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

ín

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

CARABATE'S BEGLARATION

J benefor carity that the sume which is builty to exact in the dimenticities and high of "Withow at "writene Facebars" in general containers in highlat denote therefore heatpeter", in general efficiency of the requirements for the assist of the degree of efficiency of the requirements for the assist of the degree of "OLE CONDUC AND INFORMATION OF the assist of the degree of a contract of the requirements for the assist of the degree of a contract of each high the degree of the degree of the automatic rescent of any our work correled out for a period of the automatic rescent of any our work correled out for a period of the automatics of the form the balancy. The mater the approximation of Fer. F.K. Gang, heatmen and be, S.K. Gheeb, any rescentions of Fer. 10.10 (1993) the heatment of for the experiment of the other Orbits the phase atom. Mater and the defention of the sector of the sector of the sector of the degree of the sector of t

The matter embedded in this dissectorion and here, subsidied by ne for the and of now other degree or drillogo.

33800 KR + 30A 0

10-81-10-36 - 13-1-1

OTCEPAK KARAN JULE

There is a conclusive the clove matement work by Proposition is a conclusive to the left of we know to get.

Unit S.K. Ghosh) Lesturer Civil Engel (r.J.L. Milversitz of Rootkee (*/200788-2007) (*)

(hr. P.K. Garg) Lecturer Givil, Eauge, Lapit. University of Hoorker Provist, - 207 687

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the dissertation entitled, "A Study of Various Feature Sciection Techniques in Digital Remote Sensing Analysis", in partial fulfilment of the requirements for the award of the degree of MASTER OF EMENERING INCLUMENTING FUNCTION SCIENCE SCIENCE AND PHOTOGRAMMETRIC FEATURERING, submitted in the DEPARTMENT OF CIVIL EMEINEERING, UNIVERSITY OF ROOMEKE, ROOMEKE, is an authentic record of my own work carried out for a period of about seven months during July, 1933 to January, 1934 under the supervision of Dr. P.K. Garg, Lecturer and Dr. S.K. Ghoeh, Lecturer, Department of Civil Engineering, University of Roomeke, Rookee.

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

Deepon Kumer Jhe. (DEEPAK KUHAR JHA)

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

10 4000 (Dr. S.K. Ghosh)

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

(Dr. P.K. Gar Lecturer Civil Engg. Deptt. University of Roorkee BOORKEE - 247 667

(1)

CIRCHERD LINESSON

It is a tisiture observe to emerge ap they seems of gratitude and linkelicument to Dr.P.K.Gaeg, instances and he S.K.Gaeg, instance, henceted of OP11 Depresenting, between your of the theory of and the test statements of the state of the theory of the test and the test and the state of the state of the test and the test and the state of the state of the alternative of the state of the state of the state of the alternative of the state of the state of the state of the alternative of the state of the state of the state of the alternative of the state of the state of the state of the alternative of the state of

Section and Sec Youri Julia, G.G.K.E., B.G.F.F.C. Section for mostling as ISBURGENE and advancessing Auxiliation in the Configure 1 along place as easies of any hamiltaintees in the face-flig at a the the Section, for under constants but area into accessing at a the the Section, for under constants but area to the parents.

J. was also liseding to by Proceeds the William Miss. Miss. Proceeds Barray, Mr. Statjaw, Miss. Her. Brange, Adurg and Mr. Michael Proceed one (both trappingtion and processing-potent atthemet which that was 'n '' on here ways and Philand Experiment of another that accords.

Sporter thanks must be applied by Friends Mr. Friedred. Dr. Sonfest Wr. 751 Mr. Stored Mongelik who hollood as and every and of information. Should also a device that every whey fore and Mr. France Single Office for the hor to be provided in an also thankful to all block and filled m directly on uniteractly during as course of study. (must show that

0.07

ACKNOWLEDGEMENTS

It is a distinct pleasure to express my deep sense of gratitude and indebiddness 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 ne ISROVISION and other necessary facilities in the Section. I also place on record my indebtedness to the faculty of B.S.P.E. Section, for their completent height work.

I am also thankful to sy friends Mr. K.K.Sinha, Mims. Bhawana Sharam, Mr. Sanjeev Jain, Mr. Romesh Adury and Mr. Arbind Pramad for their inspiration and encouragement without which it would have been very difficult for ms to complete this rourse.

Special thanks are due to my friends Mr. P.Rajesh, Mr. Santesh Jaa and Mr. Horraj Manglik who helped me at each and every step of this work. Thanks are also due to Mr. A.K.Nigan, Mr. Vijay Veer and Mr. Pretsp Singh (Photo tech.) for their help whenever needed. I as also thankful to all those who helped me directly or indirectly during my course of study. Decket Kumar Ba

(DEEPAK KUMAR JHA)

(11)

The purpose of system for the content of established and the cast and and the data is for evolve a trade-iff between the cast and of a standard, of information of the data and the analysis of the cast and the analysis of the cast of the analysis of the cast of the cast

over his been conclude out in the steamer. In the itree star, beache selective telecitiques have been applied to under to first end. The list, condition of two and these beens. In accust when the selective heave has been attestified the of concays why quarks telephane. By early early established the off concays why builders using the loci contraction of bends, in a wellten list builders using the loci contraction of bends, in a wellten list color-litection has been real with get attends of an and then list builders and the loci contraction of bends. In a wellten list builders are real with the been real with get build be been there are a build be to attend to the list of the been there is a heap to be a

(0.0)

SYNOPSIS

The purpose of applying feature selection techniques in multipactral data is to provide a trade-off between the cost and the accuracy of classification in order to reduce the dimensionality of satclifte 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, Paina and its surrounding area lying in the Binar State of India has been selected as the study area. Various feature selection techniques such as Divergence, Insaformed Divergence, Bhatjacharya Distance, Jeffrey Mausita Distance and Brightness Value Overlapping Index (EVOI) have been used employing digital satellite data of LMDSAT-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 ut the best combination of two and three bands. In second stage, the entire image has been classified into six classes viz. water bodies, vegotation, 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. Accure, assessment of the classification has late been

(111)

[1t is suggested thatWiheCaP4a@THMWheg while the suggested thatWiheCaP4a@THMWheg while the suggested thatWiheCaP4a@THMWheg while the suggested to a suggested to the suggeste

	BOX ENQUALBERT NOTIFIC STOLATE OF 310H	.8
	NEMOTE SENSING ANALYSIS	
:	WORLAND FUELDED JAUMAH I.S	
6	2.8 BEDING WILLING WILLING ST	
	S'N BOE OF FEATURE SELECTION TECHNIQUES	
	STREAMER TO LITERARY AND INDUCTION FOR	
a -	CLASSIFICN LUE ACTING	
	VARIABLE FEATOR: SELECTION TECHNIQUES	. 82
g^+	BUITE CONTRACT 1.8	
	PERCENTION FOR MOTOR BELLEVILLE	
: [opplagres/70 T.S.E	
	2.2 i 5 Philamatical Images Lies	
C7	or Divergence	
.1	3.2.2 Transferent Idvargence.	
	and the second sec	

(1v)

Description of the same of CONTENTS — statistic the analysis of the first of CONTENTS (Second Second Sec

 BACKGROUND soldperse endoord to surptombed i floan toot , and , d.2. UFILITY-OF FEATURE/SELECTION of the testandor of this , and the surptomber of the second surptomber of the surptomber of the second surptom

1.4 ORGANISATION OF STHE STREETS LONGIDARY IN THE STREETS OF BOTTOM

RENOTE SENSING ANALYSIS

з.

2.1 MANUAL INTERPRETATION	4
2.2 DIGITAL INTERPRETATION	4
2.3 ROLE OF FEATURE SELECTION TECHINIQUES	5
2.4 INFORMATION REQUIREMENTS OF SUPERVISED	
CLASSIFICATION ALGORITHHS	6
VARIOUS FEATURE SELECTION TECHNIQUES	
3.1 INTRODUCTION	9
3.2 FEATURE SELECTION TECHINIQUES	10
3.2.1 Divergence	11
3,2.1.1 Mathematical Properties	
of Divergence	13
3.2.2 Transformed Divergence	14
3.2.3 Bhattacharya Distance	15

至于11日日

	$g_{13}(g_{1})g_{1}$ 3: 2: 4 (Jeffreys: Natural Distance $f:\mathcal{M} \to \mathcal{K}$	16				
	3.2.5 d. Brightness Value Overlapping Index					
	1.3.3 CHOICE OF A REAFURE SELECTION TECHNIQUE	19				
	143.4 CASE STUDIES of Line 11 and 1 and 12					
1	3.4.1. Feature Selection of Multispectari					
	seitell committe Senstrig Dabar 1. Seiter	21				
Q.	3.4.2 Optimum by BandgesSelection for					
11	works to a Supervised Classification "Well B. H	22				
	12190 13131 34 ImprovementA2 30 Inol LADT Wegetal Fon					
45	Separability Using Wansformed					
1.5	Divergence PELFRESEX YURDIDA V.B	23				
	3.4.4 Selection Werseuspectral ST Bank	, Э				
tap.	Combination for Land Covery Land Use					
[21) [21]	eraded deficients of the transformation of the second second second second second second second second second s					
	easion! yJ i Natrus-Over Tapping ¹ Tixlex ² (BVOI) ¹	24				
4.	DESCRIPTION OF THE STUDY AREA AND DATA USED					
	1-4: 1- SALIENT (FEATURES) OF THE STUDY ABEA 28					
	4.1. p. Elimette et lausch egentitet	29				
: t	4.1v2 Land Use and Land Cover Classes	30				
·	de 412 DATA DEEDE CARTE SIGNE SIGNATION	31				
	4:3 FEQUIPMENT USED to mailedungand dir 3	31				
5.58	METHODOLOGY (1098) Abbel 251000 from 1					
	SOT IDENTIFICATION OF CLASSES DOLLD MILLS	37				
1502	5.2 SELECTION AND EXTRACTION OF TRAINING AREA	37				
1.4	5.3 DETERMINATION OF TRAINING DATA STATISTICS	40				

