A STUDY OF VARIOUS FEATURE SELECTION TECHNIQUES IN DIGITAL REMOTE SENSING ANALYSIS

A DISSERTATION

submitted in partial fulfilment of the
requirements for the award of the degree
of
MASTER OF ENGINEERING
in
CIVIL ENGINEERING
(With Specialization in Remote Sensing and Photogrammetric Engineering)

DEEPAK K

DEPARTMENT OF CIVIL ENGINEERING
UNIVERSITY OF ROORKEE
ROORKEE-247 667 (INDIA)

JANUARY, 1994

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the dissertation entitled, "A Study of Various Feature Selection Techniques in Digital Remote Sensing Analysis", in partial fulfilment of the requirements for the award of the degree of MASTER OF ENGINEERING in Civil Engineering with specialization in REMOTE SENSING AND PHOTOGRAMMETRIC ENGINEERING, submitted in the DEPARTMENT OF CIVIL ENGINEERING, UNIVERSITY OF ROORKEE, ROORKEE, is an authentic record of my own work carried out for a period of about seven months during July, 1993 to January, 1994 under the supervision of Dr. P.K. Garg, Lecturer and Dr. S.K. Ghosh, Lecturer, Department of Civil Engineering, University of Roorkee, Roorkee.

The matter embodied in this dissertation has not been submitted by me for the award of any other degree or diploma.

PLACE : ROORKEE

DATED : 31-01-1994 (DEEPAK KUMAR JHA)

This is to certify that the above statement made by the Deepak Kumar Jha is correct to the best of my knowledge.

(Dr. S.K. Ghosh)
Lecturer
Civil Engg. Deptt.
University of Roorkee
ROORKEE - 247 667

(Dr. P.K. Garg)
Lecturer
Civil Engg. Deptt.
University of Roorkee
ROORKEE - 247 667

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the dissertation entitled, **"A Study of Various Feature Selection Techniques in Digital Remote Sensing Analysis"**, in partial fulfilment of the requirements for the award of the degree of **MASTER OF ENGINEERING** in Civil Engineering with specialization in **REMOTE SENSING AND PHOTOGRAMMETRIC ENGINEERING**, submitted in the **DEPARTMENT OF CIVIL ENGINEERING, UNIVERSITY OF ROORKEE, ROORKEE**, is an authentic record of my own work carried out for a period of about seven months during July, 1993 to January, 1994 under the supervision of **Dr. P.K. Garg**, Lecturer and **Dr. S.K. Ghosh**, Lecturer, Department of Civil Engineering, University of Roorkee, Roorkee.

The matter embodied in this dissertation has not been submitted by me for the award of any other degree or diploma.

PLACE : ROORKEE

DATE : 31-01-1994

Deepak Kumar Jha.
(DEEPAK KUMAR JHA)

This is to certify that the above statement made by Mr. Deepak Kumar Jha is correct to the best of my knowledge.

(Dr. S.K. Ghosh)	P.K.Garg 30/1/94 (Dr. P.K. Garg)
Lecturer	Lecturer
Civil Engg. Deptt.	Civil Engg. Deptt.
University of Roorkee	University of Roorkee
ROORKEE - 247 667	ROORKEE - 247 667

ACKNOWLEDGEMENT

It is a distinct pleasure to express my deep sense of gratitude and indebtedness to Dr.R.K.Garg, Lecturer and Dr S.K.Ghosh, Lecturer, Department of Civil Engineering, University of Roorkee, for their valuable and encouraging guidance, joyful and friendly suggestions and patient review. Without their help and guidance this dissertation would have rather been impossible.

I also want to express my sincere thanks to Dr.R.S.Tiwari, Prof. and Co-ordinator, centre for Remote Sensing and R.S.P.E. Section and Sri Kamal Jain, O.C.M.C., R.S.P.E. Section for providing me ISROVISION and other necessary facilities in the section. I also place on record my indebtedness to the faculty of R.S.P.E. Section, for their constant help during this work.

I am also thankful to my friends Mr. M.K.Sinha, Miss. Poonam Sharma, Mr. Sanjeev Jain, Mr. Ramesh Adury and Mr. Arvind Prasad for their inspiration and encouragement without which it would have been very difficult for me to complete this course.

Special thanks are due to my friends Mr. P.Rajesh, Mr. Santosh Kumar and Mr. Neeraj Manglik who helped me at each and every step of this work. Thanks are also due to Mr. A.K.Nigam, Mr. Vijay Veer and Mr. Pratap Singh (Photo tech.) for their help whenever needed. I am also thankful to all those who helped me directly or indirectly during my course of study.

([illegible])

ACKNOWLEDGEMENTS

It is a distinct pleasure to express my deep sense of gratitude and indebtedness to **Dr. P. K. Garg**, Lecturer and **Dr. S. K. Ghosh**, Lecturer, Department of Civil Engineering, University of Roorkee, for their valuable and encouraging guidance, joyful and friendly suggestions, and patient review. Without their help and guidance this dissertation would have rather been impossible.

I also want to express my sincere thanks to **Dr. R. S. Tiwari**, Prof. and Co-ordinators, centre for Remote Sensing and R.S.P.E. Section and **Sri Kamal Jain**, O.C.M.E., R.S.P.E. Section for providing me ISROVISION and other necessary facilities in the Section. I also place on record my indebtedness to the faculty of R.S.P.E. Section, for their consistent help during this work.

I am also thankful to my friends **Mr. K.K.Sinha, Miss. Bhawana Sharma, Mr. Sanjeev Jain, Mr. Romesh Adury and Mr. Arbind Prasad** for their inspiration and encouragement without which it would have been very difficult for me to complete this course.

Special thanks are due to my friends **Mr. P.Rajesh, Mr. Santosh Jha and Mr. Neeraj Manglik** who helped me at each and every step of this work. Thanks are also due to **Mr. A.K.Nigam, Mr. Vijay Veer and Mr. Pratap Singh (Photo tech.)** for their help whenever needed. I am also thankful to all those who helped me directly or indirectly during my course of study.

Deepak Kumar Jha

(DEEPAK KUMAR JHA)

SYNOPSIS

The purpose of applying feature selection techniques to multispectral data is to provide a trade-off between the cost and the accuracy of classification in order to reduce the dimensionality of satellite data as well as computational time for the analysis. Feature selection undertakes the task of selecting a subset of bands from available number of bands of sensor. These bands are selected on the basis of either separability measure or degree of overlap between the classes present in the area.

In the present study, [illegible] and the surrounding area lying [illegible] has been selected as the study area. [illegible] feature selection [illegible] Transformed Divergence, Bhattacharya Distance, Jeffreys Matusita Distance and Brightness Value Overlapping Index (BVOI) have been used employing digital satellite data of LANDSAT-5 TM, 1992. The study has been carried out in two stages. In the first stage feature selection techniques have been applied in order to find out the best combination of two and three bands. In second stage, the entire image has been classified into six classes viz water bodies, vegetation, dry sand, wet sand, urban areas and boulders using the best combination of bands. In addition the classification has been performed using the worst combination of bands. Accuracy assessment of the classification has also been [illegible]

SYNOPSIS

The purpose of applying feature selection techniques in multispectral data is to provide a trade-off between the cost and the accuracy of classification in order to reduce the dimensionality of satellite data, as well as computational time for the analysis. Feature selection undertakes the task of selecting a subset of bands from available number of bands of a sensor. These bands are selected on the basis of either separability measure or degree of overlap between the classes present in the area.

In the present study, Patna and its surrounding area lying in the Bihar State of India has been selected as the study area. Various feature selection techniques such as Divergence, Transformed Divergence, Bhattacharya Distance, Jeffreys Matusita Distance and Brightness Value Overlapping Index (BVOI) have been used employing digital satellite data of LANDSAT-5 TM, 1992. The study has been carried out in two stages. In the first stage, feature selection techniques have been applied in order to find out the best combination of two and three bands. In second stage, the entire image has been classified into six classes viz. water bodies, vegetation, dry sand, wet sand, urban areas and boulders using the best combination of bands. In addition, the classification has been performed using the worst combination of bands. Accuracy assessment of the classification has also been carried using error matrix. The results show that classification accuracy improves significantly when the best combination of spectral bands is used. The BVOI technique is found to be the best when only combination of best bands is to be determined for the classification. If inter class separability is to be ascertained, then Transformed Divergence and Jeffreys Matusita Distance are the most useful techniques of feature selection.

It is suggested that the areas having same types of classes must be further examined in order to establish the combination of specific band on a National/Global basis.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

The classification of satellite data is based on certain selected measurements, known as "FEATURES". In context of image processing, the term "Feature Selection" has specialised meaning 'Features 'are not geographical features visible on the image, but are rather statistical characteristics of image data-individual band or combination of bands value, that carry information concerning systematic variation in the scene. Thus, feature selection may also be called as 'Information Extraction' i.e. isolation of the statistical components within multispectral data that are most useful in portraying the essential elements of an image (Campbell, 1987). Hence, a feature should be referred as 'Useful Information ' in the image data, rather than as a physical feature present on the earth's surface. As discussed, 'Feature Selection' is the process of isolating the most useful bands of the satellite data set for further analysis while discarding the less useful aspects (i.e. errors, noise, redundancy etc.) (Campbell, 1987).

1.2 UTILITY OF FEATURE SELECTION

In analyzing the multispectral remote sensing data, feature selection plays an important role. For example, processing the multiband satellite data by a computer, cost will be substantially high if all the available bands are taken into consideration.

Feature selection techniques provide a trade-off between the cost of classification and accuracy with considerable reduction in computation time by selecting the optimum combination of bands containing maximum information about the land cover classes (Kamat, 1978). In principle, feature selection attempts to identify and then eliminate those bands which carry repetitive information as in other bands. Resulting satellite data set now obtained, contains maximum information (Campbell, 1987).

1.3 OBJECTIVE OF THE PRESENT STUDY

Various feature selection techniques based on different mathematical relationships are available to select the optimum combination of bands from the total number of available bands of a sensor. Unfortunately, no proper assessment regarding the utility of such methods have been carried out still now. In this study, it is proposed to carry out an investigation to assess the utility of all the feature selection techniques, viz. Divergence, Transformed Divergence, Bhattacharya Distance, Jeffreys Matusita Distance and Brigntness Value Overlappng Index (BVOI).

Other objectives of this study are defined as follows :

(i) To develop computer programs of all the feature selection techniques in order to differentiate one land cover class from the others.

(ii) To identify the best two bands and best three bands of LANDSAT-TM data for land use/land cover classification using feature selection techniques in the study area, and

(iii) To assess and compare the classification accuracy of land use/land cover classes identified from best combination of TM bands.

1.4 ORGANISATION OF THE THESIS

The whole work has been presented in seven chapters as described below.

Chapter 2 deals with the role of feature selection techniques for remote sensing data analysis. Chapter 3 describes the mathematical relationship of various feature selection techniques, which can be used for the selection of best bands combination. Chapter 4 gives the information regarding the study area and the satellite data used. Chapter 5 outlines the methodology adopted and the analysis procedure of satellite data using feature selection techniques. Chapter 6 describes the analysis of satellite data and their results, which is followed by discussions, while in chapter 7 the conclusions and recommendations of the study are given.

CHAPTER 2

ROLE OF FEATURE SELECTION TECHNIQUES FOR REMOTE SENSING DATA ANALYSIS

Remote sensing data product may be used to extract the useful and update information about land classes by the following two methods of analysis :

(a) Manual interpretation

(b) Digital interpretation

2.1 MANUAL INTERPRETATION

This method of image analysis makes use of some interpretation elements, viz. tone, texture, location, pattern, shape and shadow, alongwith the analyst decision, to analyze hardcopy remote sensing data. However, this approach of analysis is not so effective because the perception of a human being is limited only to 10-15 gray level.

2.2 DIGITAL INTERPRETATION

The digital method of image analysis makes use of computer and remote sensing data to extract the useful information. The various approaches of digital classification are (Mather, 1987) :

(a) Supervised

(b) Unsupervised, and

(c) Hybrid

In Supervised classification, the location of the certain classes are known prior through aerial photographs, published maps and personal experience (Mather, 1987). The analyst demarcates specific sites of known identity on remote sensing data, known as training areas, which are then used for the classification of the entire image.

As compared to Supervised classification, Unsupervised classification does not require detailed ground reference data at the start of classification. In this technique, the computer groups the entire image into different spectral classes, depending upon their reflectance values. The analyst then assigns these spectral classes into informational classes based on ground reference data or knowledge of the area. In Hybrid classification, two or more classification algorithms may be used. For example, a Hybrid classifier having Parallelopiped and Maximum Likelihood algorithms, first uses the Parallelopiped and then transfers difficult cases to Maximum Likelihood classifier (Mather, 1987).

