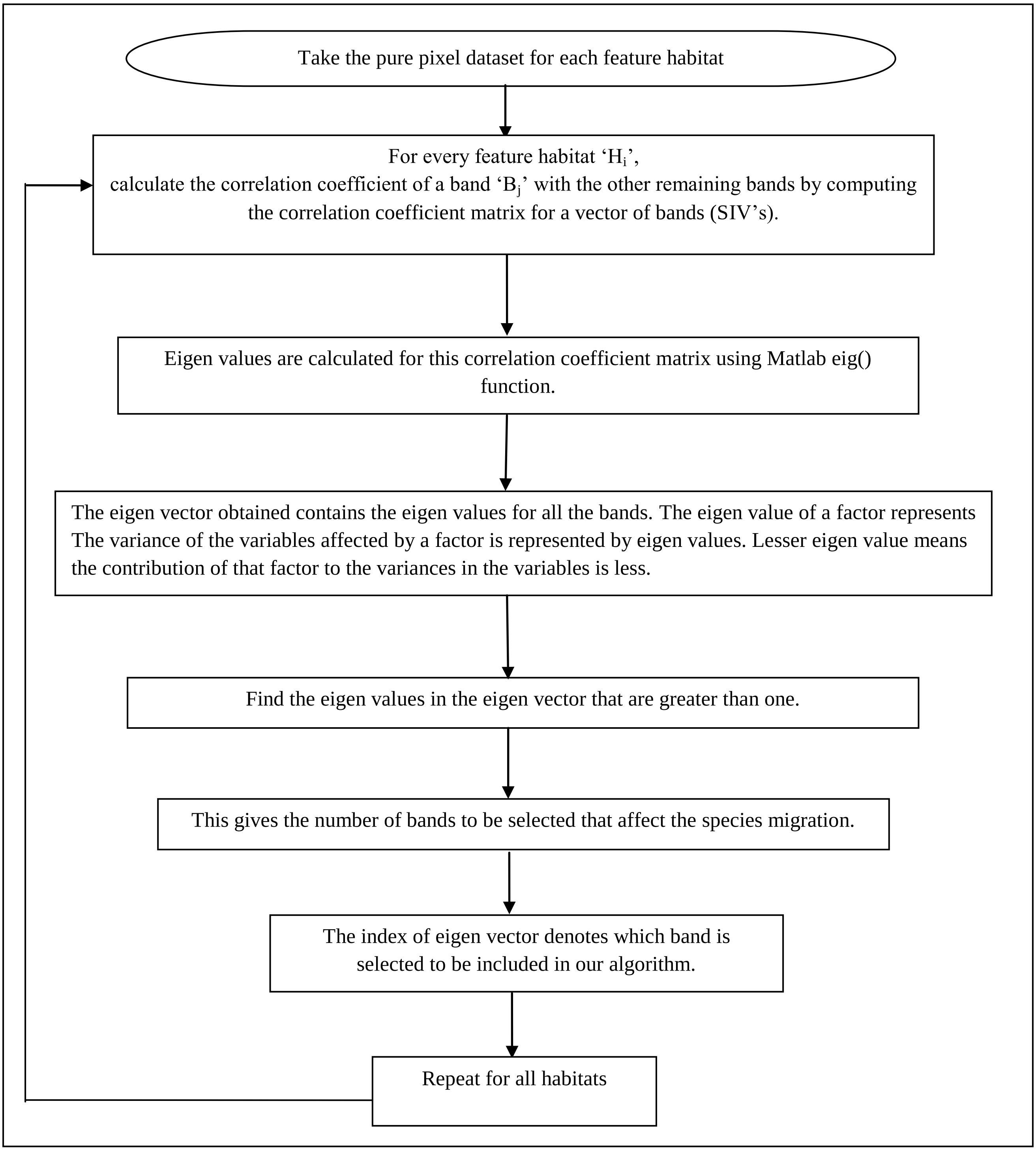

Figure 5.3.1 Flowchart for preprocessing steps

The issue is how to select the right spectral bands to characterize the problem. The eigen values for the correlation matrix are computed. The lower the eigen value, the less that factor contributes to the explanation of variances in the variables. Out of these, the number of eigen values that are greater than one are determined. This is the number of bands (SIV's) to be included in the algorithm.

In order to seek non redundant bands, correlation coefficient matrix is computed between vectors of seven bands for each feature class. Correlation quantifies the extent to which two quantitative variables, X and Y, go together. When high values of X are associated with high values of Y, a positive correlation exists. When high values of X are associated with low values of Y, a negative correlation exists. The bands with least value of correlation coefficient are chosen for further implementation in the proposed algorithm.

5.3.1.2 Algorithm II (Improved Sinusoidal BBO)

The mixed pixel dataset and pure pixel dataset for all image bands are obtained and fed to the algorithm implementing sinusoidal BBO. The algorithm is as follows -

a) For pure pixel, find correlation coefficient matrix and its eigen vector. The correlation between bands is determined using Pearson correlation coefficient. The formula for ρ is:

$$\rho_{X,Y} = \frac{\mathrm{cov}\,(X,Y)}{\rho_X \rho_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\rho_X \rho_Y} \qquad (5)$$

b) Find K i.e. the number of eigen values greater than one. Select K least correlated bands.

c) Repeat the following steps, considering K bands that are selected

    i. Determine HSI for all types of pure pixels and term it as Original_HSI using equations (1), (2), (3).

    ii. For every mixed pixel, add it to both its constituent feature class.

        1. Add mixed pixel both the sets of pure pixels forming temporary set.

2. Calculate HSI for both the temporary sets using equation (1), (2), (3). Term it as New_HSI.
3. For both temporary sets calculate the difference between this New_HSI and Original_HSI.
4. Normalize deviation and use it to determine the immigration rate, λ for each temporary set using equation (4).
5. The set for which λ is less, is a better habitat and mixed pixel is tagged to that habitat.
6. The temporary sets that were formed for the process are freed.

d) When the above mentioned procedure is completed for all the mixed pixels, the resultant set which is obtained contains all resolved pixels.

So far, we discussed the model we proposed to solve the mixed pixel problem. We also discussed the two approaches that we have used for the mixed pixel resolution problem. We showed how we adapted BBO in these approaches differently. Certain parameters were modified while adapting BBO according to our problem requirement. In the next chapter we discuss the experimental setup that was used and the results that we obtained from these two approaches.

Chapter 6.

Experiments and results

In this chapter we present the experimental setup used to find results for our mixed pixel problem. We explain the mixed pixel and pure pixel dataset for different land cover classes. This chapter also contains the results that we have obtained after applying our algorithm. The comparisons of our algorithm are depicted in the form of bar graph and line graph. The resultant image with resolved mixed pixels is also shown.

6.1Experimental setup

Our algorithm is implemented in MATLAB 7.9.0 (R2009b). The proposed algorithm has been applied to the seven-band satellite image of size 472 X 576 of the Alwar area in Rajasthan. The Alwar image is imported through the image processing toolbox and the details of the MATLAB functions used, is given in Appendix B.