	5.4 APPLICATION OF FEATUREI SELECTION (TECHINIQUES	40 ·
÷.,	> 5.4.1. Determination of Separability Indices	
911	Trailsformed	
13	Divergence, BhattacharyarDistance and	
	:	41
·	5.4.2 Determination of of Brightness Value	
	(a) entapping/Indexe1000 E.A.C	42
57	5.5 IDENTIFICATION: OF BEST-COMPNAREON OF BANDS	42
	5,8 CLASSIFICATION OF SATELLIVE GATACUSING BEST	
	Have COMBLINATION, OF BANDS of Chemody 2	42
1-5	5.7 ACCURACY ASSESSMENT COROSTOVIO	48
6.	RESULTS AND DISCUSSION ROLLOSIOS AND DISCUSSION	
	$\theta_{ij} 1_{j-j} \mathbf{RESNLTS}_{\mathrm{finit}-\mathrm{Ext}(i_{i})} = \dots = \dots = (1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,$	49
		49
b <	6.1.2 Computation of Separability Indices	
	GEPU ATAUS INS. AND Divergenner, 30 HolinarisConnied	
£S.	A HAI wergence; Baattscharya Distinate and	1
~5	Jeffreys Matualta Distance . 1 . 4	54
:42	Bala 2.1 in Heing Two Bandar Shubhation	54
10	6.1.2.2 Using Three Bands Gombination	62
18	6.1.3 Computation of Britshtmassauthed we h	
	Overlapping Index (BVOI) Y00.00200108	62
VE	6.1.4 Classification of imagenend data Accu-	
	S 2 SELECTION AND EXPROMINATION VIEW MING AREA	68
521	V38Y DRINEARA IMAGENEMIX ON UDITODAG 2 2	

7.

CONCLUSIONS AND RECOMMENDATIONS

7.1 CONCLUSIONS

7.2 LIMITATIONS OF THE STUDY

80 OKGONOXUMBE 1.1

-insten 7.3 RECOMMENDATIONS FOR FUTURE WORK at application of 82 entrant to REFERENCES ("BERLIARS" as investing immediation to 84 c. antessing, the treat features fully itself in oppositelised account lot tracked out no addicty acquired includence to one' contract. individuality state open in a set of a state of the state individual of the state o band or coabination of bands value, that carry information encounting mystemette and then in the scene. Thus, feature a looking may make be called or 'Haformation Extra that' for not tentreparation and the elemenation can be taken with the policinal that to extremely belowing on the entremeted of media of the is botto to throate conduction and a (1980). (1) superior search "useful information " in the tange data, rather than as a physical indust protect on the estimation and second feature to book is the process of industing the most sports hands of the establishe data not for conther analysis while discouling the lane production decores and an and an endowed the con-. (Vii) . Hodowill

HORIGALCE SERVERS, SO VELLEV S.I.

In enalyzing the mittaperian reachs, perchanded to the followextention place an bijection unit. For permute, none the file matrixed matrix is a complex, yet with the inderivation for the file of the percent of the second sec

CHAPTER 1

INTRODUCTION

1.1 (BACKGROUND

naves and to

The classification of Safelliffe data is based on certain selected measurements, known as "FEATURES". In "conflext" 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 suitispectral 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 satcilite 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 reacts sensing data, feature melection playm an laportant role. For example, proceeding the multisend;satellite data by a computer, cost will be substantially high if all the available bands are taken into consideration.

Feature, selection.techniques.provider a trade-off. between the cost ' of classification and accuracy with considerable reduction in computation time by selecting the optimum combination configuration containing maximum information about the land cover classes PERSONAL STATE OF MONTAVENAMED, 7:1 (Kazat. 1978). In principle, feature selection attempts to on prestrictly and then eliminate those bands which carry repetitive would hidden out information as in other bands. Resulting satellite data set now Chapter 2 deals with the role of feature anisotics tertelones obtained, contains maximum information (Campbell, 1987). for remin sensing data analysis. att cafforen E nalasd) 1.3. ORJECTIVE OF THE SPRESENT STUDY tray is atmonthelms isothermanica mathematical arelationships (ane available stor selects the hopf staff)

combination of bands from the; total "number of available bands (or a sensor. Unfortunately, no proper assessment regarding the utility)

of such_methods_bave_been_comried out_still now_finethis_study_

It is proposed to carry out an investigation to massas the difficu

of_{ell}all...the.:features.selections/dechniques.s/Mix Divergence/ Transformed Divergence. Bhathabharya.Distance/JJfffreyb/#Browsita

Distance and Brightness Value Overlappng Index (BVOI).

Other objectives of this study are defined as follows :

- To develop computer programs of all the feature selection techniques in order to differentiate one land cover class from the others.
- (11) To identify the best two bands and best three bands of LANDSAT-TH data for land use/land cover class! fication using feature selection techniques in the study area, and

(111) To assess and compare the classification accuracy of land computation tise by asincting the uptimum combinat, shaad MT words consistences busiced in the test stress was a series 1.4 ORGANISATION OF THE THESIS In principly, induces scholler streamly in 00000 Also de The whole work has been presented in seven chapters as settledged writes a the writed shadd big dual bis (11455) described below. ". Surmation as in other builds. Peopliting pulpilite data set as a Chapter 2 deals with the role of feature selection techniques contains maximum information. Comparin, 1980; 181.4 for remote sensing data analysis. Chapter 3 describes the mathematical relationship of various Fedture Se Fedture Herbidies. which gean be used for the (selection b) best bahas combination. Chapten 4 gives the Information regarding the study area and the satellite.data used. . Chapter 5 jout lines the set body adopted and the analysis procedure of satelitte data ushist feature selection (dechniques) (Chapter 58 Thearibes" the Danalystal of satellite of data meaned; theirs in esults of which so i for reasons who discussions: while sim tachabters(i79[othe930.h0nblusions[iSand recommendations of the study (ane given the the second to compare of the second to the

ilutions and Brightness Value Overlapping Index (2007).

b r diputies of this shall are defined as follows -

- (1) To develop computer programs of all ing restore reduction residences to order to differentiate one hand on or stags (resetibe uthoug).
- (1) In identify the best two hards and both there rends of the second second marking and second second second and the second second and construct a first second second

3.5

CHAPTER 2

ROLE OF FEATURE SELECTION TECHNIQUES FOR

REMOTE SENSING DATA ANALYSIS

Reacts 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 HANUAL INTERPRETATION

This method of image analysis makes use of some interpretation elements, viz. tone, texture, location, pattern, shape and shadow, alongwith the analysid decision, to analyze hardcopy remote sensing data. However, this approach of analyzis is not so effective because the perception of a human being in limited only to 10-15 grey 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 from the re-trailed to back of a regime. entire image.As. compared into provided to classification, in Unsupervised classification does not require detailed ground reference data at the start of classification, which is technique; the computer groups the entire image into different spectral classes, depending upon their reflectance values a laThe analysto then assigned these spectral classes date defenational classes based of iground reference data or knowledge of the area. cTech (Rybe-94 classification, two or more classification algorithms may be used. For example, a Hybrid classifler having Parallelopiped and Maximum Likelihood algorithms, first uses the Parallelopiped and then 2.4 TRADEGATION REQUIREMENT OF SUPERIOR ACCOUNTS ALL transfers difficult cases to Maximum Likelihood classifier

n theorem efficient mentionals subsolutions before all off (Hither, 1997). Four was dir, contained bestatete mess (contained) all 4400 unart

Feature .mekectkar/sicohnkyas/.ars/sicohnkyas/.dem/mite classification.di.empde.moning.dem by Superised Approach. In all the three algorithms of Superised supproach which will be explained in the next section, the training data weighted in the training data are collected either invident fild by from reference.data.suchu.es.published.maps.meriaf photographe and Alterpaturg.. The iterating data such or biented for the section of the sectio

located on the digital remote sensing data and training statistics computed. . The training statistics data are then used by various Senture shlerion technicken in a sentime in a sentime of the sentime from all the available bands of a sensor This of this combination of bands thus, "reduces the cost of classification of satelditeordatary computation without and storage wakes Utilitie compromising the accuracy of classification itseels to that and withus, feature shection techniques blayen incontant told for the classification of digital satelite data she found for the avaitable: bands: from as sensor: are liented in vision as a single the sense in th sensor data traintences data on Krauliatan on the you Various Feature Assistation techniques mew explained in the next chapter actual of teast welces an thread a bindel a connect and contribute a compatibility and the second company to add the Price 2.4 INFORMATION REQUIRMENTS OF SUPERVISED CLASSIFICATION ALCORITHMS callingin houtfoil mained of ester fluiting and and The Supervised classification algorithms broadly requires Cista astimu input data in the form of mean, standard deviation, minimum and naximum brightness values and Unitante addaptade intril 3 Her fuel (from each class presention the phage (Matheil/19887), prodeed The various Supervised class blicstrop algor (they are different as plug (1). Manual algorithms of Supersylvering to safet and the on an (11). Paraldslopleeding t will confide that the initial the state of the second state one of the second state of the secon has inchinimum Distance, chassified with Spaint date from Padina the mean brightness) value of theutestwing data for each chick in all spectral bands (Jensen, 1988)/ The Parallelopiped algorithm

requires the sean value as well as standard deviation of the classes of interest in all spectral bunds. The MaxHaus Likelthood classifier requires sean value and variance-covariance satrix of the training data for-ench class-in each bund (dosen, 1985).4

The Minisum Distance classifier_is_further.elaborated wis it has been used for the present study. The advantage of: Minisum Distance classifier is that it is computationally easier and faster. However, it does not produce very accurate results for class havever, it does not produce very accurate results for class haven high variance, as only the mean values are used by this classifier. In this classifier, the distance between each unknown pixel ($W_{1,j,k}$) and seen vector ($\mu_{c,k}$) is calculated and the pixel is assigned to that class whose distance is minisum. The Minisum Distance can be calculated either by using Euclidian Distance based on Phythagorean rule or Round the Block Distance (Sesin and Davis, 1978).