2.3 ROLE OF FEATURE SELECTION TECHNIQUES

Feature selection techniques are needed for the classification of remote sensing data by Supervised approach. In all the three algorithms of Supervised approach which will be explained in the next section, the training data requires as an input. The training data are collected either in the field or from reference data such as published maps, aerial photographs and literature. The training data are collected for each class to be identified from remote sensing technique. These classes are located on the digital remote sensing data and training statistics computed. The training statistics data are then used by various feature selection techniques to select an optimum subset of bands from all the available bands of a sensor. This optimum combination of bands thus reduces the cost of classification of satellite data, computation time and storage space without compromising the accuracy of classification.

Thus, feature selection techniques play an important role for the classification of digital satellite data, particularly if the available bands from a sensor are large in number e.g. LANDSAT-TM sensor data.

Various feature selection techniques are explained in the next chapter.

2.4 INFORMATION REQUIREMENTS OF SUPERVISED CLASSIFICATION ALGORITHMS

The Supervised classification algorithms broadly requires input data in the form of mean, standard deviation, minimum and maximum brightness values and variance-covariance matrix derived from each class present on the image (Mather, 1987).

The various Supervised classification algorithms are:

(i) Minimum Distance

(ii) Parallelopiped

(iii) Maximum Likelihood

In Minimum Distance classifier, the input data required is the mean brightness value of the training data for each class in all spectral bands (Jensen, 1986). The Parallelopiped algorithm requires the mean value as well as standard deviation of the classes of interest in all spectral bands. The Maximum Likelihood classifier requires mean value and variance-covariance matrix of the training data for each class in each band (Jensen, 1986).

The Minimum Distance classifier is further elaborated as it has been used for the present study. The advantage of Minimum Distance clssifier is that it is computationally easier and faster. However, it does not produce very accurate results for class having high variance, as only the mean values are used by this classifier. In this classifier, the distance between each unknown pixel ($BV_{i,j,k}$) and mean vector ($\mu_{c,k}$) is calculated and the pixel is assigned to that class whose distance is minimum. The Minimum Distance can be calculated either by using Euclidian Distance based on Phythagorean rule or Round the Block Distance (Swain and Davis, 1978).

Euclidian Distance is given by (Thomas et.al., 1987)

$$E_L = \sqrt{\sum_{k=1}^{n} (BV_{i,j,k} - \mu_{c,k})^2} \qquad \ldots(2.1)$$

Where, $BV_{i,j,k}$ = Pixel value at location (i,j) in band k

$\mu_{c,k}$ = Mean value of class c in band k

n = Number of spectral bands used

Round the Block or L1 Distance is given by (Thomas et.al. 1987) as :

$$L1 = \sum_{k=1}^{n} |(BV_{i,j,k} - \mu_{c,k})| \qquad \ldots(2.2)$$

In the present study, the Euclidian Distance has been used for the computation of minimum distance because it is most accurate and basic method to determine the distance between any two points.

CHAPTER 3

[illegible]

3.1 INTRODUCTION

Once the training [illegible] collected a judgement [illegible] portions (bands) of [illegible] discriminating each class from the others. This process is commonly called "feature selection", which eliminates those bands from the analysis procedure that provide only redundant spectral information or very little information compared to the other bands. In this way the dimensionality (i.e., the number of bands) to be processed in the complete data set can be reduced significantly. This [illegible] minimises the cost of [illegible] space. Feature Selection involves a statistical analysis [illegible] determine the degree of separability between the two classes in the training data. Combinations of bands are normally ranked by feature selection techniques according to their potential ability to discriminate each class from the others using 'n' bands at a time (Jensen, 1986).

Statistical methods of feature selection are used to quantitatively select the subset of bands from all the bands of a sensor that provide the greatest degree of statistical separability between any two classes 'c' and 'd'. The basic problem of spectral pattern recognition is (Jensen, 1986):

CHAPTER 3

VARIOUS FEATURE SELECTION TECHNIQUES

3.1 INTRODUCTION

Once the training statistics for each class in each band is collected, a judgement has to be made in order to determine those spectral bands of a sensor that are most effective in discriminating each class from the others. This process is commonly called 'Feature Selection', which eliminates those bands from the analysis procedure that provide only redundant spectral information or very little information compared to the other bands. In this way the dimensionality (i.e. the number of bands to be processed) of the complete data set can be reduced significantly. This, in turn, not only minimizes the cost of analysis but also it reduces the computational time and storage space. Feature Selection involves a statistical analysis to determine the degree of separability between the two classes in the training data. Combinations of bands are normally ranked by feature selection techniques according to their potential ability to discriminate each class from the others using 'n ' bands at a time (Jensen, 1986).

Statistical methods of feature selection are used to quantitatively select the subset of bands from all the bands of a sensor that provide the greatest degree of statistical separability between any two classes 'c' and 'd'. The basic problem of spectral recognition is (Jensen, 1986) :

"Given a spectral distribution of data in 'n' bands of remotely sensed data, finding discrimination techniques that will allow separation of major land cover classes with a minimum of error and a minimum number of bands".

This problem is demonstrated diagrammetically in Fig. 3.1, using two classes in single band data.

On examining a typical histogram shown in Fig. 3.2, it is found that there is substantial overlap betwen classes 1 and 4 in band 'a ' and between classes 3 and 4 in band 'b '. When there is an overlap, any decision rule that one could make to separate or distinguish between two classes must be concerned with the following two types of errors :

(i) A pixel may be assigned to a class to which it does not belong (an error of commission).

(ii) A pixel is not assigned to its appropriate class (an error of omission).

Both types of errors can be minimised by selecting an optimum subset of bands and applying appropriate classification techniques. If the training data for each class in each band are assumed to be normally distributed as depicted in Fig. 3.1, separability measures can be used to identify the optimum subset of bands for classification purpose.

3.2 FEATURE SELECTION TECHNIQUES

Literature reveals that the commonly used separability measures for feature selection are :

(a) Divergence

(b) Transformed Divergence

(c) Bhattacharya Distance

(d) Jeffreys Matusita Distance, and

(e) Brightness Value Overlapping Index (BVOI).

3.2.1 Divergence (DIV_{cd})

Divergence was one of the first measures of statistical separability beetween the two classes used in the machine processing of remote sensor data, and is still widely used as method of feature selection (Swain and Wacker, 1971; Haack, 1983). It addresses the basic problem of deciding combination of the best 'q' bands out of 'n' bands to be used in Supervised classification. The number of combinations 'C' of 'n' bands taken q bands at a time is defined as (Jensen, 1986) :

$$C\left(\frac{n}{q}\right) = \frac{n!}{q!\,(n-q)!} \qquad \ldots (3.1)$$

Thus, for six LANDSAT-TM bands (excluding thermal band), if the best three bands are to be used in the classification, the above equation yields 20 such combinations.

Divergence minimizes this problem by giving only one best combination. It is computed using the mean and variance-covariance matrices of class statistics as defined by the training data set. The degree of Divergence or 'separability' between two classes c and d (DIV_{cd}) can be computed as (Swain and King, 1973) :

$$DIV_{cd} = 0.5 \ Tr \left[(V_c - V_d) \ (V_d^{-1} - V_c^{-1}) \right]$$

$$+ \ 0.5 \ Tr \left[(V_c^{-1} + V_d^{-1}) (M_c - M_d)(M_c - M_d)^T \right] \qquad \ldots (3.2)$$

where,

Tr = Trace of a matrix (i.e, the sum of the diagonal elements)

c,d = Two classes used for separability analysis

Vc, Vd = The variance-covariance matrices of brightness value of the training data for the two classes, 'c' and 'd', under investigation, and

Mc, Md = Mean matrices of brightness values of the training data for the two classes c and d.

The sizes of the variance-covariance matrices 'V_c' & 'V_d' and the mean matrices 'M_c' & 'M_d' are function of the number of bands used in the training process (i.e., if three bands were trained upon, both 'V_c' & 'V_d' would be matrices of 3x3 dimension and both 'M_c' & 'M_d' would be matrices of 3x1 dimension).

Although, Divergence only provides a measure of the distance between any two class density, its use can be extended for multiclasses by taking the average of Divergence values over all possible pairs of classes, taking any two classes 'c' and 'd' at a time, while holding the subset of band 'q', constant. Another subset of bands 'q' is selected for the same 'm' classes and analyzed. The subset of bands having the maximum average Divergence is considered to be the superior set of bands for the classification purpose. Average Divergence can be expressed as (Swain and King, 1973):

$$DIV_{AVE} = \frac{\sum_{c=1}^{m-1} \sum_{d=c+1}^{m} DIV_{cd}}{C} \quad \ldots (3.3)$$

where C = Possible number of combinations for selecting a pair of classes.

Using equation (3.3), the bands subset 'q' with the highest average Divergence is selected as the most appropriate subset of bands to classify 'm' classes. The major practical problem with DIV_{cd} is that it continues to increase even after full class separability is attained.

3.2.1.1 Mathematical Properties of Divergence

The above expression then leads on to three mathematical properties of Divergence (Swain and Davis, 1978).

(i) $DIV_{cc} = 0$ The Divergence of one likelihood distribution relative to itself is zero (the classes are identical)

(ii) $DIV_{cd} > 0$ For two different likelihood functions the Divergence is always greater than zero.

(iii) $DIV_{cd} = DIV_{dc}$ Divergence is symmetrical between classes over the same n dimensional feature space.

More details about this technique are given in Swain and King (1973).

3.2.2 Transformed Divergence (DIV_{cd}^{T})

A nonlinear relationship between the classification accuracy and Divergence exists due to the unbound characteristics of this measure as shown in Fig. 3.3. Thus, a transformation has been applied to saturate DIV_{cd} measure to more closely represent the correct classification, as shown in Fig. 3.4, which yielded Transformed Divergence (DIV_{cd}^{T}) (Swain and Davis, 1978).

Transformed Divergence is calculated as :

$$DIV_{cd}^{T} = 2000\ [1 - EXP(-DIV_{cd} / 8)] \qquad \ldots (3.4)$$

This statistics gives an exponentially decreasing weight to increasing distance between the classes 'c' and 'd'. It also scales the Divergence value to lie between 0 and 2000. A Transformed Divergence value of 2000 suggests excellent class separtion. In between 1900 and 2000 provides good separation. If its value lies between 1700 and 1900, then the separation is fair, while below 1700 suggests poor separation (Swain and Davis, 1978).

The average Transformed Divergence value for a combination of bands and best bands combination for classification calculated using the process explained in section 3.2.1. The more details are given in Swain and Davis (1978).

3.2.3 Bhattacharya Distance (B_{cd})

The Bhattacharya Distance (B_{cd}) is another measure of the statistical separability between pairs of multivariate Gaussian distribution and is expressed as (Swain and King, 1973)

$$B_{cd} = \frac{1}{8}\ (M_c - M_d)^T \left[\frac{V_c+V_d}{2}\right]^{-1} (M_c - M_d) + \frac{1}{2}\ L_n\ \frac{\det\ (V_c+V_d)/2}{\sqrt{\det(V_c)}\ \sqrt{\det(V_d)}} \qquad \ldots(3.5)$$

where Mc, Md, Vc, Vd, T, c, and d are defined previously and det = Determinants of the matrix.

To select the best 'q' bands from the original 'n' bands for separability of 'm' classes, the Bhattacharya Distance is calculated between each of the $m(m-1)/2$ pairs of classes for all possible ways of choosing 'q' bands from 'n' dimensions. The best 'q' bands are those dimension whose average of the Bhattacharya Distance (B_{cd}) between $m(m-1)/2$ classes is the highest. The average of the Bhattacharya Distance can be expressed as (Swain and King, 1973);

$$B_{AVE} = \frac{2}{m(m-1)} \sum_{c=1}^{m-1} \sum_{d=c+1}^{m} B_{cd} \qquad \ldots(3.6)$$

It has been found that Bhattacharya Distance is more appropriate to inter class separability problems than the Divergence when the class probability distributions are broad. However, when the classes are well defined both the Bhattacharya

Distance and the Divergence approaches yield similar results (Thomas et. al., 1987).

The more details are given in Swain and King (1973).

