

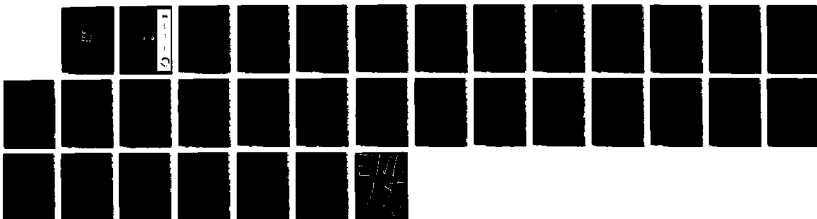
AD-A183 370

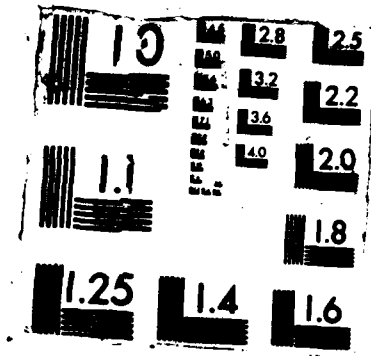
CLASSIFICATION OF SELECTED RADAR IMAGERY PATTERNS USING 1/1  
A BINARY TREE CLASSIFIER(U) ARMY ENGINEER TOPOGRAPHIC  
LABS FORT BELVOIR VA N D FOX OCT 86 ETL-8442

UNCLASSIFIED

F/O 17/9

NL





AD-A183 370

12

ETL-0442

DTIC FILE CO

Classification of selected radar imagery patterns using a binary tree classifier

Neil D. Fox

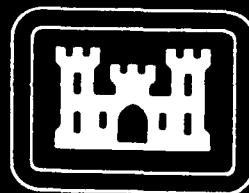
DTIC ELECTE  
AUG 06 1987  
S D  
Q4D

October 1986

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED

Prepared for  
U.S. ARMY CORP OF ENGINEERS  
ENGINEER TOPOGRAPHIC LABORATORIES  
FORT BELVOIR, VIRGINIA 22060-5546

87 8 5 011



E

T

L



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ETL-0442	2. GOVT ACCESSION NO. ADA183370	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CLASSIFICATION OF SELECTED RADAR IMAGERY PATTERNS USING A BINARY TREE CLASSIFIER	5. TYPE OF REPORT & PERIOD COVERED Research Note Dec 83 - Sep 84	
	6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) Neil D. Fox	8. CONTRACT OR GRANT NUMBER(s)	
9. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Army Engineer Topographic Laboratories Fort Belvoir, Virginia 22060-5546	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Engineer Topographic Laboratories Fort Belvoir, Virginia 22060-5546	12. REPORT DATE October 1986	
	13. NUMBER OF PAGES 31	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Binary Decision Tree Bayes Classifier Radar Imagery Feature Selection		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This report tells the results of classifying radar imagery using a binary tree classifier. It was found that this classification algorithm works well with radar imagery, which would indicate a normal (Gaussian) feature vector distribution. The number of elements in each feature vector is the limiting factor, classification time is negligible once the tree structure has been created.		

## PREFACE

This work reported on was done under DA Project 4A161102B52C, Task B, Work Unit 0015, "Automated Radar Feature Extraction."

The work was performed during the period December 1983 to September 1984 under the supervision of Dr. Frederick W. Rohde, Team Leader, Center for Physical Sciences, and Dr. Robert D. Leighty, Director, Research Institute.

COL Alan L. Laubscher, CE, was Commander and Director, and Mr. Walter E. Boge was Technical Director of the U.S. Army Engineer Topographic Laboratories during the report preparation.

Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	



## CONTENTS

	PAGE
<b>PREFACE</b>	i
<b>ILLUSTRATIONS</b>	iii
<b>INTRODUCTION</b>	1
<b>DESIGN</b>	1
<b>IMAGERY USED FOR TEST</b>	4
<b>RESULT</b>	5
<b>CONCLUSIONS</b>	5
<b>REFERENCES</b>	5
<b>APPENDIX A. Data for Creation of Decision Tree</b>	10
<b>APPENDIX B. Software Listing</b>	14

## ILLUSTRATIONS

FIGURE	TITLE	PAGE
1	Creating the Binary Decision Tree	2
2	Flowchart of Classifier	3
3	Decision Tree for Image samples from Huntsville, AL	6
4	Decision Tree for Image Samples from Elizabeth City, NC	7
5	Classified Results for Image Samples from Huntsville, AL	8
6	Classified Results for Image Samples from Elizabeth City, NC	9

# CLASSIFICATION OF SELECTED RADAR IMAGERY PATTERNS USING A BINARY TREE CLASSIFIER

## INTRODUCTION

By using hierarchical clustering, it is possible to improve the classification results of a single-stage classifier. The inefficiency of the single-stage classifier is in part due to the simultaneous use of all feature vector components. By using a decision tree, only the features best suited to separate the classes at a node are used.

A decision tree is a set of nodes that represent a set of feature vectors in the training set. The root node contains all the vectors in the set. This root node is then clustered into two new nodes (called "sons"), and hence, two new "subclasses" are formed. The subclasses are further clustered until only terminal nodes (those with a dominant class) remain.

An unknown sample enters the tree at the root node. A decision rule is used to classify the feature vector downward in the tree to either the left or right son. This process of sending the sample down the tree is repeated until it reaches a terminal node. It then is classified to be the same class as the dominant class of that node.

Feature selection is done at each node. Ideally, all combinations of feature vector components should be tested, but this becomes impractical owing to time constraints when the number of feature vector components become large. Fortunately, the number of feature vector components necessary to classify radar imagery can be small.