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

$$E_{L} = \sqrt{\sum_{k=1}^{n} (BV_{1,j,k} - \mu_{c,k})^{2}} \qquad \dots (2.1)$$

Where, BV_{1.1.k} = Pixel value at location (i,j) in band k

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

= Number of spectral bands used

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

n .		
$\sum_{\mathbf{k}=\mathbf{i}} (\mathbf{BV}_{\mathbf{i},\mathbf{j},\mathbf{k}} - \boldsymbol{\mu}_{\mathbf{c},\mathbf{k}}) $		(2.2)
the first contraction and the most off	1.01.1	a second of the

In the present study; the Euclidian Distance has been used for the computation of william distance because 11 is most accurate and basic sethod to determine the distance between any t to points. 134.0

and all the second as a second to a state

1012209999901 1.2

to tend as to according they and and the to edited and and فالموصوف المراهية ومستعيد المواجد ومراجب والمراجب والمراجب والمراجع والمراجع والمراجع والمراجع an anti-ante ten cos legi ante i fa valadi applica-- success with success of soft and risks with the set in the second s class soles believed the datar the Distance - subset better vicesion From the analysis procedure that provide only redonded specifies. who she as associated the information compared to the other stand to indem out will all network out we whill if indexid the provide the state of a fair the state of a state of the state of t to consist an initial and the protact fill relatestimetic constructions and designed on a sub-standing designed and stavione (dollations a assistation policeter) another income debenative the depression experimental behaviors the two attempts in training data. Undertailed of index are present with collide initiality starting to their prioriticly which is the object is all galar bradito add must easily done minimization the of (Jac (Japan), 1980)

Statistical satisfies of flocing schedulos are used to space the voir solar the one's divide from all the house of a open that the product the prester open of a flottest signification to be clean of the states of and 'd'. The house product devices recording to the terms town.

ē.

CHAPTER 3

VARIOUS FEATURE SELECTION TECHNIQUES

3.1 INTRODUCTION

Once the training statistics for each class in each hand 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 ' hands at a time (Jensen, 1986).

Statistical setheds 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 'o' and 'd'. The basic problem of spectral recognition is (Jensen, 1986) :

"Given a spectral distribution of data-in-data-bands of remotely sensed data, finding discrimination techniques that will allow separation of major land cover-classes with the minimum of groups and a minimum number of bands with 'unsuble second bat, in)

This problem fis: demonstrated diagrammetical tyrith Fig) 3.1, using two classes in single band data

 VID conserved 1.5.P On examining a typical histogram shown in Fig. 3.2, it is indicated to sample and and and and so south and the found that there is substantial overlap between classes 1 and 4 in satilate set at best specific the clauses used in the analysis band 'a ' and between classes 3 and 4 in band 'b ' When there in the least of found of the state of and a man of a solution of the state of the stat an overlap, any decision rule that one could make to separate or distinguish between two classes must be concerned with the bud edt te en fenideus weihinst is a biene sizte at resented it following two types of errors : tentinenat di testi in ol the second database for (1) A pixel may be assigned to a class to which it does not and the state of the first statement of the state of the belong (an error of commission). a fact amount for Seritum of and a second p (11) A pixel is not assigned to its appropriate class (an error Creating of the of 3

of omission).

...Both types of arrivations be infinitized by beleating advoptions subset.(ioff) theorem e addati deploying the populated to be addet in the subset.) Leading use. If the training data for teach to be addet in the infinite additional additional to be provide a light for the infinite additional to be provide a light for the infinite additional to be provide a light for the infinite additional to be provide a light for the infinite additional to be provide a light for the infinite additional to be provide a light for the infinite additional to be provide a light for the infinite additional to be provide a light for the infinite additional to be additin to be a

Literature reveals that the commonly used separability

to state Divergence we can add attraction determine proposition

(17) (b) Transformed Divergender with gettacht under besaus glattere

to maked (Bhattadharya:Bisbance.co has) only a to the mathematical and .

(d) Jeffreys Matusita Distance: androimum products a bus no. . .

. . . (q) Brightness (Value Sverlapping Index (BV01) (and cut)

attributer of the second state of the second s

Divergence was one of the first measures of statistical is k bon it asserts resulted getter - cellustandes of short great energy separability beetween the two classes used in the machine st when classes i and at a basis i and a to hand 'b'. When there is processing of remote sensor data, and is still widely used as in everythe an declaren of the fail one could ask anto in separate of method of feature selection (Swain and Wacker, 1971; Hanck, 1983). its data borrookee of two private but doneses its toget, its It addresses the basic problem of deciding combination of the best showing to board and served its 'q' bands out of 'n' bands to be used in Supervised the other at delate to enable of characters of the bester \$ 45. classification. The number of combinations 'C' of 'n' hands taken control on a Section and Language q bands at a time is defined as (Jensen, 1986) : $\frac{define (ac)}{C} \frac{define (ac)}{q} = \frac{define (ac)}{q(a-q)!}$... (3.1) Augine to

childrandon error character the concerned strends topological

and part is due.

. 11

$$DIV_{cd} = 0.5 \text{ Tr} \left[(v_c - v_d) (v_d^{-1} - v_c^{-1}) \right]$$

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

where,

The Frace of a matrix (i.e, the sum of the diagonal elements)

c,d = Two classes used for separability analysis or the desire Vo, Vd = The variance-covariance matrices of brighthéss value of the training data for-the-two/efasses(?eran#"dig" under these investigation, and copyrection of generated for separate investigation, and copyrection of generated for separated

Mc, Md = Mcan matrices of brightness values of the training data for the two classes cland d. which we should be $_{\rm eq} \rm VIT$

The sizes of the variance-covariance matrices "V 218 V and the mean matrices 'M' & 'M' are function of the number of bands $d_{1}^{(i)} = d_{1}^{(i)} + d_{1}^$ used in the training process (i.e., if three bands were trained 'M,' & 'M,' would be matrices of 3x1 dimension). 0 = _____VUI (1) m. The bears of the real to conserve it with Although. Divergence only provides a measure of the distance reactive to itself is zero the started between any two class density, its use can (inclush) be extended for multiclasses by taking the average of Divergence values over all 0¹⁷ 60 For two different likeligend from an possible pairs of classes, taking any two classes 'c' and 'd' at in a golf actions events at arright will a time, while holding the subset of band 'q', constant. Another to estimate another is a subset of band 'q', constant. Another subset of bands 'q' is selected for the same 'm' classes and every everyon the information and is a subset of bands 'q'. 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):

$$\frac{\sum_{i=1}^{m-1} \sum_{j=1}^{m} DIV_{cd}}{C}$$

where C = Possible number of combinations for selecting a pair of $\omega_{\rm eff}$, classes.

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

3.2.1.1 Mathematical Properties of Divergence

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

of seignrath by the contribution and (1) DIV = 0 The Divergence of one likelihood distribution . He was a serve anticipation a personal of the analysis of relative to itself is zero (the classes are a vit of sea density, the use one nal issarte of identical) $\begin{array}{c} (11) & \text{DIV}_{cd} > 0 \end{array} \quad \begin{array}{c} \text{For two different} & \text{SM} (5) & \text{SM} (1) & \text{SM$ in the advectory taking any two observer for and the start in Divergence is always greater than zero. the particular to the participation of the same that, offer sec in. the same n dimensional feature space. More details about this technique are given in Swip and King all and garest the second region of the second and we (1973).

3.2.2 Transformed Divergence (DIV_{cd}^T), Account of the state of the second state of

 $\overset{\text{def}}{\longrightarrow} (\overset{\text{def}}{\overset{\text{de}}}{\overset{\text{de}}{\overset{\text{de}}}{\overset{\text{de}}{\overset{\text{de}}}{\overset{\text{de}}{\overset{\text{de}}}{\overset{\text{de}}}{\overset{\text{de}}{\overset{\text{de}}}}{\overset{\text{de}}}{\overset{\text{de}}}}{\overset{\overset{de}}}}{\overset{\overset{de}}}{\overset{\overset{de}}}$

This statistics gives an exponentially decreasing, weight to increasing distance bytween the classes 'o', and 'd', it, also scales the Divergence value to lie between, 0, and, 2000, i.A. Transformed Divergence value of 2000 suggests accellent, class separtion. In bytween 1980, and 2000, provides good separation. If its value tips bytween 1700, and 1900, then the separation is fair.

where Mc, Md, Vc, Vd, T, c, and d, are defined previously and (5, 6) and (5, 6)

To be table the fact of "hands from the original" of bands for a septembility of " is" of anise. The "Built tacknings" is the talk of the twent"such of the given log parts of chasses for all possible ways for choosing in bands from in" discussions. The best is taken the set of the fact of the set of the "Built tack of the "bands from the set of the fact of the "all control of the "bands for the set of the "bands of the set of the "bands for the "bands for the set of the "bands of the set of the "bands for the "bands for the bands of the "bands of the "bands for the "bands for the "bands of the "bands of the "bands for the "bands for the bands of the "bands of the "bands for the "bands for the bands of the "bands of the "bands for the bands of the "bands of the "bands of the bands of the "bands for the "bands of the "bands of the bands of the bands of the bands of the set of the "bands of the bands of

$$\underset{\text{where a light by the second result of the second rescond rescond result of the second result of the second$$

It has been found that Bhattacharya Distance is more appropriate to inter class separability problems than the Divergence when the class probability distributions are beneficiary Newwer; when the classes are woll defined both the Bhattacharya

Distance and the Divergence approaches yield similarresults (Thomas et. ml., 1987).