3.2.4 Jeffreys Matusita Distance (J_{cd})

A saturating transformation applied to Bhattacharya Distance (B_{cd}), yields the Jeffreys Matusita Distance (J_{cd}), which is given by the following equation (Swain and King, 1973)

$$J_{cd} = 2\ [1 - \text{EXP}\ (-B_{cd})] \qquad \ldots(3.7)$$

This statistics also gives an exponentially decreasing weight to increasing distance between the classes c and d. Since the value of $\text{EXP}(-B_{cd})$ lies between 0 and 1, J_{cd} ranges from 0 to 2 with 2 corresponding to the largest separation. Swain et. al. (1973) observed that this saturating behaviour of J_{cd} is responsible for its utility as a feature selection criterion in multiclasses problem. For multiclasses problem the feature selection criterion is taken as (Swain and King, 1973) :

$$J_{AVE} = \frac{2}{m(m-1)} \sum_{c=1}^{m-1} \sum_{d=c+1}^{m} J_{cd} \qquad \ldots(3.8)$$

where J_{AVE} = Average Jeffreys Matusita Distance of all the classes in the number of bands used.

J_{AVE} values are then calculated for all possible combinations of bands and that subset of bands is chosen for classification which gives the maximum J_{AVE} value.

3.2.5 Brightness Value Overlapping Index (BVOI)

This recent approach to find the set of optimum spectral bands is based on the degree of overlap in brightness values between classes, called "Brightness Value Overlapping Index" (BVOI). The method is simpler as compared to those discussed earlier, mainly because of the requirement of less number of input parameters which reduce the computations time (Saha and Kudrat, 1991).

The following mathematical expressions are used for calculation of BVOI (Ma and Olson, 1989)

$$F_{j,k} = \sum^{N_{j,k}} f(x_{i,k}) \qquad \ldots (3.9)$$

$$F_{aj} = \frac{1}{M} \sum_{k=1}^{M} F_{j,k} \qquad \ldots (3.10)$$

$$F_{ta} = \sum_{j=1}^{N} F_{aj} \qquad \ldots (3.11)$$

Where $x_{i,k}$ = i^{th} brightness value within a class of band 'k'

$f(x_{i,k})$ = Frequency of brightness value $x_{i,k}$.

$N_{j,k}$ = The range of brightness value within class 'j'

$F_{j,k}$ = Cumulative frequency for class 'J' of band 'k'

M = Number of spectral bands,

F_{aj} = Average cumulative frequency over all bands of class 'J'

N = Number of classes in the data set, and

F_{ta} = Total average cumulative frequency over all classes.

If overlap does not exists amongst the classes in band 'k', then,

$$F_{tk} = \sum_{j=1}^{N} F_{j,k} = F_o = 100 \text{ percent} \quad ...(3.12)$$

$$F_{ta} = \frac{1}{M} \sum_{k=1}^{M} F_{tk} \quad ...(3.13)$$

so, F_{ta} is also the band average of F_{tk}'s

where F_o is defined as the cumulative frequency for the whole data set of a single band, which always equalled 100 percent and F_{tk} is the total cumulative frequencies over all classes in band 'k'.

If overlap exists amongst the classes in band 'k' then

$$F_{tk} = \sum_{j=1}^{N} F_{j,k} > F_o = 100\% \quad ...(3.14)$$

The degree of overlap amongst the classes is determined as

$$BVOI = F_{tk}/F_{ta} \text{ (for band 'k')} \quad ...(3.15)$$

and $BVOI = F_{ta}/N$ (for the data set). ...(3.16)

The steps involved for the computation of BVOI are summarised below.

(i) Determine the maximum and minimum brightness values of each class from training data of each spectral band.

(ii) Determine the cumulative percentage of pixels having brightness value ranging from the minimum to maximum for each class, based on the histograms of the whole data set.

(iii) Repeat steps (i) and (ii) for each class in each band.

(iv) Compute the average of the cumulative percentage of pixels in all bands for each class.

(v) Sum the cumulative percentages of pixels for all classes in each band and sum the averages of cumulative percentage of pixels in all bands for each class.

(vi) Compute the BVOI value of each band by dividing the total cumulative percentages for specified band to the sum of average of all bands for each class.

(vii) The BVOI value of given data set is determined by dividing the sum of the average of all bands for each class by the number of class.

3.3 CHOICE OF A FEATURE SELECTION TECHNIQUE

It is clearly evident that each of the reviewed feature selection technique has a number of advantages and disadvantages in terms of the computer resources, applicability to one or more dimension, ability to reliably assess separability, etc. Acknowledging the above factors and previous experience, Thomas et. al. (1987) suggested the following choices of different feature selection techniques in different situation:

(i) If the classes were tending toward true homogeneity and a limited set of bands is to be used then our choice will be the Divergence.

(ii) If the same situation as (i) applied but the whole set of bands is to be used then our choice will be the Transformed Divergence.

(iii) If the classes are less truly homogeneous and a limited set of bands is to be used then our choice will be Bhattacharya Distance.

(iv) If the same situation as in (iii) applied, but the whole set of bands is to be used then our choice will be Jefferys Matusita Distance, and

(v) The BVOI technique is applicable in all the cases as mentioned above.

However, any feature selection technique is but a 'mean-to-an-end'. The final choice is the selection of bands that are used in the classification process and the final evaluation is how good, or bad, is that classification product.

3.4 CASE STUDIES

A large volume of literature related to feature selection techniques are available. Selected case studies have been discussed here under.

3.4.1 Feature Selection of Multispectral Remote Sensing Data

The Divergence and Transformed Divergence distances were used by Kamat (1978) with the ISRO - Multispectral Scanner for determination of land use pattern in Panchmahal District of Gujarat. The data were available in five spectral bands; first three bands in visible range, fourth band in the infrared region and the fifth band in the thermal infrared region. Only first four bands were used in the study. Both the feature selection techniques were tested for four classes, viz. water, barren land, forest and vegetation.

The Divergence and Transformed Divergence distances were calculated for all pairs of classes using equation (3.2) and (3.4) for the following two cases :

(a) Selection of the best two bands from the available four bands, and

(b) Selection of the best three bands from the available four bands.

It was found that bands 1 and 2 are the best bands for first case and bands 1, 2 & 4 for the second case, for identification of all the classes. However, for separating any particular class from the others, the optimum bands combination was different, which reflected that the feature selection is considerably dependent on the classes of interest. Further, it was found that the standard deviation for barren land is more than that of other classes, which indicated that this class is not homogeneous and may have some mixed vegetation. Thus, a detailed ground truth information for these classes is required to identify the best bands.

3.4.2 Optimum Band Selection for Supervised Classification

Mausel et. al. (1990) selected an agricultural site near the town of Weslaco in Hindalgo Country, Texas to use feature selection and Supervised classification techniques for six classes, viz. cotton, cantaloupe, sorghum, Johnson grass, pigweed and bare soil. Training areas (24 Nos.) of size 7.0m x 9.2m were selected randomly. The site was imaged on 31st May and 24th July 1983 near noon on moderately sunny days from an altitude of 900m using USDA - ARS multispectral video system in order to collect spectral information in four spectral bands (0.42 to 0.43 µm, 0.52 to 0.55 µm, 0.64 to 0.67 µm, and 0.84 to 0.89 µm). The data were digitized and registered to create an eight bands multitemporal data set with a spatial resolution of 0.2 meters.

Training statistics were extracted from within each of the 24 training plots on false colour composite. Mean value and variance - covariance matrices of brightness value for each class were computed. The Divergence, Transformed Divergence, Bhattacharya Distance and Jeffreys Matusita Distance were also computed for all class pairs of one to eight bands combinations. Supervised Gaussian Maximum Liklihood classification procedure was applied to classify the entire digital image. Result clearly showed that an increase in spectral bands after three to four spectral bands hardly increases the accuracy of classification. In fact, the accuracy stablized around four spectral bands, as shown in Fig. 3.5. The CPU time required for the different number of combinations of bands are also shown in Fig. 3.6, which indicated that the CPU time increases appreciably with an increase in number of bands.

3.4.3 Improvement in Vegetation Separability Using Transformed Divergence

While analyzing different vegetation types for river Oykel, Scotland, Ghosh (1991) found that there was a confusion in the Transformed Divergence for TM data of 1984 and 1989 amongst different classes of vegetation namely light forest, dense forest and heather. The Transformed Divergence produced very poor separation between these classes, particularly for 1989 image. When synthetically generated data such as ratio, NDVI etc. is incorporated along with the raw satellite data and same training data is used, it was found that the Transformed Divergence between each class pair, as well as overall Transformed Divergence of all the classes improved significantly. For raw data of year 1989, the value of overall Transformed Divergence for best bands combination was found to be 1892, while a best combination of raw data with synthetically generated data gave an overall Transformed Divergence of 1999. This value of Transformed Divergence indicated that excellent separation between the classes can be achieved if synthetically generated data are combined with raw satellite data using the same training data. It was further observed that the classification accuracy improved from 70% to approximately 93.5%, when the raw data and enhanced data were used in combination.

3.4.4 Selection of Spectral Band Combination for Land Cover/Land Use Classification Using Brightness Value Overlapping Index (BVOI)

Saha and Kudrat (1991) applied the Brightness Value Overlapping Index technique of feature selection using digital LANDSAT - TM data of Jan. 1986 in order to classify the major land cover classes over a part of Central Gangetic alluvial plain, in Aligarh District. The major land cover classes present in the area were waste land (salt-affected and water-logged land), cropland, water bodies and built-up land. The data used for the analysis included five spectral bands (1, 2, 3, 4, & 5) as band 7 data not included due to its bad quality & noise. The training sites of different land use/land cover classes were marked on topographical maps during field visits, which were later identified on the digital image. Training statistics for different land cover classes were generated to compute BVOI values of each band. After computation of BVOI values in each band, the following four spectral bands combinations were used for classification employing Maximum Likelihood Supervised classification :

(i) Bands 1, 2 and 3

(ii) Bands 2, 3 and 4

(iii) Bands 1, 4 and 5, and

(iv) Bands 2, 4 and 5

The error matrices were also generated to assess the accuracy of classification for all the above four cases. The result showed that the combination of bands 2, 4 and 5 gave minimum BVOI values and maximum classification accuracy as 96%. Similarly, combination of bands 1, 2 and 3 produced maximum BVOI values and minimum classification accuracy as 86%.

The most vital precaution to be taken for this tehcnique is to avoid combination of bands having even small BVOI value but higher correlation(Saha and Kudrat, 1991).

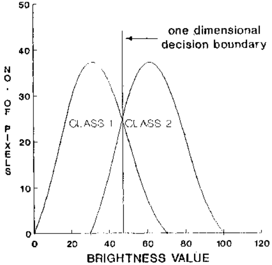


FIG. 3.1 : THE BASIC PROBLEM IN REMOTE SENSING PATTERN RECOGNITION CLASSIFICATION (SOURCE : JENSEN, 1986)

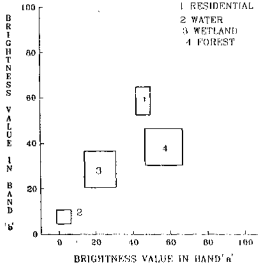
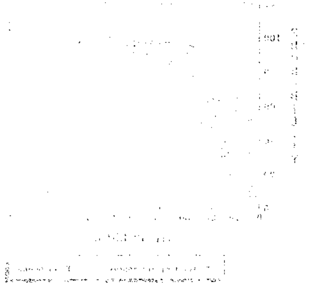


FIG. 3.2 : SCATTERPLOT OF BRIGHTNESS VALUE OF TWO BAND DATA (SOURCE : JENSEN, 1986)



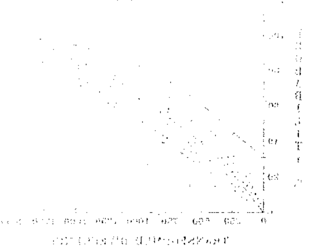


FIG. 3.4 PROBABILITY OF CORRECT CLASSIFICATION (PC) Vs. TRANSFORMED DIVERGENCE (SOURCE: SWAIN AND KING 1973)

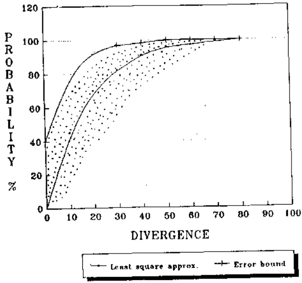


FIG. 3.3 PROBABILITY OF CORRECT CLASSIFICATION (P) Vs. DIVERGENCE (SOURCE : SWAIN AND KING , 1973)

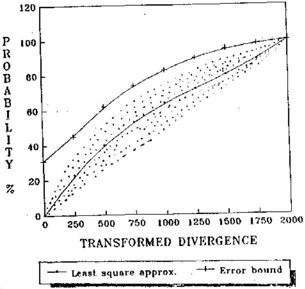
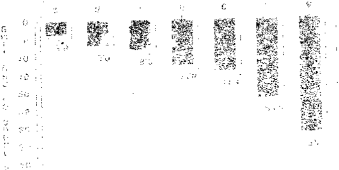
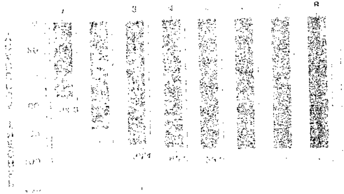


