A
Dissertation
On

"Hybrid Bio-Inspired Techniques for Land Cover Feature Extraction - A Remote Sensing Perspective"

Submitted in partial fulfillment of the requirement
For the award of Degree of

MASTER OF ENGINEERING
(Computer Technology and Application)

Submitted by
Lavika Goel
(University Roll no. 8406)

Under the Guidance of
Dr. DAYA GUPTA
(Head of Department)
Department of Computer Engineering

DEPARTMENT OF COMPUTER ENGINEERING
DELHI COLLEGE OF ENGINEERING
BAWANA ROAD, DELHI
(DELHI UNIVERSITY)
2008-2010

CERTIFICATE

DELHI COLLEGE OF ENGINEERING
(Govt. of National Capital Territory of Delhi)
BAWANA ROAD, DELHI – 110042
Date:___________

This is to certify that the thesis entitled "**Hybrid Bio-Inspired Techniques for Land Cover Feature Extraction- A Remote Sensing Perspective**" submitted by Ms. Lavika Goel of Delhi College of Engineering in partial fulfillment of the requirement for the award of the degree of **Master of Engineering** in **Computer Technology & Applications** is a record of the student's own work carried out by her under my supervision and guidance. This is a beneficial work in field of Remote Sensing using very recent Swarm Intelligence based technologies and resources. The matter embodied in this thesis has not been submitted for the award of any other degree or diploma.

(Dr. DAYA GUPTA)
H.O.D & PROJECT GUIDE
(Dept. of Computer Engineering)
Delhi College of Engineering
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I address my sincere thanks to Almighty God for giving me the inner strength to complete my dissertation. He was always there protecting and saving me.

It is a great pleasure to have the opportunity to extend my heartiest felt gratitude to everybody who helped me throughout the course of this project. It is a distinct pleasure to express my deep sense of gratitude and indebtedness to my learned supervisors Dr. Daya Gupta and Dr. V.K. Panchal for their invaluable guidance, encouragement and patient reviews.

This project has been a dedicated effort towards the development of a highly autonomous artificially intelligent system, which primarily would not have been possible at the first place without the apt guidance of the Head of Computer Engineering Department, respected Dr. Daya Gupta. Her motivation and encouragement during the project made me go further than expected in the development process and the redaction of this current report.

This project has been accomplished in collaboration with DTRL lab at DRDO. And hence, I would also like to present my heartiest felt gratitude to Dr. V. K. Panchal, Add. Director & Scientist 'G, Defence Terrain & Research Lab-DRDO who provided me the Invaluable Satellite Data for the experimental study. I would also like to thank him for his invaluable guidance & support throughout the completion of this work.

This dedication would not be complete without expressing my gratitude towards all staff members and my dear classmates for their interest and curiosity towards my project, where definitely to talk about biogeography for solving problems can sound curious. Also, I owe a special thanks to the crowd who are active in the field of Swarm Intelligence. Last but not the least, I would love to thank my family and friends, who directly and indirectly have given me moral support and for their relentless advice throughout the completion of this work.

(LAVIKA GOEL)
Master of Engineering (M.E.)
Computer Technology & Applications
Dept. of Computer Engineering
DELHI COLLEGE OF ENGINEERING
BAWANA ROAD, DELHI – 110042

ABSTRACT

Recent developments in applied and heuristic optimization methods used for feature extraction from satellite images have been strongly influenced and inspired by natural and biological system. The findings of recent studies are showing strong evidence to the fact that some aspects of biogeography can be adaptively applied to solve specific problems in science and engineering.

This project presents a hybrid biologically inspired technique called the ACO2/PSO/BBO (Ant Colony Optimization2 / Particle Swarm Optimization / Biogeography Based Optimization) Technique that can be adapted according to the database of expert knowledge for a more focused Satellite Image Classification. The hybrid classifier explores the adaptive nature of Biogeography Based Optimization Technique and therefore is flexible enough to classify a particular land cover feature more efficiently than others based on the 7-band image data distribution and hence we term our classifier as the hybrid bio-inspired pattern analysis based intelligent classifier. The thesis also presents a comparative study of the proposed classifier and the other recent Soft Computing Classifiers such as ACO, Hybrid PSO-ACO2, Hybrid ACO-BBO Classifier , Fuzzy sets, Rough-Fuzzy Tie up and the Semantic Web Based Classifiers with the traditional probabilistic classifiers such as the Minimum Distance to Mean Classifier and the Maximum Likelihood Classifier. The accuracy of the results that had been checked by obtaining the KHAT statistics show that highly accurate land cover features can be extracted effectively when the proposed algorithm is applied to the 7-Band Image.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

In remote sensing, satellite based sensors are burgeoning as a major facilitator of geo-spatial information providing different manifestations of the terrain. The satellite image is one of the main source for capturing the geo-spatial information [29]. Remote sensing with multi-spectral satellite imagery is based on the concept that different features/objects constituting the land cover reflect electro-magnetic radiations over a range of wavelengths in its own characteristics way according to its chemical composition and physical state. A multi-spectral remote sensing system operates in a limited number of bands and measures radiations in series of discrete spectral bands. The spectral response is represented by the discrete digital number spectral signatures of an object may be used for identification much like a fingerprint [29].

There are two main type of classifying techniques: *Supervised and Unsupervised classification* [29]. When spectral classes, based on numerical information, are grouped first and are then matched by the analyst to information classes, then it is termed as unsupervised classification. Clustering algorithms are used to determine the statistical structures in the data for example K-Means approach [30]. In supervised classification, the homogeneous samples of the different surface cover types of interest are used. To recognize spectrally similar areas, numerical information in all spectral bands for the pixels comprising these areas is used. For each pixel in the image a comparison is made with these signatures and defined in the class it most closely "resembles" e.g., ‘Minimum Distance to Means’ classifier [27].

A new wide range of computational algorithms have emerged from the behaviour of social insects. Social insects are usually characterized by their self organization and with the minimum communication or the absence of it. Every social insect individually is self-autonomous. They can obtain information about the environment and interact with the remote insects or environment indirectly, by stigmergy. All these features characterize Swarm Intelligence [17] which can be used for satellite image classification as the results of applying these techniques for classification are found to be more accurate.

1.1 Motivation

Satellite images contain land cover types some of which cover significantly large areas, while some (e.g., bridges and roads) occupy relatively much smaller regions. In the future, new generation of high-resolution satellite sensors will acquire enormous data quantity, so extracting features from a remote sensed image will become even more difficult. Researchers have been continuously searching for new techniques that can extract maximum information from the remotely sensed image.

Very recently the concept of swarm intelligence [10], an optimized approach of image classification of satellite multi-spectral images, has been introduced in this category. This technique with lower cost and higher degree of classification accuracy, will be able to replace high resolution high cost satellite imageries. Various swarm intelligence techniques such as the Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO) and hybrid ACO2/PSO optimization have been used for solving the problem of satellite image classification. However, out of these enormous techniques available for image classification, at present, it is not possible to state which classifier is best for all situations as the characteristic of each image and the circumstances for each study vary so greatly. Therefore, it is essential that each analyst understand the alternative strategies for feature extraction so that he or she may be prepared to select the most appropriate technique for the feature extraction task in hand.

Our proposed work is motivated by the very recently introduced swarm intelligence technique called the **Biogeography based Optimization** technique [5] and bases its fundamentals from the research paper titled " Biogeography Based Optimization" by Dan Simon (December 2008) [5] . The findings of recent studies are showing strong evidence to the fact that some aspects of biogeography [7] can be adaptively applied to solve specific problems in science and engineering. The main characteristic of the BBO technique for satellite image classification is that this technique is flexible to classify the desirable features more efficiently than the other features and hence it shows a wide range of efficiencies in classifying different features of an image. Also BBO technique improves its solutions with each iteration and thus probabilistically refines its classification at each iteration.

By the motivation of these characteristic strengths of the above swarm intelligence techniques, we formulated an algorithm which is a hybrid of the ACO, PSO and BBO techniques for Satellite Image Classification. Hence we propose a swarm intelligence based classification algorithm which integrates the ACO2/PSO technique [8] with the recently introduced population based swarm intelligence technique called the Biogeography based optimization technique [5] for building an artificially intelligent hybrid classifier.

1.2 Related Work

In remote sensing the problem of Satellite Image Classification has been solved by using the traditional classical approaches like Parellelopiped Classification [30], Minimum Distance to Mean Classification [30], Maximum Likelihood Classification [30] etc. However, these techniques show limited accuracy in information retrieval and high resolution image is needed. Also these techniques are insensitive to different degrees of variance in the spectral response data.

To provide a solution to the above problems, soft computing techniques were introduced in remote sensing for image classification. Soft computing techniques differ from the above conventional (hard) computing techniques in that, unlike hard computing, they are tolerant of imprecision, uncertainty and partial truth. In effect, the role model for soft computing is the human mind. The guiding principle of soft computing is: Exploit the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness and low solution cost. The principal constituents of soft computing techniques are fuzzy logic [26], rough set theory [23], neural network theory, probabilistic reasoning, and Swarm Intelligence Techniques [17] with the latter subsuming belief networks, genetic algorithms, chaos theory and parts of learning theory.

However, the soft computing techniques like the fuzzy classifier [26] and the rough set classifier [23] were not able to provide good result in case of ambiguity sine the main goal of these techniques was to synthesize approximation of concepts from the acquired data. Hence, these techniques did not provide very much accurate results with low spatial resolution images. Also these techniques were not able to handle the crisp and continuous data separately.

The solution to the above drawbacks was provided by the recently introduced concept of swarm intelligence [10] [17]. Our proposed work bases its fundamentals from Swarm Intelligence. This technique improves the classification of satellite multi-spectral images and is more accurate when working with low spatial resolution images. As the frontiers of space technology advance, the knowledge derived from the satellite data has also grown in sophistication. We need more accurate and efficient image classification algorithms. Also no single classifier can prove to satisfactorily classify all the basic land cover classes of a region. Hence we propose a swarm intelligence based classification algorithm which combines the strengths of the ACO2/PSO technique [8] with the characteristic strengths of the recently introduced population based swarm intelligence technique called the Biogeography based optimization technique [5] .This combination gives this classifier artificial intelligence to identify the features efficiently classified by BBO and by ACO2/PSO separately based on an analysis of the training set data distribution graph.

1.3 Problem Statement

This work focuses on proposing a hybrid algorithm inspired by biogeography for feature extraction from high resolution satellite multi-spectral images. Biogeography based optimization algorithm is basically used to find the optimal solution of a problem [5]. But land cover feature extraction is a clustering problem that requires each feature to be extracted as a cluster. The original BBO algorithm does not have the inbuilt property of clustering. But here to extract features from the image, we have tried to make the clusters of different land cover features and proposed a hybrid of this modified biogeography based algorithm with the ACO2/PSO technique for extraction of land cover features from the satellite image hence leading to the classified image.

The proposed work presents a hybrid biologically inspired technique, called the ACO2/PSO/BBO Technique that can be adapted according to the database of expert knowledge for a more focused satellite image classification. The hybrid classifier explores the adaptive nature of biogeography based optimization technique and therefore is flexible enough to classify a particular land cover feature more efficiently than others based on the 7-band data and hence can be adapted according to the application. The proposed hybrid classifier improves the efficiency of the biogeography based classifier for a more efficient, focused and refined image classification. The classifier is artificially intelligent to identify the features efficiently classified by BBO and by ACO2/PSO separately based on a similarity pattern analysis and therefore we term our classifier as the Hybrid Bio-Inspired Pattern Analysis Based Intelligent Classifier.

1.4 Scope of the work

Our approach to use a *hybrid of the ACO2/PSO and the BBO Approach for Land Cover Feature Extraction from Satellite images* is able to perform some very important targets successfully completed.

- The proposed algorithm has been applied to the 7-band cartoset satellite image of size 472 X 576 of the Alwar area in Rajasthan (since it contains a variety of land cover features) to classify the area into 5 features namely water, urban, barren, vegetation and rocky region and assign color codes to them hence solving the problem of land cover feature extraction of this region.
- We also present a comparative study of the results of applying our proposed classifier and the other recent soft computing classifiers such as ACO [19], Hybrid PSO-ACO2 [8], Hybrid ACO-BBO Classifier, Fuzzy sets [21] [22], Rough-Fuzzy Tie up [21] [22] and the Semantic Web Based Classifiers [32] when compared with the traditional probabilistic classifiers such as the Minimum Distance to Mean Classifier and the Maximum Likelihood Classifier on the 7-Band Alwar Image and observed that our approach provides a remarkable increase in accuracy of classification as compared to the simple biogeography based land cover feature extraction technique as well as the other recent and traditional image classification techniques. In fact, our hybrid classifier is able to achieve the highest efficiency amongst all of the swarm intelligence techniques used till now and that too with minimum computation and resources.
- This algorithm is made to prove itself as a good competitor to its complementary techniques with most inaccurate inputs it can handle. The system performance can be increased with better unsupervised classifications and better training sets.

1.5 Organization of thesis

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

Section 2: This section gives an introduction to the remote sensing fundamentals. It describes the process of remote sensing , the meaning of the digital number and the satellite imaging. It also describes the elements of the imaging system in detail. This section also describes the process of land cover feature extraction which forms an application of remote sensing.

Section 3: This section describes the image classification techniques. The different supervised and unsupervised traditional approaches are described here. Here the recent Soft Computing approaches for image classification are also given. Also, an introduction to the different Swarm Intelligence Techniques such as the Ant Colony Optimization, Particle Swarm Optimization, Bee Colony Optimization Techniques, and the newly introduced in this category i.e. biogeography based optimization etc. are also provided in this section.

Section 4: This section presents a brief review of the population based optimization technique, the Biogeography Based Optimization and also compares and analyses its performance as compared to other recent Swarm Intelligence Techniques for the Step Function Optimization. This section also describes an integration of the BBO with ACO to predict the deployment strategies of enemy troops in a military terrain application. This section also contains a brief review of the hybrid ACO2/PSO Optimization techniques for image classification.

Section 5: This section describes the proposed framework of the hybrid ACO2/PSO/BBO algorithm for the Land cover feature extraction from multi-spectral Satellite image. It describes the biogeography based land cover feature extraction algorithm in detail and the proposed software architecture of the integrated BBO/ACO2/PSO technique. It also presents the overall mathematical formulation of the proposed algorithm.

Section 6: This Section presents the results of the case study which is based on the experiments performed at DTRL Lab-DRDO for the classification of satellite image of Alwar area in

Rajasthan, using our proposed algorithm. Also it establishes the fact that the results obtained by our algorithm provide more accuracy then the earlier traditional approaches for image classification and also that our proposed algorithm is a significant improvement over the biogeography based land cover feature extraction algorithm.

Section 7: This section gives the details of the publications from the thesis and the conference / journal details.

Section 8: In this section the conclusion of the thesis work and the future scope of the work are presented.

Section 9: This section gives the references details of the thesis.

Appendix A: Abbreviations used

Appendix B: Introduction to MATLAB Software

Appendix C: Introduction to ERDAS software

Appendix D: Introduction to the Open Source Tool for Hybrid ACO2/PSO

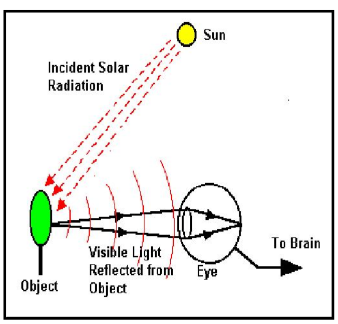
Appendix E: Introduction to ROSETTA Software

Chapter 2

LAND COVER FEATURE EXTRACTION- A REMOTE SENSING APPLICATION

Land cover refers to the surface cover on the ground, whether vegetation, urban infrastructure, water, bare soil or other. Each of these i.e. ground, vegetation, urban infrastructure, water, bare soil or other is called a feature of land cover. There are two primary methods for capturing information on land cover: field survey and thorough analysis of remotely sensed imagery. Land cover feature extraction from satellite images is an important activity for extracting geospatial information for military & civil purposes like inaccessible areas for example, hilly area, enemy areas, disaster management like tsunami, drought area etc.

2.1. Remote Sensing



Remote sensing is the science (and to some extent, art) of acquiring information about the Earth's surface without actually being in contact with it [31]. In other words, Remote sensing is the science and art of obtaining information about an object, an area or phenomenon through the analysis of the data acquired by a device that is not in contact with the object, area or phenomenon under investigation. Using various sensors, we remotely collect data that may be analyzed to obtain information about the objects, area or phenomenon being investigated. The remotely collected data can be of many

forms, including variations in force distribution, acoustic wave distributions or electromagnetic energy distributions.

Kiefer et. All [27] gave a very useful information in their book on remote sensing, from which we are presenting some remote sensing basics. Acquiring information about earth remotely is done by sensing and recording reflected or emitted energy and processing, analyzing, and applying that information. We perceive the surrounding world through our five senses. Some senses (touch and taste) require contact of our sensing organs with the objects. However, we acquire much information about our surrounding through the senses of sight and hearing, which do not require close contact between the sensing organs and the external objects. In another word, we are performing Remote Sensing all the time. Generally, remote sensing refers to the activities of recording/observing/perceiving (sensing) objects or events at far away (remote) places. In remote sensing, the sensors are not in direct contact with the objects or events being observed . The information needs a physical carrier to travel from the objects/events to the sensors through an intervening medium. The electromagnetic radiation is normally used as an information carrier in remote sensing. The output of a remote sensing system is usually an image representing the scene being observed. A further step of image analysis and interpretation is required in order to extract useful information from the image. The human visual system is an example of remote sensing in this general sense. The speed of EM energy is constant and frequency and wavelength are related as shown below:

$$\text{299,893 km/sec} = \text{WAVELENGTH} * \text{FREQUENCY}$$

2.2 Remote Sensing Process

The EM spectrum is arbitrarily segmented into major divisions. There are no natural breaks in the

EM spectrum. These separations are made for our convenience. In much of remote sensing, the process involves an interaction between incident radiation and the targets of interest.

The two basic processes involved are DATA ACQUISITIONS and DATA ANALYSIS as shown in figure 2.1. The elements of data acquisition processes are energy sources, propagation of energy through the atmosphere, energy interactions with earth surface features, retransmission of energy through the atmosphere, airborne sensors, resulting in the generation of sensor data in pictorial or digital form. In short, we use sensors to record variations in the way earth surface features reflect and emit electromagnetic energy [31]. The data analysis process involves examining the data using various viewing and interpretation devices to analyze pictorial data or a computer to analyze digital data. Reference data about the resources being studied are used when and where available to assist in the data analysis. With the aid of the reference data, the analyst extracts information about the type, extent, location and condition of various resources over which the sensor data were collected. This information is then compiled in the form of hard copy maps, tables or as computer files that can be merged with other layers of information in geographical information system. Finally the information is presented to users who apply it to their decision making process.

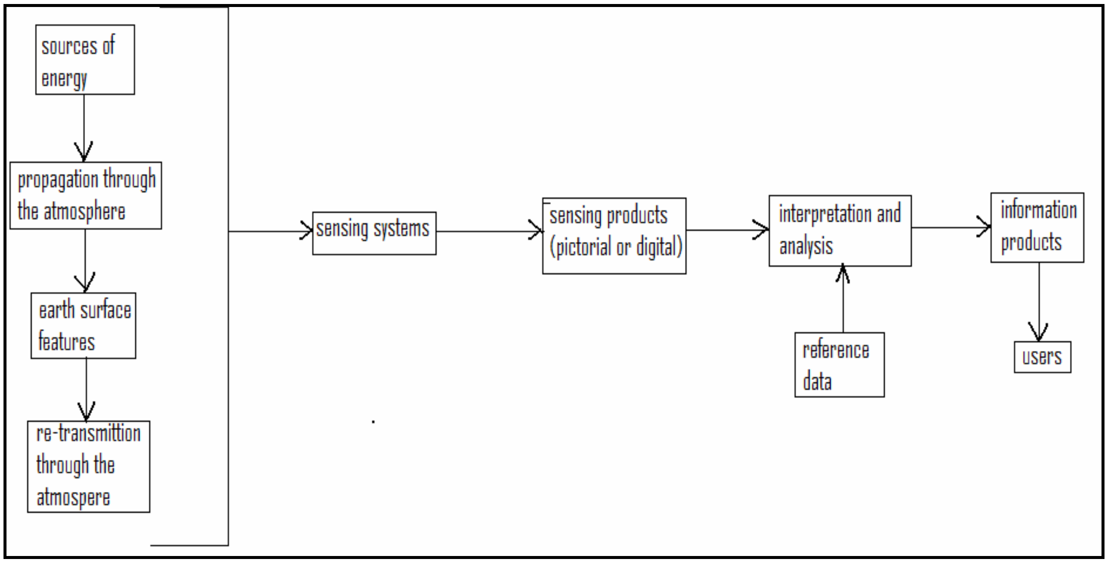


Figure 2.1 Basic processes of Remote Sensing

2.3 Satellite Image

Remote sensing affords us capability to literally see the invisible. It can be thought of as the eyes of many systems like GIS [28], providing repeated, synoptic visions of earth surfaces from the aerial or space vantage point [1-2]. Satellite remote sensing images are representation of earth surfaces as seen from space. Remote sensing with multi-spectral satellite imagery is based on the concept that different features/objects constituting the land cover reflect the electromagnetic radiations over a wide range of wavelength in its own characteristic way.

2.4. Digital Number

Remote sensing images are recorded in digital forms and then processed by the computers to produce images for interpretation purposes. Variations in the scene characteristics are represented as variations in brightness on photographic films. A particular part of scene reflecting more energy will appear bright while a different part of the same scene that reflecting less energy will appear black. Digital image consists of discrete picture elements called pixels. Associated with each pixel is a number represented as DN that depicts the average radiance of relatively small area within a scene. So, DN is defined as:

A positive integer which represents the relative reflectance or emittance of an object in a digital image.

For 8 bit images, the DN or digital number lies in the range 0-255. In a most generalized way, a digital image is an array of numbers depicting spatial distribution of a certain field parameters. Digital image consists of discrete picture elements called pixels. Associated with each pixel is a number represented as DN that depicts the average radiance of relatively small area within a scene. Normally, the range of digital numbers varies from 0-255 (8 bit image) [31].

2.5 Applying Remote Sensing for Extraction of Land Cover Features

Remote sensing is the science (and to some extent, art) of acquiring information about the earth's surface without actually being in contact with it . The process involves an interaction between incident radiation and the targets of interest. This is exemplified by the use of imaging systems where the following seven elements are involved [27] :

(a) Energy source or illumination (A): the first requirement for remote sensing is to have an energy source, which illuminates or provides electromagnetic energy to the target of interest.

(b) Radiation and the atmosphere (B): as the energy travels from its source to the target, it will come in contact with and interact with the atmosphere it passes through. This interaction may take place a second time as the energy travels from the target to the sensor.

(c) Interaction with the target (C): once the energy makes its way to the target through the atmosphere, it interacts with the target depending on the properties of both the target and the radiation.

(d) Recording of energy by the sensor (D): after the energy has been scattered by, or emitted from the target, we require a sensor (remote - not in contact with the target) to collect and record the electromagnetic radiation.