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

3.2.4 Jeffreys Matumita Distance (J_{cd})

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

$$J_{cd} = 2 [1 - EXP (-B_{cd})]$$
 ... (3.7)

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

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

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

of bands and that subset of bands is chosen for classification which gives the maximum $J_{_{\rm AUC}}$ 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" (BVDI). The method is simpler as compared to those discussed earlier, sainly 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, 1988)

$$F_{j,k} = \sum_{k=1}^{N_{j,k}} f(x_{1,k})$$
 ...(3.8)

$$F_{a,j} = \frac{1}{H} \sum_{k=1}^{H} F_{j,k}$$
 (3.10)

$$F_{ta} = \sum_{j=1}^{r} F_{a,j} \qquad \dots (3.11)$$

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

 $f(x_{1,k}) =$ Frequency of brightness value $x_{1,k}$. $N_{j,k} =$ The range of brightness value within classs

^FJ,k = Cumulative frequency for class 'J' of band 'k' M_____ Number of spectral bands,

F_{a,j} = Average cumulative frequency over all bands'of gene in sectors overclass: ij's bais over sciences with the solution (1)

 $N = - \cos \theta \mbox{ in the data set} \mbox{ and } 0$ $get e E_{ta} = 0; \quad e = Total \mbox{ average cumulative-frequency over all set over the set of the set over the set over$

If overlap does not exists hangest the classes in band 'k', then, they does not exist on down only (11) has (11 signal for all (11) only to carried exists on the contract of the operator of a stageod (vi)

$$\label{eq:Field} \begin{split} F_{1,k} &= \sum_{i} |F_{1,k}| = F_0 = 100 \mbox{ percentry and data for M/(0(12)) $$ and M/(0(12)) $$ and$$

so, Fig. is also the band advings of F_{tk} is the control with last where F_0 is defined as the control to first the information of the info

if overlap exists amongst the classes in band "k" then "

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

The steps involved for the computation of BVOI are summarised below, so its open symmetry in telescomptions in the interview of the steps

- Determine the maximum and minimum brightness values of each class from training data of each spectral band.
- (11) ..Determine --the cumulative: percentage of pixels diaving brightness value ranging from the minimum to maximum for the scholars, based on the histograms of the schole data set.
- (iii) Repeat steps (i) and (ii) for each class in each band.
- (iv) Compute the average of the cumulative percentage of pixels $[s_1 i_2 s_1 i_2 s_1]$ bands for each class $(1 + 1 + s_1 + 1) = \frac{1}{2} + s_2 + \frac{1}{2}$
 - (v) Sum the cumulative percentages of pixels for all classes in each band and sum the averages of comulative percentage of pixels in all bands for each class.
- (v)11 The BVDL selier of calves data set is determined by dividing the sum of the same of call, bands for seadicing by the must of cale of the second of the sead of the sead of the second second

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 in terms of the computer resources, applicability to one or more dimension, ability to reliably assess soparability, edec, & Acknowledging the above factors and previous experience; Thomas et. al.. (1967) suggested the following: choices of différént feature election techniques in different situation. In Action 2010

- (1) 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. A set of a set of the set
- (11) If the same altualion as (1) applied but the Whole set? of bands is to be used then our choice will be the Transformed in Obvergence. The product and there also not a finite contract of the State (1) and (1
- (111).If the classes are less truly honogeneous and a limited bet of bands is to be used then our choice will be / Békitachárýň Distance.

 (1) The same situation as in (11) applied, but the "whole"
 - set of bands is to be used then our choice will be Jefferys?
 - (v) The BVOL technique is applicable; in all "the" cases case mentioned above. A second procession of the second secon

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.

- Lagrando Marcina, and a second for case and a second for the second for the

3.4 CASE STUDIES

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

and the second state of th

3.4.1 Feature Selection of Multispectral Remote Sensing Data

The Divergence and Transformed Divergence distances were used by; Kanat.,(1978); with the ISRO - Multispectral Scanner for detempination...of sland; use; pattern in Panchamhani District of Guarat. 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 fgur classes, viz. water, barren land, forcetand 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 dust of the left of the second se

(b) Selection of the best three bands from the available four of unlimities it is bands

It was found that bands ; 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, the second process of the superscript of the Merich commentation and it's program on a bit and the state of all measure 3.4.2 Optimum Band Selection for Supervised Classification Mausel et. al. (1992) selected an agricultural site near the town of Weslaco in Hindalgo Country, Texas to use Teature selection and Supervised classification techniques for six A DE LA COMPANYA A COMPANYA classes, viz. cotton, cantaloupe; sorghum, Johnson grass, pigweed and bare soil. Training areas (24 Nos.) of size 7.0m x 9.2m were in analy and graph and grant a model's patheoner of the selected randomly. The site was imaged on 31st May and 24th July of pricedure a practical to the second (REL) mands 16061 using USDA - ARS multispectral video system in order to collect in agent, isonol Hall vision nutrings to sease a installing spectral information in four spectral bands (0.42 to 0.43 µm, 0.52 because becaused bare shorts and structure bare to 0.55 µm, 0.64 to 0.67 µm, and 0.84 to 0.89 µm). The data were and unit of the bags digitized and registered to create an eight bands multitemporal i ii 1908, offer as fore shit band rang gilori bandre dath 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. Hean value and in spengares, a hear of law of the set variance - covariance matrices of brightness value for each class 3.4.3 Improvement in Vegetation Separability Using Transformed

Die will open to fordi an who a private that and and While analyzing different vegetation types for river Oykel, Stal 1995 buy with Spin are b that a site suff spinolsess both. Scolland, Ghosh (1991) Found that there was a confusion in the to shall le an sont man cours visitables to most the Transformed Divergence for TM data of 1984 and 1989 amongst a unit or ARE matrix to both the product of a control of an different classes of vegetation namely light forest, dense forest and heather. The Transformed Divergence produced very poor - 55 per 5.84 b 0.07 per 100 83 to 5 31 met separation between these classes, particularly for 1989 image. When synthetically generated data such as ratio, NDVI etc. is production is and to less a dorange a date the enincorporated along with the raw satellite data and same training data is used, it was found that the Transformed Divergence between ach 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 (Coud to be 1892, while a best combination of raw data with mynthetically generated data gave an overall' Transformed Divergence of 1893. This value of Transformed Divergence' indicated that excellent separation between the classes—can be facilite/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 orbital 10%¹⁵ 3.4.4 Selection of Spectral Band Combination for Land Cover/Land Use Classification Using Brightness Value Overlapping Index

(BVOI)

Saha and Kudrat (1801) applied the Beiphtnems Value Overlapping Index technique of Feature selection using digital LANDSAT - TH date of Jan. 1986 in order to classify the major land cover classes over a part of Central Gangetic alluvial plain, in Allgarh District. The major land cover classes present in the area were waste land (malt-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, 8 5) as band 7 data not included due to its bad quality & noise. The training sites of different land userland 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 BWOI values of each band. After

computation of BVOF values in each band, the following four spectrul-bands combinations were used for classification employing

c (1). Bends 1, 2 cand 3, 1 for a set of set of a sequence of the sequence of the sequence of the sequence of the second set of the second second set of the second set of the second second

The error matrices were also generated to assess the accuracy of classification for all the above four-cases. The result should that the combination of bands 2, 4 and 5 gave similar BVO values and maxima classification accuracy as 98%. Similarly, combination value and maxima classification accuracy as 98%. Similarly, combination of bands 1, 2 and 3 produced maximum BVO values and minimum

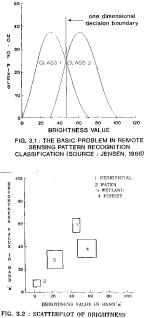
classification accuracy as 86%.

The most vital precaution to be taken for this tehenique is to avoid combination of bands having even small BVOI value but these to avoid combination of bands data (B91).

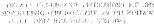
建制 医外生菌酶 网络马克伊莱 人名法

.

- - 10 -



VALUE OF TWO BAND DATA (SOURCE : JENSEN, 1986)





PRO DI PROBADITE: OF COLLECT LLANGE

 $= \left\{ \begin{array}{c} \sum\limits_{i=1}^{n} \frac{1}{2} \sum\limits_{i=1}^{n-1} \frac{1}{2} \sum\limits_{i=1}^{n$





Second and Sewageners AT 2.0 TV 924

ż

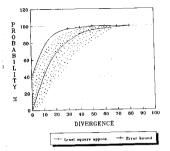
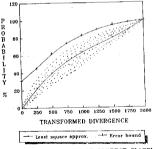
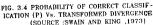


FIG. 3.3 PROBABILITY OF CORRECT CLASSIFICATION (P) Vs. DIVERGENCE (SOURCE : SWAIN AND KING , 1973)





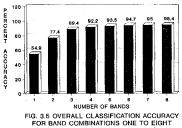
								4
· .			- 63	squ	s.t.e		1. 600	000
				100	100		l de	10.2.2
						2.1.2	10 AF 1 - 20 18/04 18/04	00
	100 M			対応				0.04
			100				開加的	Ŷ
, २३ १ ६३	5-46	14.244		5- 10-	6		1	

CHORE DECOMPORT

Charles on the transmission of the two with with the transmission of transmission of the transmission of the transmission of transmission of the transmission of transmission o



FREE AS INFORMATING THE STREET ALL CONTRACTS FREET AND ALL CONTRACTS IN CONTRACTS (SECTION CONTRACTS) CONTRACTS IN CONTRACTS



(SOURCE : MAUSEL et. al., 1990)

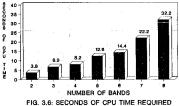


FIG. 3.6: SECONDS OF CPU TIME REQUIRED FOR CLASSIFICATION : TWO TO EIGHT BANDS (SOURCE : MAUSEL et.al., 1990)

and a start of the s The start of the start

An Arrange Constraint Cons Constraint Con

المراجع محمد المراجع ال

11.10

دینه با در است. معهد که در این ۹ در است. مهم که Horizon (۱۹۹۵) معهد میرود میرود این است. این این این این این ا میرونه این این است.

CHAPTER 4

DESCRIPTION OF THE STUDY AREA AND DATA USED

4.1 SALIENT FEATURE OF THE STUDY AREA

The study areas 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 Gauga passes through the areas approximately in West to East direction. Comprising about 236 Ka² areas extent, the study area covers a part of Indo-Cangetic plains. The whole area covered in the LAMESAT 5-TM imagery of path row no. 141-042, and Survey of India topographic sheet no. 72 $\frac{G}{2}$ at scale 1:00.000.