FIG. 3.4 PROBABILITY OF CORRECT CLASSIFICATION (P) Vs. TRANSFORMED DIVERGENCE (SOURCE :SWAIN AND KING ,1973)





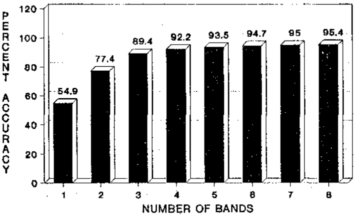


FIG. 3.5 OVERALL CLASSIFICATION ACCURACY FOR BAND COMBINATIONS ONE TO EIGHT (SOURCE : MAUSEL et. al. , 1990)

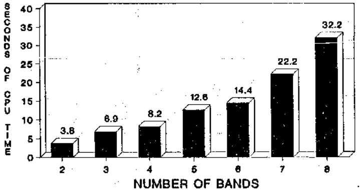


FIG. 3.6: SECONDS OF CPU TIME REQUIRED FOR CLASSIFICATION : TWO TO EIGHT BANDS (SOURCE : MAUSEL et.al.,1990)

CHAPTER 4

DESCRIPTION OF THE STUDY AREA AND DATA USED

4.1 SALIENT FEATURE OF THE STUDY AREA

The study area comprised of Patna and its surrounding, lying in the North West part of the Bihar State of India (Fig. 4.1). It lies between approximately 85°07′ E to 85°17′ E in longitude and 25°32′ N to 25°42′ N in latitude. Patna is the principal town in this region as shown in Fig. 4.1. The river Ganga passes through the area approximately in West to East direction. Comprising about 236 Km^2 area extent, the study area covers a part of Indo-Gangetic plains. The whole area covered in the LANDSAT 5-TM imagery of path row no. 141-042, and Survey of India topographic sheet no. 72 $\frac{G}{2}$ at scale 1:50,000.

The area was chosen primarily as it is agriculturally predominant with nearly all types of land use and land cover classes such as water bodies, natural vegetation, boulders, urban and sandy areas. Other factors included (i) Ground knowledge about the land use and land cover pattern, (ii) Availability of topographical map, and (iii) Easy availability of satellite data.

4.1.1 Climate

The winter season around the study area begins towards the end of the month of October and extends upto the month of February. The peak summer months are May and June. Rainy season extends from the middle of June to the end of September. The minimum temperature touches 4°C; though the maximum temperature attained is 43°C. The area generally becomes flooded in rainy seasons due to appreciable increase in the flow level of river Ganga. The river Ganga has been changing its course considerably after the construction of "Gandhi Setu".

4.1.2 Land Use and Land Cover Classes

The area is having nearly all types of land use and land cover classes. However, vegetation and the urban areas are predominent classes in the area. Vegetative areas are rapidly being converted to urban area due to fast developement in the area. Nearly 40% of the area is under vegetation and 30% of the area consists of urban areas. About 10% of the area consists of water bodies in the form of river Ganga and its tributaries, small ponds, etc. and remaining portion is covered with other land cover classes.

The six major classes were recognised by Kumar (1993), with the 1:50,000 geocoded F.C.C., toposheet and the visual interpretation followed by ground visit of the area. The same six classes have been taken for the present analysis. These are :

(i) Water bodies

(ii) Vegetation

(iii) Dry sand

(iv) Wet sand

(v) Urban areas, and

(vi) Boulders

4.2 DATA USED

The data for the analysis comprised of 512x512 pixels equivalent to roughly 236 sq. km. on ground. The entire area was imaged on 25th January 1992 by satellite LANDSAT - 5 TM, covers under path row no. 141-042. The order was given to NRSA for supply of all the bands of LANDSAT 5-TM of study area but only bands 1, 3, 4 and 7 were supplied due to their distribution policy. Although, band 2 data was highly correlated with band 1 and band 5 data with band 4. Therefore, band 2 and 5 data will not contribute significantly for feature selection as well as classification techniques in the study area. So bands 1, 3, 4 and 7 data were used for study purpose. Plate 4.1 - Plate 4.4 show the entire study area in TM bands 1, 3, 4 and 7 respectively. In addition, SOI topographic sheet 72 $\frac{G}{2}$ at a scale of 1:50,000 surveyed in 1976 has been used.

4.3 EQUIPMENT USED

(a) ISROVISION

It is a PC based image processing system devloped by ISRO, Bangalore with the aim to reduce the cost of digital image analysis.

Initially, the analysis has been planned to be carried out using toposheet only as a source of ground truth. The analysis gives poor result mainly because of time difference of satellite data (year 1992) and toposheet (year 1976). Therefore, the option of using digital image as source of ground truth has been carried forward by generating a F.C.C. using bands 7, 4 and 3 (Plate 4.5) with the help of ISROVISION.

The interactive part of the equipment has been used to delineate training areas on the image itself, as shown in plate 4.6, to extract the training data and finally classify the entire image.

(b) PC - AT

PC - AT has been used to extract the statistics of the training data and these are subsequently used for various feature selection techniques. Computer programs for the generation of training statistics and feature selection techniques have been written in FORTRAN - 77, the details of which are given in next chapter.

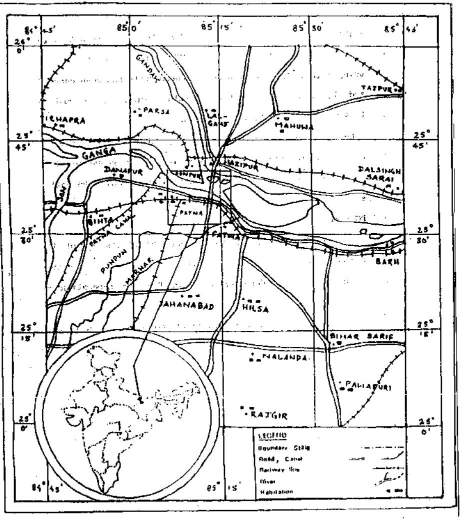


FIG. 4.1 LOCATION OF THE STUDY AREA

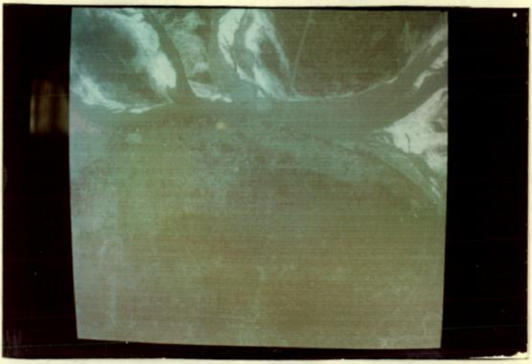


PLATE 4.1 : STUDY AREA ON LANDSAT-5 TM, BAND 1

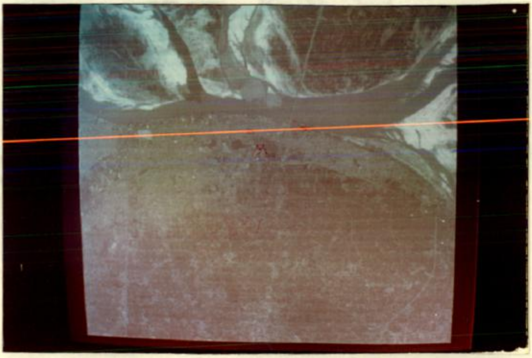


PLATE 4.2 : STUDY AREA ON LANDSAT-5 TM, BAND 3

PLATE 4.3 : STUDY AREA ON LANDSAT- 5 TM, BAND 4

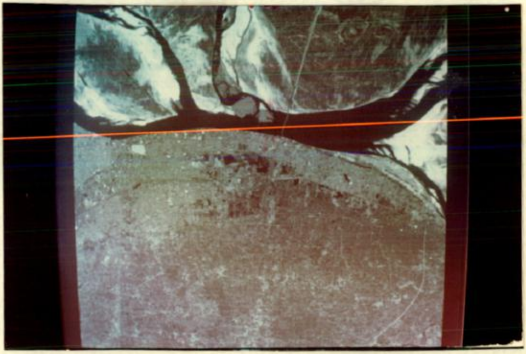


PLATE 4.4 : STUDY AREA ON LANDSAT-5 TM, BAND 7

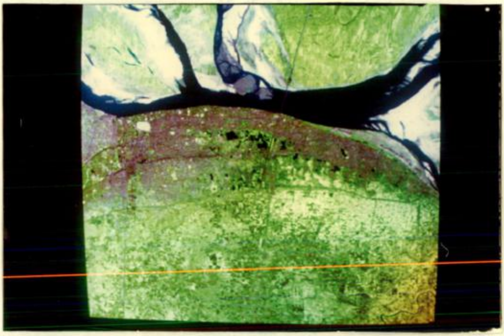


PLATE 4.5 : F.C.C OF STUDY AREA USING TM BANDS 7,4 AND 3

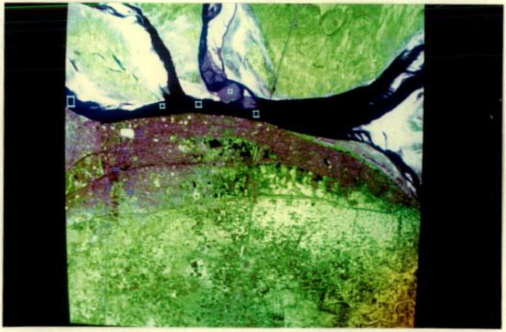


PLATE 4.6 : LOCATION OF TRAINING AREAS ON F.C.C SHOWN IN PLATE 4.5

CHAPTER 5

METHODOLOGY

Data analysis using feature selection techniques involves a systematic approach. Fig. 5.1 shows a general flow chart representing the approach to be adopted for feature selection techniques including accuracy assessment of classified data. Each step has been briefly explained in the following sections.

5.1 IDENTIFICATION OF CLASSES

For analyzing satellite data, it is necessary to identify the number of classes depending on the user's need and the level of classification.

In the present work, the following six major land use and land cover classes has been identified (Kumar, 1993) :

i) Water bodies

ii) Vegetation

iii) Dry sand

iv) Wet sand

v) Urban areas, and

vi) Boulders

5.2 SELECTION AND EXTRACTION OF TRAINING AREA

In order to perform feature selection and subsequently classification of the given data, it is necessary to select and extract the training area for each class in each band. In order to have good training data set, a minimum number of pixels are

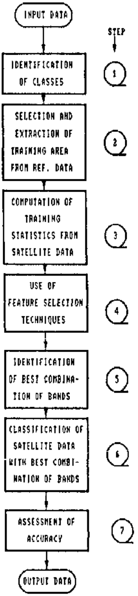


FIG. 5.1 STEPS INVOLVED IN THE ANALYSIS OF SATELLITE DATA USING FEATURE SELECTION TECHNIQUES

needed. For the same, Fitzpatrick-Lins (1980) has suggested the following relationship :

$$N = \frac{Z^2 \times P \times Q}{E^2} \qquad \ldots.(5.1)$$

where Z = Normal variate

P = Expected accuracy (%)

$Q = 100 - P$

E = Allowable error (%), and

N = Minimum number of pixels for all classes

In present study, $P = 85\%$, $Q = 15\%$, $Z \cong 2$ and $E = 5\%$ have been adopted. Using these value, minimum number of pixels (N) comes out to be around 204 for all the classes. However, the above relationship is not the only guiding factor for selecting minimum number of pixels. Lillesand and Kiefer (1973) have stated that a minimum of 10n to 100n pixels should be used in training area for each class, where n is the number of spectral bands available. It is also suggested that more the number of pixels used, the better is the statistical representation of each class.

In the present study, the above guidlines have been kept in mind while selecting the training areas for each class on LANDSAT-5 TM data. The training areas have been selected using toposheet and generating F.C.C of TM bands 7, 4 and 3 on ISROVISION. The training areas have been marked on digital F.C.C. image (plate 4.6), using mouse which activates the cursor on the display screen of ISROVISION. The co-ordinates of the opposite corners of rectangular training areas have been noted from ISROVISION, which subsequently used to extract the training data from the image using computer program written in FORTRAN-77.

5.3. DETERMINATION OF TRAINING DATA STATISTICS

The statistics for each class such as mean, standard deviation, variance, minimum and maximum brightness values have been generated for the training areas on F.C.C. image. Computer program has been developed for this purpose using a PC-AT. These statistics have been further used to generate variance-covariance matrices and correlation coefficients matrices for all the classes. All the above statistics are required as input to feature selection techniques.