6.1.1 The seven bands multispectral Alwar image

Alwar image is multi-spectral satellite image which is taken in seven different bands. These bands are Red, Green, Near-Infra Red, Middle-Infra Red, Radarsat-1, Radarsat-2, and Digital Elevation Model. The four spectral bands are in the visible bands namely: red, green, near-infrared (NIR) and middle infra-red (MIR) from LISS-III sensor of Indian Remote Sensing sat satellite Resourcesat-1.Also,two SAR images namely: low incidence S1 beam -200-270 (RS1) and other is High incidence S7 beam 450-490,(RS2) of the same area taken from Canadian Radarsat-1 satellite.

The seventh band is digital elevation model (DEM) of the area. The ground is resolution of the image from LISS-III and Radarsat-1 image is 23.5m and 10m respectively. The DEM dataset is also generated from SAR interferometry using RS1 and RS2 and have 25-meter resolution [4]. Alwar image as a grey scale image using a single band at a time is shown in figure 6.1.1 a).

Alwar region image is selected because it contains large variety of land features like water, urban, vegetation, rocky and barren. Figure 6.1.1 b) shows the Alwar image in combination of three bands at a time as a colour composite image [4].

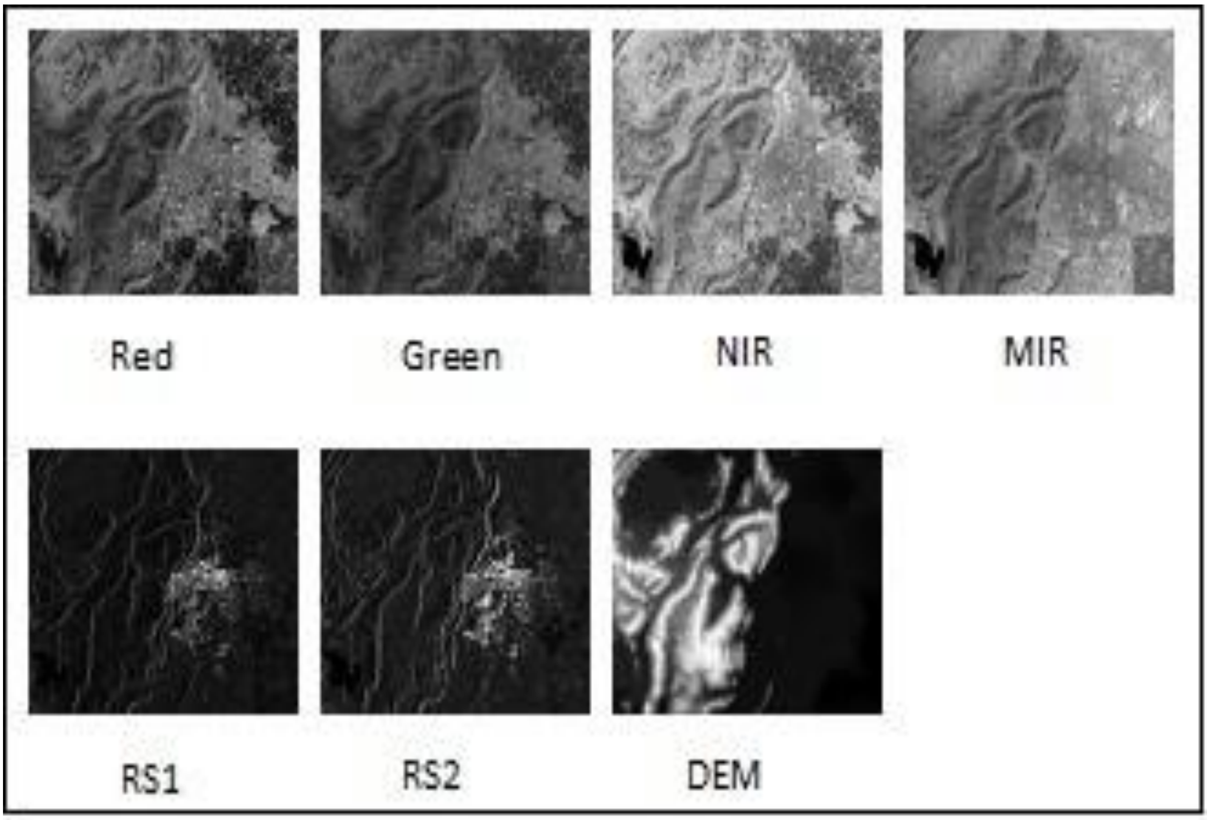

Figure 6.1.1 a) Seven bands gray images of Alwar, Rajasthan [7]

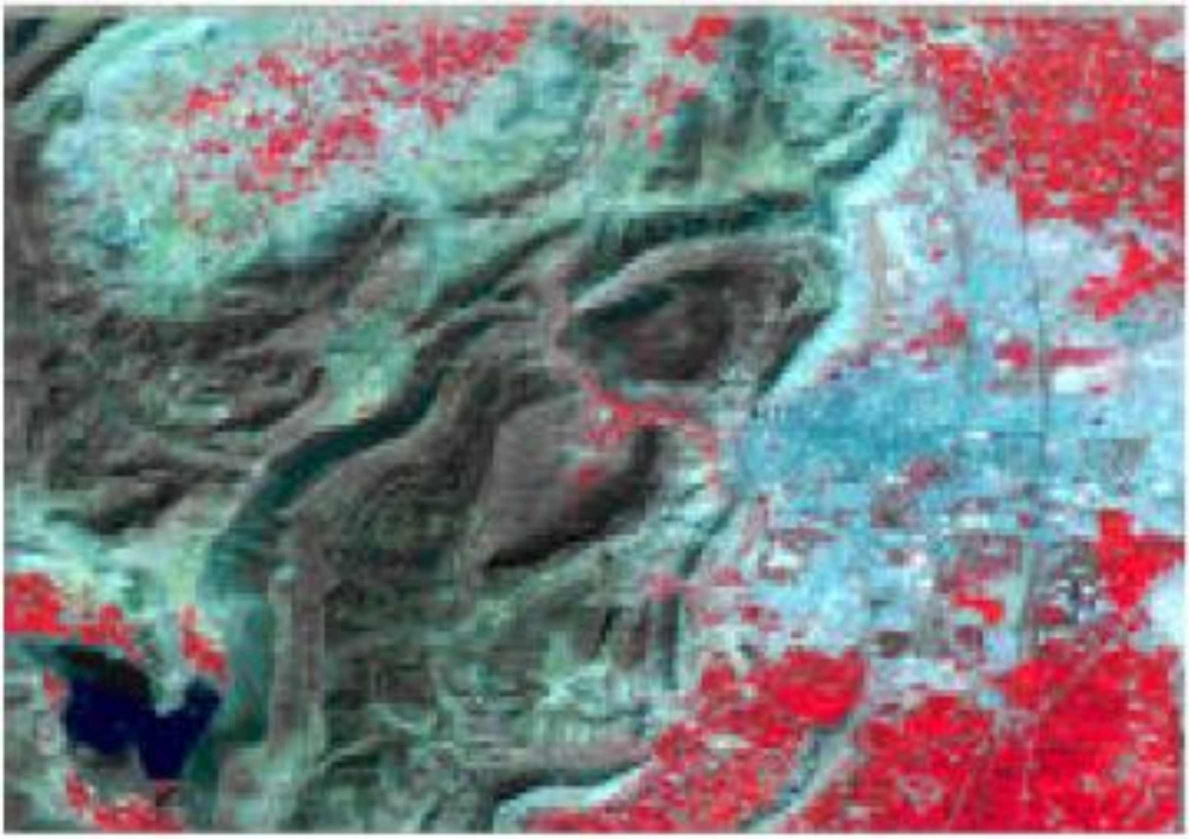

Figure 6.1.1 b) False colour satellite image of Alwar [4]

6.1.2 Training sets of image features

The training set consists of pure pixels of each feature class. It has been provided by the experts which means that that we are sure that each sample of training data is unique and belongs to only one class. The experts generate the training set from these multi-band images. These images are opened by experts of remote sensing in ERDAS IMAGINE as given in Appendix D. They manually select the pixels of different class regions and a definition colour code is set using ERDAS Signature Editor. This data set provided by the experts is in the form of digital numbers (intensity value pixel in a digital image). These sets are taken by carefully selecting the areas (pixel by pixel) from all the images and noting the DN values of the pixels in different wavelength bands.