The area was chosen primarily as it is agriculturally predeminant with mearly all types of land use and land cover classes such as water bodies, matural vegetation, boulders, urban and sandy areas. Other factors included (1) Ground knowledge about the land use and land cover pattern, (11) Availability of topographical map, and (111) Days availability of satellity edsta.

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

miniaum temperature touches 4°C; though the maximum temperature touches 4°C; though the maximum temperature of attained is 43°C. The area generally becomes flooded 'fii'rhildy'' seasons due to appreciable increase in the flow level?57 #HWF? Cangh: "The river Ganga has been changing 'its 'dourse considerably after 'the construction of "Gandhi Setu", 'or 'a 'area' d' 10006 eT touches a second setue of the setue of the second setue of the second setue.

4.112-Eand Use and Land Cover Clauses "The time of the one per star weather

The area is having nearly kill types of init the fails this is cover? Stasses... Kowiver, 'Vegetation'and' the wurban' area is preddithent 'classes 'In'the 'area... 'Vegetative' area's are' "eijidity is being' converted to urban area 'due' to 'rast' developenent' in'the' area.' New? y 40% of the area is under vegetation and 30% of the 'r area.' New? y 40% of the area is under vegetation and 30% of the 'r area.' onisits of 'urban' areas.''About' 10% 'of the 'nee's consists of urban' areas.''About' 10% 'of the 'nee's consists' and 'r on' of 'the 'r vater bödies in the form of fiver Gangai and its tributarie's. Small 'ponds' 's'th' and remaining 'portion' is covered with 'bibs' 'and cover's' classes...''s the second 's of other is a 'f' the 's' to de and 'remain' and 's' to de and 'remain' area.''

The fix major classes were recognized by Kumar (1993), with \tilde{t} the 1:50,000 geocoded F.C.C., toposheet and the visual $\frac{1252}{100}$ ($\frac{1000}{100}$), which \tilde{t} is the same fix of the area. The same fix classes have been taken for the present analysis. These are 1:

11 This is the based broken processive suffice of the based of (1) the based of the sufficiency of the sufficience of the su

server was the start of a scene with the destructed before

(v) Urban areas, and

(vi);BoulderRanger and the standard process of the standard stand

. The data for, the analysis comprised of , S128512 . pixels equivalent to roughly 236 sq. km. on, ground. The estimates and the same start of the same sta

Experience with their handpeak of the form of the state o

It is a PC based image processing system deviced by [5R0.] 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 the difference of sawing data (year 1982) and toposheet (year 1976). Therefore, the option of using digital mage as mource of ground truth has been carried forward by generating at F.G.C. using bands 7, 4 and 3 (Plate 4.6) with the help of ISROVISION.

The interactive part of the equipsent has been used to delineate training areas on the issue itself, as shown in plate 6.6, to extract the training sals and finally classify the entriinger.

(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 FORFAUN - 77, the details of which are given in next chapter.

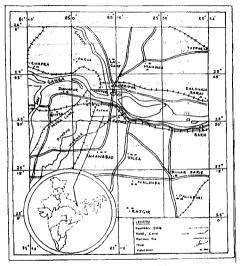


FIG. 4.1 LOCATION OF THE STUDY AREA



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

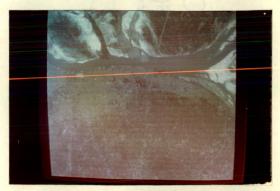


PLATE 4.2 : STUDY AREA ON LANDSAT-5 TM, BAND 3



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



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



PLATE 4.5 : F.C.C OF STUDY AREA USING TH 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 wise 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) :

- 1) Water bodies
- il) Vegetation
- 111) 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 submequently 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 shiftme number of pixels are

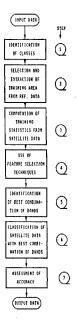


FIG. 5.1 STEPS INVOLVED IN THE ANALYSIS OF SATELLITE DATA USING FEATURE SELECTION TECHNIQUES

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

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

where

Z = Normal variate

P = Expected accuracy (%)

0 = 100 - P

E = Allowable error (%), and

N = Minimum number of pixels for all classes

In present study, P = 85%, Q = 15%, $Z \approx 2$ and E = 5% have been adopted. Using these value, sinisum number of pixels (M) comes out to be around 204 for all the classes. However, the above relationship is not the only guiding factor for welecting minisum number of pixels. Lillesand and Kiefer (1973) have stated that a sinisum 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 yeed, 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 TH data. The training areas have been selected using toposheet and generating F.C.C of TH bands 7, 4 and 3 on ISHOVISION. The training areas have been marked on digital F.C.C. image (plate 4.8), using nouse 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 within a FORTMAN-77.

5.3. DETERMINATION OF TRAINING DATA STATISTICS

The statistics for each class such as mean, standard devisition, variance, miniaux 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 teheniques, the first four viz. Divergence. Transformed Divergence. Bhattacharya Distance and Jeffreys Matualia 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 Separability Indices Using Divergence, Transformed Divergence, Enaitacharya Distance and Jeffreys Matusita Distance

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

- (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 $B\left[\frac{4}{5}c_{\mu}\right]$ ways. So, the possible combinations of two bands are (1) 1 and 3, (11) 1 and 4, (111) 1 and 7, and (1v) 4 and 7. Similarly, there are $4\left[a_{c_{\mu}}\right]$ combinations of three bands that can be selected out of four bands. These are (1) 1,3 and 4, (11) 1,3 and 7, (111) 1,4 and 7 and (1v) 3,4 and 7.

For the above two cases, the Divergence, Transformed Divergence, Bhaltacharya Distance and Jeffreys Matusita Distance have been calculated for all the 15 $\left[b_{r_{2}} \right]$ possible pairs of classes and combinations of bands using equations (3.2), (3.4), (3.5) and (3.7) respectively. In addition, the workge 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, Transformed 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 (3.8) through (3.18). A typical flow chart of this technique is shown in Fig. 5.8.

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 saverage of the above distance is minimum for each technique. In SVOI approach, those bands are taken as the best bonds whose SVU mines are alchem surgers who

5.6 CLASSIFICATION OF SATELLITE DATA USING BEST COMBINATION OF BANDS

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

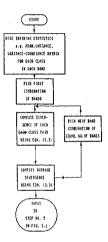


FIG: 5:2 FLOW CHART OF DIVERGENCE

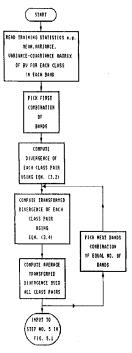


FIG:5:3 FLOW CHART OF TRANSFORMED Divergence

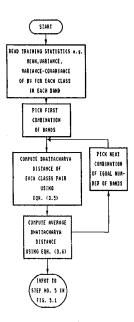


FIG: 5:4 FLOW CHART OF BHATTACHARYA DISTANCE

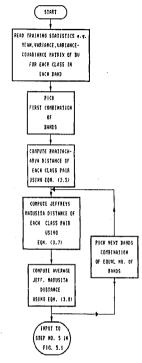


FIG:5:5 FLOW CHART OF JEFFREYS MATUSITA DISTANCE

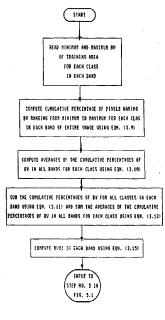


FIG: 5:6 FLOW CHART OF BUOI

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 calassified, while the non-diagonal elements give the error of calassified, while the non-diagonal elements give the arror of calassified, while the non-diagonal elements give the error of calassified, while the non-diagonal elements give the error of calassified, while the non-diagonal elements give the concrect class along each rew 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 estellite 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 mand, wet sand, urban areas and boulders have been determined from TM data, and abow in Table 6.1.

When we visually examine the minimum and maximum brightness values of different classes in different bands as shown in Table 5.2, 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

140. 01 81	ampica - 4, 4	0 cm 1 100.	or privers	In all be	apres - 545
Band:	1	з	4	7	
Mean	76.66	32.95	20.30	4.71	
Std. dev.	1.33	0.90	0.76	0.96	
Variance	1.76	0.90 0.81	0.76 0.58	0.92	
Minimum	72	30	18	3	
Maximum	81	35	22	8	
PROV THE DATE	01	55	62	0	
Band	Variance-co	variance	matrix		
1	1.76				
3	0.15	0.61			
4	0.01	0.26	0,58		
7	0.09		0.03	0.92	
	0.00	0.04	0.00	0.00	
Band	Correlation	matrix			
1	1.00				
3	0.13	1.00			
4	0.01	1.00	1.00		
7	0.07	0.05		1.00	
No. of s	cs for Vegets amples = 4, 1	Fotal No.			
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	1.83 3.36	1.43 2.06	4.42 19.58	8.04	
Minimum	69	26	44		
Haximum	79	34	66	10 26	
TRUCK I III GIII	15	54	00	2.0	
Band	Variance-co	varlance	matrix		
1	3.36				
3	1, 11	2.06			
4	-1.42		19.58		
ż		2.55		B 04	
Band	Correlation	matrix			
1	1.00				
3	0.42	1.00			
4	-0.18		1.00		
;	0.25	0.63	-0.63	1.00	
		246.	489 (100 100 100 100 100 100 100 100 100 10		