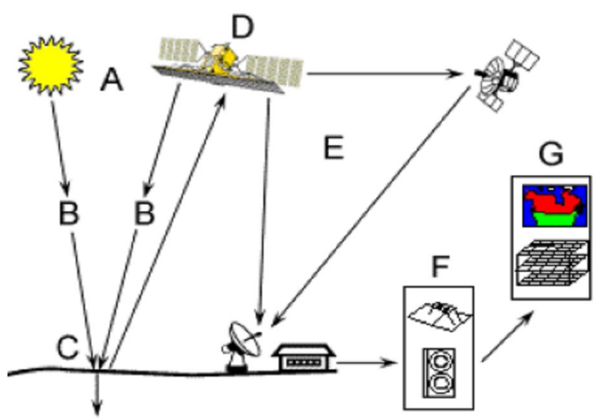


Figure 2.2. Imaging System

(e)Transmission, reception, and processing (E): the energy recorded by the sensor has to be transmitted, often in electronic form, to a receiving and processing station where the data are processed into an image (hardcopy and/or digital).

(f) Interpretation and analysis (F): the processed image is interpreted, visually and/or digitally or electronically, to extract information about the target, which was illuminated.

(g) Application (G): the final element of the remote sensing process is achieved when we apply the information we have been able to extract from the imagery about the target in order to better understand it, reveal some new information, or assist in solving a particular problem.

Remote sensing with multi spectral satellite imagery is based on the concept that different features/objects constituting the land cover reflect electro-magnetic radiations over a range of wavelengths in its own characteristics way according to its chemical composition and physical state. A multi-spectral remote sensing system operates in a limited number of bands and measures radiations in series of discrete spectral bands. The spectral response is represented by the discrete digital number. Spectral signatures of an object may be used for identification much like a fingerprint. The imaging system used in the remote sensing process is represented in figure 2.2.

2.5.1. Land Cover Feature Extraction

The objective of feature extraction is to categorize pixels in an image into land cover classes or themes. Normally multi-spectral data are used to perform the feature extraction and the spectral pattern present within the data for each pixel is used as the numerical basis for feature extraction. That is, different feature types manifest different of Digital Numbers based on their inherent spectral reflectance and emittance properties [29]. Features can be extracted in two ways from a multi-spectral data:

(a) Supervised Feature Extraction: In this type of feature extraction, the image analyst "supervises" the pixel categorization process by specifying, to computer algorithm, numerical descriptors of various land cover types present in scene. To do this, representative sample sites of known cover type called training sets are used to compile a numerical "interpretation key" that describes the spectral attributes for each feature type of interest. Each pixel in the data set is then compared numerically to each category in the interpretation key and labeled with the name of the category it “looks most like”. There are a number of strategies that can be employed to make the comparison between unknown pixels and training set pixels.

(b) Unsupervised Feature Extraction: In unsupervised approach the image data are first aggregated into natural spectral groups or clusters present in the scene by used unsupervised clustering algorithm. Then the image analyst determines the land cover identity of these spectral groups by comparing the classified image data to ground reference data.

Land Cover Mapping is a pattern classification problem, which is solved by the classification of satellite image. For this purpose the satellite image is classified by using many different techniques.

Chapter 3

IMAGE CLASSIFICATION TECHNIQUES

This chapter presents an overview of the various image classification techniques used starting from the traditional techniques to the most recent soft computing techniques. Also semantic web based image classification is added as a special instance.

3.1 Traditional Image Classification Techniques

In remote sensing the problem of Satellite Image Classification is solved by using the traditional classical approaches like Parellelopiped Classification, Minimum Distance to Mean Classification, Maximum Likelihood Classification etc. Some classical supervised classification techniques are following which are in use for satellite image classification for many years [30] .

3.1.1 Parallelopiped Classification

The Parallelopiped classifier is very simple supervised classifier that uses intervals of bounded regions of pixels' values to determine whether a pixel belongs to a class or not. The intervals' bounding points are obtained from the values of the pixels of samples for the class. Since this classifier is supervised there are two steps in its use: signature creation (training set) and classification [30].

Although parallelopiped classification is an accurate classification techniques, it is not widely used because it has several disadvantages. The most important disadvantage is that it can leave many unclassified pixels. Another disadvantage of this classification method is that it can have overlap between training pixels.

3.1.2 Minimum Distance to Mean Classification

In this classification first of all mean or average or spectral in each band for each category is determined. These values comprise mean vector for each category. By considering the two channel pixel values as position coordinates, a pixel of unknown identity may be classified by computing the distance between the value of unknown pixel and each of the category means. The distance between this pixel value and each category mean value is illustrated by dashed lines. After computing the distance the unknown pixel is assigned to the closest class, in this case "corner" [30]. If the pixel is farther than an analyst defined distance from any category mean, it would be classified as unknown.

The minimum-distance-to-means strategy [30] is mathematically simple and computationally efficient, but it has certain limitations. Most importantly, it is insensitive to different degrees of variance in the spectral response data. However this classifier is not widely used in applications where spectral classes are close to one an other in the measurement space and have high variance.

3.1.3 Gaussian Maximum Likelihood Classification

The maximum likelihood classifier [27] quantitatively evaluates both the variance and covariance of the category spectral response patterns when classifying an unknown pixel. To do this, an assumption is made that the distribution of the cloud of points forming the category training data is Gaussian (normally distributed). This assumption of normality is generally reasonable for common spectral response distributions. In the Dimensional graph the vertical axis is associated with the probability of a pixel value being a member of one of the classes. The resulting bell-shaped surfaces are called probability density functions, and there is one such function for each spectral category. The probability density functions are used to classify an unidentified pixel by computing the probability of the pixel value belonging to each category.

The major disadvantage of this technique is that the maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. The basic equation assumes that these probabilities are equal for all classes, and that the input bands have normal distributions. The maximum likelihood algorithm assumes that the histograms of the bands of data have normal distributions, however this is not true and hence not applicable in all situations.

3.2 Soft Computing Techniques

Soft computing differs from conventional (hard) computing in that, unlike hard computing, it is tolerant of imprecision, uncertainty and partial truth. In effect, the role model for soft computing is the human mind. The guiding principle of soft computing is: Exploit the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness and low solution cost. The principal constituents of soft computing are fuzzy logic [26], rough set theory [23], neural network theory, probabilistic reasoning, and Swarm Intelligence Techniques [17] with the latter subsuming belief networks, genetic algorithms, chaos theory and parts of learning theory. Our project uses Rough Set Theory Toolkit [25] for unsupervised classification in the initial stage of the Biogeography Based Optimization technique for image classification and therefore rough set theory is described in detail below. Also the concept of Swarm Intelligence which forms the basis of our proposed work, and the various swarm intelligence techniques used, are presented in the later section.

3.2.1 Rough Set Theory

Rough Set theory was developed by Zdzislaw Pawlak [23] in the early 1980's. It deals with the classificatory analysis of data tables. The data can be acquired from measurements or from human experts. The main goal of the rough set analysis is to synthesize approximation of concepts from the acquired data.

The rough set philosophy is founded on the assumption that with every object of the universe of discourse we associate some information (data, knowledge) [24]. For example, if objects are patients suffering from a certain disease, symptoms of the disease form information about patients. Objects characterized by the same information are indiscernible (similar) in view of the available information about them. The indiscernibility relation generated in this way is the mathematical basis of rough set theory. Any set of all indiscernible (similar) objects is called an elementary set, and forms a basic granule (atom) of knowledge about the universe. Any union of some elementary sets is referred to as crisp (precise) set – otherwise the set is rough (imprecise, vague) .

In the rough set approach, it is assumed that any vague concept is replaced by a pair of precise concepts called the lower and the upper approximation of the vague concept. The lower approximation consists of all objects that surely belong to the concept and the upper approximation contains all objects that possibly belong to the concept. Obviously, the difference between the upper and the lower approximation constitutes the boundary region of the vague concept. Approximations are two basic operations in rough set theory.

One of the main objectives of rough set data analysis is to reduce data size. Various notions such as indiscernibility, rough set, reduct are used to approximate inconsistent information and to exclude redundant data.

3.2.1.1 Information System

A data set is represented as a table, where each row represents a case, an event, a patient, or simply an object. Every column represents an attribute (a variable, an observation, a property, etc.) that can be measured for each object; the attribute may be also supplied by a human expert or user. This table is called an information system. More formally, it is a pair $\ddot{I} = (U, A)$, where U is a non-empty finite set of objects called the universe and A is a non-empty finite set of attributes such that $a: U \rightarrow V_a$ for every a €A. The set V_a is called the value set of a.

In many applications there is an outcome of classification that is known. This is a *posteriori* knowledge is expressed by one distinguished attribute called *decision attribute*; the process is known as supervised learning. Information systems of this kind are called decision systems. A decision system is any information system of the form $\ddot{I}= (U, A\ U \{d\})$, where d € A is the decision attribute. The elements of A are called conditional attributes or simply conditions. The decision attribute may take several values though binary outcomes are rather frequent.

3.2.1.2 Indiscernibility

A decision system (i.e. a decision table) expresses all the knowledge about the model. This table may be unnecessarily large in part because it is redundant in at least two ways. The same or indiscernible objects may be represented several times, or some of the attributes may be superfluous.

The notion of equivalence is recalled first. A binary relation $R \subseteq X * X$ which is reflexive (i.e. an object is in relation with itself xRx), symmetric (if xRy then yRx) and transitive (if xRy and yRz then xRz) is called an equivalence relation. The equivalence class of an element $x \in X$ consists of all objects $y \in X$ such that xRy.

Let $Ï = (U, A)$ be an information system, then with any $B \subseteq A$ there is associated an equivalence relation $IND_Ï(B)$:

$IND_Ï(B) = \{(x; x') \in U^2 \mid \forall a \in B, a(x) = a(x')\}$

$IND_Ï(B)$ is called the *B-indiscernibility relation*. If $(x; x') \in IND_Ï(B)$ then objects x and x' are indiscernible from each other by attributes from B. The equivalence classes of the *B-indiscernibility relation* are denoted $[x]_B$.

3.2.1.3. Set Approximation

An equivalence relation induces a partitioning of the universe. These partitions can be used to build new subsets of the universe. Subsets that are most often of interest have the same value of the outcome attribute. It may happen however that a concept cannot be defined in a crisp manner. In other words it is not possible to induce a crisp (precise) description of some objects from the table. It is here that the notion of rough set emerges. Although we cannot define these objects crisply it is possible to delineate the objects that certainly have a positive outcome, the objects that certainly do not have a positive outcome and finally the objects that belong to a boundary between the certain cases. If this boundary is non-empty the set is rough.

Let $\ddot{I} = (U, A)$ be an information system and let $B \subseteq A$ and $X \subseteq U$. We can approximate X using only the information contained in B by constructing the *B-lower* and *B-upper approximations of* X, denoted $\underline{B}X$ and $\overline{B}X$ respectively, where $\underline{B}X = \{x \mid [x]_B \subseteq X\}$ and $\overline{B}X = \{x \mid [x]_B \cap X \neq \Phi\}$ The objects in $\underline{B}X$ are certain members of X, while the objects in $\overline{B}X$ are possible members of X. The set $BN_B(X) = \overline{B}X - \underline{B}X$ is called the *B-boundary region of X*, and thus consists of those objects that we cannot decisively classify into X. The set $U-\overline{B}X$ *is* called the *B-outside region of X* and consists of those objects that certainly do not belong to X. A set is said to be *rough* (respectively, *crisp*) if the boundary region is non-empty (respectively, empty).

3.2.1.4 Reducts

The other dimension of reducing data is to keep only those attributes that preserve the indiscernibility relation and consequently set approximation. The rejected attributes are redundant since their removal cannot worsen the classification. There is usually several such subsets of attributes and those which are minimal are called Reducts. Given an information system $\ddot{I} = (U, A)$ the definitions of this notion is as follows. A reduct of $\ddot{I}$ is a minimal set of attributes $B \subseteq A$ such that $IND_{\ddot{I}}(B) = IND_{\ddot{I}}(A)$. In other words, a reduct is a minimal set of attributes from A that preserves the partitioning of the universe and hence the ability to perform classifications as the whole attribute set A does.

3.2.1.5 Discretization

In the discretization of a decision table $\ddot{I} = (U, A \cup \{d\})$, where $V_a = [v_a, w_a)$ is an interval of reals, we search for a partition P_a of V_a for any $a \in A$. Any partition of V_a is defined by a sequence of the so-called cuts $v_1 < v_2 < \ldots < v_k$ from V_a. Hence any family of partitions $\{P_a\}_{a \in A}$ can be identified with a set of cuts. In the discretization process we search for a set of cuts satisfying some natural conditions.

Discretization process can be described as one that returns a partition of the value sets of conditional attributes into intervals. The partition is done in such a way that if the name of the interval containing an arbitrary object is substituted for any object instead of its original value in $\ddot{I}$ a consistent decision is also obtained. In this way the size of the value attribute sets in a decision system is reduced.

The discretization step determines how coarsely we want to view the world. An example could be heart-beat rate at rest. Although the parameter is already expressed as discrete value (i.e. a natural number), medical doctors will usually not distinguish among, say 68 or 72 beats per minute, and classify it as normal. On the other hand, 48 to 56 beats per second is considered low, while 120 to 140 beats will be very fast.

3.2.2. Semantic Web Based Classification (Ontology): The ontology design process is provided by Gupta 2008 [32] [33] described briefly here:

Design Process of Ontology: The process is summarized as follows:

i. **Expert Analysis/ Domain Analysis:**

First step in ontology design process is to analysis the domain for which we are going to design ontology. The expert will cover the following main issues regarding ontology: Ontology scope and Knowledge source. In our study scope of our geo ontology is to classify a satellite image with maximum accuracy.

ii. **Tool and Languages/ Design Structure*:***

The ontology development tools such as Protégé, SWOOP. .Several ontology languages are available like Resource Description Framework (RD),RDFS, DAML+OIL, OWL. We have made use of RDF /XML [Antoniou 2002] language for geo-ontology construction. Basically the RDF[data model is not different from classic conceptual modeling approaches such as Entity-Relationship or Class diagrams, as it is based upon the idea of making statements about resources, in particular, Web resources, in the form of subject-predicate-object expressions [33].

iii. **Ontology Design/ Creation*:***

Ontology design covers the design of framework for ontology by user, expert and designer to represent knowledge in an efficient way [33].

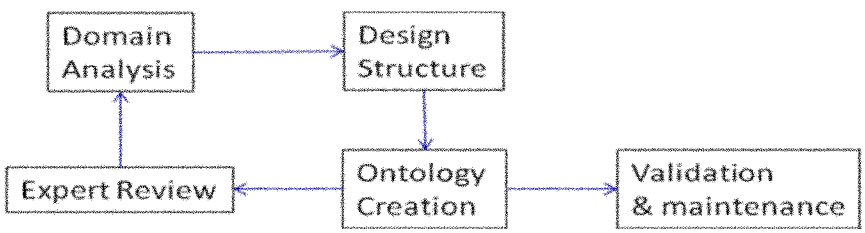


Figure 3.1 : Ontology Design Process

iv. **Validation and Maintenance**: Ontology is checked for validation. It is checked to determine if it is balanced or not.

3.2.3. Swarm Intelligence

Insects that live in colonies, ants, bees, wasps, termites have fascinated naturalists as well as poets for many years. Every single insect in a social insect colony seems to have its own agenda, and yet an insect colony looks so organized. The seamless integration of all individual activities does not seem to require any supervisor. The study of social insects provides us with powerful tool to transfer knowledge about social insects to the field of intelligent system design.

Swarm intelligence is any attempt to design algorithm or distributed problem-solving devices inspired by collective behavior of social insect colonies or other animal societies. This field deals with how social insects collectively perform some specific tasks in daily life (that include finding food, building or extending nest, efficiently feeding the brood, responding to external challenges, spreading alarm etc.), modeling their behavior, and using the model as a basis upon which artificial variations can be developed, either by tuning the model parameters beyond the biological relevant range or by adding non-biological features to the model. The social insect colony is a distributed system comprising of direct or indirect interactions among relatively simple (social) agents that can solve the problems in a very flexible and robust way: flexibility allow adaptation to changing environment, while robustness endows the colony with the ability to function even though some individuals may fail to perform their tasks [17].

3.2.3.1. Self-Organization in Social Insects

How do simple (unsophisticated) agents cooperate with each other to result in coherent functional global patterns to emerge is answered by theory of self-organization. Self-organization is a set of dynamical mechanisms whereby structures appear at the global level of a system from interactions among its lower-level components. The rules specifying the interactions among the system's constituent units are executed on the basis of purely local information, without reference to the global pattern, which is an emergent property of the system rather than a property imposed upon the system by an external ordering influence. Self-organization relies on four basic ingredients:

(a) Positive Feedback: They are simple "rules of thumb" that promote the creation of structure. For example, recruitment to a food source is a positive feedback that relies on trial laying and trial following in some ant species, or dancing in bees.

(b) Negative Feedback: Negative feedback counterbalances positive feedback and helps to stabilize the collective pattern: it may be the form of saturation, exhaustion, or competition.

(c) Amplification of fluctuations: Not only do structures emerge despite randomness, but randomness is often crucial, since it enables the discovery of new solutions, and fluctuations can act as seeds from which structures nucleate and grow.

(d) Multiple Interactions: Self-organization in social insects often requires interactions among insects: such interactions can be direct or indirect. Direct interactions are the "obvious" interactions: antennation, trophallaxis (food or liquid exchange), mandibular contact, visual contact, chemical contact (the odor of nearby nestmates), etc. Indirect interactions are more subtle: two individuals interact indirectly when one of them modifies the environment and the other responds to the new environment at a later time. Such an interaction where the environment serves as a medium of communication is called stigmergy. In addition to, or in combination with, self-organization, stigmergy is the other most important concept of swarm intelligence.

3.2.3.2. Swarm Intelligence Techniques

The soft computing techniques like the fuzzy classifier and the rough set classifier which have been used recently are not able to provide good result in case of ambiguity and also result in inaccuracy with low spatial resolution. Also these are not able to handle the continuous and the crisp data separately. Hence in our proposed work we base our fundamentals from Swarm Intelligence, an optimized approach of image classification of satellite multi-spectral images. Swarm Intelligence provides a good number of accuracy even with low spatial resolution image. This technique, with lower cost and higher degree of classification accuracy, will be able to replace high resolution high cost satellite imageries.

The various swarm intelligence techniques are described below-

(a) Ant Colony Optimization: Ant Colony Optimization is a new computational paradigm, an analogy with the way ant colonies function .The various algorithm proposed in this field are inspired by the brood sorting, division of labor, cemetery organization, food foraging behavior [11] [14].

(b) Particle Swarm Optimization: Particle Swarm Optimization is a population-based stochastic optimization algorithm based on the simulation of the social behavior of birds within a flock. In PSO, the potential solutions are modeled by particles in multidimensional space that have a position and a velocity. These particles fly through hyperspace and have two essential reasoning capabilities: their memory of their own best position (pbest) and knowledge of the global or their neighborhood's best (gbest). Particles communicate good positions to each other and adjust their own position and velocity based on these good positions [20].

(c) Hybrid ACO2/PSO Optimization: Unlike a conventional PSO the hybrid PSO-ACO algorithm can directly cope with the nominal attributes, without converting nominal values into numbers in a pre-processing phase. The hybrid PSO-ACO given by Nicholas and Frietas [12] uses sequential covering approach for rule extraction. After that they also proposed a new modified version PSO-ACO2 directly deals with both the continuous and nominal attribute-values.

(d) Bee Colony Optimization*:* Bee Colony Optimization is a population-based search algorithm that mimics the food foraging behavior of swarms of honey bees. The Algorithm is based on the dancing behavior of honey bees while moving from the nectar source to their hives [34].

(e) Bacterial Foraging Optimization Algorithm: The Bacteria Foraging Optimization Algorithm is based upon search and optimal foraging decision making capabilities of the E.Coli bacteria. The coordinates of a bacterium represent an individual solution of the optimization problem. Such a set of trial solutions converges towards the optimal solution following the foraging group dynamics of the bacteria population [35].

(f) Biogeography Based Optimisation: BBO is an optimization technique that does not involve reproduction or the generation of “children” [5]. This clearly distinguishes it from reproductive strategies such as Genetic Algorithms and Evolutionary Strategies.

1) BBO also clearly differs from ACO, because ACO generates a new set of solutions with each iteration. BBO, on the other hand, maintains its set of solutions from one iteration to the next, relying on migration to probabilistically adapt those solutions [5].

2) BBO has the most in common with strategies such as PSO and DE. In those approaches, solutions are maintained from one iteration to the next, but each solution is able to learn from its neighbours [5]

3) BBO can be contrasted with PSO and DE in that BBO solutions are changed directly via migration from other solutions . That is, BBO solutions directly share their attributes with other solutions [5].

4) GA solutions “die” at the end of each generation, while PSO and BBO solutions survive forever [5].

5) PSO solutions are more likely to clump together in similar groups, while GA and BBO solutions do not necessarily have any built-in tendency to cluster [5].

It is these differences between BBO and other population- based optimization methods that prove to be its strength.

Hence we in this thesis we propose a swarm intelligence based classification algorithm which combines the strengths of the ACO2/PSO technique with the Biogeography based optimization technique .This combination gives this classifier artificial intelligence to identify the features efficiently classified by BBO and by ACO2/PSO based on a similarity pattern analysis and therefore we term our classifier as the **Hybrid Bio-Inspired Pattern Analysis Based Intelligent Classifier.**

Chapter 4

A BRIEF REVIEW OF BBO AND HYBRID PSO/ACO2 OPTIMIZATION

4.1 Biogeography

The Science of biogeography can be traced to the work of nineteenth century naturalists such as Alfred Wallace and Charles Darwin [6]. Until the 1960s, biogeography was mainly descriptive and historical. In the early 1960s, Robert Macarthur and Edward Wilson began working together on mathematical models of biogeography, their work culminating with the classic 1967 publication ‘The Theory of Island Biogeography’. Their interest was primarily focused on the distribution of species among neighboring islands. They were interested in mathematical models for the extinction and migration of species. Since Macarthur and Wilson’s work [7], biogeography has become a major area of research. The application of biogeography to engineering is similar to what has occurred in the past few decades with genetic algorithms, neural networks, fuzzy logic, particle swarm optimization, and other areas of computer intelligence.

Mathematical models of biogeography [5] describe how species migrate from one island to another, how new species arise, and how species become extinct. The term “island” here is used descriptively rather than literally. That is, an island is any habitat that is geographically isolated from other habitats. Geographical areas that are well suited as residences for biological species are said to have a high habitat suitability index [5]. Features that correlate with HSI include such factors as rainfall, diversity of vegetation, diversity of topographic features, land area, and temperature. The variables that characterize habitability are called suitability index variables. SIVs can be considered as the independent variables of the habitat, and HSI can be considered as the dependent variable.

Habitats with a high HSI tend to have a large number of species, while those with a low HSI have a small number of species [5]. Habitats with a high HSI have many species that emigrate to nearby habitats, simply by virtue of the large number of species that they host. Habitats with a high HSI have a low species immigration rate because they are already nearly saturated with species. Therefore, high HSI habitats are more static in their species distribution than low HSI habitats. By the same token, high HSI habitats have a high emigration rate; the large number of species on high HSI islands has many opportunities to emigrate to neighboring habitats. (This does not mean that an emigrating species completely disappears from its home habitat; only a few representatives emigrate, so an emigrating species remains extant in its home habitat, while at the same time migrating to a neighboring habitat). Habitats with a low HSI have a high species immigration rate because of their sparse populations. This immigration of new species to low HSI habitats may raise the HSI of the habitat, because the suitability of a habitat is proportional to its biological diversity. However if a habitat's HSI remains low, then the species that reside there will tend to go extinct, which will further open the way for additional immigration. Due to this, low HSI habitats are more dynamic in their species distribution than high HSI habitats [5]. Figure 4.1 illustrates a model of species abundance in a single habitat. The immigration rate λ and the emigration rate μ are functions of the number of species in the habitat.