5.4 APPLICATION OF FEATURE SELECTION TECHNIQUES

The four bands of data have been analysed using all the five feature selection techniques as described in chapter-3. Among all these five tehcniques, the first four viz. Divergence, Transformed Divergence, Bhattacharya Distance and Jeffreys Matusita Distance have been found to follow the same pattern. Thus, these have analyzed simultaneously. The fifth technique (BVOI) has been analyzed separately due to its different behaviour when compared with the other techniques.

5.4.1 Determination of Separatability Indices Using Divergence, Transformed Divergence, Bhattacharya Distance and Jeffreys Matusita Distance

In these feature selection techniques, the following two modes of analysis have been accomplished :

(a) Selection of the best two bands from the available four bands, and

(b) Selection of the best three bands from the available four bands.

The two bands can be selected out of four bands by $6\left({}^{4}C_{2}\right)$ ways. So, the possible combinations of two bands are (i) 1 and 3, (ii) 1 and 4, (iii) 1 and 7, and (iv) 4 and 7. Similarly, there are $4\left[{}^{4}C_{2}\right]$ combinations of three bands that can be selected out of four bands. These are (i) 1,3 and 4, (ii) 1,3 and 7, (iii) 1,4 and 7 and (iv) 3,4 and 7.

For the above two cases, the Divergence, Transformed Divergence, Bhattacharya Distance and Jeffreys Matusita Distance have been calculated for all the 15 $\left[{}^{6}C_{2}\right]$ possible pairs of classes and combinations of bands using equations (3.2), (3.4), (3.5) and (3.7) respectively. In addition, the average of all the separability measures for each band combination have been calculated. Figures 5.2, 5.3, 5.4 and 5.5 show the typical flow charts for the computation of Divergence, Tranformed Divergence, Bhattacharya Distance and Jeffreys Matusita Distance respectively. These flow charts are self explainatory.

5.4.2 Determination of Brightness Value Overlapping Index.

One BVOI technique requires only minimum and maximum brightness values for each class in each band from the training data. Using these minimum and maximum brightness values, the BVOI values in different bands have been calculated using equations (3.9) through (3.16). A typical flow chart of this technique is shown in Fig. 5.6.

5.5 IDENTIFICATION OF BEST COMBINATION OF BANDS

After computation of the averages of Divergence, Transformed Divergence, Bhattacharya Distance and Jeffreys Matusita Distance for each combination of bands, the best combination of bands has been determined whose average of the above distance is minimum for each technique. In BVOI approach, those bands are taken as the best bands whose BVOI values are minimum successively.

5.6 CLASSIFICATION OF SATELLITE DATA USING BEST COMBINATION OF BANDS

The Minimum distance classifier based on Euclidian Distance, as described in section 2.2, has been used for the classification of the image. This classifier requires only mean brightness value for each class in each band computed from the trainng data. In this classification algorithm, the pixel of an image is labelled to that class to which it is nearest. Based on this approach, the entire image has been classified into various classes using the

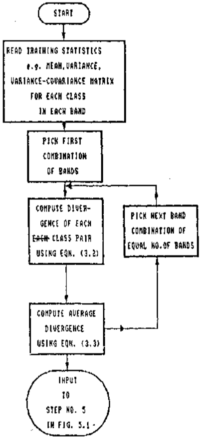


FIG. 5.2 FLOW CHART OF DIVERGENCE

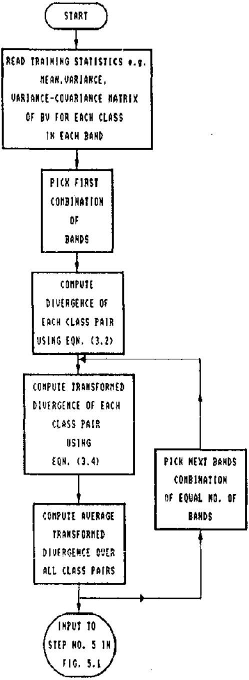


FIG. 5.3 FLOW CHART OF TRANSFORMED DIVERGENCE

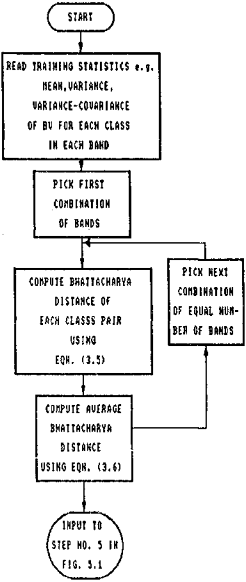


FIG. 5.4 FLOW CHART OF BHATTACHARYA DISTANCE

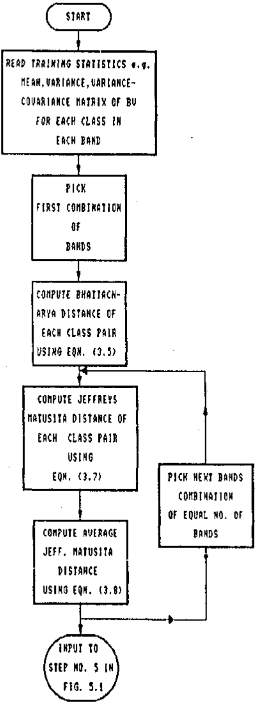


FIG.5.5 FLOW CHART OF JEFFREYS MATUSITA DISTANCE

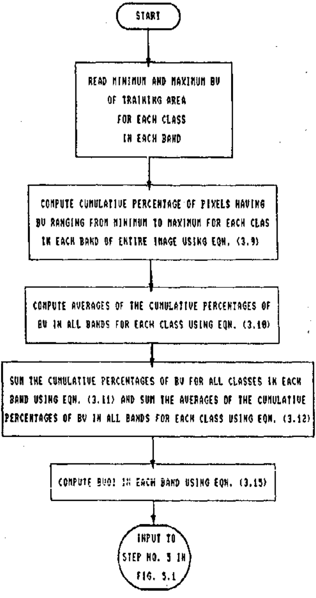


FIG: 5:6 FLOW CHART OF BVOI

best combinations of bands. The same process has been repeated for the worst combinations of bands in order to compare the result and assess the utility of feature selection techniques.

5.7 ACCURACY ASSESSMENT

The accuracy of the classified image has been assessed by generating a confusion matrix or error matrix from the training data. The diagonal elements of the matrix represent the pixels correctly classified, while the non-diagonal elements give the error of omission and error of commission. The errors of omission have been computed by adding the number of pixels assigned to incorrect class along each row of each class. Similarly, errors of commission have been computed by adding the number of pixels assigned to incorrect class along column for each class. The overall accuracy has been obtained by dividing the sum of the diagonal elements of the error matrix by the total number of pixels.

The results obtained from TM data are discussed in next chapter.

CHAPTER 6

RESULTS AND DISCUSSIONS

6.1 RESULTS

The results obtained from the analysis of TM data are described in the following sections :

6.1.1 Computation of Training Statistics

As stated earlier, for Supervised classification of satellite data using feature selection techniques, the input parameters such as mean, variance, minimum and maximum brightness values etc. for each class in each band, and variance-covariance matrix, as well as correlation coefficients matrix in different bands are required. The statistical information for all the land use and land cover classes, viz. water bodies, vegetation, dry sand, wet sand, urban areas and boulders have been determined from TM data, and shown in Table 6.1.

When we visually examine the minimum and maximum brightness values of different classes in different bands as shown in Table 6.2, it is found that not even a single band is suitable for separation of one class from the others. For example, if we are interested to find out the bands which can separate the vegetation from other classes then it is found that bands 4 and 7 are suitable for separation of vegetation from water bodies; bands 1,3 and 7 are suitable for separation of vegetation from dry sand and wet sand; band 4 is suitable for separation of vegetation from urban areas and bands 3 and 4 are suitable for

TABLE 6.1 Univariate and Multivariate Training Statistics for Six Land Cover Classes Using Four Bands of Thematic Mapper Data.

A. Statistics for Water Bodies.

No. of samples = 4, Total No. of pixels in all samples = 343

Band:	1	3	4	7
Mean	76.66	32.95	20.30	4.71
Std.dev.	1.33	0.90	0.76	0.96
Variance	1.76	0.81	0.58	0.92
Minimum	72	30	18	3
Maximum	81	35	22	8

Band	Variance-covariance matrix			
1	1.76			
3	0.15	0.81		
4	0.01	0.26	0.58	
7	0.09	0.04	0.03	0.92

Band	Correlation matrix			
1	1.00			
3	0.13	1.00		
4	0.01	0.26	1.00	
7	0.07	0.05	0.04	1.00

B. Statistics for Vegetation.

No. of samples = 4, Total No. of pixels in all samples = 258

Band:	1	3	4	7
Mean	73.00	29.16	57.08	14.60
Std.dev.	1.83	1.43	4.42	2.84
Variance	3.36	2.06	19.58	8.04
Minimum	69	26	44	10
Maximum	79	34	66	26

Band	Variance-covariance matrix			
1	3.36			
3	1.11	2.06		
4	-1.42	-3.35	19.58	
7	1.30	2.55	-7.92	8.04

Band	Correlation matrix			
1	1.00			
3	0.42	1.00		
4	-0.18	-0.53	1.00	
7	0.25	0.63	-0.63	1.00

C. Statistics for Dry Sand.
No. of samples = 3, Total No. of pixels in all samples = 214

Band:	1	3	4	7
Mean	98.37	58.84	54.40	68.16
Std. dev.	2.63	2.58	1.51	4.52
Variance	6.94	6.65	2.27	20.47
Minimum	88	48	50	49
Maximum	105	62	58	74

Band	Variance-covariance matrix			
1	6.94			
3	5.38	6.65		
4	2.39	2.91	2.27	
7	9.14	9.26	4.63	20.47

Band	Correlation matrix			
1	1.00			
3	0.79	1.00		
4	0.60	0.75	1.00	
7	0.77	0.79	0.68	1.00

D. Statistics for Wet Sand.
No. of samples = 3, Total No. of pixels in all samples = 204

Band:	1	3	4	7
Mean	89.42	49.17	45.86	54.92
Std. dev.	2.96	2.47	2.25	3.87
Variance	8.75	6.08	5.04	15.01
Minimum	84	43	36	36
Maximum	98	54	50	42

Band	Variance-covariance matrix			
1	8.75			
3	5.86	6.08		
4	1.75	2.51	5.04	
7	5.37	5.92	4.87	15.01

Band	Correlation matrix			
1	1.00			
3	0.80	1.00		
4	0.28	0.45	1.00	
7	0.47	0.62	0.56	1.00

E. Statistics for Urban Areas.

No. of samples = 4, Total No. of pixels in all samples = 211

Band:	1	3	4	7
Mean	76.20	34.24	30.80	25.89
Std. dev.	1.83	1.34	2.36	2.48
Variance	3.37	1.79	5.57	6.17
Minimum	72	32	25	19
Maximum	82	38	37	32

Band	Variance-covariance matrix			
1	3.37			
3	1.33	1.79		
4	1.60	1.55	5.57	
7	1.77	1.62	2.58	6.17

Band	Correlation matrix			
1	1.00			
3	0.54	1.00		
4	0.37	0.49	1.00	
7	0.39	0.49	0.44	1.00

F. Statistics for Boulders.

No. of samples = 3, Total No. of pixels in all samples = 102

Band:	1	3	4	7
Mean	78.75	37.30	37.49	31.01
Std. dev.	2.25	1.18	3.59	4.82
Variance	5.05	1.39	12.92	23.23
Minimum	74	34	31	23
Maximum	83	41	42	39

Band	Variance-covariance matrix			
1	5.05			
3	0.83	1.39		
4	-4.20	0.14	12.92	
7	-5.63	-0.06	15.28	23.23

Band	Correlation matrix			
1	1.00			
3	0.31	1.00		
4	-0.52	0.03	1.00	
7	-0.52	-0.01	0.88	1.00

TABLE 6.2: Range of Brightness Value of Various Classes for Different TM Bands from Training Data

CLASS		TM BANDS			
		1	3	4	7
1. Water bodies	Minimum	72	30	18	03
	Maximum	81	35	22	08
2. Vegetation	Minimum	69	26	44	10
	Maximum	79	34	66	26
3. Dry sand	Minimum	88	48	50	49
	Maximum	105	62	58	74
4. Wet sand	Minimum	84	43	36	36
	Maximum	98	54	50	42
5. Urban areas	Minimum	72	32	25	19
	Maximum	82	38	37	32
6. Boulders	Minimum	74	34	31	23
	Maximum	83	41	42	39

separating vegetation from boulders. So, not even a single band is available to discriminate vegetation from the other classes. So, feature selection techniques are needed for the determination of those bands which are most effective in separating all the classes from each other. These bands can subsequently be used for the better classification.