and:	1		4	7	
lean	98.37 2.63 6.94 88 105	58.84	54.40	68.16	
itd. dev.	2.63	2.58	1.51	4.52	
ariance	6.94	6.65	2.27	20.47	
inimum	88	48	50	49	
aximum	105	62	58	74	
Band	Variance-	ovarianc	e matrix		
1	6.94				
3	5.38	6.65			
4		2.91			
7	9.14	9.26	4.63	20.47	
Band	Correlation	on matrix			
1	1.00				
з	0.79	1.00			
4	0.60	0.75	1.00		
7	0.77	0.79	0.68	1.00	
No. of sa	s for Wet Sa mples = 3, T	nd. otal No.	of pixels	in all samp	ples =
No. of sa	s for Wet Sa mples = 3, T	nd. otal No.	of pixels	in all samp	ples =
No. of sa Band:	s for Wet Sa mples = 3, T 1	nd. otal No. 3	of pixels	in all samp 7	ples =
No. of sa Band: Mean	s for Wet Sa maples = 3, T 1 	nd. otal No. 3 49.17	of pixels 4 45.86	in all samp 7 54.92	ples =
No. of sa Band: Mean	s for Wet Sa maples = 3, T 1 	nd. otal No. 3 49.17	of pixels 4 45.86	in all samp 7 54.92	ples =
No. of sa Sand: Mean	s for Wet Sa maples = 3, T 1 	nd. otal No. 3 49.17	of pixels 4 45.86	in all samp 7 54.92	ples =
No. of sa Sand: Mean Std. dev. Mariance Hinimum	s for Wet Sa maples = 3, T 1 	nd. otal No. 3 49.17 2.47 5.08 43	of pixels 4 45.86 2.25 5.04 36	in all samp 7 54.92 3.87 15.01 36	ples =
No. of se Band: Hean Std.dev. Variance Hinimum Maximum	s for Wet Sa mples = 3, T 1 89.42 2.96 8.75 84 98	nd. otal No. 3 49.17 2.47 6.08 43 54	45.86 2.25 5.04 36 50	in all samp 7 54.92 3.87 15.01 36	ples =
No. of se Band: Mean Std.dev. Variance Minimum Maximum	s for Wet Sa maples = 3, T 1 89.42 2,96 8.75 84 98 .Variance-c	nd. otal No. 3 49.17 2.47 6.08 43 54	45.86 2.25 5.04 36 50	in all samp 7 54.92 3.87 15.01 36	ples =
No. of sa Band: Hean Std.dev. /ariance flnimum faximum Band 1	s for Wet Sa mples = 3, T 99.42 2.96 8.75 84 98 Variance-c 8.75	nd. otal No. 3 49.17 2.47 6.08 43 54 54	45.86 2.25 5.04 36 50	in all samp 7 54.92 3.87 15.01 36	ples =
No. of se Band: Mean Std.dev. Variance Minimum Maximum Band 1 3	s for Wet Sa mples = 3, T 1 89,42 2,96 8,75 84 98 Variance-c 8,75 5,86	nd. otal No. 3 49.17 2.47 5.08 43 54 ovariance 6.08	4 45.86 2.25 5.04 36 50 matrix	in all samp 7 54.92 3.87 15.01 36 42	ples =
No. of sa Band: Hean Std.dev. /ariance flnimum faximum Band 1	s for Wet Sa mples = 3, T 1 89,42 2,96 8,75 84 98 Variance-c 8,75 5,86	nd. otal No. 3 49.17 2.47 6.08 43 54 54	4 45.86 2.25 5.04 36 50 matrix	in all samp 7 54.92 3.87 15.01 36 42	ples =
No. of se Band: Hean Std.dev. /ariance 41nimum Haximum Band 1 3 4	s for Wet Sa mples = 3, T 1 89,42 2,96 8,75 84 98 Variance-c 8,75 5,86	nd. otal No. 3 49.17 2.47 5.08 43 54 ovariance 6.08 2.51 5.92	of pixels 45.86 2.25 5.04 36 50 matrix 5.04 4.87	in all samp 7 54.92 3.87 15.01 36 42	ples =
No. of se Sand: Hean Std. dev. /ariance dinimum Haximum Band 1 3 4 7	s for Wet Sa sples = 3, T 1 89,42 2,96 8,75 84 38 Variance-c 5,86 1,75 5,86 1,75 5,37 Correlati 1,00	nd. otal No. 3 49.17 2.47 8.08 43 54 ovariance 6.08 2.51 5.92 on matrix	of pixels 4 45.86 2.25 5.04 36 50 matrix 5.04 4.87	in all samp 7 54.92 3.87 15.01 36 42	ples =
No. of se Band: Mean Std.dev. Varlance Hinimum Maximum Band 1 3 4 7 Band 1	s for Wet Sa sples = 3, T 1 89,42 2,96 8,75 84 38 Variance-c 5,86 1,75 5,86 1,75 5,37 Correlati 1,00	nd. otal No. 3 49.17 2.47 8.08 43 54 ovariance 6.08 2.51 5.92 on matrix	of pixels 4 45.86 2.25 5.04 36 50 matrix 5.04 4.87	in all samp 7 54.92 3.87 15.01 36 42	ples =
No. of se Band: Mean Std. dev. Variance Minimum Maximum Band 1 3 4 7 Band	s for Wet Sa maples = 3, T 99.42 2.96 8.75 84 38 .Variance-c 5.86 1.75 5.37 Correlati 0.80 0.80 0.28	nd. otal No. 3 49.17 2.47 6.08 43 54 ovariance 6.08 2.51 5.92	of pixels 4 45.86 2.25 5.04 36 50 matrix 5.04 4.87	in all sam 7 54.92 3.87 15.01 36 42 15.01	ples =

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

E. Statistics for Urban Areas.

No. of sa	aples = 4, 1	otal No.	of pixels	in all	samples = 21
Band:	1	з	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	
Max1mum	82	36	37	32	
Band	Variance-c	ovariance	e matrix		
1	3,37				
з	1.33	1.79			
4	1,60	1.55	5.57		
7	1.77	1.62	2.58	6.17	
Band	Correlation	a 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.