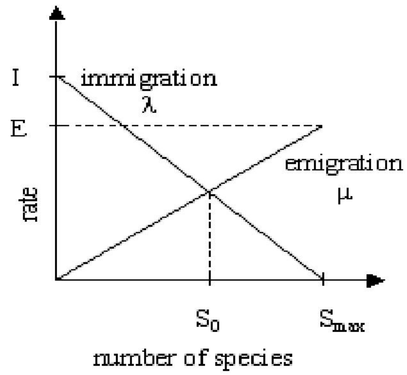


Figure 4.1: Species model of a single habitat

Consider the immigration curve. The maximum possible immigration rate to the habitat is *I*, which occurs when there are zero species in the habitat. As the number of species increases, the habitat becomes more crowded, fewer species are able to successfully survive immigration to the habitat, and the immigration rate decreases. The largest possible number of species that the habitat can support is S_{max}, at which point the immigration rate becomes zero.

Now consider the emigration curve. If there are no species in the habitat then the emigration rate must be zero. As the number of species increases, the habitat becomes more crowded, more species are able to leave the habitat to explore other possible residences, and the emigration rate increases. The maximum emigration rate is *E*, which occurs when the habitat contains the largest number of species that it can support.

The equilibrium number of species is S_0, at which point the immigration and emigration rates are equal. However, there may be occasional excursions from S_0 due to temporal effects. Positive excursions could be due to a sudden spurt of immigration (caused, perhaps, by an unusually large piece of flotsam arriving from a neighboring habitat), or a sudden burst of speciation (like a miniature Cambrian explosion). Negative excursions from S_0 could be due to disease, the introduction of an especially ravenous predator, or some other natural catastrophe. It can take a long time in nature for species counts to reach equilibrium after a major perturbation.

The immigration and emigration curves in Fig. 4.1 are shown as straight lines but, in general, they might be more complicated curves. Nevertheless, this simple model gives us a general description of the process of immigration and emigration. The details can be adjusted if needed.

4.2 Biogeography Based Optimization Algorithm

Biogeography is nature's way of distributing species, and is analogous to general problem solutions. Suppose that we are presented with a problem and some candidate solutions. The problem can be in any area of life (engineering, economics, medicine, business, urban planning, sports, etc.), as long as we have a quantifiable measure of the suitability of a given solution. A good solution is analogous to an island with a high HSI [5] [1], and a poor solution represents an island with a low HSI. High HSI solutions resist change more than low HSI solutions. By the same token, high HSI solutions tend to share their features with low HSI solutions. (This does not mean that the features disappear from the high HSI solution; the shared features remain in the high HSI solutions, while at the same time appearing as new features in the low HSI solutions. This is similar to representatives of a species migrating to a habitat, while other representatives remain in their original habitat). Poor solutions accept a lot of new features from good solutions. This addition of new features to low HSI solutions may raise the quality of those solutions. This new approach to problem solving is called biogeography-based optimization [5].

The biogeography theory of the previous section can be applied to optimization problems with a discrete domain. The theory of biogeography can be formalized in the following way:

(a) Migration: Suppose that we have a problem and a population of candidate solutions that can be represented as vectors of integers. Each integer in the solution vector is considered to be an SIV. Further suppose that we have some way of assessing the goodness of the solutions. Those solutions that are good are considered to be habitats with a high HSI, and those that are poor are considered to be habitats with a low HSI. HSI is analogous to "fitness" in other population-based optimization algorithms (GAs, for example). High HSI solutions represent habitats with many species, and low HSI solutions represent habitats with few species. It is assumed that each solution (habitat) has an identical species curve (with $E= I$ for simplicity), but the value represented by the solution depends on its HSI. S_1 in Fig.4.2 represents a low HSI solution, while S_2 represents a high HSI solution. S_1 in Fig.4.2 represents a habitat with only a few species, while S_2 represents a habitat with many species. The immigration rate λ_1 for S_1 will, therefore, be higher than the immigration rate λ_2 for S_2. The emigration rate μ_1 for S_1 will be lower than the emigration rate μ_2 for S_2 [5] [1].

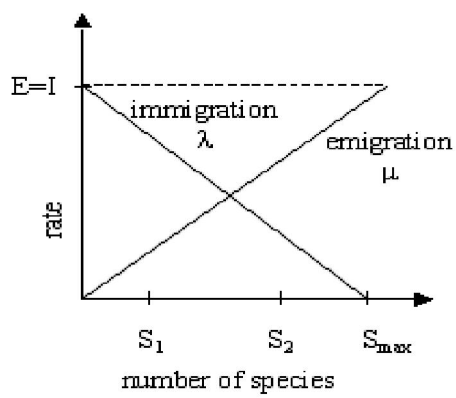


Figure 4.2. Illustration of two candidate solutions to some problem. *S*1 is a relatively poor solution while *S*2 is a relatively good solution.

The emigration and immigration rates of each solution are used to probabilistically share information between habitats [2] [5]. With probability P_{mod}, each solution is modified based on other solutions. If a given solution is selected to be modified, then its immigration rate λ is used to probabilistically decide whether or not to modify each suitability index variable in that solution. If a given SIV in a given solution S_i is selected to be modified, then the emigration rates μ of the other solutions is used to probabilistically decide which of the solutions should migrate a randomly selected SIV to solution S_i [2].

(b) Mutation: Cataclysmic events can drastically change the HSI of a natural habitat. They can also cause a species count to differ from its equilibrium value (unusually large flotsam arriving from a neighboring habitat, disease, natural catastrophes, etc.). A habitat's HSI can, therefore, change suddenly due to apparently random events. This is modeled in BBO as SIV mutation [5], and species count probabilities are used to determine mutation rates.

Each population member has an associated probability, which indicates the likelihood that it was expected *a priori* to exist as a solution to the given problem. Very high HSI solutions and very low HSI solutions are equally improbable. Medium HSI solutions are relatively probable. If a given solution has a low probability, then it is surprising that it exists as a solution. It is, therefore, likely to mutate to some other solution. Conversely, a solution with a high probability is less likely to mutate to a different solution. This mutation scheme tends to increase diversity among the population. Without this modification, the highly probable solutions will tend to be more dominant in the population. The mutation approach makes low HSI solutions likely to mutate, which gives them a chance of improving. It also makes high HSI solutions likely to mutate, which gives them a chance of improving even more than they already have. So, mutation (a high risk process) is used on both poor solutions and good solutions. Those solutions that are average are hopefully improving already, and so mutating them is avoided (although there is still some mutation probability, except for the most probable solution).

The steady state value for the probability of the number of each species to exist is given by Ma and Chen, 2009 [4] as:

$$P_k = \begin{cases} P_0 = \dfrac{1}{1+\sum_{k=1}^{n} \dfrac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k}}, & k=0 \\ P_k = \dfrac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k\left(1+\sum_{k=1}^{n} \dfrac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k}\right)}, & 1\le k\le n \end{cases} \quad \text{... (4.1)}$$

The largest possible number of species that the habitat can support is *n*. It is necessary that $\mu_k \neq 0$ for all 'k' for this limiting probabilities to exist. This condition is also sufficient.

(c) BBO Definitions and Algorithm: In order to formalize the BBO Algorithm, some definitions are provided as follows. We use R to refer to the set of real numbers, Z to refer to the set of integers, and Φ to refer to the empty set.

Definition I: A habitat $H \in SIV^m$ is a vector of *m* integers that represents a feasible solution to some problem [5].

Definition II: A suitability index variable $SIV \in C$ is an integer that is allowed in a habitat. C (is subset of Z^q) is the set of all integers that are allowed in a habitat [5].

Definition III: A Habitat Suitability Index $HSI: H \rightarrow R$ is a measure of the goodness of the solution that is represented by the habitat [5].

Definition IV: An ecosystem H^n is a group of n habitats. The size n of an ecosystem is constant. Future work could allow variable-sized ecosystems, just as some flavors of GAs allow for variable population sizes [5].

Definition V: Immigration rate λ *(HSI):R→R* is a monotonically nonincreasing function of HSI. λ_i is proportional to the likelihood that SIVs from neighboring habitats will migrate into habitat H_i and for linear immigration and emigration model it can be given by the following formula [5]:

$$\lambda_k = \mathrm{I}\left(1 - \frac{k}{n}\right) \qquad \text{… (4.2)}$$

Definition VI: Emigration rate μ *(HSI): R→R* is a monotonically nondecreasing function of HSI. μ_i is proportional to the likelihood that SIVs from habitat H_i will migrate into neighboring habitats. For linear immigration and emigration model it can be given by the following formula [5]:

$$\mu_k = \frac{\mathrm{E}k}{n} \qquad \text{… (4.3)}$$

In practice, we assume that λ and μ are linear with the same maximum values. However, these assumptions are made only for mathematical convenience, and better performance might be attainable if these assumptions are relaxed.

Definition VII: Habitat modification Ω (λ, μ): $H^n \rightarrow H$ is a probabilistic operator that adjusts habitat H based on the ecosystem H^n. The probability that H is modified is proportional to its immigration rate λ, and the probability that the source of the modification comes from H_j is proportional to the emigration rate μ_j [5].

Habitat modification can loosely be described as following algorithm. From this algorithm we note that elitism can be implemented by setting $\lambda = 0$ for the p best habitats, where p is a user-selected elitism parameter.

Select H_i with probability α λ_i

If H_i is selected

For $j = 1$ to n

Select H_j with probability α μ_j

If H_j is selected

Randomly select an SIV σ from H_j

Replace a random SIV in H_i with σ

end

end

end

Definition VIII: *Mutation M* (λ, μ): H → H is a probabilistic operator that randomly modifies habitat SIVs based on the habitat's a priori probability of existence. A habitat's probability of existence is computed from λ and μ as shown in formula (4.1). For classic BBO, the mutation rate m is inversely proportional to the solution probability, which is defined by:

$$m = m_{\max}\left(1 - \frac{P}{P_{\max}}\right) \quad \ldots (4.4)$$

where m_{max} is a user-defined parameter.

Mutation can be described as follows [4]:

Use λ_i and μ_i to compute the probability Pi

Compute mutation probability m_i

Select SIV s in H_i with probability based on m_i

If H_i(SIV) is selected

Replace H_i (SIV) with a randomly generated SIV s

End

As with habitat modification, elitism can be implemented by setting the probability of mutation selection P_i to zero for the p best habitats. From the above definition we see that mutation must be constrained to result in an HSI that satisfies the SIV constraints.

Definition IX: An ecosystem transition function $\Psi = (m, n, \lambda, \mu, \Omega, M): H^n \rightarrow H^n$ is a 6-tuple that modifies the ecosystem from one optimization iteration to the next. An ecosystem transition function can be written as follows [5].

$$\Psi = \lambda^n \circ \mu^n \circ \Omega^n \circ HSI^n \circ M^n \circ HSI^n \qquad \ldots (4.5)$$

In other words, the ecosystem transition function begins by computing the immigration and emigration rates of each habitat. Then habitat modification is performed on each habitat, followed by an HSI recalculation. Finally mutation is performed, followed again by an HSI recalculation for each habitat.

Definition X: *A* biogeography-based optimization algorithm BBO= (I, Ψ, T) is a 3-tuple that proposes a solution to an optimization problem. $I: \Phi \rightarrow \{H^n, HSI^n\}$ is a function that creates an initial ecosystem of habitats and computes each corresponding HSI. Ψ is the ecosystem transition function defined earlier, and $T: H^n \rightarrow \{true, false\}$ is a termination criterion. I could be implemented with random number generators, heuristic solutions to the optimization problem, or some other problem-dependent procedure. T could depend on the number of Ψ iterations, or the HSI of the best habitat, or some other problem-dependent criterion [5].

A BBO algorithm can be described as follows:

$I: \Phi \rightarrow \{H^n, HSI^n\}$

While not T *(where* $T{:}H^n \rightarrow \{true, false\}$*)*

$\Psi = \lambda^n \circ \mu^n \circ \Omega^n \circ \mathrm{HSI}^n \circ \mathrm{M}^n \circ \mathrm{HSI}^n$

End

The BBO algorithm [5] can be informally described with the following algorithm.

(i) Initialize the BBO parameters. This means deriving a method of mapping problem solutions to SIVs and habitats (see Definitions I and II), which is problem dependent. We also initialize the maximum species count $S_{\max}$ and the maximum migration rates E and I (as shown in fig. 4.2) and an elitism parameter. Note that the maximum species count and the maximum migration rates are relative quantities. That is, if they all change by the same percentage, then the behavior of BBO will not change. This is because if E, I, and $S_{\max}$ change, then the migration rates μ, λ, and the species count S will change by the same relative amount for each solution

(ii) Initialize a random set of habitats, each habitat corresponding to a potential solution to the given problem. This is the implementation of the I operator described in Definition X.

(iii) For each habitat, map the HSI to the number of species S, the immigration rate λ, and the emigration rate μ (see Figure 4.2 and Definitions V and VI).

(iv) Probabilistically use immigration and emigration to modify each non-elite habitat as discussed above, and then recompute each HSI (see Definition VII).

(v) For each habitat, update the probability of its species count using (4.1). Then mutate each non-elite habitat based on its probability, and recompute each HSI (see Definition VIII).

(vi) Go to step (iii) for the next iteration. This loop can be terminated after a predefined number of generations or after an acceptable problem solution has been found. This is the implementation of the *T* operator described in Definition X.

Note that after each habitat is modified (steps (ii), (iv), and (v)), its feasibility as a problem solution should be verified. If it does not represent a feasible solution, then some method needs to be implemented in order to map it to the set of feasible solutions.

4.3 Differences between BBO and Other Population-Based Optimization Algorithms

3) BBO does not involve reproduction or the generation of "children" [5]. This clearly distinguishes it from reproductive strategies such as GAs and evolutionary strategies.
4) BBO also clearly differs from ACO, because ACO generates a new set of solutions with each iteration. BBO, on the other hand, maintains its set of solutions from one iteration to the next, relying on migration to probabilistically adapt those solutions [5].
5) BBO has the most in common with strategies such as PSO and DE. In those approaches, solutions are maintained from one iteration to the next, but each solution is able to learn from its neighbours [5].
   - PSO represents each solution as a point in space, and represents the change over time of each solution as a velocity vector. However, PSO solutions do not change directly.
   - DE changes its solutions directly, but changes in a particular DE solution are based on differences between other DE solutions.
   - Also, DE is not biologically motivated.

4) BBO can be contrasted with PSO and DE in that BBO solutions are changed directly via migration from other solutions . That is, BBO solutions directly share their attributes with other solutions [5].

5) BBO has certain features in common with other biology based algorithms. Like GAs and PSO, BBO has away of sharing information between solutions [5].

6) GA solutions "die" at the end of each generation, while PSO and BBO solutions survive forever [5].

7) PSO solutions are more likely to clump together in similar groups, while GA and BBO solutions do not necessarily have any built-in tendency to cluster [5].

It is these differences between BBO and other population- based optimization methods that may prove to be its strength.

4.4 Performance and Computation Time Comparison

Dan Simon compared the performance of BBO on fourteen benchmark functions with seven other population - based methods [5]. These population - based methods were Ant Colony Optimization , Differential Evolution , Evolutionary Strategy, Genetic Algorithm, Probability-based Incremental Learning, Particle Swarm Optimization and Stud Genetic Algorithm. The benchmarks that were used were functions that are representative of those used in literature for comparison of optimization methods. These benchmark functions were Ackley, Fletcher-Powell, Griewank, Penalty #1, Penalty #2, Quartic, Rastrigin, Rosenbrock, Schewefel 1.2, Schewefel 2.21, Schewefel 2.22, Schewefel 2.26, Sphere, and Step. Simon noticed the average performance of each algorithm on these function and found that BBO and SGA both performed the best (on average) on seven of the 14 benchmarks. Also, when the best performance of each algorithm was noticed, SGA was the most effective at finding function minima when multiple runs are made, performing the best on seven of the 14 benchmarks. BBO was the second most effective, performing the best on four of the benchmarks, while ACO performed the best on three of the benchmarks.

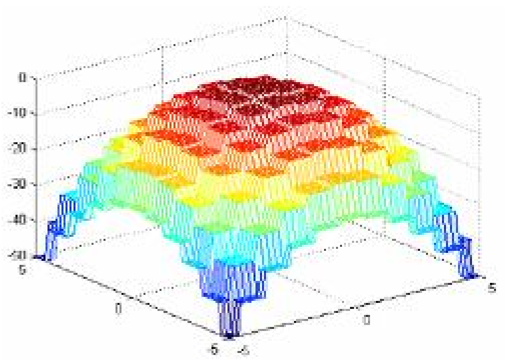
The computational time requirements of the eight optimization methods were similar. The average computational time of the optimization methods was collected as applied to the 14 benchmarks discussed in this section. PBIL was the quickest optimization method. BBO was the fifth fastest of the eight algorithms. However, it should be noted that in the vast majority of real-world applications, it is the fitness function evaluation that is by far the most expensive part of a population-based optimization algorithm. In spite of the caveats, the benchmark results shown are promising for BBO, and indicate that this new paradigm might be able to find a niche among the plethora of population-based optimization algorithms.

Table 4.1. Performance Analysis results of BBO over 100 Monte Carlo Simulations

	ACO	*BBO*	*DE*	*ES*	*GA*	*PBIL*	*PSO*	*SGA*
Acklay	182	**100**	146	197	197	232	192	103
Fletcher	1013	**100**	385	494	415	917	799	114
Griewank	162	117	272	696	516	2831	1023	**100**
Penalty 1	2.2E7	1.2E4	9.7E4	1.3E6	2.5E5	2.8E7	2.1E6	**100**
Penalty 2	5.0E5	715	5862	4.2E4	1.1E4	5.4E5	6.4E4	**100**
Quartic	3213	262	1176	7008	2850	4.8E4	8570	**100**
Rastrigin	454	**100**	397	536	421	634	470	134
Rosenbrock	1711	102	253	716	428	1861	516	**100**
Schwefel 1.2	202	**100**	391	425	166	606	592	110
Schwefel 2.21	161	**100**	227	162	184	265	179	146
Schwefel 2.22	688	**100**	290	1094	500	861	665	142
Schwefel 2.26	108	118	137	140	142	177	142	**100**
Sphere	1347	**100**	250	910	906	2785	1000	109
Step	248	112	302	813	551	3271	1161	**100**

Each algorithm had a population size of 50, an elitism parameter of 2 and ran for 50 generations. We ran 100 Monte Carlo simulations of each algorithm on each benchmark to get representative performances. Table 4.1 above shows the average performance of BBO over 100 Monte Carlo Simulations .

4.4.1 Performance Analysis Results of BBO as compared to the other recent Swarm Intelligence Techniques for Step Function Optimization viewed graphically



$$f(x) = \sum_{i=1}^{n} [\text{floor}(x_i + 0.5)]^2$$

Step Function : separable, irregular, unimodal

Fig 4.3. Step Function

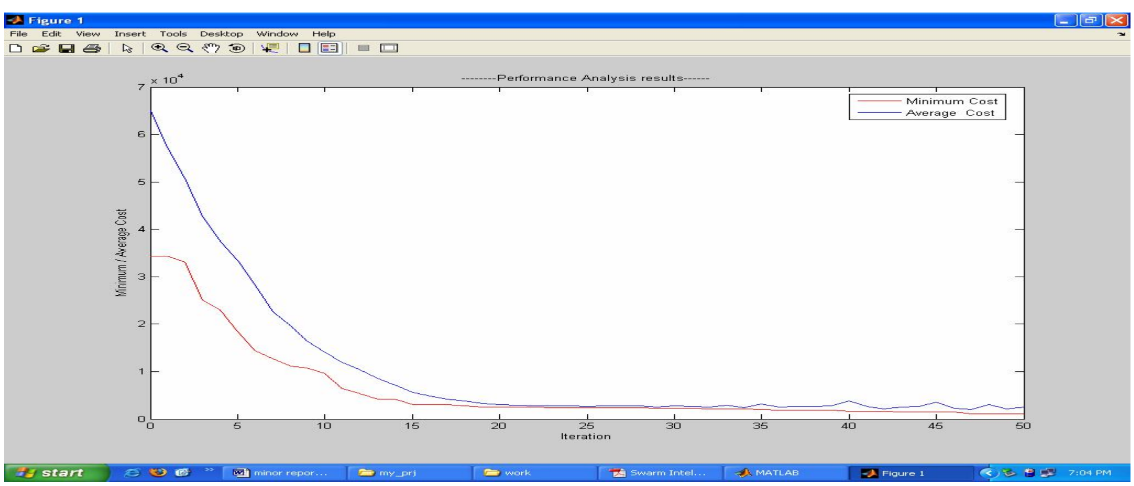


Fig. 4.4. Biogeography Based Optimization

From the graph, the following information can be observed.

- The best and mean of Iteration # 50 are 596 and 1995.76
- 0 duplicates in final population.
- 50 legal individuals in final population.
- Best solution = 906 944 945 960 965 967 970 971 973 979 1001 1005 1006 1010 1011 1026 1062 1084 1114 1119
- Diversity measure = 9072

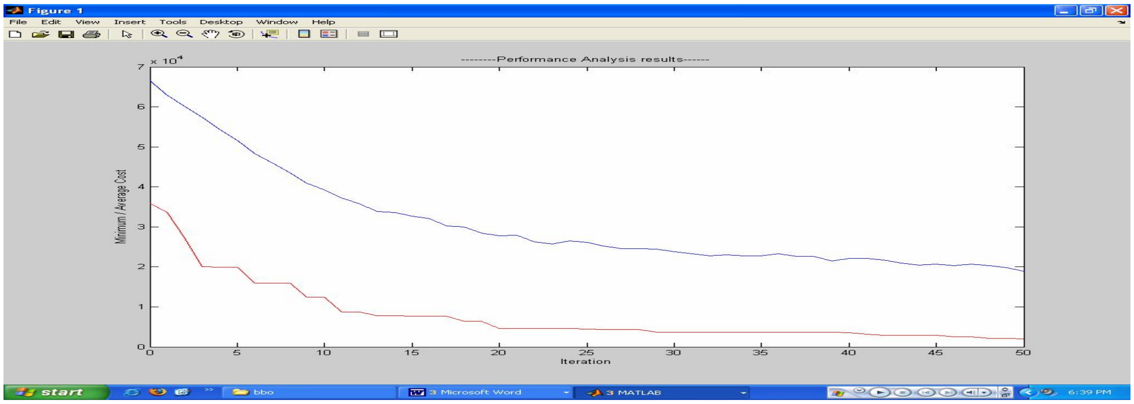


Fig 4.5. Ant Colony Optimization

From the graph, the following information can be observed.