The dataset of the problem can be in any format such as Excel Sheet, Access sheet or in Text file. We have used the dataset in Excel sheet format. The training data was provided for each of the five major land use types- water, vegetation, urban, barren and rocky. The data is in tabular format similar to the data in relational database, comprising of attribute columns and entity rows. Each row value corresponds to information of a particular pixel. The attribute values in each band column denote the DN value of pixel. The Decision according to the DN values is also set in the table as seen in the last column attribute. Few portions of our training set for barren, urban, vegetation, water and rocky is as shown next in the following table1, table 2, table 3, table 4 and table 5 respectively.

Table 1 - Training set for barren class

	A	B	C	D	E	F	G	H
1	RED	GREEN	NIR	MIR	RS1	RS2	DEM	DECISION
2	115	91	182	126	20	15	30	BARREN
3	111	90	173	131	17	34	15	BARREN
4	121	91	182	118	26	17	40	BARREN
5	125	98	188	128	25	21	27	BARREN
6	130	94	186	128	24	33	25	BARREN
7	138	106	212	137	32	22	115	BARREN
8	170	127	234	159	37	45	100	BARREN
9	156	123	204	142	27	57	78	BARREN
10	113	86	190	113	24	28	53	BARREN
11	113	88	199	113	16	38	36	BARREN
12	117	88	201	118	29	24	30	BARREN
13	115	94	175	131	35	36	15	BARREN
14	166	130	223	151	32	45	112	BARREN
15	168	130	226	149	26	29	114	BARREN
16	109	82	179	113	12	21	51	BARREN

Table 2 - Training set for urban class

	A	B	C	D	E	F	G	H
1	RED	GREEN	NIR	MIR	RS1	RS2	DEM	DECISION
2	128	106	184	142	22	35	15	URBAN
3	128	120	155	118	18	36	15	URBAN
4	117	109	157	122	19	21	15	URBAN
5	123	109	166	137	10	13	15	URBAN
6	109	103	160	131	20	26	15	URBAN
7	109	99	162	133	27	33	15	URBAN
8	117	101	166	115	22	13	15	URBAN
9	115	101	164	137	17	35	15	URBAN
10	101	88	149	115	15	20	15	URBAN
11	121	103	181	144	16	28	15	URBAN
12	146	130	186	146	31	34	15	URBAN
13	142	127	166	124	24	24	15	URBAN
14	123	112	153	124	24	24	15	URBAN
15	125	111	168	126	15	23	15	URBAN
16	115	107	159	126	33	38	15	URBAN

Table 3 - Training set for vegetation class

	A	B	C	D	E	F	G	H
1	RED	GREEN	NIR	MIR	RS1	RS2	DEM	DECISION
2	21	35	65	192	24	33	9	VEGETATION
3	23	37	73	192	18	33	14	VEGETATION
4	27	37	82	170	30	30	13	VEGETATION
5	21	35	69	252	20	22	11	VEGETATION
6	19	37	85	234	29	37	10	VEGETATION
7	17	32	69	246	25	37	10	VEGETATION
8	29	40	76	186	23	21	14	VEGETATION
9	15	25	60	237	13	26	12	VEGETATION
10	15	28	64	219	20	25	12	VEGETATION
11	13	30	64	228	28	24	12	VEGETATION
12	11	28	60	245	14	23	13	VEGETATION
13	15	33	64	245	13	30	13	VEGETATION
14	15	33	82	255	15	20	11	VEGETATION
15	29	40	74	186	17	32	15	VEGETATION
16	21	38	78	188	22	25	15	VEGETATION

Table 4 - Training set for water class

	A	B	C	D	E	F	G	H
1	RED	GREEN	NIR	MIR	RS1	RS2	DEM	DECISION
2	25	30	25	10	180	92	15	WATER
3	25	30	25	10	157	60	15	WATER
4	25	30	21	20	8	2	30	WATER
5	21	27	20	20	2	5	30	WATER
6	23	30	16	14	3	4	30	WATER
7	21	24	14	10	1	2	30	WATER
8	23	25	14	12	1	1	30	WATER
9	23	27	12	10	2	1	30	WATER
10	21	28	12	10	2	1	30	WATER
11	21	25	10	10	1	0	30	WATER
12	23	27	10	12	3	4	30	WATER
13	23	28	10	10	4	4	30	WATER
14	21	24	10	5	0	2	30	WATER
15	23	25	10	7	3	2	30	WATER
16	23	27	9	7	2	4	30	WATER

Table 5 - Training set for rocky class

	A	B	C	D	E	F	G	H
1	RED	GREEN	NIR	MIR	RS1	RS2	DEM	DECISION
2	62	49	135	91	44	40	94	ROCKY
3	84	64	160	102	20	25	165	ROCKY
4	52	45	129	85	15	29	107	ROCKY
5	91	69	171	106	10	46	123	ROCKY
6	87	67	168	104	8	21	157	ROCKY
7	76	59	157	95	9	47	114	ROCKY
8	70	51	159	95	11	46	127	ROCKY
9	82	59	159	100	7	9	173	ROCKY
10	84	67	171	95	5	57	118	ROCKY
11	74	61	140	84	51	20	216	ROCKY
12	85	67	168	102	14	88	137	ROCKY
13	64	56	115	78	14	22	122	ROCKY
14	93	69	138	85	13	36	180	ROCKY
15	56	46	104	69	21	26	117	ROCKY
16	76	53	153	98	12	22	206	ROCKY

6.1.3 Mixed pixel dataset

For our algorithm we use standard mixed pixel dataset acquired at DRDO, New Delhi. The dataset of mixed pixels includes pixels belonging two feature classes. This dataset consists of 250 mixed pixels that correspond to five groups. These groups are urban-barren, barren-vegetation, urban-vegetation, rocky-vegetation and water-vegetation.

The portions of mixed pixel dataset used in our algorithm is shown in the following table 6,table 7 and table 8 denoting data of urban-barren class, urban-vegetation class and barren-vegetation class respectively.