	amples = 3, T		of pixels	in all	samples	= 102
Band:	1	3	4	7		
Mean	78.75			31.01		
Std.dev.			3.59			
Variance	5.05	1.39	12.92	23.23		
Minimum	74	34	31	23		
Maximum	83	41	42	39		
Band	Variance-co	variance	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		
~~~~~~~~						

		- 1, · · ·					
		TM BANDS					
CLASS .		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		

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

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.

5.1.2 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, Bhaitacharya Distance and Jefferys Matumita 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. Tranformed Divergence, Bhattacharya Distance and Jeffreys Matuslia Distance is obtained for combination of bands 4 and 7. For this combination the average Divergence is 353.62, while, average Transformed Divergence is 1958.52. The corresponding Shuttacharya Distance and Jeffreys Matumita 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 some jn Table 6.4.

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

Sl.no.	Classes	DIVed	DIVed	Bed	Jcd
1.	1 - 2	6.90	1065.77	3.64	1.95
2.	1 - 3	201.83	2000.00	166.41	2.00
з.	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	1532.86	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	108.66	2.00
14.	4 - 6	21.03	1855.61	91.47	2.00
15.	5 - 6	2.00	443.05	1.36	1.49
Ave.:		53.77	1468.78	108.10	1.81

(A) Bands 1 and 3.

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	DIVed	DIVcd ^T	Bed	Jcd
1.	1 - 2	57.91	1998.56	3.88	1.96
2.	1 - 3	263.98	2000.00	165.72	2.00
з.	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 - 8	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 ,

(C) Bands 1 and 7.

51.no.	Classes	DIVed	DIVcd ^T	Bed	Jcd
1.	1 - 2	15.51	1712.31	3.66	1.95
2.	1 - 3	496.04	2000.00	166.84	2.00
з.	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.	S - 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	s. 00
13.	4 - 5	75.02	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

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

# (D) Bands 3 and 4.

51. no.	Classes	DIVed	DIVed ^T	Bed	Jcd
1.	1 - 2	115.64	2000.00	2.28	1.79
2.	1 - 3	611.44	2000.00	121.04	2.0
з.	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	1.2
6.	2 - 3	19.17	1817.95	344.95	2.0
7.	2 - 4	67.31	1999.56	152.55	2.0
8.	2 - 5	63.50	1999.29	4.62	1.9
9.	2 - 6	92.85	1999.98	12.75	2.0
10.	3 - 4	13.00	1606.43	75.00	2.0
11.	3 - 5	205.61	2000.00	213.84	2.0
12.	3 - 6	158.78	2000.00	233.09	2.0
13.	4 - 5	81.17	1999.92	77.39	2.0
14.	4 - 6	44.19	1992.02	39.41	2.0
15.	5 - 6	13.56	1632.80	3.02	1.9
Ave.:		124.51	1907.90	102.11	1.8

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

(E) Bands 3 and 7.

Sl.no.	Classes	DIVcd	DIVcd	Bed	Jcd
1.	1 - 2	31.19	1959.46	2.00	1.73
2.	1 - 3	1143.00	2000.00	121.72	2.00
з.	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
в.	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 .

(F) Bands 4 and 7.

51.no.	Classes	DIVed	DIVed ^T	Bcd	Jcd
1.	1 ~ 2	325.06	2000.00	190.15	5.00
2.	1 - 3	2345.08	2000.00	134.47	2.00
з.	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
в.	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

Sl.no.	Classes	DIVed	DI Vod ^T	Bed	Jođ
1.	1 - 2	4 & 7	3 & 4 4 & 7	4 & 7	4 & 7
2.	1 - 3	4 & 7	A11	1 & 7	A1 1
а.	1 - 4	4 & 7	1&7 3&4 3&7 4&7	3 & 47	A11
4.	1 - 5	4 & 7	4 & 7	4 & 7	4 & 7
5.	1 - 6	4 & 7	387 487	4 & 7	48.7
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	387	A11
8.	2 - 5	3 & 4	384	4 & 7	4 & 7
<b>B</b> .	2 - 6	3 & 4	3 & 4	4 & 7	A11
10.	3 - 4	4 & 7	4 & 7	1 & 3	A11
11.	3 ~ 5	1 & 7	A11	1 & 7	A11
12.	3 - 6	3 & 4	1 & 7 3 & 4 3 & 7 4 & 7	1&3	A11
13.	4 - 5	4 8 7	4 & 7	4 & 7	A11
14.	4 - 6	1 8 7	1 & 7	1 & 4	A11
15.	5 - 6	1 & 7	1 & 7	1&7	1 & 7

TABLE 5.4: Best Combination of Two TM Bands for Separating Any Two Classes from Each Other Using Different Feature Selection Inconduces

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 EVOI technique requires only minimum and maximum brightness values of training data for each class, and are shown in Table 5.2. 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 5.7.

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

S1.no.	Classes	DIVed	DIVed ^T	Bed	Jed
1.	1 - 2	162.73	2000.00	6.20	2.00
2.	1 ~ 3	664.78	2000.00	166.93	2.00
з.	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

(A) Bands 1 . 3 and 4.

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

(B) Bands 1 , 3 and 7.

51.no.	Classes	DIVcd	DIVed ^T	Bcd	Jed
1.	1 - 2	35.85	1977.36	2.36	1.81
2.	1 - 3	1150.87	2000.00	167.29	2.00
з.	1 - 4	583.90	2000.00	61.85	2.00
4.	1 - 5	32.76	1966.69	1.52	1.56
б.	1 - 6	138.26	2000.00	3.62	1.95
б.	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	S. 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

(C) Bands 1 , 4 and 7.

Sl.no.	Classes	DIVed	DIVed ^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
з.	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.	5 - 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 ,

(D) Bands 3 , 4 and 7.

S1.no.	Classes	DIVed	DJVcd ^T	Bed	Jcd
1.	1 - 2	419.08	2000.00	194.18	2.00
2.	1 - 3	2441.18	2000.00	142.42	2.00
э.	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.	S - 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	156.35	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	28.99	1946.56	3.72	1.95
Ave.:		458.94	1994.93	158.20	2.00

Sl.no.	Classes	DIVed	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	A11	1,4 8 7	A11
э.	1 - 4	3,4 8 7	A11	3,4 & 7	A1 1
4.	1 - 5	3,4 & 7	3,4 & 7 1,3 & 7	3.4 & 7	3.4 8
5.	1 - 8	3,4 & 7	All Except 1, 3 & 4	3,4 & 7	3,4 & 7
6.	5 - 3	3,4 & 7	All Except 1,3 & 4	1,4 & 7	A11
7.	2 - 4	3,4 & 7	All Except 1,3 & 4	1.4 & 7 1.4 & 7	A11
8.	2 - 5	1,3 & 4	1,3, & 4	1,4 & 7	All Except 1,3 & 4
9.	2 - 6	3,4 8,7	3,4 & 7	3,4 & 7	A11
10.	3 - 4	1,4 & 7	1,4 & 7	1,4.8 7	A11
11.	3 - 5	3,4 & 7	A11	3.4 & 7	A11
12.	3 - 6	3,4 & 7	A11	1,4 & 7	A1 1
13.	4 - 5	3.4 & 7	3,4 & 7	3,4 & 7	A11
14.	4 - 6	3.4 & 7	3,4 § 7	1,4 & 7	A11
15.	5 - 6	1,3 & 7	1,3 & 7	1,3 & 7	1.3 & 7

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

All - 1-8-3, 1/8 4, 1/8 7, 3 8 7, 3 8 7, and 4 8 7. Class - 1 - Water bodies, 2 - Vegetation, 3 - Dry sand, 4 - Wet sand, 5 - Urban areas, 6 - Boulders

CLASSES	CUMULATIV			BRIGHTNESS TM BANDS	VALUE
CLASSES	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	05,95	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	

TABLE 6.7 : Brightness Value Overlapping Index (BVOI)

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.

## 8.1.4 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 TH3,4 and 7 while worst combination is TH1,3 and 4.

The results of the best and worst two bands combination are shown in Tables 5.8 and 5.9, respectively. The overall classification accuracy is 65.86% and 80.00% respectively. Similarly, the results of the best and the worst three bands combination are shown in Tables 6.10 and 5.11 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 cleasefication 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 state two bands combination.

Men all the four available bands are used for classification (Table 5.12), the overall classification accuracy is 07.07%, which is a small isproveent in classification accuracy (0.67%), but the CPU time required to classify the image using four bands is substantially larger (4 min. 58 sc.), i.e. 50% more computer time is required. Thus, it can be seen that CPU time increases substantially with little isprovement in accuracy as the number of bands increases from three to four. The result affired with the result obtained by Mausel et. al. (1960) 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 :

- All the feature selection techniques give similar results for prediction of best two and three bands combinations.
- 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.
- 111) 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 5.6.
- 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 88.86% for best bands combination of two bands (4 and 7). The classification accuracy has been reduced to 80% when

TABLE 5.8 : Details of Classified Image and Accuracy Assessment Using Best Combination of Bands 4 and 7

(A) ERBOR MATRIX:

ACTUAL	-	INT	ERPRE	TED C	LASS				
CLASS	1	2	3	4	5	6	TOTAL	ACC. (%)	OM. (%)
۱.	343	0	0	0	0	0	343	100.00	0.00
2.	Ð	190	40	28	0	с	258	73.64	26.36
з.	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	D	0	0	188	23	511	89.10	10.90
6,	0	·0	0	0	28	74	102	72.55	27.45
TOTAL COM. (%)						38.8			
COM. (%) OVERA AVERA	UL CL GE CL % Er	8.85 ASSIF	ICATI	ON AC	13.0 CURAC	Y: 86 Y: 84 H X	86 % 76 %	of Comm	ssion
COM. (%) OVERA AVERA OM ACC.	EL CL GE CL X Er	8.85 ASSIF ASSIF	ICATI ICATI	ON AC ON AC Sslor the	13.0 CURAC	Y: 86 Y: 84 H. – X	86 % 76 %	of Commi	ssion
COM. (%) OVERA AVERA OM ACC.	:0.0 LL CL GE CL X Er - X A ETAIL	8.85 ASSIF ASSIF	ICATI ICATI	ON AC ON AC Sslor the	13.0 CURAC CURAC CURAC	Y: 86 Y: 84 M X	86 % 76 %	of Commi	sslon 6
COM. (%) OVERA AVERA OM ACC. (B) D	: 0. 0 LL CL GE CL % Er - % A ETAIL	8.85 ASSIF ASSIF ror c ccure S OF	TICATI TICATI	ON AC ON AC ON AC	CURAC CURAC CURAC L, CO Class	Y: 86 Y: 84 H X	.86 % .76 % Error		
COM. (%) OVERA AVERA OM ACC. (B) D CLASS	ELL CL GE CL X Er - X A ETAIL	8.85 ASSIF ASSIF For c ccure S OF	TICATI TICATI TICATI TICATI TICATI TICATI	ON AC ON AC Salor the SIFICA	CURAC CURAC CURAC Class TION 3 12	Y: 86 Y: 84 H X ; 3	.86 % .76 % Error	5 205	6

CLASS - 1 - Water bodies , 2 - Vegetation , 3 - Dry sand , 4 - Wet sand , 5 - Urban areas , 8 - Boulders

TABLE 8.9 : Details of Classified Image and Accuracy Assessment Using Worst Useful Combination of Bands 1 and 3

1 26 0 0 0 0 0 0	2 17 171 0 0 0 0	3 45 170 18 1 0	4 0 42 24 160 10	5 0 20 26 170	6 0 0 0	TOTAL 343 258 214 204	ACC. (%) 95.04 66.28 79.43	он. (%) 4. 96 33. 72 20. 57
0 0 0 0 0 26	171 0 0	45 170 18 1	42 24 160 10	0 20 26	0 0 0	258 214	66.28 79.43	33.72
0 0 0 0 226	0 0 0	170 18 1	24 160 10	20 26	0	214	79.43	
0 0 0 226	0 0	18 1	160 10	26	0			20.57
0 0 826	0	1	10			204	-	
0				170	~~		78.43	21.57