- The best and mean of Iteration # 50 are 1254 and 15008.94
- 0 duplicates in final population.
- 50 legal individuals in final population.
- Best solution = 873 931 938 941 946 951 952 996 999 1014 1017 1026 1036 1044 1064 1090 1116 1125 1151 1154

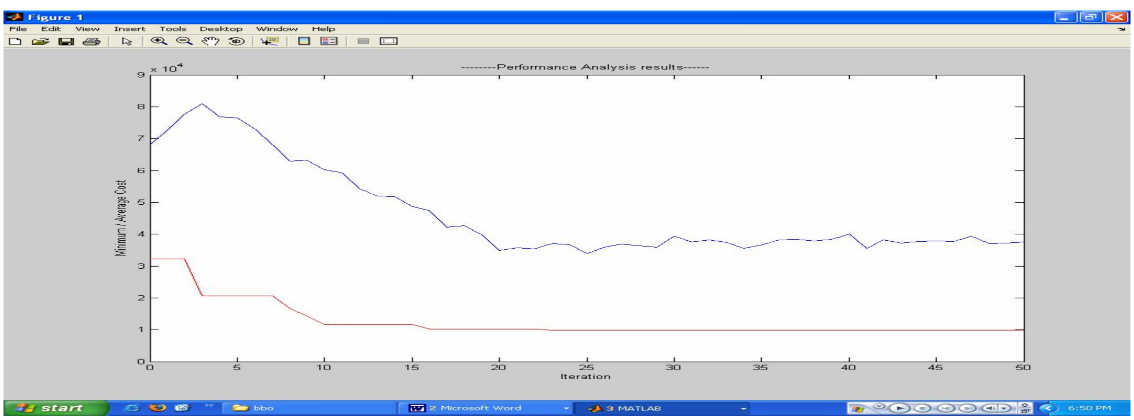


Fig 4.6 Particle Swarm Optimization

From the graph, the following information can be observed.

- The best and mean of Generation # 50 are 10498 and 39523.04
- 0 duplicates in final population.
- 50 legal individuals in final population.
- Best solution = 553.76629 621.09703 659.58609 692.40551 743.39938 833.62197 867.93232 894.72014 931.3915 952.70316 976.45594 1007.036 1018.436 1049.2832 1170.1313 1200.3639 1206.2147 1270.5336 1271.6669 1312.7829

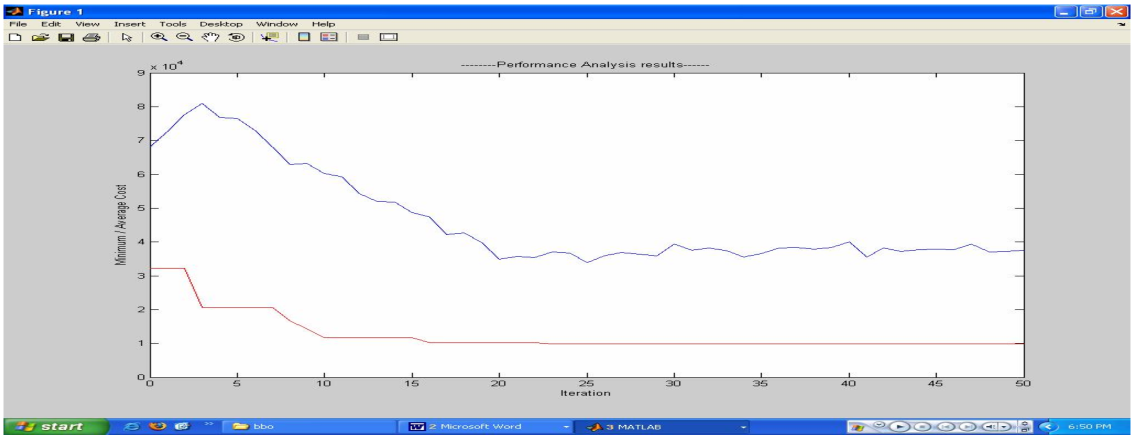


Fig 4.7. Genetic Algorithm

From the graph, the following information can be observed.

- The best and mean of Generation # 50 are 5457 and 9257.08
- 0 duplicates in final population.
- 50 legal individuals in final population.
- Best solution = 724 816 828 888 907 907 908 919 922 930 930 954 960 983 1015 1090 1175 1230 1337 1370

From the above results, it is clear that the performance of BBO is better than the other techniques discussed since BBO required the minimum time in finding the best and the average minimal point for the Step function Optimization compared to the other algorithms and hence it is an optimal algorithm.

4.4.2 Performance of BBO for the Turbofan Sensor Selection Problem

Table 4.2 below shows the results of the optimization methods on the sensor selection problem [5]. We see that BBO performs the best in terms of both average performance and best performance. We see that the performances of the methods are comparable, but BBO with probability-based mutation is clearly better than BBO without mutation. Mutation can be detrimental for large population sizes, but with small population sizes mutation helps increase diversity and increases the changes for a good solution.

These simulation results should not be taken to mean that BBO is “better” than other population-based optimization algorithms. Such a general statement would be an over-simplification, especially in view of the no free lunch theorem . However, the results presented here show that BBO provides better performance than most of the other algorithms we tested for the particular benchmarks that we examined. The results shown here indicate that BBO is at least competitive with other population-based optimization algorithms, and could provide a valuable tool for practical problems.

Table 4.2. Performance of BBO for the Turbo Fan Selection Problem

	ACO	BBO	DE	ES	GA	PBIL	PSO	SGA
Mean Minimum	8.22	**8.01**	8.06	8.15	8.04	8.18	8.14	8.02
Best Minimum	8.12	**7.19**	7.60	8.05	8.02	8.08	8.06	8.02

4.5 An Integration of BBO with ACO in Military Applications

This section presents a hybrid ACO-BBO for predicting the deployment strategies of enemy troops in the war theatre and finding the shortest and the best feasible path for attack on the enemy base station. The hybrid algorithm begins by predicting the most suitable destination for the enemy troops to position their forces, for which it uses the Population Based Optimization Technique i.e. Biogeography Based Optimization and thereafter finding the shortest and the best feasible path for attacking the enemy base station using the Ant Colony Optimization Technique, thus combining the strengths of both the techniques. Hence, the algorithm can be used to improve ACO approach which is currently used to predict the enemy troop mobility since it lacks the ability to predict the destination and can only find a suitable path to the given destination ,leading to coordination problems and target misidentification which can lead to severe casualties. The algorithm can be of major use for the commanders in the battlefield who have been using traditional decision making techniques of limited accuracy for predicting the destination. Using the Hybrid ACO-BBO technique can help in enabling the commanders in the battlefield for intelligent preparation of the battlefield by automating the process of assessing the likely base stations of the enemy and the ways in which these can be attacked , given the environment and the terrain considerations.

4.5.1 Modified Approach to Biogeography Based Optimization

Biogeography is the study of geographic distribution of biological organisms and species. It is analogous to general problem solutions. BBO is an application of biogeography to optimization problems. It is modeled after the emigration and immigration of species between habitats. In the original BBO, the author [5] has described the migration of the species from habitats with high HSI to the habitats with low HSI. In our proposed work, we consider a collection of ideal SIVs as the feature habitat against with which we compare our candidate base stations which are put in the universal habitat. We use HSI for comparison between the two solutions. HSI is the measure of similarity threshold between the habitats. We also consider an isolated habitat which consists of zero species i.e. this habitat has no solution class . SIVs are the set of features possessed by the habitat which is an indication of the suitability of the terrain for the enemy troops to deploy their forces. After each iteration , the candidate base station's HSI is determined. After the completion of all the iterations, we are left with the base station with the maximum HSI which corresponds to the maximum collection of SIVs from the feature Habitat . At the end of the algorithm, we calculate the shortest path to this Base station for attacking the base station of the enemy predicted by our modified Algorithm.The complete framework for the above discussed hybrid ACO-BBO algorithm is represented in the figure 4.8 below.

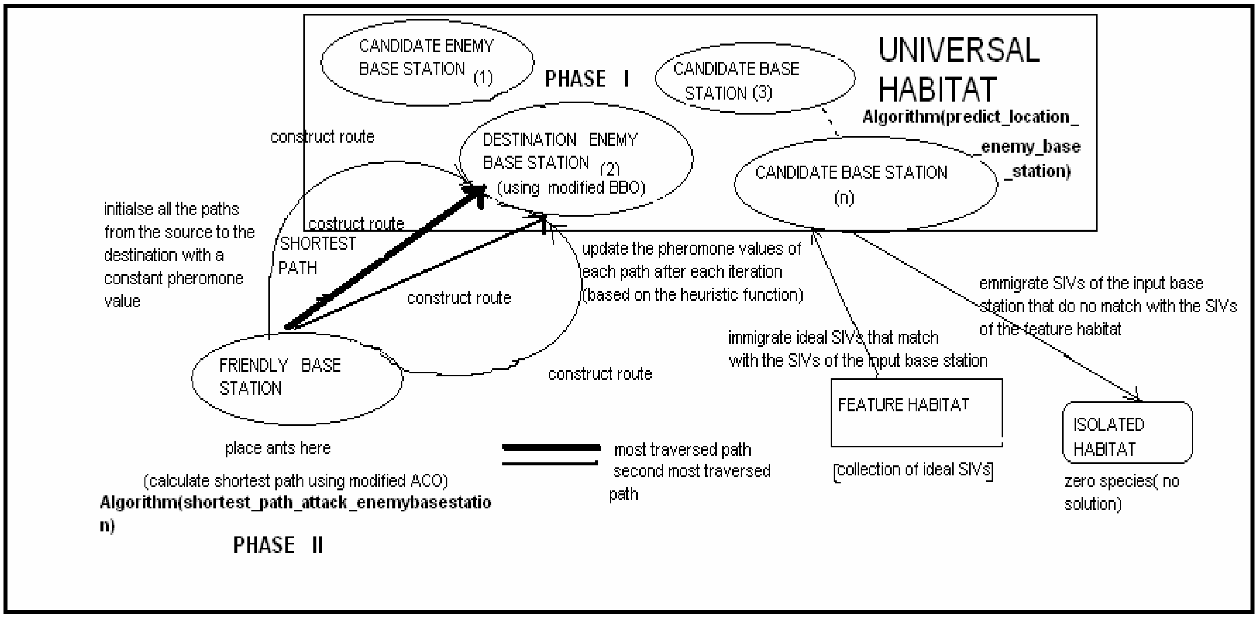


Figure 4.8. Overall Framework of the Hybrid ACO-BBO Algorithm

4.6. Hybrid ACO2/PSO Optimization

Parpinelli, Lopes and Freitas [15] were the first to propose Ant Colony Optimization for discovering classification rules, with the system *Ant-Miner*. They find out that an ant-based search is more flexible, robust and optimized than traditional approaches. Their method uses a heuristic value based on entropy measure. The goal of Ant-Miner is to extract classification rules from data [13] . Ant Miner follows a sequential covering approach to discover a list of classification rules from the given data set. It covers all or almost all the training cases. Each classification rule has the form IF <term1 AND term2 AND…> Then <CLASS>. Ant miner requires the discrimination method as a pre-processing method and it is suitable only for the nominal attributes. Mostly real-world classification problems are described by nominal or discrete values and continuous attributes. There is a limitation with Ant-Miner that it is able to cope only with nominal attributes in its rule construction process . So that discretization of continuous attributes is done in a preprocessing step. Fernando, Freitas, and Johnson proposed an extension to Ant-Miner, named cAntMiner, which was able to cope with the continuous values as well [9].

The Ant Miner and cAntMiner has already been a significant approach for data mining, but an extremely large amount of computation is required with the problem of unusually large amount of attributes and classes. The “standard” binary/discrete PSO algorithm does not deal with categorical values in a natural fashion when compared to ACO. In particular, the standard PSO for copng with binary attributes represents a particle by a bit string, where each binary value such as true or false is encoded as 1 or 0. Sousa et al. extended the standard binary PSO to cope with multi-valued categorical attributes, developing a Discrete PSO algorithm for discovering classification rules [16].

Unlike a conventional PSO the hybrid PSO-ACO algorithm can directly cope with the nominal attributes, without converting nominal values into numbers in a pre-processing phase. The hybrid PSO-ACO given by Nicholas and Frietas [12] uses sequential covering approach for rule extraction. After that they also proposed a new modified version PSO-ACO2 directly deals with both the continuous and nominal attribute-values.

Both the original PSO/ACO algorithm and the new modified version PSO/ACO2 uses a sequential covering approach to discover one classification-rule-at-a-time [18]. The new version given by Nicholas and Freitas [12] can be understood as follows-

1. Initially Rule_Set is empty(_)

2. For Each class of cases Trs = {All training cases}

3. While (Number of uncovered training cases of class A > Maximum uncovered cases per class)

4. Run the PSO/ACO algorithm for finding best nominal rule

5. Run the standard PSO algorithm to add continuous terms to Rule, and return the best discovered rule Best_Rule

6. Prune the discovered Best_Rule

7. Rule_Set = Rule_Set[Best_Rule]

8. Trs = Trs – {training cases correctly covered by discovered rule}

9. End of while loop

10. End of for lop

11. Order these rules in Rule_Set by descending Quality

It is necessary to estimate the quality of every candidate rule (decoded particle). A measure must be used in the training phase in an attempt to estimate how well a rule will perform in the testing phase. Given such a measure it becomes possible to optimize a rule's quality (the fitness function) in the training phase and this is the aim of the PSO/ACO2 algorithm. In PSO/ACO the Quality measure used was Sensitivity * Specificity [18] . Where TP, FN, FP and TN are, respectively, the number of true positives, false negatives, false positives and true negatives associated with the rule.

Sensitivity Specificity = TP / (TP + FN) TN / (TN + FP)

Equation 1: Original Quality Measure

Later it is modified as -

Sensitivity Precision = TP / (TP + F7) TP / (TP + FP)

Equation 2: Quality Measure on Minority Class

This is also modified with using Laplace correction as;

Precision = 1 + TP / (1+ k + TP + FP)

Equation 3: New Quality Measure on Minority Class [4]

Where k is the number of classes.

So, PSO/ACO1 attempted to optimize both the continuous and nominal attributes present in a rule antecedent at the same time, whereas PSO/ACO2 takes the best nominal rule built by PSO/ACO2 and then attempts to add continuous attributes using a standard PSO algorithm.

Chapter 5

FROM BBO TO HYBRID ACO/PSO/BBO FOR LAND COVER FEATURE EXTRACTION

The findings of recent studies are showing strong evidence to the fact that some aspects of biogeography can be adaptively applied to solve specific problems in science and engineering. Our proposed Algorithm combines the strengths of the modified BBO technique for land cover feature extraction with the hybrid ACO2/PSO Technique for a more refined image classification. Biogeography based optimization algorithm is basically used to find the optimal solution of a problem. But land cover feature extraction is a clustering problem that requires each feature to be extracted as a cluster. The original BBO algorithm does not have the inbuilt property of clustering. But here to extract features from the image, we have tried to make the clusters of different land cover features and proposed a hybrid of this modified biogeography based algorithm with the ACO2/PSO technique for extraction of land cover features from the satellite image.

The proposed work presents a hybrid biologically inspired technique, called the ACO2/PSO/BBO (Ant Colony Optimization2/ Particle Swarm Optimization / Biogeography Based Optimization) Technique that can be adapted according to the database of expert knowledge for a more focussed satellite image classification. The hybrid classifier explores the adaptive nature of biogeography based optimization Technique and therefore is flexible enough to classify a particular land cover feature more efficiently than others based on the 7-band data and hence can be adapted according to the application. The proposed hybrid classifier improves the efficiency of the biogeography based classifier for a more efficient, focussed and refined image classification. And, hence we name our classifier as the hybrid bio-inspired pattern anlysis based intelligent classifier. The proposed algorithm has been applied to the 7-Band cartoset satellite image of size 472 X 576 of the Alwar area in Rajasthan since it contains a variety of land cover features.

We also present a comparative study of the results of applying our proposed classifier and the other recent soft computing classifiers such as ACO, Hybrid PSO-ACO2, Hybrid ACO-BBO Classifier, Fuzzy sets, Rough-Fuzzy Tie up and the Semantic Web Based Classifiers when compared with the traditional probabilistic classifiers such as the Minimum Distance to Mean Classifier and the Maximum Likelihood Classifier on the 7-Band Alwar Image.

The proposed framework for the Hybrid ACO2/PSO/BBO is described in the following sections.

5.1. Dataset used

Our objective is to use the proposed hybrid algorithm as an efficient Land cover classifier for satellite image. We have taken a multi-spectral, multi resolution and multi-sensor image of size 472 X 576 of the Alwar area in Rajasthan. The satellite image for 7 different bands is taken. These bands are Red, Green, Near-Infra Red, Middle-Infra Red, Radarsat-1, Radarsat-2, and Digital Elevation Model. The expert has the attribute set *P* from the Indian Remote Sensing (IRS-P6) satellite optical band image set *i.e.* Green (G) ,Red (R), Near-Infrared and Middle-Infra-red bands. The ground resolution of these images is 23.5m and is taken from LISS -III, sensor and is provided to us courtesy of DTRL (Remote sensing) Lab, DRDO, Delhi.

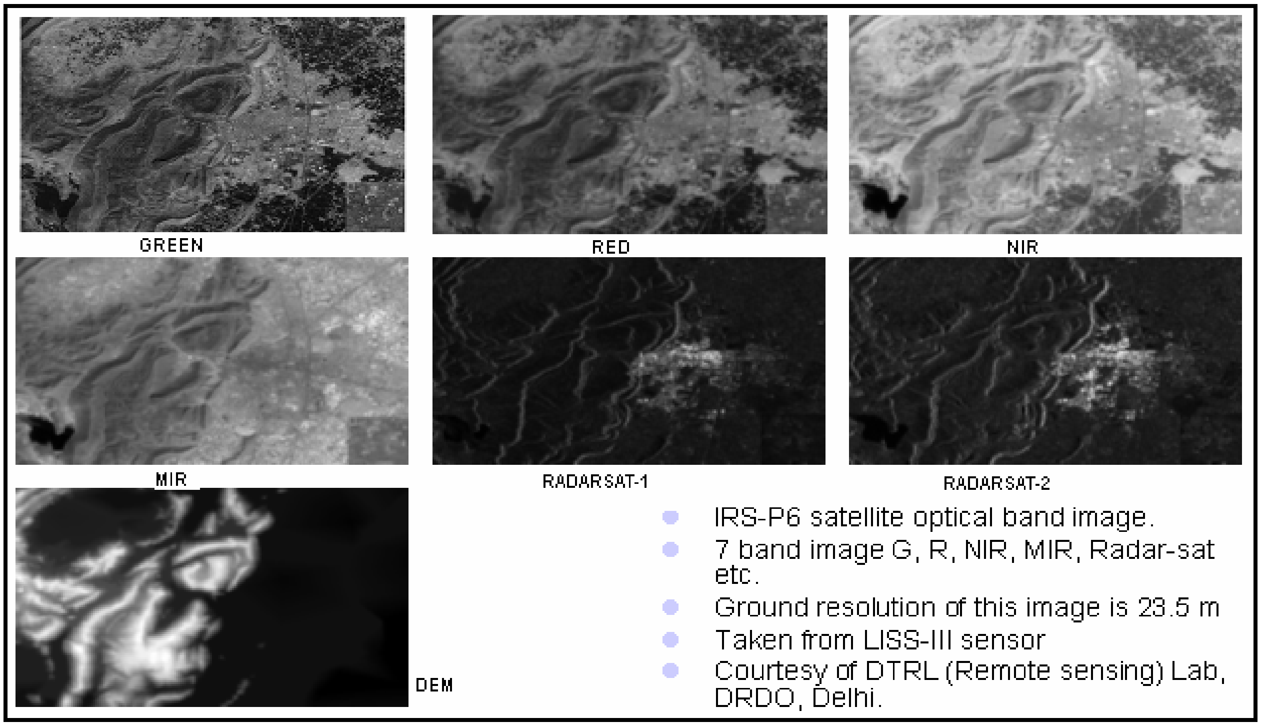


Fig 5.1. 7-Band Satellite Image of Alwar Area in Rajasthan

The land cover classification with independent attributes set x consists of two sets of radarsat microwave images radarsat-1 and radarsat-2 and digital elevation model data. The r1 dataset is from low incidence angle between 20^o-27^o and ground resolution of 10m. The high incidence angle for r2 dataset is between 45^o-49^o in the c-band. The dem i.e. D dataset is also generated from SAR interferometery using r1 and r2 and have 25 meter resolution. These are chosen because the characteristics / observed behavior of the visual bands and of the microwave bands are typically not related.

5.2 Biogeography Based Land Cover Feature Extraction

5.2.1 Defining BBO parameters : The BBO parameters of the Biogeography Based Land Cover Feature Extraction algorithm are defined as follows [3]:

Definition I: In our study, each of the multi-spectral band of image represents one Suitability Index Variable (SIV) of the habitat Further, since image in each band is a gray image, SIV$\in$ C is an integer and C $\subset$[0,255].

Definition II: A habitat H $\in$ SIV^m where m is 7 because we have 7 –band data.

Definition III: Initially there exists a Universal habitat that contains all the species to be migrated. Also we have as many other habitats as the number of features to be extracted from the image. In our image, we want to extract 5 features –water, vegetation, urban, rocky and barren from the image. Each of these features is represented by a habitat. So the ecosystem H^6 is a group of 6 habitats (one Universal habitat and five feature habitat).

Definition IV: Since we are using an image of almost 2.5 lakh pixels, considering each pixel as a species would make the algorithm extremely slow. So we has used rough set theory to obtain the random clusters of pixels(by using discretization and partitioning concept of rough set theory) and each of the resulting cluster will be considered as mixed species that migrate from one habitat to another. These species can also be termed as 'elementary classes' of a habitat.

Definition V: We have used standard deviation of pixels as Habitat Suitability Index to help in feature extraction.

Definition VI: In the BBO algorithm, Simon proposed the migration of SIV values from a high HSI habitat to a low HSI habitat. The shared features remain in the high HSI solutions, while at the same time appearing as new features in the low HSI solutions.

In Biogeography Based Land Cover Feature Extraction algorithm, rather than moving SIV, we are moving species altogether from a Universal habitat to feature habitat. The species does not remain shared: it is removed from the Universal habitat and migrated to feature habitat.

Definition VII: Maximum Immigration rate and Maximum Emigration Rate are same and equal to number of species initially in Universal habitat. The algorithm follows a linear curve (E=I). Number of species and thus the Maximum Immigration Rate and Maximum Emigration rate can vary in each iteration. Maximum species count (S_{max}) and the maximum migration rates are relative quantities. That is, if they all change by the same percentage, then the behavior of BBO will not change .

Definition VIII: Since mutation is not an essential feature of BBO, it is not required in the proposed algorithm. Elitism, too, is an optional parameter; it has not been used in the Biogeography Based Land Cover Feature Extraction algorithm.

So the assumptions [3] made in the work can be summarized as follows:

(a) Initially it is considered that there exists a Universal habitat consisting of all the species.

(b) It has been assumed that species migrate from one habitat to another habitat as a mixed population.

(c) Also for the feature extraction purpose, in each generation, all the habitats have been considered exactly once.

(d) In next generation only the unclassified pixels have been used. However the total population is still present.

(e) Also, we have used NIR and MIR band of the image for partitioning the image (done using rough set theory).

5.2.2. Algorithm for Biogeography Based Land Cover Feature Extraction

The following Biogeography based algorithm can be used to extract land cover features from the satellite image [3].

Input: Multi-spectral satellite image

Output: Extracted features from the image

Algorithm (BiogeographybasedLandCoverFeatureExtraction)

(i) Get the multi-spectral satellite image.

(ii) Cluster the pixels of image randomly (using rough set theory) and consider each cluster as a species of Universal habitat.

(iii) Consider other habitats one for each land cover feature-Water, Urban, Rocky, Barren and Vegetation -having members produced by experts.

(iv) Define HSI, $S_{max,}$ immigration rate (λ) and emigration rate (μ).

(v) Calculate HSI of each of the feature habitat.

(vi) (a) Select a species from the Universal habitat and migrate it to one of the feature habitat.

(b) Recalculate the HSI of feature habitat after the migration of the species to it.

(vii) If the recalculated HSI is within the threshold, then:

(a) Absorb the species in feature habitat.

(b) Go to step (viii)

Else if any unconsidered feature habitat is left then:

(a) Migrate the species to that feature habitat and recalculate the HSI after migration.