Table 6 – Mixed pixels of urban and barren class

	A	B	C	D	E	F	G	H
1	RED	Green	NIR	MIR	RS1	RS2	DEM	Decision
2	144	122	192	140	29	65	14	U-B
3	148	119	199	142	34	27	14	U-B
4	148	119	199	142	30	19	14	U-B
5	125	106	168	129	34	24	14	U-B
6	113	104	137	96	18	86	14	U-B
7	132	117	177	129	18	117	14	U-B
8	128	112	181	131	21	36	14	U-B
9	136	112	190	133	27	49	14	U-B
10	134	119	168	122	30	74	14	U-B
11	150	125	206	140	45	70	14	U-B
12	146	123	201	146	28	27	14	U-B
13	140	112	193	138	20	60	14	U-B
14	132	115	184	137	35	41	14	U-B
15	121	106	164	120	21	49	14	U-B
16	142	117	192	140	29	51	14	U-B

Table 7 - Mixed pixels of urban and vegetation class

	A	B	C	D	E	F	G	H
1	RED	Green	NIR	MIR	RS1	RS2	DEM	Decision
2	78	83	117	146	84	47	12	U-V
3	93	85	122	111	45	38	12	U-V
4	87	86	120	109	53	34	12	U-V
5	78	83	117	146	44	37	12	U-V
6	93	85	122	111	40	62	12	U-V
7	87	86	120	109	75	54	11	U-V
8	74	70	115	128	111	101	10	U-V
9	50	61	113	151	165	87	10	U-V
10	82	78	109	106	90	167	9	U-V
11	68	64	120	133	132	97	10	U-V
12	58	61	109	126	41	57	12	U-V
13	93	86	124	109	42	80	12	U-V
14	93	86	124	109	19	63	12	U-V
15	109	101	138	122	59	56	9	U-V
16	107	96	137	124	56	45	12	U-V

Table 8 - Mixed pixels of barren and vegetation class

	A	B	C	D	E	F	G	H
1	RED	Green	NIR	MIR	RS1	RS2	DEM	Decision
2	111	104	179	177	13	11	15	B-V
3	123	106	179	173	10	20	15	B-V
4	109	101	162	182	12	17	15	B-V
5	68	75	140	224	11	12	15	B-V
6	87	86	168	170	13	17	15	B-V
7	70	75	153	182	14	20	15	B-V
8	76	85	160	181	16	18	15	B-V
9	119	109	170	160	19	15	15	B-V
10	97	88	151	168	14	16	15	B-V
11	84	78	146	142	14	23	15	B-V
12	74	70	149	153	10	20	15	B-V
13	54	57	117	160	13	20	15	B-V
14	89	85	162	162	16	14	15	B-V
15	76	78	157	170	26	17	15	B-V
16	72	74	138	162	17	17	14	B-V

6.2 Results of mixed pixel resolution

To show the efficiency of the proposed biogeography based algorithm we have plotted the number of mixed pixel correctly resolved against the total number of mixed pixels. We have plotted bar graph to easily compare our result. Initially, result obtained from the first algorithm employing sinusoidal migration is compared with the linear approach. Next the results obtained from the second algorithm, which employs band selection technique are compared with the linear migration approach.

6.2.1 Including all seven bands – First approach

Using our first approach based on sinusoidal BBO, comparison of the resolved pixels with the linear approach [13] is shown in figure 6.2.1. The graph shows better performance of sinusoidal BBO over linear BBO.

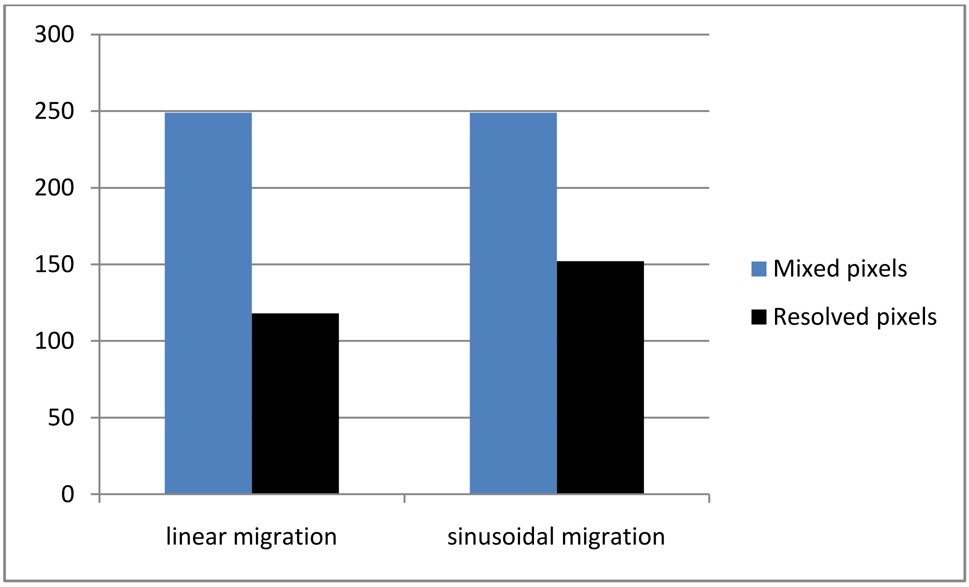

Fig.6.2.1 Comparison of linear and sinusoidal approach

6.2.2 Including non redundant bands – Second approach

Using our improved sinusoidal algorithm which was designed selecting data of non redundant bands produced results as shown in figure 6.2.2 a). The non redundant bands selected through our algorithm for particular group of mixed pixels are as follows: for water-vegetation DEM, NIR, MIR bands were used; for rocky-vegetation DEM, NIR, MIR bands were used; for barren-vegetation DEM, RS2, MIR bands were used; for urban-vegetation RS1, DEM, MIR bands were used; for urban-barren RS1, RS2, MIR bands were used. The graphical comparison depicts better performance of sinusoidal migration with non redundant bands. Figure 6.2.2 b) shows the portion of Alwar image obtained after mixed pixels have been resolved.