326	0	8			30	211	80.57	19.43
			13	20	69	102	67.24	32.36
TOP	9.0 ASSII ASSII	27.4 TICAT	35.4 ION AC	CURAC CURAC COM. ~ Class	30.3 Y: 94 Y: 7 X E	7.89 %	Commissio	on ,
	1		2	3		4	5	6
	41		82	12	3	164	205	246
:	2495	56 11	04359	105	12	13710	63544	45063
:	9.1	52 :	39.91	4.	01	5.23	24.24	17.19
	CL TOP X A AIL	CLASSIF TOP of C % Accurs AILS OF 1 41 : 2499	CLASSIFICAT: TOP OF OBISS X Accuracy of ALLS OF CLASS 1 41 : 24956 11	CLASSIFICATION AC ror of Omlesion , % Accuracy of the ALLS OF CLASSIFICA 1 2 41 82 ; 24956 104359	CLASSIFICATION ACCURAC           TOP of Dalssion , COM - 3           X Accuracy of the Class           AILS OF CLASSIFICATION           1         2           41         82           21         24955           104359         105	CLASSIFICATION ACCURACY: 7           TOP of Dalssion , COM % E           % Accuracy of the Class.           AILS OF CLASSIFICATION :           1         2           41         82           41         82           1         23           41         82           1         82           1         105	X Accuracy of the Class. AILS OF CLASSIFICATION : 1 2 3 4 41 82 123 164 : 24955 104359 10512 13710	CLASSIFICATION ACCURACY: 77.89 X TOTO OF DALESSION, COM X Error of Commission X Accuracy of the Class. AILS OF CLASSIFICATION : 1 2 3 4 5 41 82 123 164 205 : 24955 104359 10512 13710 63544

(A) ERROR MATRIX:

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

(A) ERROR MATRIX:

1. 34 2. () 3. () 4. () 5. () 6. () 70TAL: 34: COH. : 0 (%) 0VERALL AVERACE 0N ACC	1 2 343 0 0 258 0 0 0 0 0 0 0 0 0 0 143 258 0.0 0.0 143 258 0.0 0.0	0 0 206 204 0.0 2.	0 0 18 197 8 205	94 116	343 258 214 204	93.36	0.00 0.00 3.74 2.94 6.64
2. ( 3. ( 4. 1 5. 1 6. 1 OTAL: 34: COH. :0 (%) OVERALL AVERACE OM ACC	0 258 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 206 6 0 19 0 0 0 0 206 204 0.0 2.	0 0 197 8 205	0 2 6 14 94	258 214 204 211 102	100.00 96.26 97.06 93.36	0.00 3.74 2.94 6.64
3. ( 4. 1 5. 1 6. 1 0TAL: 34: COH. : 0 (%) 0VERALL AVERAGE 0H ACC	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	206 6 0 19 0 0 0 0 206 204 0.0 2.	0 197 197 8 205	2 6 14 94	214 204 211 102	96.26 97.06 93.36	3.74 2.94 6.64
4. 1 5. 1 6. 1 OTAL: 34: COH. : 0 (%) OVERALL AVERAGE ON	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 19 0 0 0 0 206 204 0.0 2.	197 197 8	6 14 94 116	204 211 102	97.06 93.36	2.94 6.64
5. 1 6. 0 0TAL: 34: COH. : 0 (%) 0VERALL AVERAGE 0H ACC	0 0 0 0 0 0,0 0.0	0 0 0 0 206 204 0.0 2.	197 8 205	14 94 116	211 102	93.36	6.64
6. I OTAL: 34: COH. : 0 (%) OVERALL AVERAGE OM ACC	0 0 043 258 0.0 0.0	0 0 206 204 0.0 2.	205	94 116	102		
OTAL: 34: COH. : 0 (%) OVERALL AVERAGE OM ACC	143 258 0.0 0.0	206 204 0.0 2.	205	116		92.16	7.84
COH. :0 (%) OVERALL AVERAGE OM ACC	0.0 0.0	0.0 2.			1332		
AVERAGE OM ACC	L CLASSI	ICATION					
	Error of Accurat	cy of the	Class		- % Er	ror of Co	mmissio
CLASS :	: 1	2	3		4	5	6
GRAY :	41	82	12	3	164	205	246
PIXELS	5 : 2530	4 10770	12 83	55	13011	7709 :	30063
% AREA	N : 9.6	5 41.09	э э.	19	4,96	29.64	11.47
ELAPSED	TIME T	N CLASSIF	ICATIO	N : 0	0:03:14	. 28	
CLASS -				- Vea	etation	, 3 - Dr	y sand

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

ACTOAL		INT	ERPRE	TED (	LASS				
CLASS	1	2	3	4	5	6	TOTAL	ACC. (%	) OM. (%)
1.	342	1	0	D	0	0	343	99.71	0.29
2.	D	237	10	10	1	0	258	91.86	8.14
з.	0	0	188	26	o	0	214	87.85	12.15
4.	0	0	0	172	10	22	204	84.31	15.69
5.	0	o	0	10	173	28	211	81.99	18.01
б.	Ð	0	0	в	13	81	102	79.41	20.59
TOTAL:									
(%)									
OVERAL							9.56 % 7.52 %		
OVERAL	Eri - Ad	ASSIF	Cm1s y of	DN A	CURAC , COP Cless.	Y: 8	7.52 %	of Commi	ssion ,
OVERAJ AVERAJ OM. ACC.	GE CI Eri - Ac	ASSIF	Cm1s y of	DN A	CURAC , COP Cless.	Y: 8 1	7.52 %	of Commi	ssion , 6
OVERAI AVERAI OM ACC. (B) D	GE CI Eri - Ac ETAII	ASSIF	Cm1s Om1s of CLASS	CN A slon the SIFIC	CORAC , COP Class. ATION	£¥: 8 4. − 3	7.52 % Error		
OVERAD AVERAD OM. ACC. (B) DI CLASS	GE CI - Ac ETAII	ASSIF	Cm1s Cm1s CLASS	SIFIC	CORAC , COP Class. ATION 3 12:	.₩: 8 4 :	7.52 % Error 4	5 205	6

(A) ERBOR MATRIX:

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

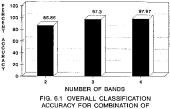
TABLE 6.12 : Details of Classified Image and Accuracy assessment Using all the Four TM Bands (Bands 1 3, 4 and 7). 

(A) ERROR MATRIX:

	-			TED C	-				
CLASS	1	2	Э	4	5	6	TOTAL	ACC. (%	) OM. (%)
1.	343	0	0	0	0	0	343	100.0	0 0.00
г.	0	258	0	0	0	0	258	100.0	0 0.00
з.	0	0	206	6	0	г	214	96.2	6 3.74
4.	0	0	0	202	0	2	204	99.0	2 0.98
5.	0	D	0	0	197	14	211	93.3	6 6.64
6.	0	0	o	0	3	99	102	97.0	6 2.94
(%)	LL CL	0.0 ASSIF	ICATI		CURA	CY: 97			
ON.	- Er	ror a	f Cm i	ssion		эм. –		of Comm	ission ,
ON. ACC. (B) D	- Er - Ad ETAIL	ror o curac S OF	f Cmi y of CLASS	ssion the C	lass Tion	)m, - :		of Conm	lssion ,
DN. ACC. (B) D	- Er - Ac ETAIL	S OF	f Cmi y of CLASS	ssion the C SIFIC/ 2	i , Ci lass TION	эм, -	Error 4	5	lssion , 6
ON. ACC. (B) D CLASS GRAY	- Er - Ac ETAIL	S OF	f Cmi y of CLASS	ssion the C SIFICA 2	TION 3 12:	эм, – : з	Error 4 164	5 205	6 246
ON. ACC. (B) D CLASS GRAY PIXEL	- Er - Ac ETAIL : : S:	ror o curac S OF 1 41 25529	f Cmi y of CLASS E 114	ssior the C SIFICA 2 2 2 2 2 3 2	1 , 0 1855 110N 110N 3 12: 8299	эм, - : : 3	Error 4 164	5	6 246
ON. ACC. (B) D CLASS GRAY PIXEL	- Er - Ac ETAIL : : S:	ror o curac S OF 1 41 25529	f Cmi y of CLASS E 114	ssior the C SIFICA 2 2 2 2 2 3 2	1 , 0 1855 110N 110N 3 12: 8299	эм, - : : 3	Error 4 164 2788	5 205	6 246 25995

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



BANDS (TWO TO FOUR)

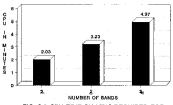


FIG. 6.2 CPU TIME (IN MIN.) REQUIRED FOR CLASSIFICATION : TWO TO FOUR BANDS



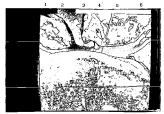
1 - Water bodies, 2- Vegetation, 3- Dry mand, 4- Wet sand S- Urban areas, 6- Boulders

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, 5- Boulders



## CHAPTER - 7

## CONCLUSIONS AND RECOMMENDATIONS

### 7.1 CONCLUSIONS

In this study, various feature selection techniques have been discussed in order to deternine the best combination of bands, to be used subsequently in classification of the digital activitie data. Digital analysis of multispectral LANSAT-5 TM data has been carried out to assess the utility of various feature selection techniques. Viz. Divergence. Transformed Divergence. Banttacharya Distance, Jeffreys Matusita Distance. From the results of this study. He following concessions are drawn :

- 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.
- 11) The Divergence and Bhattacharya Distance may not be an officient 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 maturation value is isocriant.
- 111) 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 minisum brightness values for each class in each band from training areas which results in a considerable saving of CFU time.
  - v) The errors of omission and commission can be reduced significantly, which in turn increases the classification accuracy, 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.
- vil) 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

- The results of the present study are only valid for the IM data in given study area.
- (11) The time difference between the satellite data and the reference data may incorporate error in the result.
- (111) 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 TH 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 require may be obtained.

- 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.
- (11) 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.
- (111) 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 satelite data.

### REFERENCES

- Campbell, J.B. (1997), "Introduction to Remote Sensing", The Guilford press, New York, pp. 269-273.
- Flizpatrick 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.8, pp. 169-173.
- Ghosh, S.K.(1981), "Hiver 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 Stratshjelyde, Scotland, U.K., pp. 137-143.
- Hanck, B.N. (1983), "An Analysis of Thematic Mapper Simulation Data for Urban Environments", Remote Sensing of Environment, Vol. 13, pp. 265-275.
- Jensen, J.R. (1986), "Introductory Digital Image Processing ", Prentice Hall, Englewood Cliffs, New Jersey, pp. 178-231.
- Kamat, D.S. (1978). "Selection of Multispectral Data in Remote Sensing Application", Satellite for Earth Observation (SED): Workshop on Digital Processing of Remately Sensed Data, ISROspace Application Centre, Almendhad, pp. 183-185.
- Kumar, A. (1993), "Study of Planform Changes in a Selected Reach of an Alluvial River Using Resolt Sensing". M.E. Thesis, W.R.D.T.C., University of Roorkee, Roorkee.

- Lillesand, T.M. and Klefer, R.W. (1979), "Remote Sensing and Image Interpretation", John Willey & Sons, New York, pp. 461-477.
- 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.
- Mather, P.H. (1987), "Computer Processing of Remotely Sensed Images", John Wiley & Sons, New York.
- Mausel, P.W., Kramber, W.J. and Lee, J.K. (1980), "Optimum Band Selection for Supervised Classification of Multispectral Data", Photogrammetric Engineering and Remote Sensing, Vol.56, No. 1, pp. 55-50.
- Saha, S.K. and Kudrat, M. (1991), "Selection of Spectral Band Combination for Land Use/Land Cover Classification Using a Brightness Value Overlapping Index (BV01)", Photonirvachak Vol. 19, No.3, pp. 141-147.
- Sheffield, C. (1965), * Selection of Band Combinations from Multispectral Data", Photogrammetric Engineering and Remote Sensing", Vol. 51, No.6, pp. 681-687.
- Swain, P.H. and Davis, S.M.(1978), "Remote Sensing : The Quantitative Approach", McGraw - Hill, New York, pp. 164-175.
- Swain, P.H. and King, R.C. (1973), "Two Effective Feature Selection Criteria for Multispectral Remote Sensing", LARS Technical Report 042873.

- Swain, P.H. and Wacker A.C.(1971), "Comparison of Divergence and B - Distance in Feature Selection", West Lafayette, Ind : Laboratory for the Application of Remote Sensing, Technical Report 020071.
- Thomas, I.L., Bening, V.M. and Ching, N.D.(1987), "Classification of Remotely Sensed Images", Adam Hilger, Bristol, U.K., pp. 83-101.