(b) Go to step (vii).

Else:

(a) Use rough set theory to discretize the species (as it contained mixed pixels) and make random clusters which are considered as separate species.

(b) Add these new species to Universal habitat.

End If

(viii) If no species is left in Universal habitat then:

Stop the process

Else:

Go to step (vi).

End If

The detailed working of the above algorithm is given as following:

Step (i): A 7-band satellite image of Alwar region is taken as an input image for Biogeography based feature extraction. The image has 257712 pixels and each pixel is of 25m by 25m spatial resolution.

The image is divided into 20 equal parts and each part corresponds to a class. Each pixel in a class has x and y coordinates along with the 7-band values of each pixel. Each class is considered separately and the following steps are applied to each class.

Step (ii): A class is discretized using the Rosetta software based on the discretization Algorithm. The class is then partitioned based on the MIR and NIR bands to obtain the elementary classes (random clusters). These elementary classes are now considered as species and they reside in the Universal habitat.

Step (iii): For our process we have considered five feature habitats namely Urban, Rocky, Vegetation, Water and Barren land along with the Universal habitat. These feature habitats consist of training sets as species. The training set consists of those pixels whose feature is already known. The training sets are provided by the expert.

Step (iv): The HSI is the standard deviation of all the pixels in the elementary class and is calculated on each SIV (band) of the image separately. S_{max} is the maximum number of species and its value varies in each iteration. Immigration rate (λ) and the emigration rate (μ) of a habitat having k species can be defined by formula (4.2), (4.3) defined in chapter 4.

Step (v): The feature habitats initially contain the training pixels of the corresponding feature and HSI is calculated on those training pixels. For example, water habitat initially contains pixels of water and standard deviation of all the training pixels of water is calculated in each of seven bands.

Step (vi): Each species is taken from the Universal habitat and is migrated to each habitat one at a time. The HSI of the habitat is recalculated after migrating the species to it.

Step (vii): (a) The recalculated HSI after the migration of species is compared with the original HSI of the habitat (that contained training pixels only). If the compared HSI is within the threshold i.e., it varies from -1 to +1 from the original HSI then it means that the habitat is suitable for that species and hence it can migrate to it. Figure 5.2 shows the comparison of HSI of training sets' pixels with recalculated HSI after migrating the species.

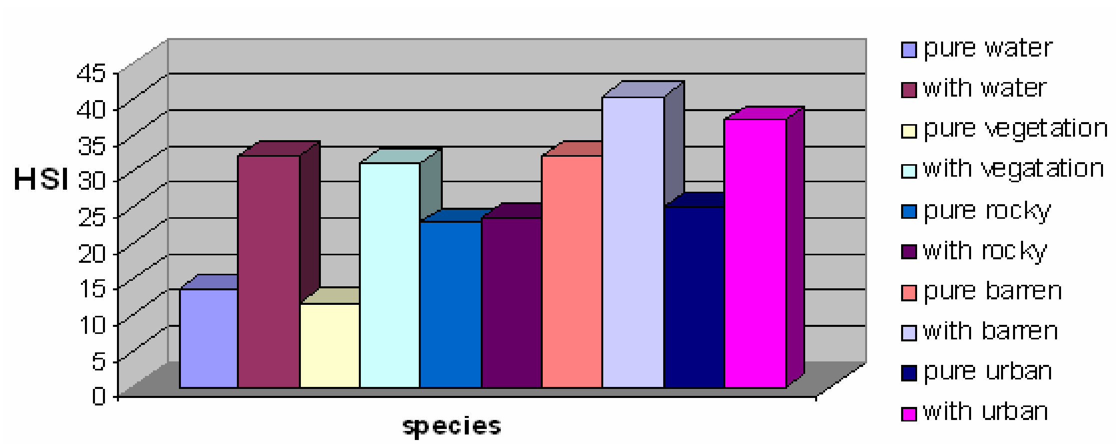


Figure 5.2. HSI matching in habitats

(b) If the recalculated HSI is not within the threshold for any of the original habitat then the particular elementary class contained a mixture of species and it is again partitioned into elementary classes using Rosetta software. These classes are migrated into the Universal habitat and considered as unclassified species and further be used for the next generation.

Step (viii): The checking of all the habitats results in the completion of first generation and then we move towards the second generation for further refinement following the similar process. The maximum number of generations can correspond to the situation when we are left with no more unclassified class (species) in the Universal habitat.

5.2.3. Deriving a mathematical formulation of Biogeography based feature extraction algorithm

The process of Biogeography Based Land Cover Feature Extraction is divided into three steps:

- The first step considers a class and concatenates it with various training sets (i.e. water, vegetation, rocky, barren and urban). These classes and training sets are saved as excel sheets containing x coordinate, y-coordinate, DN values of all the bands. After concatenation each result is stored in a different sheet.
- The next step is to use a Heuristic procedure to decide which land cover property each class belongs to. This is done (in MatLab) by comparing the mean of the Standard Deviation for each of these classes (defined as the Fitness Function) with the Standard Deviation of the Feature Habitat class, using a specific threshold value .

  Therefore, Fitness function = Difference of the mean of the Standard Deviation for each of these classes .

Feature Habitat class = class which contains the standard training set pixels of the 7-Band Image of the Alwar region for comparison.

- In the final step, this function decides which value of mean of standard deviation has minimum difference from the original class. i.e. HSI = Standard Deviation for each of the classes .

If this value is within the threshold then that class (species) will migrate to that habitat. If not it can migrate to other class .This can be mathematically represented as below –

Let x_i represent one of the 20 Rosetta classified rough set classes [30] i.e. the universal habitat and y_i represent the training set gray level values i.e. the feature habitat for the i^{th} band of the 7-band image for each of the 5 land cover features to be extracted.

Then,

If

$$\left| \left[\Sigma\sigma_{xi}/n \right]_{UH} - \left[\Sigma\sigma_{yi}/n\right]_{FH} \right| 6_{j=1} < \text{threshold} ,$$

where,

UH = Universal Habitat

FH = Feature Habitat

Then, the feature is decided as ‘j’ i.e. the said Equivalence class corresponds to the feature ‘j’.

Else,

j =1 which means that the feature ‘j’ is treated as unclassified .

If it belongs to no class it can simply move to the universal habitat and divides itself to a number of classes which then choose their habitats .

5.3. Hybrid ACO2/PSO Land Cover Feature Extraction

In the hybrid ACO2/PSO technique, we have a multi-spectral multi-band satellite image for classification. By these multi-band images the experts generate the training set. Experts of remote sensing open these images in ERDAS IMAGINE [37]. They manually select the pixels of different class regions and set a definition colour code using ERDAS Signature Editor. This data set provided by the experts in the form of digital numbers (intensity value pixel in a digital image) which are taken with help of ERDAS . These sets are taken by carefully selecting the areas (pixel by pixel) from all the images and noting the DN values of the pixels. This dataset can be in any Excel Sheet, Access sheet or in Text file. The Decision according to the DN values is also set in the table.

From these DN values rules are extracted by applying the PSO/ACO2 algorithm. To apply the Modified Hybrid ACO/PSO2 the Open Source Tool for Hybrid ACO/PSO2 is used [8]. Before applying these algorithms for rule-set generation we change the training set into the unified Attribute Relation File Format, i.e., .arff format . We obtain the rule sets from the Open Source Tool for PSO/ACO2. The PSO/ACO2 tool window for extracting the classification rules is as shown below in fig 5.3.

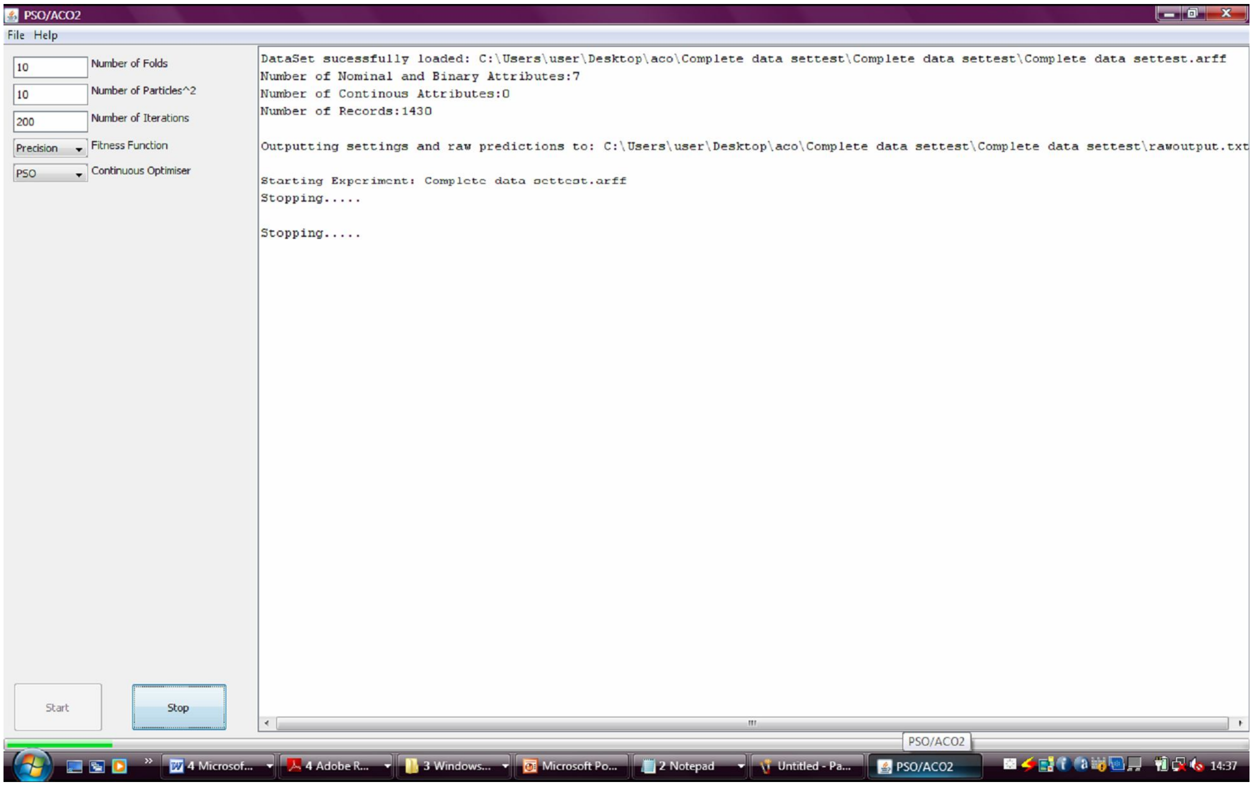


Fig 5.3.Hybrid PSO/ACO2 toolkit window for extracting classification rules

But these rule set do not provide the classified image. So we code these rules in MATLAB file [37]. In MATLAB coding the multi-band Images are read and then for each pixel of image the DN values are checked according to the rule set contained. If the pixel validates any rule then the decision of rule is set to be the class of pixel and the pixel is set to a specified colour. On executing this MATLAB file we obtain the *Final Classified Image*.

The algorithm for the PSO/ACO2 image classification [8] can therefore be summarised as follows -

1. Retrieve the original multi-spectral multi-band image in .tiff format.

2. Generate the training set using ERDAS software, according to the different classes by experts.

3. Convert this training data set into the unified Attribute Relation File Format.

4. Apply the Hybrid PSO/ACO2 on the training data set obtained at step 3, by open Source Tool to get rule set and go to step 5.

5. For each rule in the rule set perform the following actions—

– For each pixel of original image

– if the pixel validates the rule

Then

the pixel is set to be in the region decided by the rule and the color code is assigned.

Else

Move to next rule.

- Finally the classified image is obtained in .jpeg, jpg, .tiff or any other image format.

The region matching or classification coding is done in Matlab and it will give the output classified image on executing. The Matlab window which contains the snippet of the Matlab Coding of region matching for Alwar Area using rule set of hybrid PSO/ACO2 approach is given in fig 5.4.

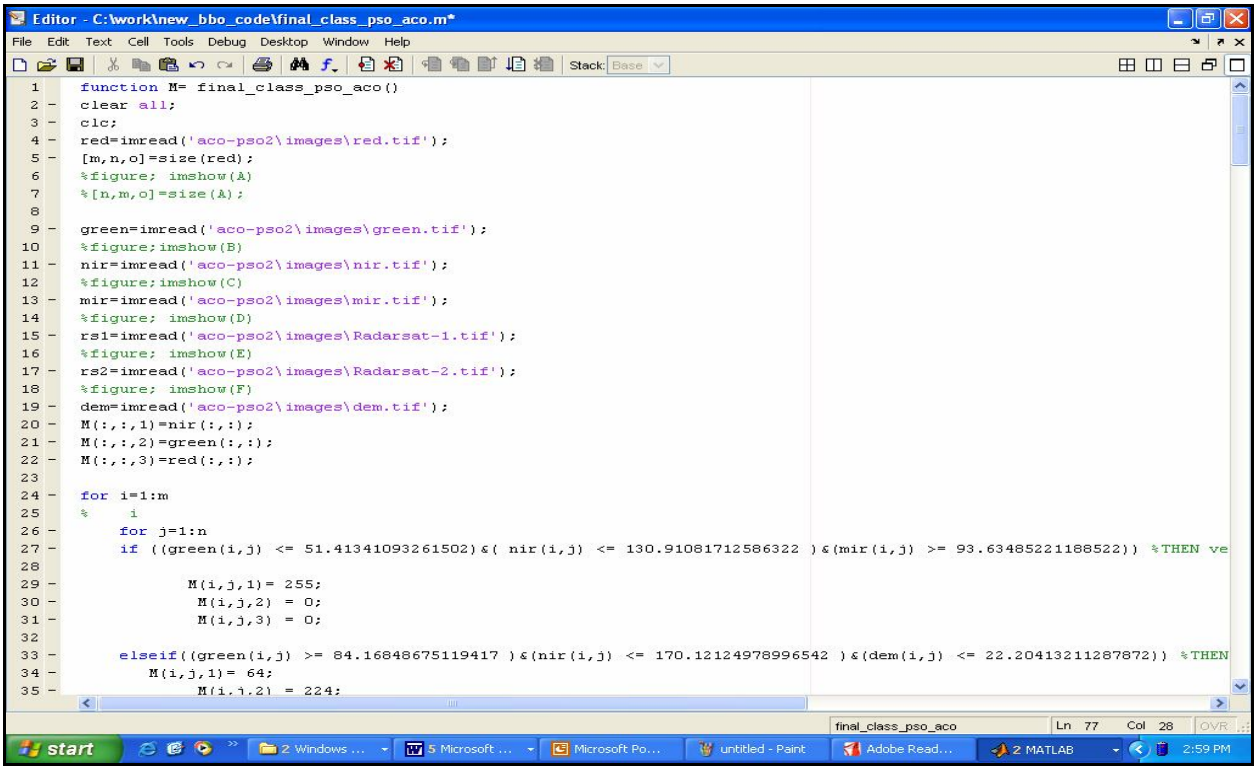


Fig 5.4. Matlab window containing the code for region matching for PSO/ACO2

5.4 From BBO to Hybrid ACO2/PSO/BBO based Land Cover Feature Extraction

5.4.1. Proposed Software Architecture

In this project we have implemented BBO algorithm combined with the ACO2/PSO technique for the purpose of features extraction from a satellite image. The proposed software architecture of our Bio-Inspired Pattern Analysis based Intelligent classifier is given below in figure 5.5.

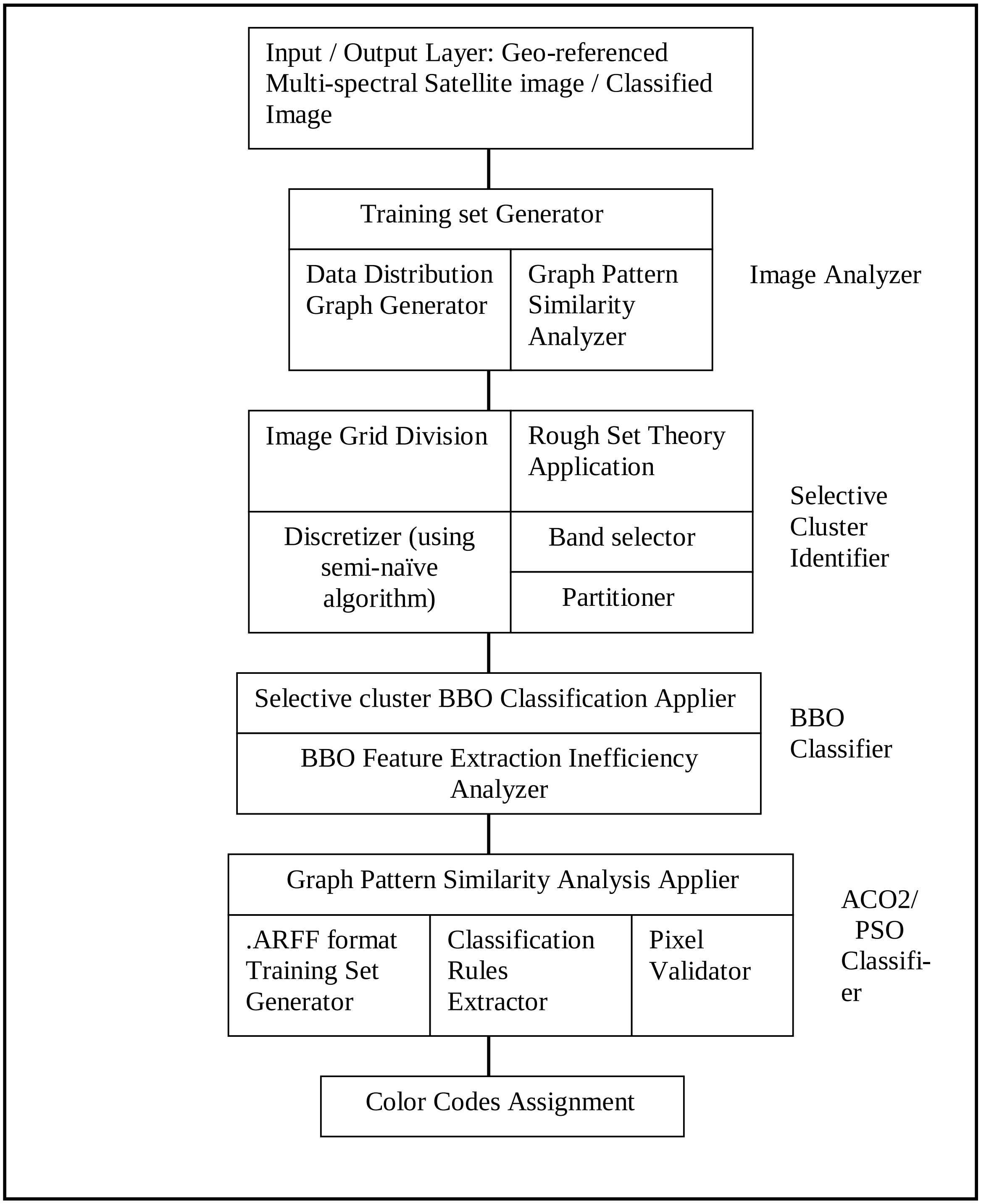


Figure 5.5. Software architecture of the proposed intelligent classifier

The above layers of the software architecture of our proposed hybrid classifier is explained step by step below-

Input layer: A high resolution multi-spectral satellite image is taken as input. For our illustration, the image used is the 7-band cartoset satellite image of size *472 X 576* of the Alwar Region in Rajasthan.

Image Analyzer Layer: This layer is responsible for analysis of the input satellite image to exploit the Data distribution of the training set and feed its output to the next layer i.e. selective cluster identifier layer.

(i) Training set Generation Layer: In this layer, we generate the training set using ERDAS software, according to the different classes by experts.

(ii) Data Distribution Graph Generation Layer: Now, using this training set, we plot the data distribution graph between the average of the Standard Deviations of each land cover feature viz water, urban, rocky, vegetation and barren (plotted on the y-axis) for each of the 7-Bands of the image i.e. Red, Green, NIR, MIR, RS1, RS2 and DEM (plotted as the x-axis) as shown in figure 5.6. On the x-axis, each of the 7-bands of the image are represented as integers from 4-10 for Red, Green, NIR, MIR, RS1, RS2 and DEM bands respectively. We are not considering integers 1-3 in the graph plot on the x-axis since these represent the other attributes of the image namely pixel no. and x and y coordinates which are not required at this stage.

(iii) Graph Pattern Similarity Analysis Layer: From the graph, it can be observed that the minimum difference between the average standard deviations of the NIR and the MIR bands of the Alwar Image is achieved in particularly two land cover features, those of water and urban area, both of which exhibit the same graph pattern in the NIR and the MIR bands i.e.

$|\text{ average of standard deviation of NIR band} \sim \text{average of standard deviation of the MIR band }|_{\text{lowest}} = \{\text{water, urban}\}$

Hence, it can be concluded that these are the two features that will be most efficiently classified by our hybrid algorithm which works in the NIR and MIR bands.

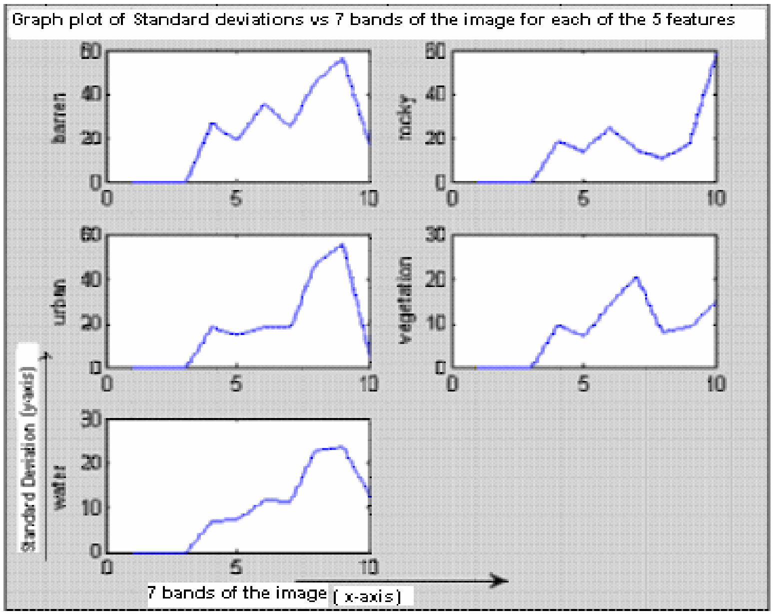


Figure 5.6. Graph plot of the Standard Deviations of each Land Cover feature v/s each of the 7-Bands in which the Alwar Image is viewed.

Selective Cluster Identifier Layer: Based on the observation of the adaptive nature of BBO algorithm for Land Cover Feature Extraction , we observe that BBO is able to classify some particular feature's pixels with greater efficiency than the other features based on the band that we select for creation of equivalence classes in Rosetta . The functionality of this layer is described in detail below-

(i) Image Grid Division Layer: In this layer, the satellite image is divided into n clusters, in our case n = 20. In the biogeography based classification layer, basically we identify to which cluster of the image BBO technique will show good classification efficiency and hence should be applied to.

(ii) Rough Set Theory Application Layer: For the purpose, we use rough set theory toolkit i.e. *Rosetta software* [25] for dicretizing each of the 20 clusters using the semi-naïve Algorithm & then partition each of them based on the band which is able to classify the particular feature that we want to extract from the image. This is because BBO follows an adaptive strategy and will classify the desired feature most efficiently since it has its corresponding band selected in the partitioning step.