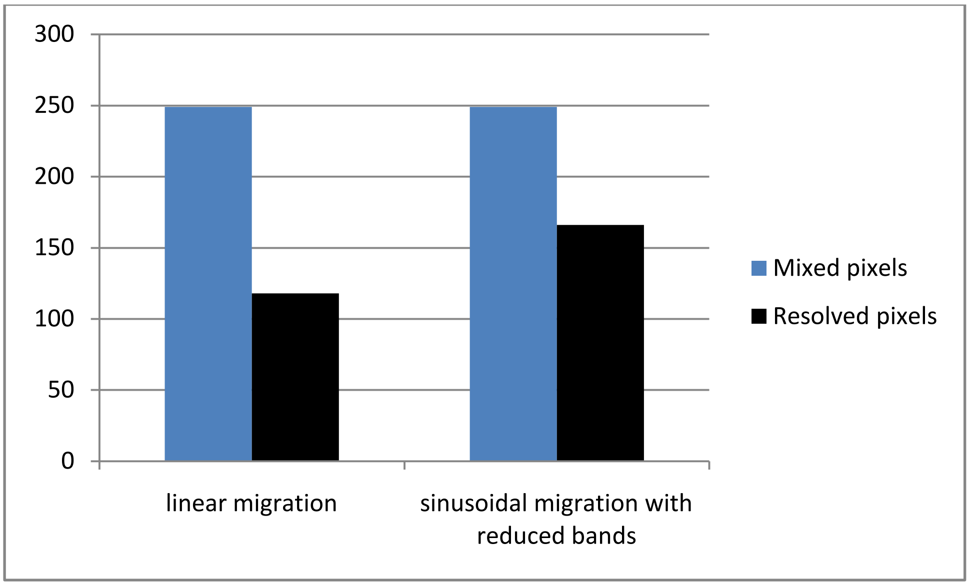

Fig.6.2.2 a) Comparison of linear and reduced band sinusoidal approach

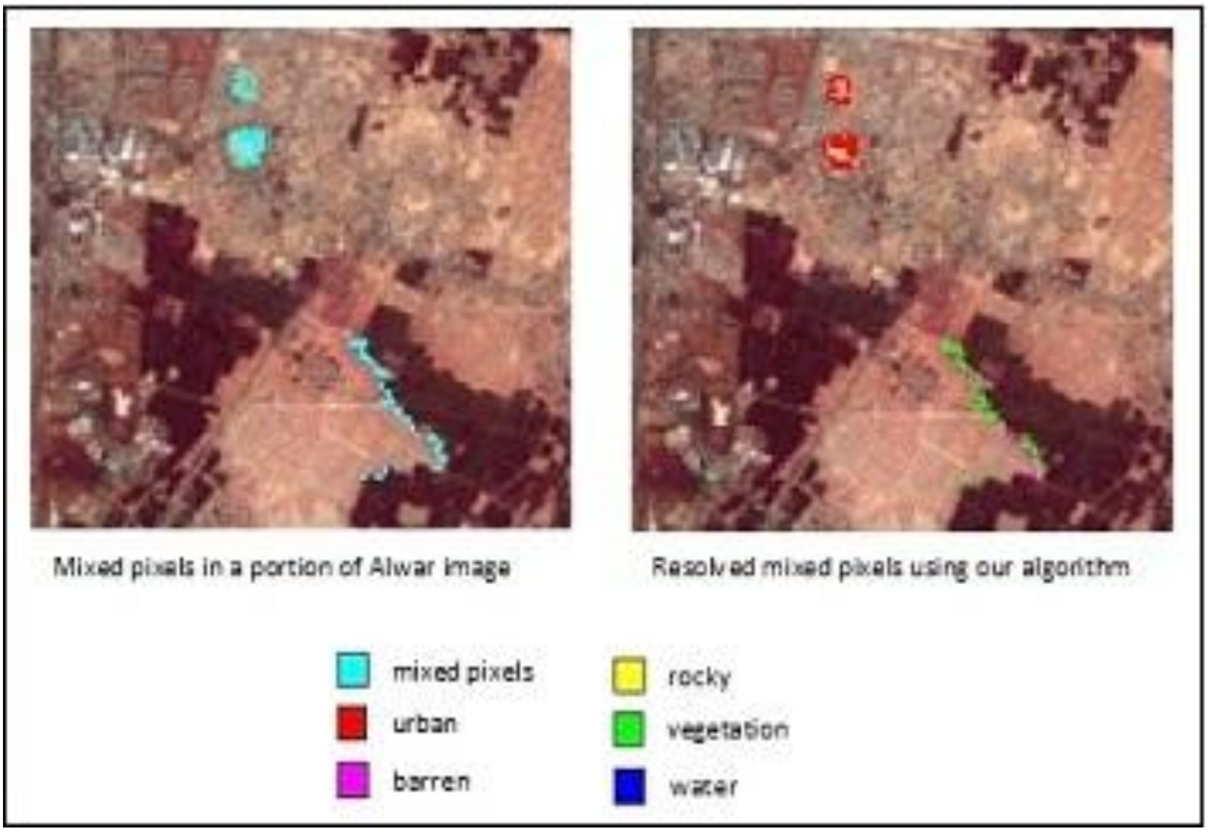

Fig.6.2.2 b) Portion of Alwar image with mixed pixels and resolved mixed pixels

6.3 Evaluation and analysis

A non linear migration curve has improved the resolution of mixed pixels in comparison with the linear migration model. Hence, more optimal classification results are obtained and accurate analysis of the information extracted from the image can be done.

The output produced from the algorithm is justified by comparing with PSO/ACO classified image. This image is taken as a standard image for comparison of linear and sinusoidal migration model. The bar graph in figure 6.3 a) shows the comparison of Linear migration approach and our two approaches – Sinusoidal BBO and Improved sinusoidal BBO. Results from these graph clearly shows that our proposed algorithm outperforms linear algorithm.

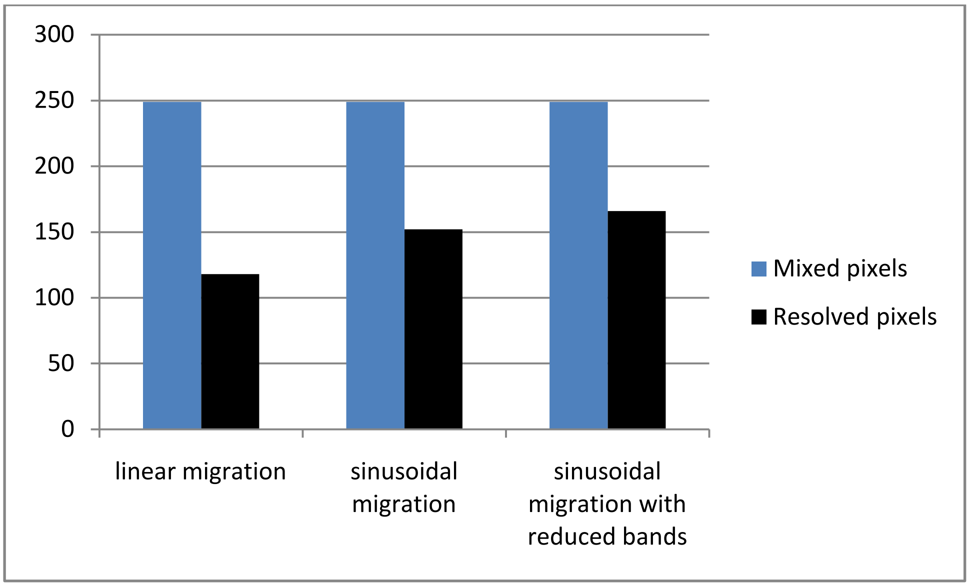

Figure 6.3 a) Graphical analysis of linear, sinusoidal and reduced band sinusoidal approaches

In addition to above comparisons, proposed algorithm has been further analysed by allowing it an option to assign any class to the mixed pixel. The result obtained for set of urban-barren, barren-vegetation and urban-vegetation mixed pixels is depicted in the following figure 6.3 b), figure 6.3 c) and figure 6.3 d) respectively. The proposed algorithm assigns the mixed pixel to a class which is one of its constituent. Therefore even without supplying the subset of features to choose from, the proposed algorithm correctly identifies the regions to which the mixed pixel can belong.