6.1.2 Computation of separability Indices using Divergence, Transformed Divergence, Bhattacharya Distance and Jeffreys Matusita Distance

All the separability Indices mentioned above, have been computed separately for all possible two bands and three bands combination and described below :

6.1.2.1 Using Two Bands Combination

The Divergence, Transformed Divergence, Bhattacharya Distance and Jefferys Matusita Distance have been computed as described in chapter 5 for two bands combinations between fifteen pairs of classes. In addition, the averages of all these indices have been computed for each two bands combination. The details are shown in Table 6.3.

It has been found that the maximum value of the average of Divergence, Transformed Divergence, Bhattacharya Distance and Jeffreys Matusita Distance is obtained for combination of bands 4 and 7. For this combination the average Divergence is 353.82, while, average Transformed Divergence is 1966.32. The corresponding Bhattacharya Distance and Jeffreys Matusita Distance is 125.35 and 1.98 respectively. Hence, all the four techniques discussed above, results in bands 4 and 7 as best two bands combination, for the separation of all the classes, to be used for classification. However, in some of the cases, the best bands combination for inter class separability are different, according to different techniques as shown in Table 6.4.

TABLE 6.3: Different Separability Measures for Six Land Cover Classes Using Two Bands Combination

(A) Bands 1 and 3.

Sl. no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	6.90	1065.77	3.64	1.95
2.	1 - 3	201.83	2000.00	166.41	2.00
3.	1 - 4	72.16	1999.76	60.61	2.00
4.	1 - 5	0.11	27.36	0.32	0.55
5.	1 - 6	3.50	707.93	0.88	1.17
6.	2 - 3	167.22	2000.00	365.34	2.00
7.	2 - 4	68.44	1999.62	164.72	2.00
8.	2 - 5	4.84	907.76	4.52	1.98
9.	2 - 6	11.63	1532.66	16.67	2.00
10.	3 - 4	11.16	1504.16	78.41	2.00
11.	3 - 5	120.93	2000.00	279.02	2.00
12.	3 - 6	73.67	1999.80	281.55	2.00
13.	4 - 5	41.15	1988.32	106.66	2.00
14.	4 - 6	21.03	1655.61	91.47	2.00
15.	5 - 6	2.00	443.05	1.36	1.49
Ave.:		53.77	1468.78	108.10	1.81

Class - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand , 4 - Wet sand , 5 - Urban areas , 6 - Boulders

(B) Bands 1 and 4.

Sl.no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	57.91	1998.56	3.88	1.96
2.	1 - 3	263.98	2000.00	165.72	2.00
3.	1 - 4	111.42	1998.99	59.23	2.00
4.	1 - 5	1.78	398.88	0.21	0.38
5.	1 - 6	15.68	1718.33	2.39	1.82
6.	2 - 3	13.61	1635.28	362.19	2.00
7.	2 - 4	37.61	1980.43	161.50	2.00
8.	2 - 5	23.64	1895.80	0.96	1.23
9.	2 - 6	27.33	1934.30	16.73	2.00
10.	3 - 4	9.92	1421.42	78.17	2.00
11.	3 - 5	125.88	2000.00	279.33	2.00
12.	3 - 6	60.60	1998.97	281.35	2.00
13.	4 - 5	40.92	1987.99	106.85	2.00
14.	4 - 6	16.97	1760.29	91.63	2.00
15.	5 - 6	8.07	1270.96	2.81	1.88
Ave.:		54.35	1733.34	107.53	1.82

Class - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand ,
4 - Wet sand , 5 - Urban areas , 6 - Boulders

(C) Bands 1 and 7.

Sl.no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	15.51	1712.31	3.66	1.95
2.	1 - 3	496.04	2000.00	166.84	2.00
3.	1 - 4	219.98	2000.00	60.45	2.00
4.	1 - 5	3.88	768.53	0.14	0.26
5.	1 - 6	24.61	1907.72	2.42	1.82
6.	2 - 3	303.68	2000.00	365.26	2.00
7.	2 - 4	136.79	2000.00	164.15	2.00
8.	2 - 5	10.77	1479.50	4.20	1.97
9.	2 - 6	28.66	1944.40	16.72	2.00
10.	3 - 4	16.64	1750.29	78.15	2.00
11.	3 - 5	209.41	2000.00	278.95	2.00
12.	3 - 6	142.57	2000.00	279.97	2.00
13.	4 - 5	76.02	1999.90	106.64	2.00
14.	4 - 6	49.27	1995.77	89.55	2.00
15.	5 - 6	62.26	1999.65	57.44	2.00
Ave.:		119.74	1837.20	111.64	1.87

Class - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand ,

4 - Wet sand , 5 - Urban areas , 6 - Boulders

(D) Bands 3 and 4.

Sl.no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	115.64	2000.00	2.28	1.79
2.	1 - 3	611.44	2000.00	121.04	2.00
3.	1 - 4	296.06	2000.00	249.96	2.00
4.	1 - 5	12.31	1570.81	0.90	1.19
5.	1 - 6	73.13	1999.79	0.96	1.23
6.	2 - 3	19.17	1817.95	344.95	2.00
7.	2 - 4	67.31	1999.56	152.55	2.00
8.	2 - 5	63.50	1999.29	4.62	1.98
9.	2 - 6	92.85	1999.98	12.75	2.00
10.	3 - 4	13.00	1606.43	75.00	2.00
11.	3 - 5	205.61	2000.00	213.84	2.00
12.	3 - 6	158.78	2000.00	233.09	2.00
13.	4 - 5	81.17	1999.92	77.39	2.00
14.	4 - 6	44.19	1992.02	39.41	2.00
15.	5 - 6	13.56	1632.80	3.02	1.90
Ave.:		124.51	1907.90	102.11	1.87

Class - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand ,

4 - Wet sand , 5 - Urban areas , 6 - Boulders

(E) Bands 3 and 7.

Sl.no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	31.19	1959.46	2.00	1.73
2.	1 - 3	1143.00	2000.00	121.72	2.00
3.	1 - 4	571.92	2000.00	47.33	2.00
4.	1 - 5	24.88	1910.85	0.24	0.43
5.	1 - 6	111.80	2000.00	2.82	1.88
6.	2 - 3	505.86	2000.00	346.99	2.00
7.	2 - 4	261.99	2000.00	254.25	2.00
8.	2 - 5	30.08	1953.45	5.93	1.99
9.	2 - 6	31.99	1963.32	12.88	2.00
10.	3 - 4	20.85	1852.41	74.89	2.00
11.	3 - 5	170.40	2000.00	213.42	2.00
12.	3 - 6	134.29	2000.00	132.95	2.00
13.	4 - 5	56.84	1998.36	77.15	2.00
14.	4 - 6	28.19	1941.02	39.36	2.00
15.	5 - 6	10.73	1477.11	3.71	1.95
Ave.:		208.93	1937.07	102.38	1.87

Class - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand ,

4 - Wet sand , 5 - Urban areas , 6 - Boulders

(F) Bands 4 and 7.

Sl.no.	Classes	DIVcd	$\overline{DIVcd}^{T}$	Bcd	Jcd
1.	1 - 2	325.06	2000.00	190.15	2.00
2.	1 - 3	2345.08	2000.00	134.47	2.00
3.	1 - 4	1235.19	2000.00	84.85	2.00
4.	1 - 5	211.89	2000.00	2.15	1.77
5.	1 - 6	414.93	2000.00	41.60	2.00
6.	2 - 3	19.47	1824.57	343.86	2.00
7.	2 - 4	47.57	1994.77	126.12	2.00
8.	2 - 5	29.60	1950.55	147.89	2.00
9.	2 - 6	42.22	1989.79	147.70	2.00
10.	3 - 4	36.12	1978.12	28.63	2.00
11.	3 - 5	311.10	2000.00	275.44	2.00
12.	3 - 6	158.28	2000.00	138.63	2.00
13.	4 - 5	83.20	1999.94	150.33	2.00
14.	4 - 6	28.71	1944.75	64.66	2.00
15.	5 - 6	16.93	1812.33	3.71	1.95
Ave.:		353.82	1966.32	125.35	1.98

Class - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand ,

4 - Wet sand , 5 - Urban areas , 6 - Boulders

TABLE 6.4: Best Combination of Two TM Bands for Separating Any Two Classes from Each Other Using Different Feature Selection Techniques

Sl.no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	4 & 7	3 & 4 4 & 7	4 & 7	4 & 7
2.	1 - 3	4 & 7	All	1 & 7	All
3.	1 - 4	4 & 7	1 & 7 3 & 4 3 & 7 4 & 7	3 & 47	All
4.	1 - 5	4 & 7	4 & 7	4 & 7	4 & 7
5.	1 - 6	4 & 7	3 & 7 4 & 7	4 & 7	4 & 7
6.	2 - 3	2 & 7	1 & 3 1 & 7 3 & 7	1 & 3	All Except 4 & 7
7.	2 - 4	3 & 7	1 & 7 3 & 7	3 & 7	All
8.	2 - 5	3 & 4	3 & 4	4 & 7	4 & 7
9.	2 - 6	3 & 4	3 & 4	4 & 7	All
10.	3 - 4	4 & 7	4 & 7	1 & 3	All
11.	3 - 5	1 & 7	All	1 & 7	All
12.	3 - 6	3 & 4	1 & 7 3 & 4 3 & 7 4 & 7	1 & 3	All
13.	4 - 5	4 & 7	4 & 7	4 & 7	All
14.	4 - 6	1 & 7	1 & 7	1 & 4	All
15.	5 - 6	1 & 7	1 & 7	1 & 7	1 & 7

All - 1 & 3 , 1 & 4 , 1 & 7 , 3 & 4 , 3 & 7 , and 4 & 7.
Class - 1 -Water bodies , 2 - Vegetation , 3 - Dry sand , 4 - Wet sand , 5 - Urban areas , 6 - boulders

6.1.2.2 Using Three Bands Combination

The different separability measures have been also computed for three bands combinations between all the fifteen possible pairs of classes. The averages of all these measures have been computed subsequently and shown in Table 6.5. The result shows that the bands 3,4 and 7 are the best combination of three bands according to different techniques, viz. Divergence, Transformed Divergence, Bhattacharya Distance and Jeffreys Matusita Distance. These three bands can be used to achieve the best performance in classification. The best combination of bands for inter class separability has been again found to be different as depicted in Table 6.6.

6.1.3 Computation of Brightness Value Overlapping Index (BVOI)

The BVOI technique requires only minimum and maximum brightness values of training data for each class, and are shown in Table 6.2. Using these minimum and maximum brightness values of all the classes in different bands, the BVOI values for all the bands have been computed as described in chapter 5, and are shown in Table 6.7.

The BVOI values of bands 1,3,4 and 7 have been found to be 1.36, 0.99, 0.87 and 0.78 respectively. Band having less BVOI value contains less overlap between different classes and therefore, considered as the best band. Hence, this technique yields bands 4 and 7 as the best two bands and bands 3,4 and 7 as the best three bands to be used for classification. The bands combinations 1 and 3, and 1,3 and 4 have been found to be the worst combination of two and three bands respectively.

TABLE 6.3: Different Separability Measures for Six Land Cover Classes Using Three Bands Combination

(A) Bands 1 , 3 and 4.

Sl.no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	162.73	2000.00	6.20	2.00
2.	1 - 3	664.78	2000.00	166.93	2.00
3.	1 - 4	312.34	2000.00	61.69	2.00
4.	1 - 5	12.78	1595.20	1.71	1.64
5.	1 - 6	78.75	1999.90	3.63	1.94
6.	2 - 3	200.32	2000.00	365.00	2.00
7.	2 - 4	71.53	1999.73	164.74	2.00
8.	2 - 5	68.31	1999.60	4.99	1.99
9.	2 - 6	93.60	2000.00	17.59	2.00
10.	3 - 4	13.70	1639.17	82.62	2.00
11.	3 - 5	225.80	2000.00	279.54	2.00
12.	3 - 6	175.64	2000.00	282.73	2.00
13.	4 - 5	81.25	1999.92	115.10	2.00
14.	4 - 6	52.19	1997.06	92.07	2.00
15.	5 - 6	13.70	1639.70	3.03	1.90
Ave.:		155.15	1924.69	109.84	1.96

1 - Water bodies , 2 - vegetation , 3 - Dry sand ,

4 - Wet sand , 5 - Urban areas , 6 - Boulders

(B) Bands 1 , 3 and 7.