(iii) Band selection Layer: Depending on our application, in other words depending on which feature we want to extract from the image most efficiently, we can choose the band for partitioning. For example, if we want to extract the barren area more efficiently, we choose the green band and for rocky region extraction we choose the MIR Band. The RS-1 and RS-2 bands are used to extract the urban area and also for extracting the edges of rocky region from the 7-band image. However, the drainages of rocky region are best viewed in the Red band and water and vegetation pixels are best viewed in NIR and MIR Bands. For our illustration, we have chosen the NIR and MIR band of the 7-band image since we want to extract the water pixels effectively and clearly identify the water body in the image and these are the bands in which the water feature is particularly more highlighted and best viewed.

(iv) Discretization and Partitioning Layer: Therefore, we use the NIR and the MIR bands for discretization and partitioning step in the semi –naïve algorithm used for creating rough set equivalence classes, thus creating Equivalence classes for each of the clusters. This is what is termed as Unsupervised Classification .Each of these resultant classes are put in the Universal Habitat.

BBO Classifier Layer:

(i) Selective Cluster BBO Classification Applier Layer: Hence, we apply BBO Technique on those clusters of the satellite image which show the maximum classification efficiency which is due to the fact that these are the clusters which predominantly show the presence of the feature that is most efficiently classified by the BBO Algorithm. The procedure for applying the biogeography based land cover feature extraction technique was described in section 5.2.

In our illustration, BBO will be able to extract the water pixels most efficiently since we choose the NIR and MIR bands for partitioning. And therefore, we apply BBO on the k^{th} cluster of the Alwar image since this is the cluster which gives the maximum classification efficiency because it predominantly shows the presence of water body in the image .

(ii) BBO Feature Extraction Inefficiency Analysis Layer: However, it is also observed that BBO shows a wide range of efficiencies in classifying the rest of the features. In other words, BBO is not able to classify all the 5 features with good classification efficiency and shows poor performance on some of the remaining features. In fact, in our illustration, it shows the poorest performance in classifying the urban pixels . Hence we proceed for the ACO2/PSO classification to improve the image classified by BBO.

ACO2/PSO Classifier:

(i) Graph Pattern Similarity Analysis Application Layer: Now, we know from the data distribution graph plotted in the Image Analyzer Layer that our hybrid algorithm will show similar efficiency (i.e. maximum) for the urban pixels too as for the water pixels since their graph pattern in the NIR and MIR bands is similar, therefore, we then apply ACO2/PSO Technique on the rest of the image .

In our case, BBO performed better than ACO2/PSO only on the water pixels and hence we applied BBO on the k^{th} cluster . Although BBO gave good performance on vegetation pixels too (since in NIR and MIR bands these are the 2 features that are best viewed) ,however, the efficiency was no better than ACO2/PSO classification and hence we did not apply BBO on the vegetation pixels.

(ii) Training set in .arff format Generator Layer: In order to apply ACO2/PSO Classifier, we take the training set for the 7-Band Alwar image in .arff format as input to generate the classification rules from it.

(iii) Classification Rules Extractor Layer: For the purpose, we use the Hybrid ACO2/PSO open source Tool and then apply the extracted classification rules on each of the remainder clusters of the image.

(iv) Pixel Validator Layer: On each of the pixel of the image, we check for pixel validation on each pixel in the cluster & thus obtain a refined classification of the image . The above procedure of applying ACO2/PSO generated rules for classifying the image is as described in section 5.3.

Color Codes Assignment Layer: Finally color codes are assigned for each pixel of the image corresponding to the classified feature of each pixel. The color codes for water, urban, vegetation, rocky and barren features are assigned as blue, red, green, yellow and black color respectively.

Output Layer: Final ACO2/PSO/BBO Classified image obtained in .jpeg, .tiff or any other image format.

5.4.2. Mathematical Formulation of the ACO2/PSO/BBO Algorithm

Therefore, the working of our proposed hybrid algorithm can be summarized in the form of the following equation in fig 5.7.

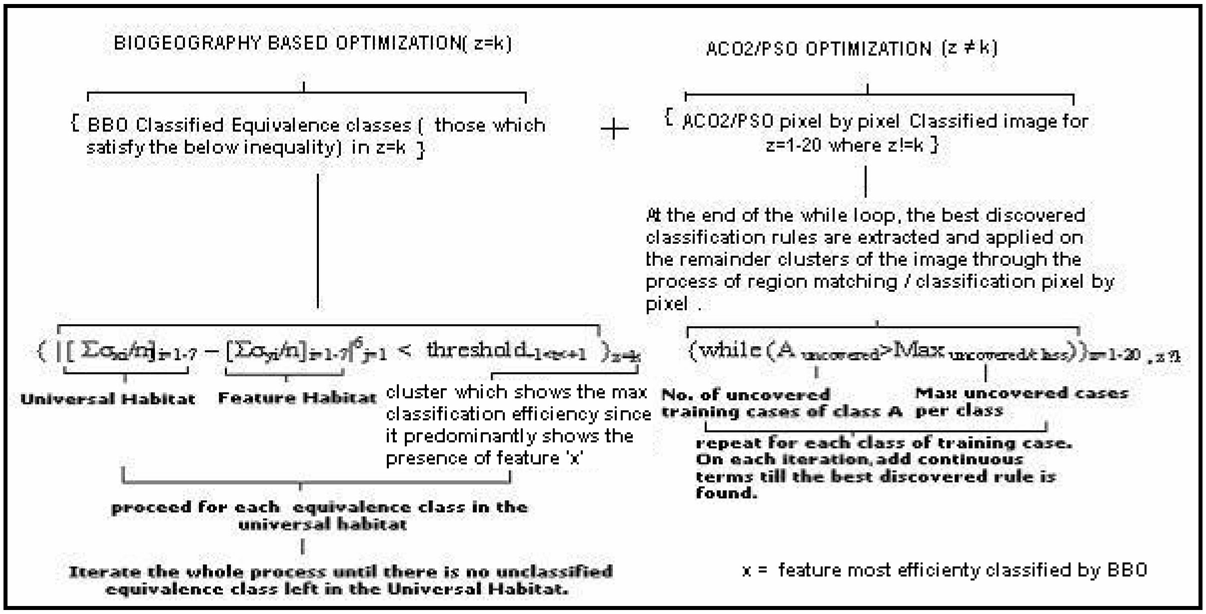


Fig 5.7. Mathematical formulation of the ACO2/PSO/BBO Algorithm

where

the Universal Habitat contains the rough set classified equivalence classes and the feature habitat consists of the expert generated training set of the original alwar image in 7-bands.

And,

Z=input cluster of the 7-band image (z ranges from 1-20)

J= feature under consideration (j ranges from 1-6 , j=1 for unclassified feature)

X=feature most efficiently classified by BBO

K=cluster which shows the maximum classification efficiency since it predominantly shows the presence of feature ‘x’

The above equation can be explained as follows-

(i) For z=k , we proceed in the following manner for the BBO Optimizer -i.e. for each i^{th} band where 'i' ranges from 1-7 , we calculate the difference in the standard deviation of the i^{th} band of the Universal Habitat and the i^{th} band of the Feature Habitat containing the expert generated training set of the image. If this difference is the minimum for the feature 'j' and also less than the pre-specified threshold value of,$-1 < t < +1$, then that particular equivalence class is classified as the feature 'j' else j=1(unclassified).The process is repeated for each equivalence class until there is no equivalence class left in the universal habitat and the whole process is iterated till there is no unclassified Equivalence class left.

(ii) For z=1-20, where $z \neq k$, we use the ACO2/PSO Optimization, wherein the training set for the 7-Band Alwar image in .arff format is used as input to generate rules from it using the open source Tool for each class of training case and on each iteration, we add continuous terms till the best discovered rule is found. The classification rules are then applied on the remainder of the clusters checking for pixel validation on each of them.

Hence, we obtain a more refined classified image with an improved Kappa coefficient (Kappa coefficient is a way of measuring the efficiency of a classification algorithm. It is described in detail later in section 6.3.2) which is much better than the Kappa Coefficient we get when we apply the original BBO Algorithm on the 7-Band Image.

This in turn leads us to the improved flexible Hybrid version of the BBO Algorithm for Satellite Image Classification which will classify the particular feature chosen by the band used in the unsupervised classification , most efficiently ,which is in turn based on the expert knowledge and the band information contained in the training set of the particular area. Thus, we have efficiently exploited the properties of the BBO technique to adapt itself to a more focused classification which upon integrating with the ACO2/PSO Technique makes an advanced classifier. Hence, we have obtained a hybrid algorithm which can be adapted to incorporate the expert knowledge for a more flexible, efficient and refined classification, thus resulting in what we term as the hybrid bio-inspired pattern analysis based intelligent classifier. The detailed internal architecture of this Hybrid ACO2/PSO/BBO based classifier is illustrated by means of a flowchart in fig. 5.8.

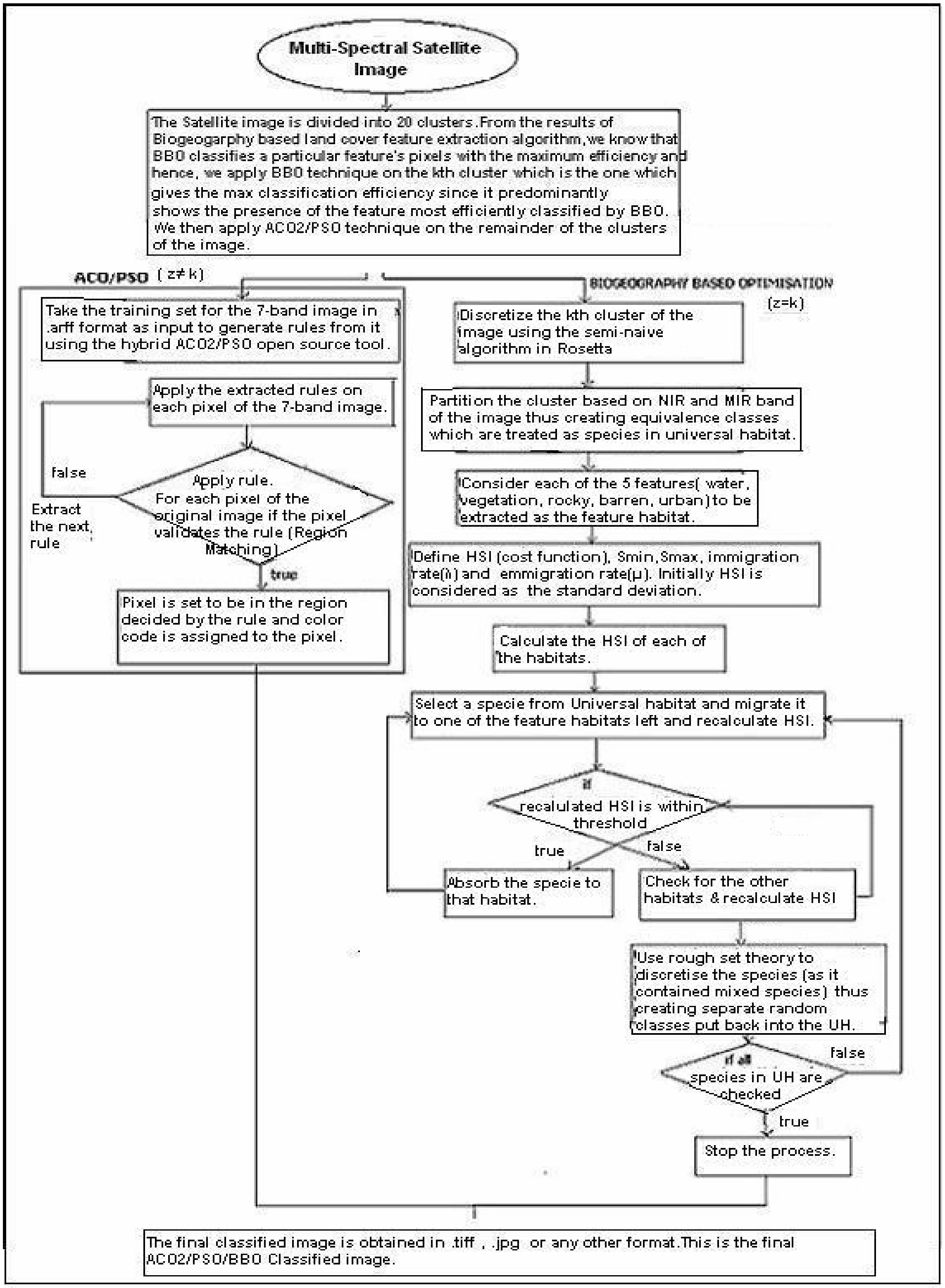


Fig 5.8. Detailed Internal architecture of the Hybrid ACO/PSO/BBO classifier

Chapter 6

RESULTS AND DISCUSSION

The algorithm proposed in the previous chapter has been implemented in Matlab 7.0 [37] .The algorithm has been applied to 7-band image of Alwar City (in Rajasthan, India) the dataset of which was given in figure 5.1 of section 5.1.

6.1 Classification results of Biogeography based land cover feature extraction

6.1.1. Elementary Classes

After the image has been read, it is divided into 20 parts. Each part is called a class. Further pixels in each class are grouped into random clusters, called elementary classes. The elementary classes are obtained with the help of ROSETTA software version 1.4.41 [25].

Rosetta is ROugh SET Toolkit for Analysis. ROSETTA is a toolkit for analyzing tabular data within the framework of rough set theory. The image data is first read in ROSETTA as shown in figure 6.1.

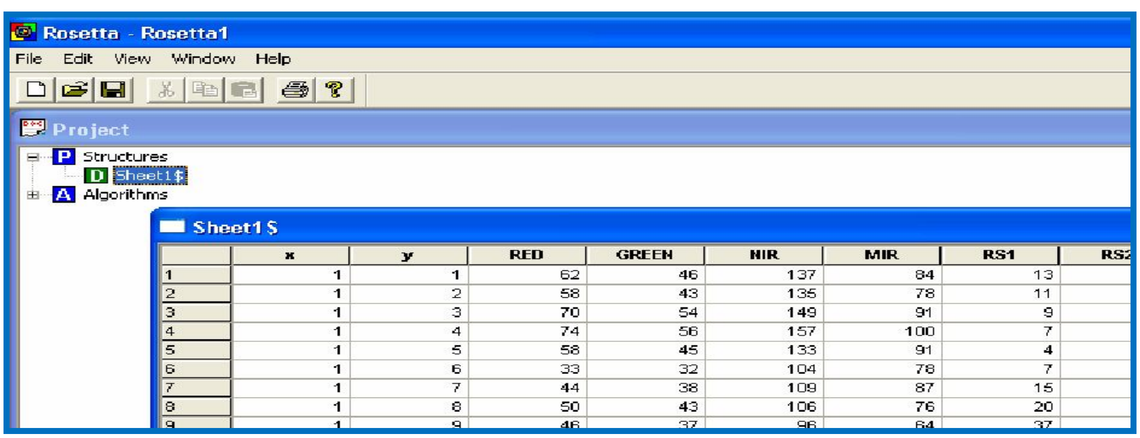
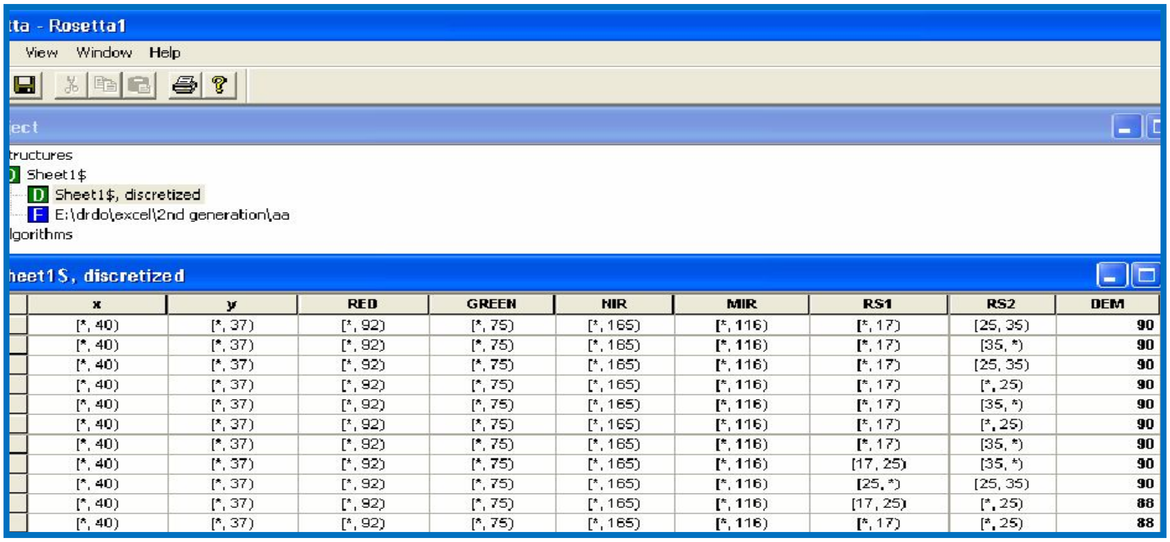


Figure 6.1. Input Image Pixels in ROSETTA software

The elementary classes are obtained by using discretization followed by partition process in ROSETTA. Input read data is further discretized by using naïve discretization algorithm defined in ROSETTA. In the resulting data, DN value of each pixel is replaced by the discretized interval as shown in figure 6.2.

•


x	y	RED	GREEN	NIR	MIR	RS1	RS2	DEM
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[*, 17)	[25, 35)	90
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[*, 17)	[35, *)	90
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[*, 17)	[25, 35)	90
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[*, 17)	[*, 25)	90
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[*, 17)	[35, *)	90
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[*, 17)	[*, 25)	90
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[*, 17)	[35, *)	90
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[17, 25)	[35, *)	90
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[25, *)	[25, 35)	90
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[17, 25)	[*, 25)	88
[*, 40)	[*, 37)	[*, 92)	[*, 75)	[*, 165)	[*, 116)	[*, 17)	[*, 25)	88

Figure 6.2. Discretized Image Data in ROSETTA software

This discretized data is further partitioned on the basis of NIR and MIR band to obtain the final elementary classes. The band used for partitioning may affect the final results. The elementary classes obtained in ROSETTA are shown in the figure 6.3. Each elementary class consists of the row index of each pixel. The pixel's DN values corresponding to these indexes are obtained before applying the proposed algorithm.

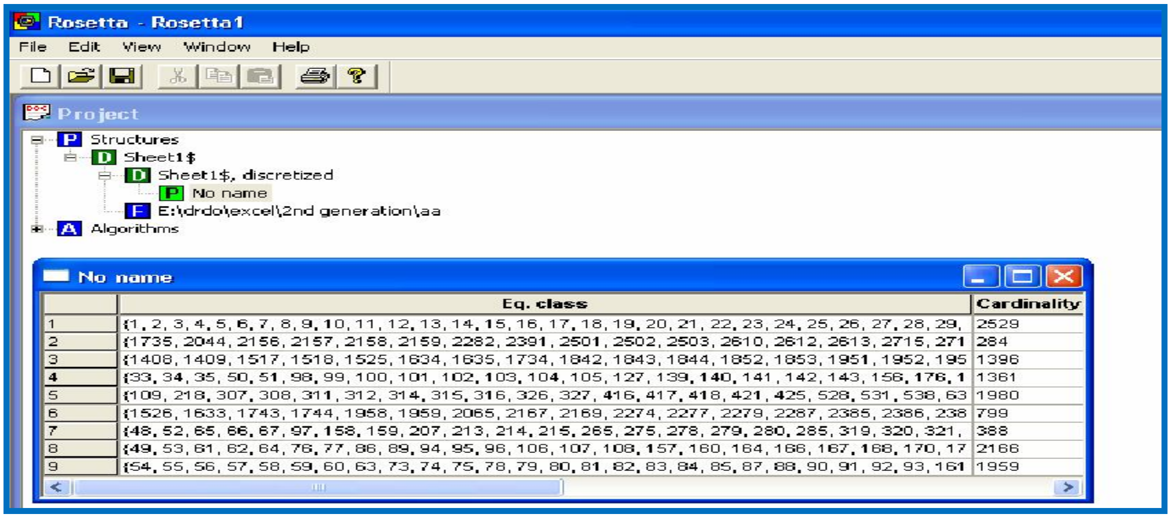


Figure 6.3. Elementary Classes in ROSETTA software

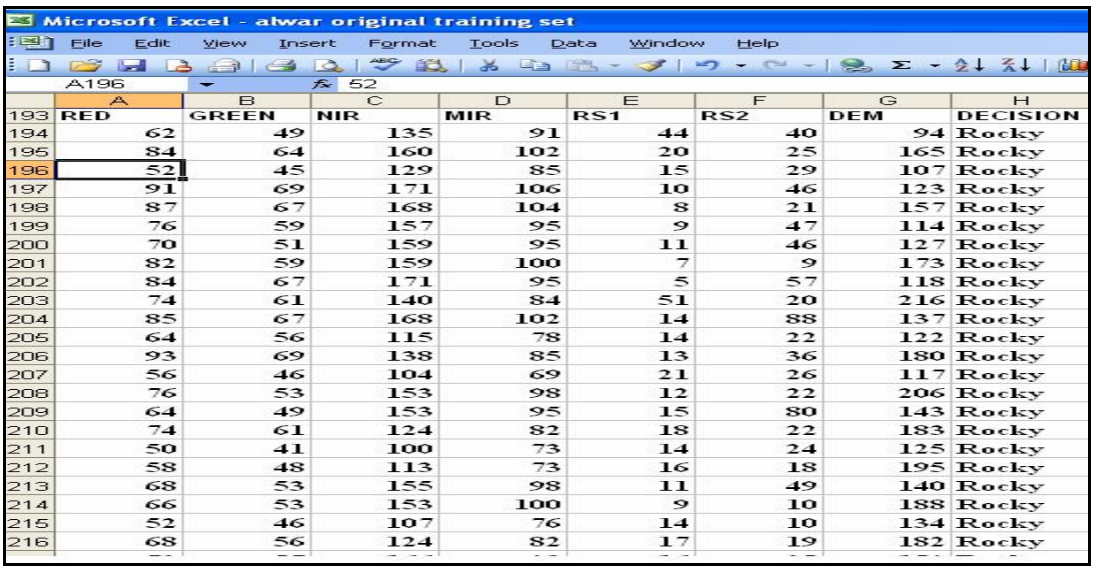
The elementary classes obtained by ROSETTA are put in the Universal Habitat and each elementary class is treated as a species (but it could be a mixture of species).

6.1.2. Training Sets in Feature Habitats

Each feature habitat initially consists of training sets. Training sets of the image are provided by the experts, which help us to classify the whole image. These are basically some of the pixels of the image, for which an expert is sure about that this pixel is of which land cover region. The training data set consisting of 680 locations on the image is collected taking all the 7-band images together in order to avoid any ambiguity for confirmed training sites in the two data sets p and x.

The knowledge resident with the training set is assumed to be the one obtained from the expert, duly verified from ground checks and is confined to the r, g, n, m bands. This data set can be represented in a tabular form similar to that of a relational database tables. Rows of the table represent the training pixels and the DN values in the columns related to the 7-bands viz., r, g, n m, rs1, rs2 and dem. The table has, therefore, total 7 attributes- 4 attributes (r, g, n and m), termed as attributes set p attributes and 3 attributes (rs1, rs2, dem) termed as x attributes.

The training sets used in this study have been stored in the Microsoft excel sheets, with one row representing one pixel and 7-columns representing 7-bands of the image. Decision attribute describes the feature to which the each pixel (row) belongs. Figure 6.4 shows the training set of rocky region of the Alwar image, stored in an Excel sheet. These training set pixels are read in the corresponding feature habitat in the beginning of the algorithm.