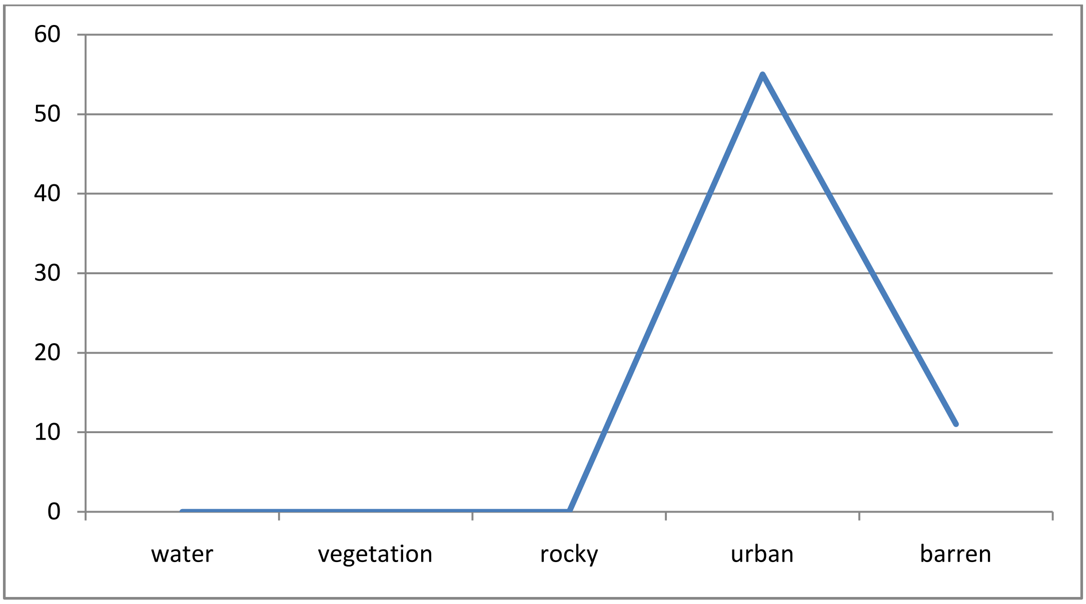

Figure 6.3 b) Urban – Barren Mixed pixels

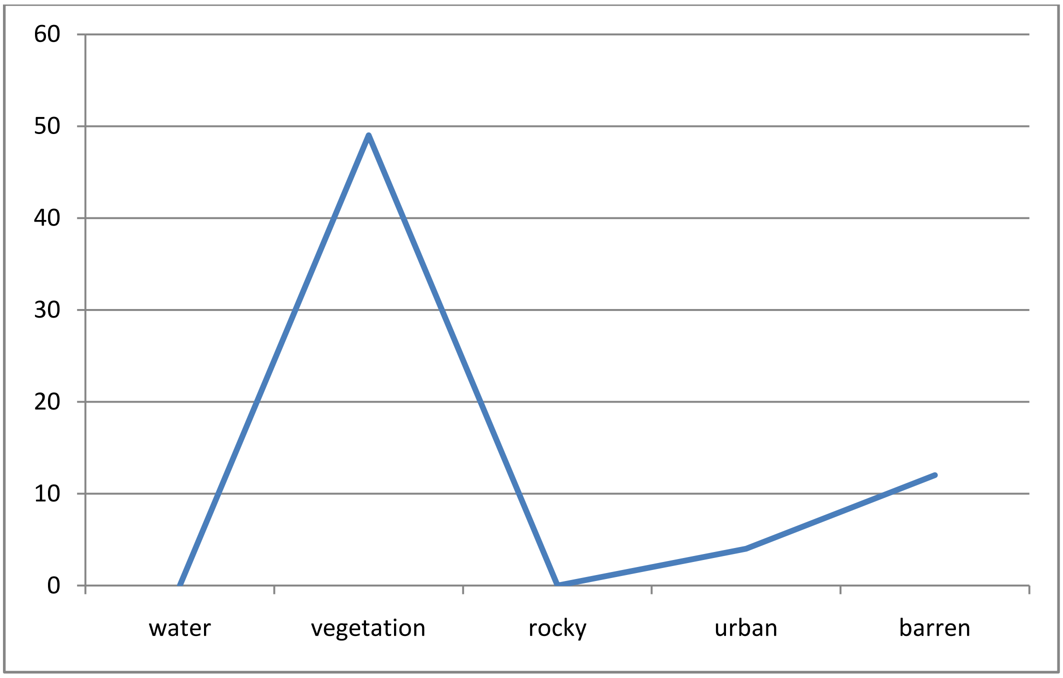

Figure 6.3 c) Barren – Vegetation Mixed pixels

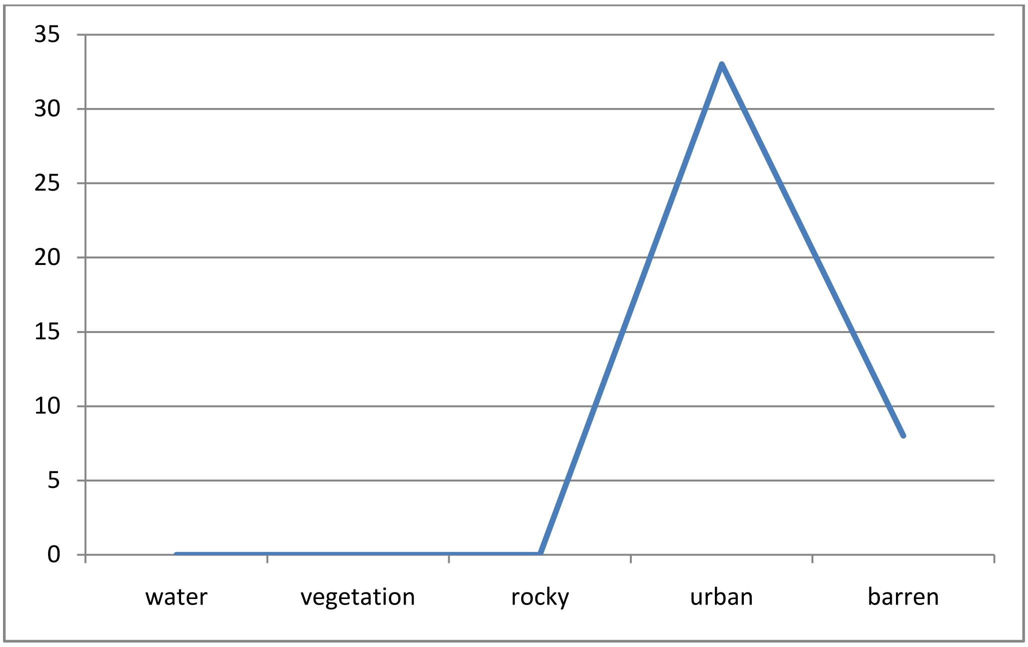

Figure 6.3 d) Urban – Vegetation Mixed pixels

Chapter 7.

Conclusion and future scope

In this work we tried to develop an improved technique to solve the mixed pixel resolution problem by using sinusoidal migration concept of BBO. To improve further this technique we have adopted redundant band removal techniques based on their correlation values.

We followed two approaches. The algorithm used in first approach resolves mixed pixels using sinusoidal migration model. Linear migration cannot take into account other constraints such as evolution of particular species, and population size [2]. This limitation is overcome by employing immigration rates to be obtained from a sinusoidal curve. Depending on an intended analytical task, some features may become irrelevant. These irrelevant (redundant) features affect not only the efficiency of desired analysis but also the performance and accuracy. Therefore, to improve it further, second algorithm is designed where concept of selection of bands containing unique information for a particular land class is used.

The least correlated bands have been chosen to reduce computational overhead and memory space without compromising on the efficiency of the algorithm. The number of bands to be selected are determined from the eigen vector values of the correlation coefficient matrix. The most expensive part of a population based optimization algorithm is the calculation of fitness function. Mean of standard deviation of different image bands of the image has been used as fitness function here.

Future work can be conducted on studying the effect of different fitness functions instead of mean of standard deviation. Functions that reduce the computational complexity and at the same time maintains the performance should be incorporated. Species in a habitat would prefer to migrate to a habitat closer to its original habitat. This factor can be taken care of while selecting immigrating habitats. Apart from images acquired through satellite remote sensing, this technique for mixed pixel resolution may also be applied to industrial images or medical imagery. The algorithm may further be extended to hyperspectral imagery.