Sl. no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	35.85	1977.36	2.36	1.81
2.	1 - 3	1150.87	2000.00	167.29	2.00
3.	1 - 4	583.90	2000.00	61.85	2.00
4.	1 - 5	32.76	1966.69	1.52	1.56
5.	1 - 6	138.26	2000.00	3.62	1.95
6.	2 - 3	567.30	2000.00	365.59	2.00
7.	2 - 4	267.98	2000.00	164.93	2.00
8.	2 - 5	30.87	1957.81	6.04	1.99
9.	2 - 6	32.27	1964.58	18.08	2.00
10.	3 - 4	27.41	1934.98	78.42	2.00
11.	3 - 5	213.59	2000.00	279.26	2.00
12.	3 - 6	144.11	2000.00	282.30	2.00
13.	4 - 5	79.83	1999.91	106.78	2.00
14.	4 - 6	54.17	1997.71	91.94	2.00
15.	5 - 6	69.47	1999.66	64.68	2.00
Ave.:		228.58	1986.58	113.11	1.96

Class - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand ,

4 - Wet sand , 5 - Urban areas , 6 - Boulders

(C) Bands 1 , 4 and 7.

Sl.no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	336.91	2000.00	206.56	2.00
2.	1 - 3	2410.46	2000.00	167.76	2.00
3.	1 - 4	1250.95	2000.00	88.18	2.00
4.	1 - 5	214.42	2000.00	2.16	1.77
5.	1 - 6	418.63	2000.00	43.62	2.00
6.	2 - 3	310.63	2000.00	365.81	2.00
7.	2 - 4	241.10	2000.00	165.14	2.00
8.	2 - 5	25.32	1915.57	155.36	2.00
9.	2 - 6	47.30	1994.59	152.81	2.00
10.	3 - 4	37.22	1980.92	79.39	2.00
11.	3 - 5	312.02	2000.00	280.26	2.00
12.	3 - 6	159.03	2000.00	283.19	2.00
13.	4 - 5	88.94	1999.97	158.09	2.00
14.	4 - 6	57.32	1998.45	93.29	2.00
15.	5 - 6	68.95	1999.64	4.74	1.98
Ave.:		398.61	1992.61	149.76	1.98

Class - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand ,

4 - Wet sand , 5 - Urban areas , 6 - Boulders

(D) Bands 3 , 4 and 7.

Sl.no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	419.08	2000.00	194.18	2.00
2.	1 - 3	2441.18	2000.00	142.42	2.00
3.	1 - 4	1275.91	2000.00	88.61	2.00
4.	1 - 5	227.15	2000.00	12.46	2.00
5.	1 - 6	509.10	2000.00	54.14	2.00
6.	2 - 3	519.25	2000.00	357.15	2.00
7.	2 - 4	274.34	2000.00	154.73	2.00
8.	2 - 5	64.50	1999.36	156.36	2.00
9.	2 - 6	93.63	1999.98	314.39	2.00
10.	3 - 4	36.12	1978.12	176.07	2.00
11.	3 - 5	371.59	2000.00	324.63	2.00
12.	3 - 6	343.83	2000.00	144.82	2.00
13.	4 - 5	156.48	2000.00	178.23	2.00
14.	4 - 6	123.09	2000.00	71.09	2.00
15.	5 - 6	28.98	1946.56	3.72	1.95
Ave.:		458.94	1994.93	158.20	2.00

Class - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand ,

4 - Wet sand , 5 - Urban areas , 6 - Boulders

TABLE 6.6: Best Combination of Three TM Bands for Separating Any Two Classes from Each Other Using Different Feature Selection Techniques

Sl. no.	Classes	DIVcd	$DIVcd^T$	Bcd	Jcd
1.	1 - 2	3,4 & 7	All Except 1,3 & 7	1,4 & 7	All Except 1,3 & 7
2.	1 - 3	3,4 & 7	All	1,4 & 7	All
3.	1 - 4	3,4 & 7	All	3,4 & 7	All
4.	1 - 5	3,4 & 7	3,4 & 7 1,3 & 7	3,4 & 7	3 ,4 & 7
5.	1 - 6	3,4 & 7	All Except 1, 3 & 4	3,4 & 7	3,4 & 7
6.	2 - 3	3,4 & 7	All Except 1,3 & 4	1,4 & 7	All
7.	2 - 4	3,4 & 7	All Except 1,3 & 4	1,4 & 7 1,4 & 7	All
8.	2 - 5	1,3 & 4	1,3, & 4	1,4 & 7	All Except 1,3 & 4
9.	2 - 6	3,4 & 7	3,4 & 7	3,4 & 7	All
10.	3 - 4	1,4 & 7	1,4 & 7	1,4 & 7	All
11.	3 - 5	3,4 & 7	All	3,4 & 7	All
12.	3 - 6	3,4 & 7	All	1,4 & 7	All
13.	4 - 5	3,4 & 7	3,4 & 7	3,4 & 7	All
14.	4 - 6	3,4 & 7	3,4 & 7	1,4 & 7	All
15.	5 - 6	1,3 & 7	1,3 & 7	1,3 & 7	1,3 & 7

All - ~~1 & 3~~ , 1,3& 4 , 1,3& 7 , ~~3 & 4~~ , 3,4& 7 , and 1,4 & 7.
Class - 1 -Water bodies , 2 - Vegetation , 3 - Dry sand , 4 - Wet sand , 5 - Urban areas , 6 - Boulders

TABLE 6.7 : Brightness Value Overlapping Index (BVOI)

CLASSES	CUMULATIVE PERCENTAGE OF BRIGHTNESS VALUE DISTRIBUTION IN TM BANDS				
	1	3	4	7	AVERAGE
1. Water bodies	63.40	47.58	07.45	11.67	32.53
2. Vegetation	83.00	76.50	49.90	68.98	69.60
3. Dry sand	05.85	05.26	21.35	06.45	09.73
4. Wet sand	07.60	05.63	50.50	02.95	16.67
5. Urban area	65.00	37.25	05.85	42.38	37.62
6. Boulders	41.18	22.03	35.20	19.35	29.44
Total:	266.03	194.25	170.25	151.78	195.59
BVOI:	1.36	0.99	0.87	0.78	

BVOI of Data Set = 195.59/4 =48.90

The Best Two Bands are : 4 and 7.

The Worst Two Bands are : 1 and 3.

The Best Three Bands are : 3 , 4 and 7.

The Worst Three Bands are : 1 , 3 and 4.

6.1.4 Classification of Image and its Accuracy Assessment

In order to assess the utility of feature selection techniques, the whole image (512 pixels x 512 pixels) has been classified using Minimum Distance classifier. Further, a comparison on the basis of classification for the best and the worst bands combinations have been also undertaken.

The analysis has been done considering two band combination, the best bands combination is TM4 and TM1, while the worst bands combination is TM1 and TM3. Similarly, for three bands combination, the best combination is TM3,4 and 7 while worst combination is TM1,3 and 4.

The results of the best and worst two bands combination are shown in Tables 6.8 and 6.9, respectively. The overall classification accuracy is 86.86% and 80.00% respectively. Similarly, the results of the best and the worst three bands combination are shown in Tables 6.10 and 6.11 respectively.

It can be clearly seen that the analysis is certainly not acceptable for the worst two bands combination, since the overall accuracy is less than 85%, as already stated in section 5.2, while for the best two band combination, the overall classification accuracy is just above the acceptable standard of 85%. In the 3 bands combinations, the overall classification accuracy, both for the best and the worst two bands combination is much higher that the best two bands combination.

When all the four available bands are used for classification (Table 6.12), the overall classification accuracy is 97.97%, which is a small improvement in classification accuracy (0.67%), but the CPU time required to classify the image using four bands is substantially larger (4 min. 58 sec.), i.e. 50% more computer time is required. Thus, it can be seen that CPU time increases substantially with little improvement in accuracy as the number of bands increase from three to four. The result affirmed with the result obtained by Mausel et. al. (1990) that best three to four bands are suitable for classification along with proper saving in

CPU time. This study further enhanced their statement that only best three bands are suitable for classification along with considerable saving in CPU time.

6.2 DISCUSSIONS

The results are discussed as below :

i) All the feature selection techniques give similar results for prediction of best two and three bands combinations.

ii) In most of the cases, the best bands combination for separating any two classes has been found to be same as that for separating all the classes.

iii) For two different classes, the addition of an extra band never decreases the class separability.

iv) The best three bands combination also includes both the bands of best two bands combination.

v) For inter class separability, the first four feature selection techniques produce different results in some of the cases as depicted in Table 6.6.

vi) The BVOI technique is unable to determine the best band combination for inter class separability.

vii) The overall classification accuracy has been found to be 86.86% for best bands combination of two bands (4 and 7). The classification accuracy has been reduced to 80% when

TABLE 6.8 : Details of Classified Image and Accuracy Assessment Using Best Combination of Bands 4 and 7

(A) ERROR MATRIX:

ACTUAL	INTERPRETED CLASS								
CLASS	1	2	3	4	5	6	TOTAL	ACC.(%)	OM.(%)
1.	343	0	0	0	0	0	343	100.00	0.00
2.	0	190	40	28	0	0	258	73.64	26.36
3.	0	18	182	14	0	0	214	85.05	14.95
4.	0	0	0	180	0	24	204	88.23	11.77
5.	0	0	0	0	188	23	211	89.10	10.90
6.	0	0	0	0	28	74	102	72.55	27.45
TOTAL:	343	208	222	222	216	121	1332		
COM. (%)	:0.0	8.85	18.0	18.9	13.0	38.8			

OVERALL CLASSIFICATION ACCURACY: 86.86 %
AVERAGE CLASSIFICATION ACCURACY: 84.76 %

OM. - % Error of Omission , COM.- % Error of Commission
ACC. - % Accuracy of the Class.

(B) DETAILS OF CLASSIFICATION :

CLASS :	1	2	3	4	5	6
GRAY :	41	82	123	164	205	246
PIXELS :	25234	95238	8766	13806	75076	44014
% AREA :	9.63	36.33	3.35	5.27	28.64	16.79

ELAPSED TIME IN CLASSIFICATION :00:02:02.34

CLASS - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand , 4 - Wet sand , 5 - Urban areas , 6 - Boulders

TABLE 6.9 : Details of Classified Image and Accuracy Assessment Using Worst Useful Combination of Bands 1 and 3

(A) ERROR MATRIX:

ACTUAL CLASS	INTERPRETED CLASS 1	2	3	4	5	6	TOTAL	ACC. (%)	OM. (%)
1.	326	17	0	0	0	0	343	95.04	4.96
2.	0	171	45	42	0	0	258	66.28	33.72
3.	0	0	170	24	20	0	214	79.43	20.57
4.	0	0	18	160	26	0	204	78.43	21.57
5.	0	0	1	10	170	30	211	80.57	19.43
6.	0	0	0	13	20	69	102	67.24	32.36
TOTAL:	326	188	234	249	216	99	1332		
COM. (%) :	0.0	9.0	27.4	35.4	21.3	30.3			

OVERALL CLASSIFICATION ACCURACY: 80.00 %
AVERAGE CLASSIFICATION ACCURACY: 77.89 %

OM. - Error of Omission , COM. - % Error of Commission , ACC. - % Accuracy of the Class.

(B) DETAILS OF CLASSIFICATION :

CLASS :	1	2	3	4	5	6
GRAY :	41	82	123	164	205	246
PIXELS :	24956	104359	10512	13710	63544	45063
% AREA :	9.52	39.81	4.01	5.23	24.24	17.19

ELAPSED TIME IN CLASSIFICATION :00:02:04.48

CLASS - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand , 4 - Wet sand , 5 - Urban areas , 6 - Boulders

TABLE 6.10 : Details of Classified Image and Accuracy Assessment Using Best Combination of Bands 3, 4 and 7

(A) ERROR MATRIX:

ACTUAL CLASS	INTERPRETED CLASS 1	2	3	4	5	6	TOTAL	ACC.(%)	OM.(%)
1.	343	0	0	0	0	0	343	100.00	0.00
2.	0	258	0	0	0	0	258	100.00	0.00
3.	0	0	206	6	0	2	214	96.26	3.74
4.	0	0	0	198	0	6	204	97.06	2.94
5.	0	0	0	0	197	14	211	93.36	6.64
6.	0	0	0	0	8	94	102	92.16	7.84
TOTAL:	343	258	206	204	205	116	1332		
COM. (%) :	0.0	0.0	0.0	2.9	3.9	19.0			

OVERALL CLASSIFICATION ACCURACY: 97.30 %
AVERAGE CLASSIFICATION ACCURACY: 96.47 %

OM. - Error of Omission , COM.(%) - % Error of Commission , ACC. - Accuracy of the Class.