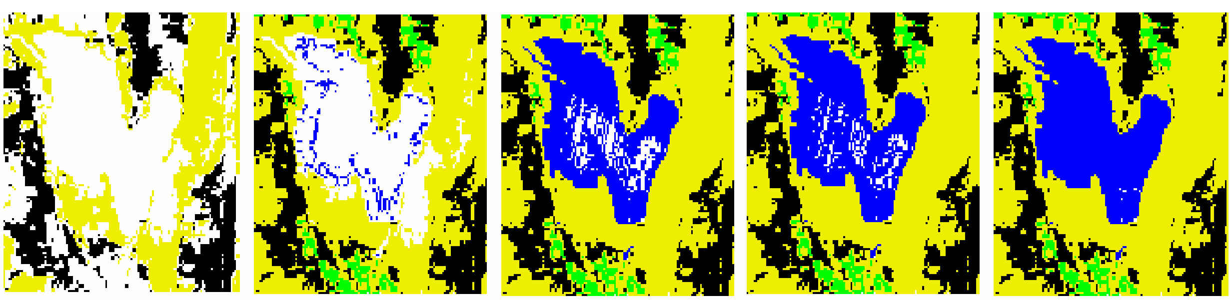
	A	B	C	D	E	F	G	H
193	RED	GREEN	NIR	MIR	RS1	RS2	DEM	DECISION
194	62	49	135	91	44	40	94	Rocky
195	84	64	160	102	20	25	165	Rocky
196	52	45	129	85	15	29	107	Rocky
197	91	69	171	106	10	46	123	Rocky
198	87	67	168	104	8	21	157	Rocky
199	76	59	157	95	9	47	114	Rocky
200	70	51	159	95	11	46	127	Rocky
201	82	59	159	100	7	9	173	Rocky
202	84	67	171	95	5	57	118	Rocky
203	74	61	140	84	51	20	216	Rocky
204	85	67	168	102	14	88	137	Rocky
205	64	56	115	78	14	22	122	Rocky
206	93	69	138	85	13	36	180	Rocky
207	56	46	104	69	21	26	117	Rocky
208	76	53	153	98	12	22	206	Rocky
209	64	49	153	95	15	80	143	Rocky
210	74	61	124	82	18	22	183	Rocky
211	50	41	100	73	14	24	125	Rocky
212	58	48	113	73	16	18	195	Rocky
213	68	53	155	98	11	49	140	Rocky
214	66	53	153	100	9	10	188	Rocky
215	52	46	107	76	14	10	134	Rocky
216	68	56	124	82	17	19	182	Rocky

Figure 6.4. Training Set of Rocky Region

The initial stage of the algorithm can be interpreted as the Universal habitat contains species to be migrated (obtained using rough set theory [23]) and the feature habitats contain training pixels provided by expert. After calculating the HSI on training pixels, each species in the Universal habitat is migrated to the feature habitat and the HSI of the habitat is recalculated after migration. If the change in HSI is above the threshold value, the suitability of the species in other feature habitat is checked. The species is absorbed in that habitat where the recalculated HSI is very close to the original HSI calculated by training pixels.

Fig 6.5 shows the 16th cluster of the Alwar image after running multiple iterations of Biogeography based land cover feature extraction algorithm on it with the threshold value taken as -1 to +1. The yellow color represents rocky area, green color represents vegetation area, black color represents barren area and red color represents the urban area. The white color in the image represents the unclassified pixels of image in the first iteration. It is analogous to those species that are not absorbed in any feature habitat and are put back to the Universal habitat. These species represent the unclassified pixels that can be further classified in the next iteration of algorithm.

By taking threshold value from -1 to +1, we see that no water body (represented in blue color) is extracted in the first iteration (fig.6.5 (a)) and water pixels remain unclassified (represented in white color). With the increasing iteration (fig. 6.5(b), 6.5(c), 6.5(d), 6.5(e)), water body is clearly extracted at the end of the fifth iteration.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3 (d) Iteration 4 (e) Iteration 5

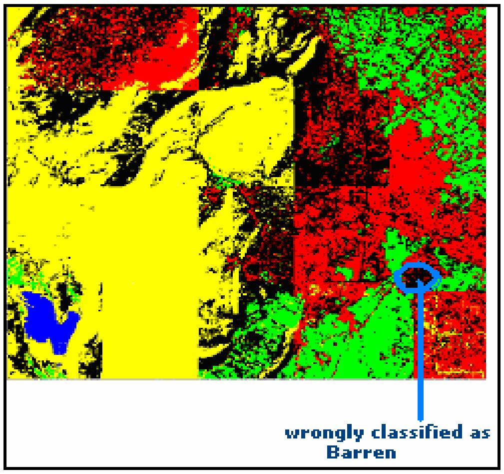
Figure 6.5. Water Body extraction (in blue color) from the Alwar image with each iteration. White color portion represents the unclassified image.

After five iterations of the algorithm, each species is absorbed in the corresponding habitat and the Universal habitat becomes empty. This terminates the biogeography based feature extraction algorithm on the 16^{th} cluster. The number of iterations can increase further if we decrease the threshold range to -0.5 to 0.5. Lower is the threshold value, more are the chances that the recalculated HSI does not lie in threshold value, and hence more are the chances that species will not get absorb in any feature habitat and will be put back to the Universal habitat. So more iterations will be required to migrate species to the correct habitat. So the threshold range has inverse relation with the number of iterations. However, as the number of iterations of algorithm increase, we get more refined results.

6.2. Hybrid ACO2/PSO/BBO Classification : Results and Discussion

Based on the results obtained on applying the BBO algorithm to the 7-Band Image of Alwar region for Land Cover Feature Extraction, we observe that we are able to classify water pixels with the highest efficiency i.e. 99% efficiency and these are the pixels best viewed in the NIR and MIR bands in the BBO Technique & hence, we apply BBO Technique on the 16th cluster of the Satellite image of Alwar region (z=16) as shown in figure 6.5 since this is the cluster which gives the maximum classification accuracy and also predominantly shows presence of water body in the Alwar Image .The cluster also shows a significant improvement in the classification upon each iteration since after each iteration a significant amount of unclassified pixels are reduced to classified ones as is evident from figure 6.5. However, BBO shows poor efficiency , in fact the poorest, in classifying the urban pixels as shown in fig. 6.6. Here the encircled region in the BBO Classified Image shows that BBO wrongly classifies the urban pixels as barren ones which will also be reflected from Table 6.1 in the accuracy assessment section 6.3 where BBO classifies 91 urban pixels wrongly out of 190 total urban pixels .

Therefore, in order to classify the urban pixels efficiently, we then apply ACO2/PSO Technique [8] on the remainder of the clusters of the image ($z \neq 16$) by taking the training set for the 7-Band Alwar image in .arff format as input to generate rules from it using the open source Tool as described in section 5.3 and then applying them on the remainder of the clusters checking on pixel validation for each pixel in the cluster & thus obtain a more refined classification of the image with an improved Kappa coefficient of 0.9818 (calculated in accuracy assessment section 6.3) which is much better than the Kappa Coefficient of 0.6715 we get, when we apply the original BBO Algorithm on the 7-Band Image . This in turn leads us to the improved Hybrid version of the BBO Algorithm for Satellite Image Classification where both the urban and the water features are classified with the highest efficiency i.e. almost 100% with no omission errors followed by rocky with only 1 omission error (column wise error) and thereafter barren and vegetation features ,respectively. After applying the proposed algorithm to the 7-band of Alwar Image, the classified image is obtained in figure 6.7. From the figure, it is clearly shown that our proposed ACO2/PSO/BBO based intelligent classifier is able to correctly classify the encircled region as urban which was wrongly classified by the simple BBO Classifier. The yellow, black, blue, green, red color represents rocky, barren, water, vegetation, urban region respectively. As the threshold limit of HSI matching is lowered, the species do not get absorbed in the feature habitat and return to universal habitat. Those species are further discretized and classified in next iterations (generation).



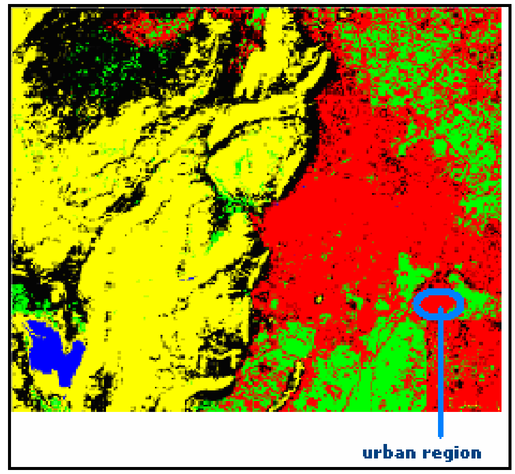


Fig 6.6. Classified image after applying BBO (with Kappa Coefficient=0.6715)

Fig 6.7. Hybrid ACO2/PSO/BBO Classified Image (Kappa Coefficient=0.98182)

From the figures 6.7 and 6.8 , it is evident that the Hybrid ACO2/PSO-BBO Technique produces a more refined image as compared to the BBO classified image.

6.3. Accuracy Assessment of the proposed Algorithm

A classification is not complete until its accuracy is assessed . Accuracy assessment is a general term for comparing the classification to geographical data that are assumed to be true, in order to determine the accuracy of the classification process. Usually, the assumed-true data are derived from ground truth data. It is usually not practical to ground truth or otherwise test every pixel of a classified image. Therefore, a set of reference pixels is usually used. Reference pixels are points on the classified image for which actual features are (or will be) known. The reference pixels are randomly selected. Accuracy assessment is an important step in the classification process. The goal is to quantitatively determine how effectively pixels were grouped into the correct feature classes in the area under investigation.

6.3.1. Error Matrix

Now we proceed to calculate the classification accuracy of our proposed algorithm using the classification *error matrix* . Error matrices compare, on category-by category basis, the relationship between known reference data (ground truth) and the corresponding results of an automated classification. We took 150 vegetation pixels, 190 Urban pixels, 200 Rocky pixels, 70 water pixels, 170 barren pixels from the training set and the error matrix obtained is shown in Table 6.3 . In Table 6.3, the training set pixels that are classified into proper land cover categories are located along the major diagonal of the error matrix (running from upper left to lower right). All the non diagonal elements represent the error of omission and commission. Omission errors correspond to the nondiagonal column elements Commission errors are represented by nondiagonal row elements .

The error matrix's interpretation along column suggests how many pixels are classified correctly by the algorithm. For e.g. in the first column, out of total 150 vegetation pixels, 142 pixels were correctly classified as vegetation, 2 vegetation pixels were classified wrongly as Barren, and 5 as Urban ,a total of only 7 misclassified pixels. However, From the simple BBO Error matrix given in Table 6.1, it is found that 17 vegetation pixels were wrongly classified as Barren pixels ,and 6 as water, which amounts to a total of 23 wrongly classified pixels. Also, the diagonal elements (diagonal elements indicate the no. of correctly classified pixels in that category) of Table 6.3 (Hybrid ACO2/PSO-BBO Classifier) vary drastically when compared to those in Table 6.1 (simple BBO Classifier). For example, in table 6.3, whole 190 out of 190 pixels were correctly classified as Urban pixels whereas simple BBO Classifier in table 6.1 could only classify 88 pixels correctly as urban pixels and it classified 91 pixels wrongly as barren ones. This indicates a drastic improvement in the accuracy of the results obtained by BBO through the use of the Hybrid Technique which is able to classify , in particular the urban and the water pixels, with almost 100% efficiency (with no omission errors) ,which was what was also reflected earlier, from the data distribution graph plotted . The error matrix for the hybrid ACO-BBO Classifier is also calculated and presented in Table 6.2 from which it is reflected that the our proposed classifier is an improvement over the hybrid ACO2/BBO based intelligent Classifier too which is a hybrid of the cAntMiner and the BBO based land cover feature extraction algorithm.

Table 6.1. Error matrix when only BBO is applied
Kappa coefficient = 0.6715

	Vegetation	Urban	Rocky	Water	Barren	Total
Vegetation	127	9	0	0	2	138
Urban	0	88	1	0	32	121
Rocky	6	2	176	1	17	202
Water	0	0	3	69	0	72
Barren	17	91	20	0	119	247
Total	150	190	200	70	170	780

	Vegetation	Urban	Rocky	Water	Barren	Total
Vegetation	145	0	0	0	0	145
Urban	2	188	0	0	9	199
Rocky	0	0	194	0	0	194
Water	1	0	3	70	0	74
Barren	1	2	2	0	157	162
Total	149	190	199	70	166	774

Table 6.2. Error Matrix when Hybrid ACO/ BBO technique is applied.
Kappa Coefficient=0.96699

	Vegetation	Urban	Rocky	Water	Barren	Total
Vegetation	142	0	0	0	0	142
Urban	5	190	0	0	0	195
Rocky	0	0	198	0	3	201
Water	0	0	0	70	0	70
Barren	2	0	1	0	163	166
Total	149	190	199	70	166	774

Table 6.3. Error matrix when Hybrid ACO2/PSO/BBO is applied.
Kappa Coefficient=0.9818

6.3.1.1. Producer's Accuracy

Several other descriptive measures can be obtained from error matrix. The accuracy of individual category can be calculated by dividing the number of correctly classified pixels in each category by either the total number of pixels in corresponding row or column. Producer's accuracies(as shown in table 6.4) result from dividing the number of correctly classified pixels in each category (on the major diagonal) by the number of training set pixels used for that category(the column total). This figure indicates how well the training pixels of a given cover type are classified.

The results in the table 6.4 show that water and urban pixels have been extracted perfectly followed by rocky pixels which show 99% efficiency.

Table 6.4 Producer's Accuracy

Feature	Accuracy Calculation	Producer's Accuracy
Vegetation	142/150	95%
Urban	190/190	100%
Rocky	198/200	99%
Water	70/70	100%
Barren	163/170	96%

6.3.1.2. User's Accuracy

User's Accuracies (as shown in table 6.5) is computed by dividing the number of correctly classified pixels in each category by the total number of pixels that were classified in that category (the row total).This figure is a measure of commission error and indicates the probability that a pixel classified into a given category actually represents that category on the ground.

Table 6.5 User's Accuracy

Feature	Accuracy Calculation	Producer's Accuracy
Vegetation	142/142	100%
Urban	190/195	97%
Rocky	198/201	98.5%
Water	70/70	100%
Barren	163/166	98%

6.3.2 KHAT Statistics

A further point to be made about interpreting classification accuracies is the fact that even a completely random assignment of pixels to classes will produce percentage correct values in the error matrix. In fact, such a random assignment could result in a surprisingly good apparent classification result. The $\hat{k}$ ("KHAT") statistic is a measure of the difference between the actual agreement between reference data and an automated classifier and the chance agreement between the reference data and a random classifier. Conceptually, $\hat{k}$ can be defined as

$$\hat{k} = \frac{observed\ accuracy - chance\ agreement}{1 - chance\ agreement} \quad \text{… (6.1)}$$

This statistic serves as an indicator of the extent to which the percentage correct values of an error matrix are due to "true" agreement versus "chance" agreement. As true agreement (observed) approaches 1 and chance agreement approaches 0, its value approaches 1. This is the ideal case. In reality, its value usually ranges between 0 and 1. For example, a value of 0.98 of KHAT statistics can be thought of as an indication that an observed classification is 98 percent better than one resulting from chance. A KHAT coefficient with the value of 0 suggests that a given classification is no better than a random assignment of pixels. In cases where chance agreement is large enough, KHAT coefficient can take on negative values-an indication of very poor classification performance. (Because the possible range of negative values depends on specific matrix, the magnitude of negative values should not be interpreted as an indication of relative classification performance). The principle advantage of computing KHAT coefficient is the ability to use this value as a basis for determining the statistical significance of any matrix or the differences among matrices. The KHAT coefficient incorporates the non-diagonal elements of error matrix (and hence error of omission and commission) as a product of the row and column marginal.

The KHAT statistic is computed as:

$$\hat{k} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \cdot x_{+i})} \qquad \ldots (6.2)$$

Where,

r = number of rows in the error matrix

x_{ii} = the number of observations in row i and column i (on the major diagonal)

x_{i+} = total of observations in row i (shown as marginal total to right of the matrix)

x_{+i} = total of observations in column i (shown as marginal total at bottom of the matrix)

N = total number of observations included in matrix.

To illustrate the computation of KHAT for the error matrix in table 4.1,

$$\sum_{i=1}^{r} x_{ii} = 142 + 190 + 198 + 70 + 163 = 763$$

$$\sum_{i=1}^{r} x_{i+} . x_{+i} = (149*142) + (190*195 + (199*201)+ (70*70) + (166*166) = 1, 30{,}663$$

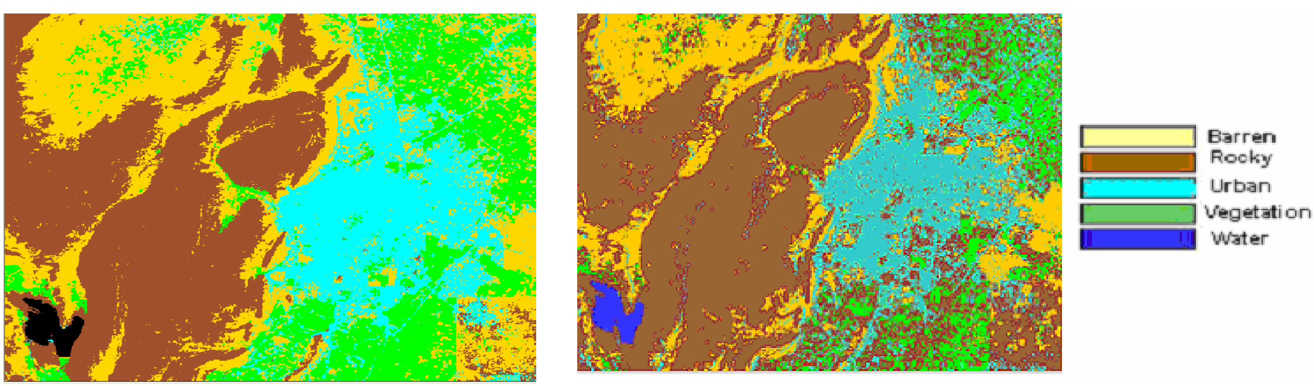
$$\text{'k' (kappa)} = (780*763 - 130663) / (780*780 - 130663)$$

$$= 0.98182$$

Hence the KHAT statistics of the proposed hybrid ACO2/PSO/BBO based land cover feature extractor (or the hybrid bio-inspired pattern analysis based intelligent classifier) is 0.9818. This value is a substantial agreement between the ideal classifier and the proposed algorithm. The Kappa (K) coefficient of the Alwar image is 0.9818 which indicates that an observed classification is 98.82% better than one resulting from chance.

6.4. Classification comparison of Hybrid ACO2/PSO/BBO Classifier with the traditional probabilistic classifiers

Figure 6.8 compares the Hybrid ACO2/PSO/BBO classifier with the Minimum Distance Classifier & Maximum Likelihood Classifier. A comparison of the Kappa Coefficients of the Hybrid ACO2/PSO/BBO Classifier with the Traditional Classifiers is given in Table 6.6.



Minimum Distance Classifier (Kappa Coefficient=0.7364)

Maximum Likelihood Classifier (Kappa Coefficient=0.7525)

Fig 6.8. A comparison with the Traditional Probabilistic Classifiers

Table 6.6. A comparison of Hybrid ACO2/PSO/BBO Classifier with traditional classifiers.

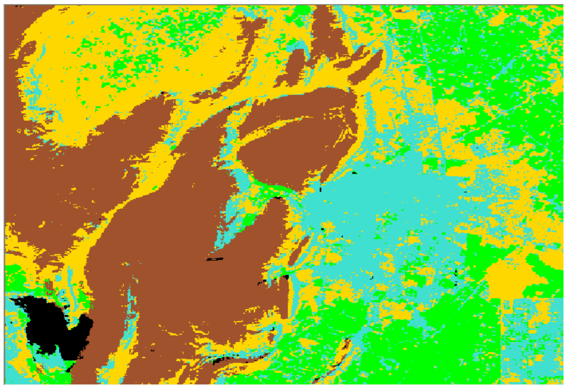
Minimum Distance Classifier(MDC)	Maximum likelihood Classifier(MLC)	Biogeography Based Optimization (BBO)	Hybrid ACO-BBO Classifier	Hybrid ACO-BBO Classifier
0.7364	0.7525	0.6715	0.96699	0.98182

6.5. Classification Results of other Soft Computing Techniques used for Satellite Image Classification

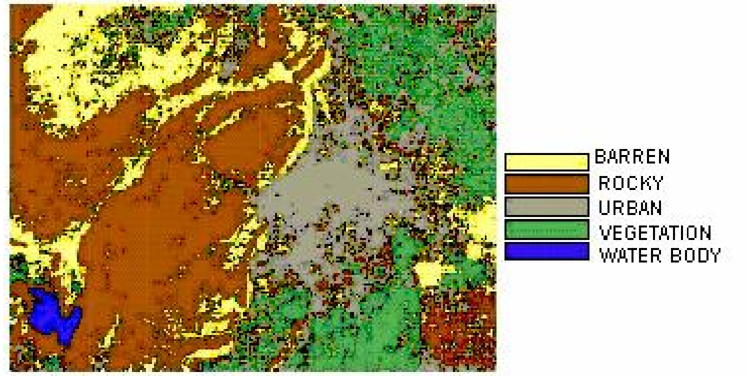
From the above discussion, it is evident that the Hybrid ACO2/PSO/BBO based classifier is a much efficient classifier as compared to the traditional probabilistic classifiers such as the MDMC and MLC. However, this Hybrid ACO/PSO/BBO technique also produces comparable results with the other recent soft computing classifiers. The Satellite image classification results of the recent soft computing classifiers when applied on the same Alwar Image , are shown below. For the purpose, Rough Sets, Fuzzy Sets, Rough-Fuzzy Tie-up, Ant Colony Optimization, Particle Swarm Optimization and Biogeography Based Optimization methods and their hybrids such as the hybrid ACO-BBO Classifier, Rough –Fuzzy Tie Up and hybrid ACO2/PSO Classifier are analyzed in the paper . Also, semantic-based image classification is added, as a special instance. Fig 6.9(a) shows the Fuzzy Classification of Alwar region which has a Kappa –Coefficient of 0.9134. Fig 6.9(b) presents the results of an integrated Rough –Fuzzy Tie Up Approach which has a Kappa Coefficient of 0.9700. Fig 6.9(c) applies the cAntMiner Algorithm on the Alwar Region which has a Kappa Coefficient of 0.964. Fig 6.9(d) shows the result of applying the hybrid ACO-BBO Technique on the Alwar Image which has a Kappa-Coefficient of 0.96699. Fig 6.9 (e) applies the Hybrid

ACO2/PSO Classifier which has a Kappa Coefficient of 0.975. Fig 6.9 (f) presents the results of the Semantic Web Based Classifier on the image with a Kappa Coefficient of 0.9881.