Chapter 8.

Publications from the thesis

Conference Name: Global Congress on Intelligent Systems (GCIS 2013)

Paper Title: "Efficient resolution of Mixed Pixels using Bio-Inspired Heuristics"

Authors: Daya Gupta, Lavika Goel, Tanu Varshney and V.K. Panchal

Status: Paper submitted for review in the journal.

Location: Hong Kong, China

Conference date: 3-4 December, 2013

Publisher/Proceedings: If the paper is accepted for the conference, its conference proceedings will be published by the CPS which will submit the conference proceedings in the IEEE Xplore and submit the conference proceedings to Ei Compendex and ISTP for indexing.

References

[1] Simon, D., “Biogeography-based optimization”, IEEE Transactions on Evolutionary Computation, vol. 12, no. 6, pp. 702-713, December 2008.

[2] Haiping Ma, “An analysis of the equilibrium of migration models for biogeography-based optimization”, Information Sciences, 3444–3464, 2010.

[3] Daya Gupta, Lavika Goel, V.K.Panchal, “Embedding Expert Knowledge to Hybrid Bio-Inspired techniques- An Adaptive Strategy towards Focussed Land Cover feature Extraction”, International Journal of Computer Science & Information Security, (IJCSI), vol.8, no.2, 2010.

[4] V.K. Panchal and Nitish Gupta, “Swarm Intelligence for Mixed Pixel Resolution”, IEEE Geo Sciences and Remote Sensing Symposium (IGRASS), pp- 2801-2804, 2011.

[5] V.K. Panchal , Samiksha goel, Mitul Bhatnagar, “Biogeography Based Land Cover Feature Extraction” , in VIII International Conference on Computer Information Systems and Industrial Management (CISIM 2009) ,Coimbatore, India, December 2009.

[6] A. L. Choodarathnakara, Dr. T. Ashok Kumar, Dr. Shivaprakash Koliwad, Dr. C. G. Patil, “Mixed Pixels: A Challenge in Remote Sensing Data Classification for Improving Performance”, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 9, November 2012.

[7] Lavika Goel, Daya Gupta, V.K. Panchal, "Hybrid bioinspired techniques for land cover feature extraction: A remote sensing perspective," Applied soft computing, Elsevier publications, Vol. 12(2), pp. 832-849, 2011.

[8] P. Fisher, “The pixel: a Snare or a Delusion”, International Journal of Remote Sensing, 1997, Vol. 18, pp. 679-685.

[9] Giles M. Foody, "Relating the Land-Cover Composition of Mixed Pixels to Artificial Neural Network Classification Output", American Society for Photogrammetry and Remote Sensing, 1996.

[10] Marco Dorigo and Thomas Stutzle, "Ant Colony Optimization" IEEE Transaction on Evolutionary Computation, Vol 6, Issue 4, pp-358-365, 2002.

[11] Kennedy, J.; Eberhart, R., 1995. Particle Swarm Optimization, Proceedings of IEEE International Conference on Neural Networks(IV), pp. 1942–1948.

[12]URL:http://essedesignguide.org/index.php?option=com_content&task=view&id=165&Itemid=9

[13] Giles M. Foody," Status of land cover classification accuracy assessment", Remote Sensing of Environment 80, 185– 201, 2002

[14]URL:ftp://ftp.wsl.ch/downloads/babst/Fernerkundung_WS2012/literatur/remote_sensing.pdf

[15] J.M. Sotoca, F. Pla, A.C. Klaren, "Unsupervised band selection for multispectral images using information theory", Proceedings of the 17th International Conference on Pattern Recognition 1051-4651/04 IEEE, 2004.

[16] URL: http://www.crisp.nus.edu.sg/~research/tutorial/optical.htm

[17] D.Du, D. Simon, and M. Ergezer，"Biogeography-Based Optimization Combined with Evolutionary Strategy and Immigration Refusal," IEEE International Conference on Systems, Man, and Cybernetics. San Antonio, TX, pp. 1023-1028, October 2009.

[18] T.M. Lillesand, R.W. Kiefer & J.W. Chipman, ―Remote Sensing & Image Interpretations‖ 5th edition, Wiley & Sons Pvt. Ltd., pp 586-592, 2005, England.

[19] V.K. Panchal, D. Bhugra, S. Goel, V. Singhania," Study on the behaviour of bbo over natural terrain features", IEEE Electronics Computer Technology(ICET) 3rd International Conference , Vol 4, pp 28-32, 2011.

[20] V.K. Panchal, P. Singh, N. Kaur, H. Kundra," Biogeography based Satellite Image Classification", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6 , No.2, 2009.

[21] URL: https://www.e-education.psu.edu/natureofgeoinfo/book/export/html/1608

[22] A. Bhattacharya, P.K. Chattopadhyay, "Solving complex economic load dispatch problems using biogeography-based optimization," Expert Systems with Applications, vol..37,pp. 3605–3615, 2010.

[23] Pei Liang, YanChunyu," Study on Mixed Pixel Classification Method of Remote Sensing Image based on Fuzzy Theory", Urban Remote Sensing Joint Event, 2009

[24] Lavika Goel, Daya Gupta,V.K Panchal," Information Sharing in Swarm Intelligence Techniques: A Perspective Application for Natural Terrain Feature Elicitation", International Journal of Computer Applications (0975 – 8887) Volume 32– No.2, October 2011

[25] Haiping Ma, Minrui Fei, Zhiguo Ding, Jing Jin, "Biogeography-Based Optimization with Ensemble of Migration Models for Global Numerical Optimization", WCCI 2012 IEEE World Congress on Computational Intelligence June, 10-15, 2012 - Brisbane, Australia

[26] P. Bosdogianni, Maria Petrou and Josef Kittler, "Mixed pixel classification in Remote sensing", Image and Signal Processing for Remote Sensing, Proc. SPIE Vol. 2315, p. 494-505,1994.

[27] N.R. Pal and J.C. Bezdek , Dept. of Computer. Sci., Univ. of West Florida, "Validity for the fuzzy cmeans model." IEEE Transaction Of Fuzzy System, Vol 3, pp-370-379, Aug 1995.

[28] Jos´e Mart´ınez Sotoca, Filiberto Pla, and Jos´e Salvador S´anchez, "Band Selection in Multispectral Images by Minimization of Dependent Information", IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 37, no. 2, march 2007

[29] Aaron Moody, Sucharita Gopal, and Alan H. Strahle,"Artificial Neural Network Response to Mixed Pixels in Coarse-Resolution Satellite Data", Remote sensing environment, 58:329-343 (1996) ©Elsevier Science Inc., 1996

[30] Holden, A.A. Freitas, A hybrid particle swarm/ant colony algorithm for the classification of hierarchical biological data, in: IEEE Swarm Intelligence Symposium (SIS-05), pp. 100–107, 2005.

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

[32] Settle, J.J. and Drake, N.A., "Linear mixing and the estimation of ground cover proportions", International Journal of Remote Sensing, Vol. 14, pp. 1159-1177, 1993.

[33] Ichoku, C and Karnieli, A. "Areview of mixture modelling techniques for sub-pixel land cover estimation", Remote Sensing Reviews, Vol.13, pp 161-186, 1996.