(B) DETAILS OF CLASSIFICATION :

CLASS :	1	2	3	4	5	6
GRAY :	41	82	123	164	205	246
PIXELS :	25304	107702	8355	13011	7709	30063
% AREA :	9.65	41.09	3.19	4 .96	29.64	11.47

ELAPSED TIME IN CLASSIFICATION : 00:03:14.28

CLASS - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand , 4 - Wet sand , 5 - Urban areas , 6 - Boulders

TABLE 6.11 : Details of Classified Image and Accuracy Assessment Using Worst Combination of Bands 1 , 3 and 4

(A) ERROR MATRIX:

ACTUAL	INTERPRETED CLASS								
CLASS	1	2	3	4	5	6	TOTAL	ACC.(%)	OM.(%)
1.	342	1	0	0	0	0	343	99.71	0.29
2.	0	237	10	10	1	0	258	91.86	8.14
3.	0	0	188	26	0	0	214	87.85	12.15
4.	0	0	0	172	10	22	204	84.31	15.69
5.	0	0	0	10	173	28	211	81.99	18.01
6.	0	0	0	8	13	81	102	79.41	20.59
TOTAL:	342	238	198	226	197	131	1332		
COM. (%) :	0.0	0.4	5.0	23.9	12.2	38.2			

OVERALL CLASSIFICATION ACCURACY: 89.56 %
AVERAGE CLASSIFICATION ACCURACY: 87.52 %

OM. - Error of Omission , COM. - Error of Commission ,
ACC. - Accuracy of the Class.

(B) DETAILS OF CLASSIFICATION :

CLASS :	1	2	3	4	5	6
GRAY :	41	82	123	164	205	246
PIXELS :	22974	109296	7868	12866	63167	45973
% AREA :	8.76	41.69	3.0	4.91	24.10	17.54

ELAPSED TIME IN CLASSIFICATION :00:03:16.56

CLASS - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand ,
4 - Wet sand , 5 - Urban areas , 6 - Boulders

TABLE 6.12 : Details of Classified Image and Accuracy assessment Using all the Four TM Bands (Bands 1 3, 4 and 7).

(A) ERROR MATRIX:

ACTUAL	INTERPRETED CLASS								
CLASS	1	2	3	4	5	6	TOTAL	ACC. (%)	OM. (%)
1.	343	0	0	0	0	0	343	100.00	0.00
2.	0	258	0	0	0	0	258	100.00	0.00
3.	0	0	206	6	0	2	214	96.26	3.74
4.	0	0	0	202	0	2	204	99.02	0.98
5.	0	0	0	0	197	14	211	93.36	6.64
6.	0	0	0	0	3	99	102	97.06	2.94
TOTAL:	343	258	206	208	200	117	1332		
COM. (%) :	0.0	0.0	0.0	2.9	1.5	15.4			

OVERALL CLASSIFICATION ACCURACY: 97.97 %
AVERAGE CLASSIFICATION ACCURACY: 97.62 %

OM. - Error of Omission , COM. - Error of Commission ,
ACC. - Accuracy of the Class.

(B) DETAILS OF CLASSIFICATION :

CLASS :	1	2	3	4	5	6
GRAY :	41	82	123	164	205	246
PIXELS:	25528	114574	8299	12788	74959	25995
% AREA :	9.73	43.71	3.17	4.98	28.59	9.92

ELAPSED TIME IN CLASSIFICATION : 00:04:58.24

CLASS - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand , 4 - Wet sand , 5 - Urban areas , 6 - Boulders

the worst combination of two bands 1 and 3 is used for classification. Similarly, the classification accuracy using best three bands and worst three bands, have been found to be 97.30% and 89.56% respectively. Hence, the improved accuracy using best combination of bands, identified by feature selection techniques indicates its advantage and utility in digital remote sensing analysis.

viii) Increase in spectral bands for classification after three bands hardly increases the accuracy as shown in Fig. 6.1. On the other hand, the CPU time required for classification considerably increases with an increase in spectral bands as shown in Fig. 6.2.

ix) Errors of omission and commission for all the classes decreases with an increase in spectral bands.

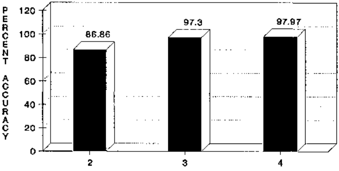


FIG. 6.1 OVERALL CLASSIFICATION ACCURACY FOR COMBINATION OF BANDS (TWO TO FOUR)

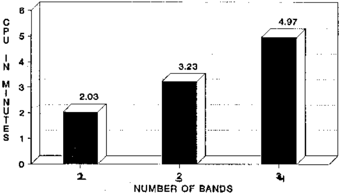
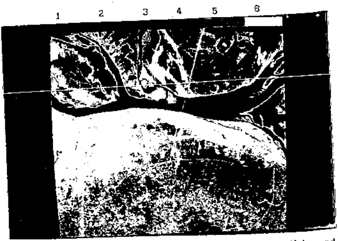


FIG. 6.2 CPU TIME (IN MIN.) REQUIRED FOR CLASSIFICATION : TWO TO FOUR BANDS

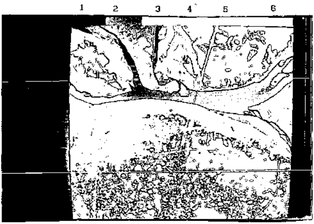
1 - Water bodies, 2- Vegetation, 3- Dry sand, 4- Wet sand
5- Urban areas, 6- Boulders

PLATE 8.1 : CLASSIFIED IMAGE, USING BANDS 4 AND 7



1 - Water bodies, 2- Vegetation, 3- Dry sand, 4- Wet sand
5- Urban areas, 6- Boulders

PLATE 8.2 : CLASSIFIED IMAGE, USING BANDS 3, 4 AND 7



1 - Water bodies, 2- Vegetation, 3- Dry sand, 4- Wet sand
5- Urban areas, 6- Boulders

PLATE 6.3 : CLASSIFIED IMAGE, USING BANDS 1,3,4 AND 7

CHAPTER - 7

CONCLUSIONS AND RECOMMENDATIONS

7.1 CONCLUSIONS

In this study, various feature selection techniques have been discussed in order to determine the best combination of bands, to be used subsequently in classification of the digital satellite data. Digital analysis of multispectral LANDSAT-5 TM data has been carried out to assess the utility of various feature selection techniques, viz. Divergence, Transformed Divergence, Bhattacharya Distance, Jeffreys Matusita Distance. From the results of this study, the following conclusions are drawn :

i) Under similar conditions, the result of each feature selection technique from the TM data is same for separating all the classes of interest, in the study area.

ii) The Divergence and Bhattacharya Distance may not be as efficient as other techniques, because their values of separability increase for each class pair even after full separability between the classes have been attained . Both these techniques, however, may give more precise measurements of the statistical distance between the classes, because they do not have a limit and thus, could be used in those studies where actual separability without a saturation value is important.

iii) The Tranformed Divergance and Jeffreys Matusita Distance show almost similar results for predicting the best combination of bands to separate all the classes of interest. Both these methods consider the limit at which full separability is attained, so can be used more effectively for classification purpose. However, out of these two methods, one should prefer Transformed Divergence technique due to its computational efficiency.

iv) BVOI technique is best amongst all the feature selection techniques because it considers only the degree of overlap in brightness values between the classes, which is a major hurdle in the process of classification. Another advantage of this technique is that it requires only maximum and minimum brightness values for each class in each band from training areas which results in a considerable saving of CPU time.

v) The errors of omission and commission can be reduced significantly, which in turn increases the classification accurary, by using best combination of bands.

vi) The errors of omission and commission for classes boulders and urban areas are relatively high in comparison to the other classes, which reflect that these classes may contain some mixed pixels. A detailed ground truth information of such classes, therefore, will be of immense use to predict best combination of bands.

vii) The combination of bands after three bands increases the classification accuracy marginally, whereas, the CPU time increases almost linearly.

7.2 LIMITATIONS OF THE STUDY

(i) The results of the present study are only valid for the TM data in given study area.

(ii) The time difference between the satellite data and the reference data may incorporate error in the result.

(iii) The classification of the image has been carried out using mean classifier, the implementation of other classifiers may change the classification accuracy.

(iv) The accuracy of the classified image was based on certain pixels of each class, which may not represent the true classification accuracy of the entire scene.

(v) The analysis has been carried out using only four bands of LANDSAT-5 TM due to non-availability of data in other bands.

7.3 RECOMMENDATIONS FOR FUTURE WORK

In the present study, an attempt has been made to study the utility of various feature selection techniques, using the readily available satellite data within the time constraints. However, there is a need for further study with the following modifications so that more refined results may be obtained.

(i) Multitemporal data of different seasons and date from other sensors may give more correct information regarding the best feature selection techniques for a particular area.

(ii) Comparative performance of various feature selection techniques for different terrains must be evaluated, so that users may select best combination of bands for best classification of different land use and land cover classes.

(iii) The reference data such as field data and published maps must be collected for the same period as that of satellite data, so that training areas represent true ground condition to assess the accuracy of classification.

It is expected that , the further study with the above modifications, if carried out, may help in obtaining the results of a particular area, which may provide more refined classification of digital satellite data.

REFERENCES

1. Campbell, J.B.(1987), "Introduction to Remote Sensing", The Guilford press, New York, pp. 269-273.
2. Fitzpatrick - Lins, K. (1980), "The Accuracy of Selected Land Use and Land Cover Maps at Scales of 1:250,000 and 1:100,000", Journal of Research, U.S. Geological Survey, Vol.6, pp. 169-173.
3. Ghosh, S.K.(1991), "River Basin Management Using Remote Sensing Data Analysis : An Assessment of Data Input to Hydrological Modelling as Derived from Satellite Data", Ph.D. Thesis, University of Stratchlclyde, Scotland, U.K., pp. 137-143.
4. Haack, B.N.(1983), "An Analysis of Thematic Mapper Simulation Data for Urban Environments", Remote Sensing of Environment, Vol. 13, pp. 265-275.
5. Jensen, J.R.(1986), "Introductory Digital Image Processing ", Prentice Hall, Englewood Cliffs, New Jersey, pp.176-231.
6. Kamat, D.S.(1978), "Selection of Multispectral Data in Remote Sensing Application", Satellite for Earth Observation (SEO): Workshop on Digital Processing of Remotely Sensed Data, ISRO-space Application Centre, Ahmedabad, pp. 163-185.
7. Kumar, A. (1993), "Study of Planform Changes in a Selected Reach of an Alluvial River Using Remote Sensing". M.E. Thesis, W.R.D.T.C., University of Roorkee, Roorkee.

8. Lillesand, T.M. and Kiefer, R.W. (1979), "Remote Sensing and Image Interpretation", John Willey & Sons, New York, pp. 461-477.

9. Ma, Z. and Olson, C.E.(1989), "A Measurement of Spectral Overlap among Cover Types", Photogrammetric Engineering and Remote Sensing, Vol. 55, No.10, pp. 1441-1444.

10. Mather, P.M.(1987), "Computer Processing of Remotely Sensed Images", John Wiley & Sons, New York.

11. Mausel, P.W., Kramber, W.J. and Lee, J.K.(1990), "Optimum Band Selection for Supervised Classification of Multispectral Data", Photogrammetric Engineering and Remote Sensing, Vol.56, No.1, pp. 55-60.

12. Saha, S.K. and Kudrat, M.(1991), "Selection of Spectral Band Combination for Land Use/Land Cover Classification Using a Brightness Value Overlapping Index (BVOI)", Photonirvachak Vol. 19, No.3, pp. 141-147.

13. Sheffield, C.(1985), " Selection of Band Combinations from Multispectral Data", Photogrammetric Engineering and Remote Sensing", Vol. 51, No.6, pp. 681-687.

14. Swain, P.H. and Davis, S.M.(1978), "Remote Sensing : The Quantitative Approach", McGraw - Hill, New York, pp. 164-175.

15. Swain, P.H. and King, R.C.(1973), "Two Effective Feature Selection Criteria for Multispectral Remote Sensing", LARS Technical Report 042673.

16. Swain, P.H. and Wacker A.G.(1971), "Comparison of Divergence and B - Distance in Feature Selection", West Lafayette, Ind : Laboratory for the Application of Remote Sensing, Technical Report 020871.

17. Thomas, I.L., Bening, V.M. and Ching, N.D.(1987), "Classification of Remotely Sensed Images", Adam Hilger, Bristol, U.K., pp. 83-101.