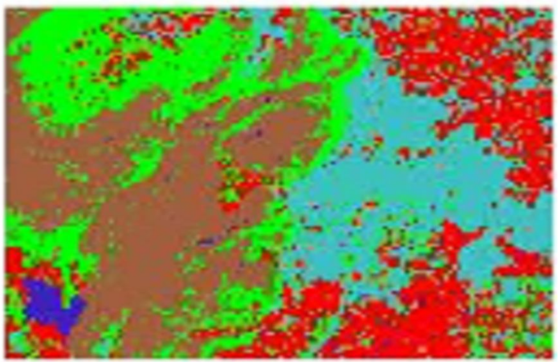
The Table 6.7 below compares the Kappa Coefficients of the Soft Computing Classifiers v/s the Traditional Probabilistic Classifiers .From the Table, it is clearly reflected that Soft Computing Classifiers are much more refined & efficient than the Probabilistic Classifiers . Also, it is reflected from the results of classification that our proposed hybrid classifier is the best known classifier from amongst all the present soft computing classifiers, its performance being only slightly less than the semantic web based classifier.



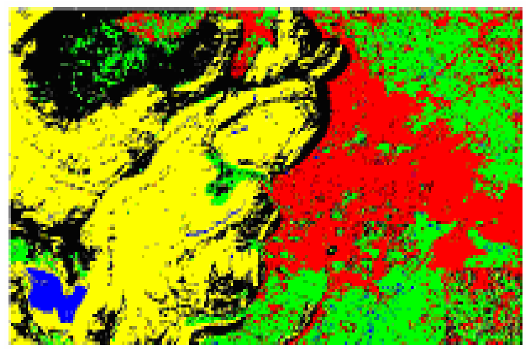
(a) Fuzzy Classification of Alwar Region
(Kappa Coefficient=0.9134)



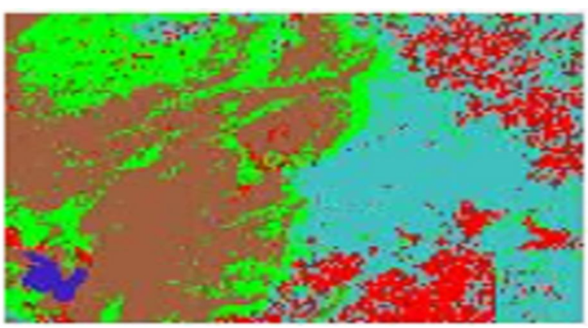
(b) Rough-Fuzzy Tie Up
(Kappa Coefficient=0.9700)



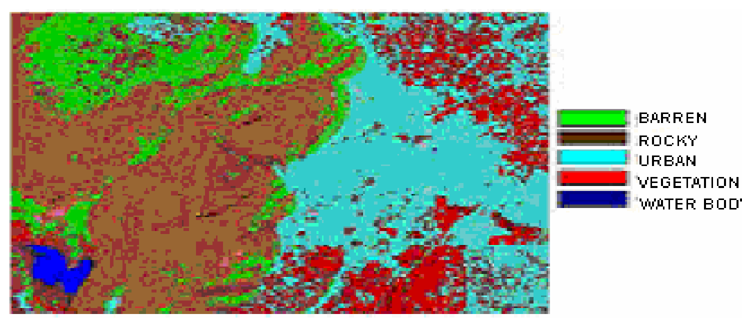
(c) cAntMiner Algorithm
(Kappa Coefficient=0.964)



(d) Hybrid ACO-BBO Algorithm
(Kappa Coefficient=0.96699)



(e) Hybrid ACO2/PSO Algorithm
(Kappa Coefficient=0.975)



(f) Semantic Web Based Classifier
(Kappa Coefficient=0.9881)

Fig 6.9. Classified Images of Alwar Region after applying various Soft Computing Techniques

Table 6.7. Kappa Coefficient (k) of Soft Computing Classifiers v/s Probabilistic Classifiers

Minimum Distance Mean Classifier (MDMC)	Maximum Likeli-hood Classifier (MLC)	Fuzzy set	Rough-Fuzzy Tie up	cAnt-Miner	Hybrid ACO2/ PSO	Semantic Web Based Classifier	Biogeo-graphy Based Classifier	Hybrid ACO-BBO Classifier	Hybrid ACO2/ PSO/BBO Classifier
0.7364	0.7525	0.9134	0.9700	0.964	0.975	0.9881	0.6715	0.96699	0.98182

(Probabilistic Classifiers) (Soft Computing Classifiers)

Technology Growth →

Chapter 7

PUBLICATIONS FROM THESIS

During the period of working over this project we interacted with International community working on Artificial Intelligence , in particular the founder of the key technique addressed in the thesis, (called the Biogeography Based Optimization) Dan Simon (Professor, Cleveland State University). We discussed our approach for representing knowledge with him and collected reviews and worked over the suggestions sent to us. Two research papers have been accepted in International conferences/journals for presentation and also published in their proceedings. Third paper has been communicated to the Applied Soft Computing Journal of Elsevier publications and is currently under review.

First paper focuses on proposing a hybrid of the Ant Colony Optimization with the BBO technique for predicting the deployment strategies of enemy troops in a military terrain application. This paper can be of major benefit to the commanders in the battlefield for intelligent preparation of the battlefield and is also under practical implementation at DTRL, DRDO. The second paper embeds expert knowledge into the bio-inspired classifier and presents an adaptive approach towards focused satellite image classification. This paper is currently under experimental study by a Ph.D student working in collaboration with DTRL, DRDO and is likely to be used as the base paper for an ongoing research at DRDO. The third paper presents our hybrid bio-inspired pattern analysis based intelligent classifier for focused land cover feature extraction and this paper is currently under review by Dan Simon, the pioneering author of the BBO technique . He is a senior member of IEEE and also professor in Cleveland State University, Ohio, USA.

7.1 Details of Conference:

1. Conference Name*:* International Multi Conference on Intelligent Systems & Nanotechnology (IISN-2010) **(Published).**

URL: http://www.istk.org

Paper Title: “Hybrid ACO-BBO Approach for Predicting the Deployment Strategies of Enemy Troops in a Military Terrain Application”

Authors: Lavika Goel, Dr. V. K. Panchal, Dr. Daya Gupta, Col. Rajiv Bhola

Location: Klawad, Haryana.

Conference date: February 26-28, 2010

Publisher/Proceedings: The paper is included in the conference proceedings, which has an ISBN number. The proceedings will also be submitted for several database indexes.

7.2. Details of journal

1. Journal name: International Journal of Computer Science & Information Security (IJCSIS Vol 8, No. 2, May 2010), ISSN: 1947-5500 , USA. **(Published)**

URL: http://sites.google.com/site/ijcsis/

Paper Title: “Embedding Expert Knowledge to Hybrid Bio-Inspired Techniques- An Adaptive Strategy Towards Focussed Land Cover Feature Extraction”

Authors: Lavika Goel, Dr. V. K. Panchal, Dr. Daya Gupta.

Location: USA.

Publisher/Proceedings: IJCSIS Publications, ISSN: 1947-5500 [Copyright © 2010 -2011, IJCSIS, USA] . IJCSIS Publications Indexed @ *Google Scholar,* @ *SCIRUS* , @ *ScientificCommons* and @ *DOAJ.*

2. **Journal Name:** International Journal of Applied soft Computing , The Official Journal of the World Federation on Soft Computing (WFSC), Elsevier Publications. **(Communicated)**

URL:[**http://www.elsevier.com/wps/find/journaldescription.cws_home/621920/description**](http://www.elsevier.com/wps/find/journaldescription.cws_home/621920/description)

Paper Title: " Multi-Spectral Satellite Image Classification using Improved Hybrid Bio-Inspired Techniques for extraction of Land Cover Features"

Authors: Lavika Goel, Dr. V. K. Panchal, Dr. Daya Gupta.

Location: Headquartered in Amsterdam, Netherlands.

Publisher/Proceedings: Elsevier Publications, ISSN: 1568-4946 , **Impact Factor:** 1.909, Copyright © 2010, Elsevier B.V.

Chapter 8

CONCLUSION AND FUTURE SCOPE

8.1 Conclusion

In this thesis, we have proposed a hybrid bio-inspired pattern analysis based intelligent classification algorithm as an efficient land cover classifier for the satellite image. Many of the land cover features are identified much more clearly when proposed algorithm is used. By using this approach we are able to classify the satellite image according to different areas like water, urban, vegetation, barren and rocky region with different colors assigned to each feature's pixels. It can be concluded from the results that the theory of biogeography can be efficiently used to design classificatory algorithm whereas originally it was used for solving the optimization problems only. The accuracy of the results is checked by creating an error matrix and Kappa (KHAT) coefficient has been generated from error matrix. The Kappa (KHAT) –coefficient has been used as a measure of elicited knowledge. This also provides us a mechanism to compare the knowledge content with other paradigms of supervised classification.

It can also be concluded that discrepant uncertainties inherent in satellite remote sensing images for geospatial features classification can be taken care of by use of soft computing techniques effectively. For the purpose, Rough Sets, , Fuzzy Sets, Rough-Fuzzy Tie-up, Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO) and Biogeography Based Optimization (BBO) methods are analyzed in the paper . Also, semantic-based image classification is added, as a special instance. Decision system required for any supervised classification can be made consistent and free from indecisive regions by using this spectrum of methods. The Landcover Classification is taken as a case study. It is perceived, from this research, that Kappa coefficient, a well founded metric for assessing accuracy of classification in remote sensing community, may be used for comparative study of the results from soft computing methods.

This thesis presents a novel approach wherein BBO can be combined with ACO/PSO to solve the Image Classification problems in remote sensing for feature extraction from high resolution multi-spectral satellite images .BBO can be used for further refinement of the image classified by simple ACO algorithms such as the cAntMiner Algorithms , since BBO refines its solutions probabilistically after each iteration unlike ACO/PSO which produces new solutions with each iteration and also it is particularly flexible to incorporate the expert knowledge for a more focused image classification. Hence using a combination of the two techniques i.e. the ACO2/PSO and BBO Technique, can be of major benefit.

8.2 Future Scope

The future scope of the research includes proposing certain modification to the algorithm so that the Kappa coefficient can be improved further. Also, an unsupervised version of the Biogeography based satellite image classification may be explored by including the context information of clusters used in generation of species granules and then evaluating the HSI. The current system is implemented using a simple heuristic technique namely standard deviation; the system performance can be increased by using other heuristic functions.

Further, in this thesis we have used rough set theory to obtain the elementary classes (or species). Since the band used in partitioning affect the final results, the rough set theory can be replaced by other swarm based clustering techniques, in this way the proposed feature extraction algorithm will become independent of the bands and hence, the algorithm will become completely 'swarm based'.

The results presented are preliminary and there is a lot of scope for improvement to develop this algorithm as an ideal classifier. In future, the algorithm efficiency can be further improved by lowering the threshold value used in BBO algorithm thus leading to more iterations and refined results. Also, we can further divide the image into more clusters so that a more accurate comparison can be made and the decision about which of the two techniques to be applied on the particular cluster , can be further streamlined. The system performance can be further increased by using better unsupervised classifications and better training sets.

Chapter 9

REFERENCES

1. Simon, D. "A Probabilistic Analysis of a Simplified Biogeography-Based Optimization Algorithm", submitted to IEEE Transactions on Evolutionary Computation for review (currently available at: http://academic.csuohio.edu/simond/bbo/ simplified/bbosimplified.pdf), 2009.
2. Simon, D., Du, D. and Ergezer, M., "Biogeography - Based Optimization Combined with Evolutionary Strategy and Immigration Refusal", to be published in the proceedings of IEEE International Conference on System ,Man and Cybernetics, San Antonio, U.S.A. (currently available at: http://embeddedlab.csuohio.edu/ BBO/BBO_Papers/mbbo.pdf), 2009.
3. 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.
4. Ma, H. and Chen, X., "Equilibrium Species Counts and Migration Model Tradeoffs for Biogeography-Based Optimization", to be published in the proceedings of 48th IEEE Conference on Decision and Control, Shanghai, China (currently available at: http://embeddedlab.csuohio.edu/BBO/BBO_Papers/HaipingMa.pdf), 2009.
5. Simon, D., "Biogeography-based optimization", IEEE Transactions on Evolutionary Computation, vol. 12, no. 6, pp. 702-713, December 2008.
6. Darwin, C. , "The Origin of Species", Sixth Edition, Gramercy Books, New York, pp.343-396, 1995.

7. Macarthur, R. and Wilson, E., “The Theory of Biogeography”, First Edition, Princeton University Press, New Jersey, pp.19-67, 1967.
8. Shelly Bansal, Daya Gupta, V.K. Panchal ,Shashi Kumar, “Remote Sensing Image Classification by Improved Swarm Inspired Techniques” in International Conference on Artificial Intelligence and Pattern Recognition (AIPR-09), Orlando, FL, USA ,July 13-16,2009 .
9. Fernando E.B. Otero, Alex A. Freitas, and Colin G. Johnson, cAnt-Miner: An Ant Colony Classification Algorithm to Cope with Continuous Attributes in Springer-Verlag Berlin, Heidelberg 2008.
10. Omkar, S.N., Manoj, K.M., Mudigere, D. and Muley, D., “Urban Satellite Image Classification using Biologically Inspired Techniques”, In Proceedings of IEEE International Symposium on Industrial Electronics, Vigo, Spain, pp. 1767 – 1772, 2007.
11. Piatrik, T. and Izquierdo, E., “Image Classification using an Ant Colony Optimization approach”, In Proceedings of First International Conference on Semantic and Digital Media Technologies, Athens, Greece, pp.159-168, 2006.
12. N. Holden and A.A. Freitas. A hybrid particle swarm/ant colony algorithm for the classification of hierarchical biological data. In: Proc. 2005 IEEE Swarm Intelligence Symposium (SIS-05), pp. 100-107, IEEE, 2005.
13. Song, Q., Guo, P. and Jia, Y., “Ant Colony Optimization algorithm for remote sensing image classification using combined features”, In Proceedings of International Conference on Machine Learning and Cybernetics, Kunming, China, vol. 6, pp. 3478-3483, 2005.
14. M. Dorigo and T. Stuetzle. Ant Colony Optimization. MIT Press, 2004.

15. R.S. Parpinelli, H.S. Lopes and A.A. Freitas. Data Mining with an Ant Colony Optimization Algorithm, IEEE Trans. On Evolutionary Computation, special issue on Ant Colony algorithms, pp. 321-332, Aug 2002.
16. J. Kennedy and R. Mendes, Population structure and particle swarm performance. Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2002), Honolulu, Hawaii USA. 2002.
17. Bonabeau, E., Dorigo, M. and Theraulaz, G., "Swarm Intelligence from Natural to Artificial System", First Edition, Oxford University Press, U.S.A., pp.1-24, 1999.
18. D.J. Hand. Construction and Assessment of Classification Rules. Wiley, 1997.
19. Dorigo, M., Maniezzo, V. and Colorni, A., "The Ant System: Optimization by a colony of cooperating agents", IEEE Transactions on Systems, Man, and Cybernetics–Part B, vol.26, no.1, pp.1-13, 1996.
20. Kennedy, J. and Everhart, R.C.,"Particle Swarm Optimization", In Proceedings of the IEEE International Conference on Neural Networks,1995, Perth, Australia, vol. 4, pp. 1942-1948, 1995.
21. Panchal VK, Singhal Naresh, Kumar Shashi, Bhakna Sonam, Rough-Fuzzy Sets Tie-Up for Geospatial Information Proceedings of International Conference on Emerging Scenarios in Space Technology and Applications (ESSTA2008), SSTA2008), Chennai, India, vol-I, 2008.
22. Pal, S.K. and Skowron, A. , "Rough Fuzzy Hybridization: A new Trend in Decision Making", First Edition, Springer-Verlag publishers, New York, pp.3-23, 1999.
23. Pawlak, Z., "Rough Set Theory and its Applications to Data Analysis",Cybernetics and Systems 29(7): 661-688 , 1998.

24. Pawlak, Z. , “Rough set approach to knowledge-based decision support”, European Journal of Operational Research, vol. 99, no. 1, pp. 48-57, 1997.
25. Ǿhrn, A. and Komorowski, J., ROSSETA “ A Rough Set tool kit for analysis of data” ,in roc.3rd International Joint Conference on information Sciences, Vol ,Durham,NC,March 1997.
26. Zadeh,L.A.,"Fuzzy sets," Information and Control, vol. 8, no. 3, pp.338- 353, June 1965.
27. Thomas M. Lillesand , Ralph W. Kiefer, “Remote Sensing and Image Interpretation”, 6th Edition, Wiley, 2008.
28. DI Kaichang, LI Deren, LI Deyi, Remote Sensing Image Classification With Gis Data Based On Spatial Data Mining Techniques in proceedings of IEEE, 2007.
29. Long, W., III; Srihann, S, Geoscience and Remote Sensing Symposium, Unsupervised and supervised classifications: Land cover classification of SSC image: unsupervised and supervised classification using ERDAS Imagine, IGARSS '04. Proceedings, Volume 4, 20-24, Sept 2004.
30. Lillesand, T.M., Kiefer, R.W. and Chipman, J.W. , “Remote Sensing and Image Interpretation”, Fifth Edition, Wiley & Sons Ltd., England, pp.586-592, 2003.
31. Canada Center for Remote Sensing, “Fundamentals of Remote sensing”, pg no. 1-164.
32. Sonal Kumar, Daya Gupta, ,V.K. Panchal, Shashi Kumar, “Enabling Web Services For Classification Of Satellite Images”, in 2009 International Conference on Semantic Web and Web Services (SWWS'09), Orlando, FL, USA ,July 13-16,2009.
33. Gupta, Daya, Kartar, J, “ Knowledge Representation with ontology and Relational database to RDF converter”, In the proceedings of EISWT-08, ISBN : 978-1-60651-003-02, pp 141-146, 2008.

34. Teodorović, D. and dell'orco M. , "Bee colony optimization: A cooperative learning approach to complex transportation problems", In Proceedings of 16th Mini-EURO Conference on Artificial Intelligence Methods in Transportation, Poznan, Poland, pp. 51-60, 2005.

35. Passino, K.M., "Biomimicry of Bacterial Foraging for Distributed Optimization and Control", IEEE Control Systems Magazine, vol. 22, no. 3, pp.52 –67, 2002.

36. ERDAS,/http://www.ERDAS.com/

37. MATLAB, http://www.mathworks.com/

Appendix A

Abbreviations

DN - Digital number

ACO- Ant Colony Optimization

BBO- Biogeography Based Optimization

PSO- Particle Swarm Optimization

DE- Differential Evolution

MDMC- Minimum Distance to Mean Classifier

MLC- Maximum Likelihood Classifier

ES- Evolutionary Strategy

BCO- Bee Colony Optimization

GIS-Geographic Information Systems

SC-Soft Computing

AS-Ant System

TSP- Travelling Salesman Problem

SAR- Synthetic Aperture Radar

GA- Genetic Algorithm

FL- Fuzzy Logic

RS- Rough set theory

NN- Neural Network theory

PR-Probabilistic Reasoning

FCM- Fuzzy c-means

ROSETTA- Rough set toolkit Theory Toolkit

NIR- Near Infra-Red

MIR- Middle Infra-Red

RS1- Radarsat-1

RS2- Radarsat-2

DEM- Digital Elevation Model

LISS-Linear Imaging Self- Scanning Memory

RCBBO- Real Coded Biogeography Based optimization

HSI- Habitat Suitability Index

SIV- Suitability Index Variables

DPSO- Discrete Particle Swarm Optimization

Appendix B

An Introduction to MATLAB

URL: http://mathworks.com/

MATLAB is a high performance language for technical computing. It integrates computation, visualization and programming in an easy-to-use environment where problems and solutions are expresses in familiar mathematical notation. Typical uses include:

1. Math and computation
2. Algorithm, simulation and prototyping
3. Modeling, simulation and prototyping
4. Data analysis, exploration and visualization
5. Scientific and engineering and visualization
6. Application development, including graphical user interface building

MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. This allows you to solve many technical computing problems, especially those with matrix and vector formulations in a fraction of the time it would take to write a program in a scalar non interactive language such as C or FORTRAN.

The name MATLAB stands for Matrix Laboratory. MATLAB was originally written to provide easy access to matrix software developed by the LINPACK and EISPACK projects.

Today MATLAB uses software developed by the LAPACK and ARPACK projects, which together represent the state-of-the-art in software for matrix computation.

MATLAB has evolved over a period of years with input from many users. In university environments, it is the standard instructional tool for introductory and advanced courses in mathematics, engineering and science. In industry, MATLAB is the tool of choice for high-productivity research, development and analysis.

MATLAB features a family of application-specific solutions called toolboxes. Very important to most users of MATLAB toolboxes allow you to learn and apply specialized technology. Toolboxes are comprehensive collections of MATLAB functions (M-files) that extend the MATLAB environment to solve particular classes of problems. Areas in which toolboxes are available include signal processing, control system neural networks, fuzzy logic, wavelets, simulation and many others.

FILE TYPES

MATLAB can read and write several types of files. There are mainly five different types of files used in MATLAB which is used for storing data or programs.

- **M-FILES**- They are the standard ASCII files, with a .m extension to the file name. There are basically two types of files and they are SCRIPT and FUNCTION file. In general, mostly MATLAB files are saved as M-FILES.
- **MAT-FILES**- They are the binary data-files, with a .mat extension to the filename. These files are created when you save the MATLAB data with the save command. The data which you save in MATLAB can only be read by mat lab as it save in a

special format.

• **FIG-FILES-** They are the binary figure-file, with a .fig extension to the filename. Such files are created by saving a figure in this format by using the save and save as option in it. These files basically create all kind of information which is used for again recreating a figure and can be opened by filename .fig.

• **P-FILES**- These are the compiled M-File , with a .p extension to the filename. These file can be executed directly without using any compiler and parsed in it. These files are created with the P-CODE command.

• **MEX-FILES-** These are MATLAB-callable Fortran and C program , with the .mex extension to the filename. Use of these file require some experience in MATLAB and lot of patience in it .

Appendix C

An Introduction to ERDAS

URL: http://www.ERDAS.com/

Overview

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

Appendix D

An Introduction to Open Source Tool for PSO/ACO2

URL: http://www.PSOACO2.sourceforgenet.com/

PSO/ACO2 uses a hybrid Particle Swarm Optimization/Ant Colony Optimization 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 optimizer itself is sound, but it may produce sub-optimal results.

Version 0.95 feature creeps into Differential Evolution support, that very popular optimizer. Code adapted from Java implementation by Mikal Keenan and Rainer Storn. Note that it probably wont work well with a mix of continuous and nominal attributes, due to problems integrating the PSO/ACO2 nominal optimizer 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. With tic-tac-toe precision works really well because there are "smaller" patterns, with breast-cancer sensitivity works better. Overall precision seems to perform better.

Appendix E

An Introduction to Rosetta : A Rough Set Toolkit

URL: http://www.lcb.uu.se/tools/rosetta/

An overview of the ROSETTA GUI is given by Øhrn . The following ROSETTA GUI details are worth noting:

• A decision system can be read into a new ROSETTA project by selecting Open... from the main File menu, and will be placed immediately below the root of the Structures node in the project tree.

• Branches in the project tree can be expanded or collapsed by left-clicking on the "_" or "_" symbols next to the icons.

• Right-clicking on an icon in the project tree brings up a pop-up menu for that object. In the following, the symbol "I" will be used to denote menu navigation.

• Left-clicking twice on an icon in the project tree can be used as a shortcut for viewing that object in detail.

• Grayed columns in views of decision systems indicate that the corresponding attributes are "masked away" and subsequently ignored by the ROSETTA kernel in any analysis steps. Missing values are indicated by the string Undefined.

• Rules can be sorted directly in their views by right-clicking the column to sort by.

• To rename an object, first left-click once on its icon to select it. Then left-click once more on the icon's label. The icon's label is edited directly in place.

• To view progress messages and warnings, select Messages from the main View menu. A document has to be present before the View menu appears.