[34] URL: http://www.nrsc.gov.in/#

[35] H. M. Horwitz, R.F. Nalepka, P. D. Hyde, and J.P. Morgenstern, "Estimating the proportions of objects within a single pixel resolution elements of a multispectral scanner," International Symposium of Remote Sensing of Environment, pp. 1307-1320, 1971.

[36] G. M. Foody, D. P. Cox, "Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions", International Journal of Remote Sensing, Vol. 15, No. 3, pp. 619-631, 1994.

[37] M.S. klein Gebbinck, T.E. Schouten," Decomposition of mixed pixels", Image and Signal Processing for Remote Sensing, pp. 104-115,1995.

[38] Lavika Goel, Daya Gupta,V.K Panchal,"Biogeography based Anticipatory Computing Framework for Intelligent Battle Field Planning", 12th International Conference on Hybrid Intelligent Systems (HIS), 2012.

Appendix A

Abbreviations Used

ANN – Artificial Neural Network

ACO – Ant Colony Optimization

BBO–Biogeography Based Optimization

DEM – Digital Elevation Model

DN – Digital Number

EA – Evolutionary Algorithm

EM – Electromagnetic

FL – Fuzzy Logic

HSI – Habitat Suitability Index

IFOV – Instantaneous Field Of View

MDC – Minimum Distance Classifier

MIR – Middle Infra-Red

MLC – Maximum Likelihood Classifier

NIR – Near Infra-Red

PSO – Particle Swarm Optimization

RS1 – Radarsat-1

RS2 – Radarsat-2

SI – Swarm Intelligence

SIV – Suitability Index Variable

Appendix B

Introduction to MATLAB

MATLAB is a mathematical and graphical software package; it has numerical, graphical, and programming capabilities. It has built-in functions to do many operations, and there are toolboxes that can be added to augment these functions (e.g., for signal processing). There are versions available for different hardware platforms, and there are both professional and student editions.

Basic Matlab concepts

When the MATLAB software is started, a window is opened: the main part is the Command Window. In the Command Window, there is a statement that says:

In the Command Window, you should see:

>>

The >> is called the prompt. In the Student Edition, the prompt appears as:

EDU>>

In the Command Window, MATLAB can be used interactively. At the prompt, any MATLAB command or expression can be entered, and MATLAB will immediately respond with the result. It is also possible to write ***programs*** in MATLAB, which are contained in ***script files*** or M-files.

There are several commands that can serve as an introduction to MATLAB and allow one to get help:

- **info** will display contact information for the product
- **demo** has demos of several options in MATLAB
- **help** will explain any command;
- **help help** will explain how help works

- **helpbrowser** opens a Help Window

A script is a sequence of MATLAB instructions that is stored in a file and saved. The contents of a script can be displayed in the Command Window using the **type** command. The script can be executed, or run, by simply entering the name of the file (without the .m extension). To create a script, click File, then New, then M-file. A new window will appear called the Editor. To create a new script, simply type the sequence of statements (notice that line numbers will appear on the left).

Image Representation in Matlab

In MATLAB a binary and gray-scale image is represented by one 2-dimensional array, whereas a color image are represented by a 3-dimensional array (one 2-dimensional array for each of the color planes or color channels red, green and blue). A gray-scale or binary pixel consists of one data value. A color pixel consists of 3 data values (each for one of the color channels). The most common data types of the individual pixels are:

uint8 : unsigned integer: data range 0..255

double : double precision float: data range 0.0 ... 1.0

Basic MATLAB functions

Image information

imfinfo('foo.ext')	displays information on image format etc. of the file foo.ext
imformats	displays an overview of all MATLAB image formats
whos img	displays information about the array img: size, data type, etc.

Reading, writing and displaying images

myImg = imread('foo.ext')	reads file foo.ext into array myImg
imwrite(anImg, 'foo.ext')	writes the image anImg to the file foo.ext.
imshow(myImg)	displays the image myImg as gray-scale, binary or colour image depending on the data type of myImg
figure(n)	opens a new window with number n
imshow()	displays the image within this window

Basic image processing functions

img = uint8(zeros(512,1024))	creates a black image with width 1024 and height 512 of type uint8
img = double(zeros(512,1024))	creates a black image with width 1024 and height 512 of type double
img = double(ones(512,1024))	creates a white image with width 1024 and height 512 of type double
red = myImg(:,:,1)	stores the red component of myImg (rgb-image) in array red
green = myImg(:,:,2)	stores the green component of myImg in array green
blue = myImg(:,:,3)	stores the blue component of myImg in array blue
mx = max(myImg(:)))	computes the maximum value of an 2-d array
mi = min(myImg(:)))	computes the minimum value of an 2-d array

Appendix C

Introduction to PSO/ACO Open Source Tool

PSO/ACO2 uses a hybrid Particle Swarm Optimization/Ant Colony Optimisation algorithm to generate classification rules. It takes data sets of the form of ARFF. It supports binary, nominal and continuous attributes. This algorithm still needs to consider rule interaction properly. The PSO/ACO2 optimiser itself is sound, but it may produce sub-optimal results. Version 0.95 feature creeps into Differential Evolution support, that very popular optimiser. Code adapted from Java implementation by Mikal Keenan and Rainer Storn. Note that it probably won't work well with a mix of continuous and nominal attributes, due to problems integrating the PSO/ACO2 nominal optimiser and DE code.

In version 0.9 you can now set a test set, as well as performing the standard cross-validation procedure. Open a training set using File > Open..., then set an optional test set using the File > Open Test Set. This program was developed at the University of Kent at Canterbury England, by Nicholas Holden, under the supervision of Alex A Freitas. It was pulled directly out of experimental code and given a front end, so expect spaghetti code and major bugs. While using this tool you can select the number of iterations the algorithm should run for, for each rule as 100 also. For lots of continuous attribute you will see improvements with 200 iterations. You can select the number of particles, the default is 10^2, which actually equates to 100 in the algorithm. You can select whether to use Precision or Sensitivity*Specificity (as with Ant-Miner) as the fitness function for particles. These will work better or worse depending on the data set. Overall precision seems to perform better.

Appendix D

Introduction to ERDAS

ERDAS is pleased to provide ERDAS IMAGINE® version 8.4. Many private and commercial users who need to extract and interpret information from imagery recognize ERDAS IMAGINE as a must have. With ERDAS IMAGINE 8.4, ERDAS' latest, most advanced release of ERDAS IMAGINE , production workflows are enhanced and simplified like never before. As an example, the Batch Wizard streamlines repetitive procedures such as importing; reprojecting, and exporting large numbers of files at once, using a wizard approach to record and "re-play" commonly used procedures. Also featured is the IMAGINE Expert Classifier™ – a tool for graphically building and executing geographically aware, rules-based expert systems. This tool can be used to build decision support systems, classifiers for high-resolution imagery, GIS analysis techniques, etc. These can then be distributed to other users for use with their own data.

Key Features Summary

• IMAGINE Expert Classifier

• Direct read and edit of ESRI's Shapefiles

• Direct read of ESRI's SDE data

• Enhanced and expanded native raster file handling

• Re-projection of raster data on-the-fly

• Batch processing wizard

• Enhanced Viewer functionality

• Improved print versatility on Windows NT

• International 2-byte font support in Annotation layers

• Support for ERDAS IMAGINE .img files larger than 2 GB