Since an assumption of hierarchical classes is made, a Bayes classifier was chosen to take advantage of the normal distribution one expects in such a system. Also, only the mean vector and covariance matrix need be computed for each node.<sup>1</sup>

## DESIGN

The feature vector components chosen for the tree are

1. Covariance.
2. Skewness.
3. The number of lines detected by a Hough transform.
4. The average pixel value.
5. The number of pixels over a threshold value.

The clustering was done with a K-means algorithm, (k=2).

The decision rule<sup>2</sup> is

$$\underline{X} \in \omega_i \text{ iff } d_1(\underline{X}) > d_2(\underline{X}), \quad i=1, 2 \text{ and } j=2, 1$$
$$d_i(\underline{X}) = \ln p(\omega_i) - \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (\underline{X} - \underline{\mu}_i)^T \Sigma_i^{-1} (\underline{X} - \underline{\mu}_i)$$

Where  $\omega_i$  is class  $i$ ,

$\underline{X}$  is the unknown sample,

$\Sigma_i$  is the covariance matrix of the  $i^{\text{th}}$  class, and

$\underline{\mu}_i$  is the mean vector of the  $i^{\text{th}}$  class,

$p(\omega_i)$  is the a priori probability of  $\underline{X} \in \omega_i$ .

<sup>1</sup> Jack K. Mui and King-Sun Fu, "Automated Classification of Nucleated Blood Cells Using a Binary Tree Classifier," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 5, Sep. pp. 429-442, 1980.

<sup>2</sup> J. T. Tou and R. C. Gonzales, "Pattern Recognition Principles", Addison-Wesley Publishing Company, Reading, Massachusetts, 1974.



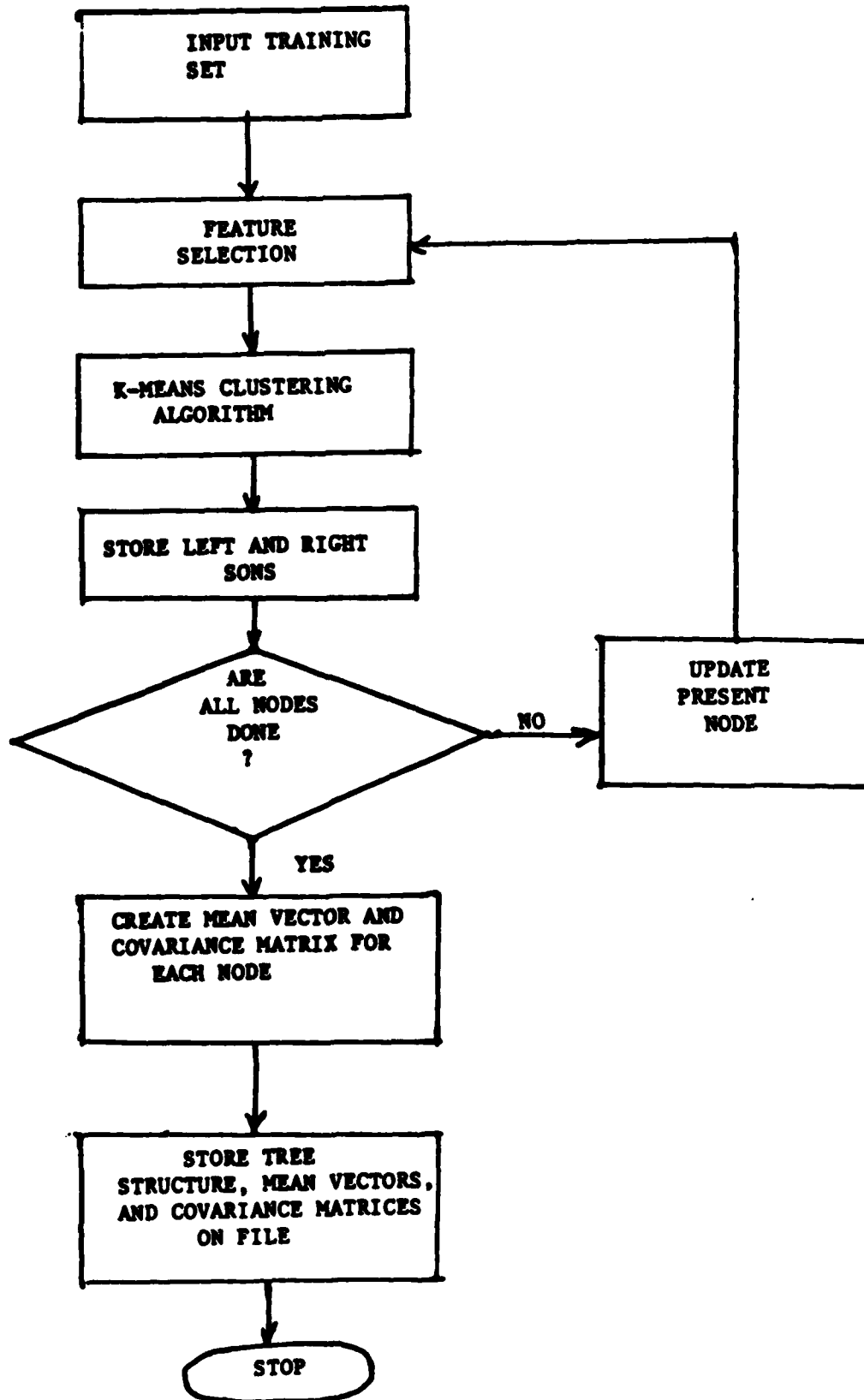


FIGURE 1. Creating the Binary Decision tree.

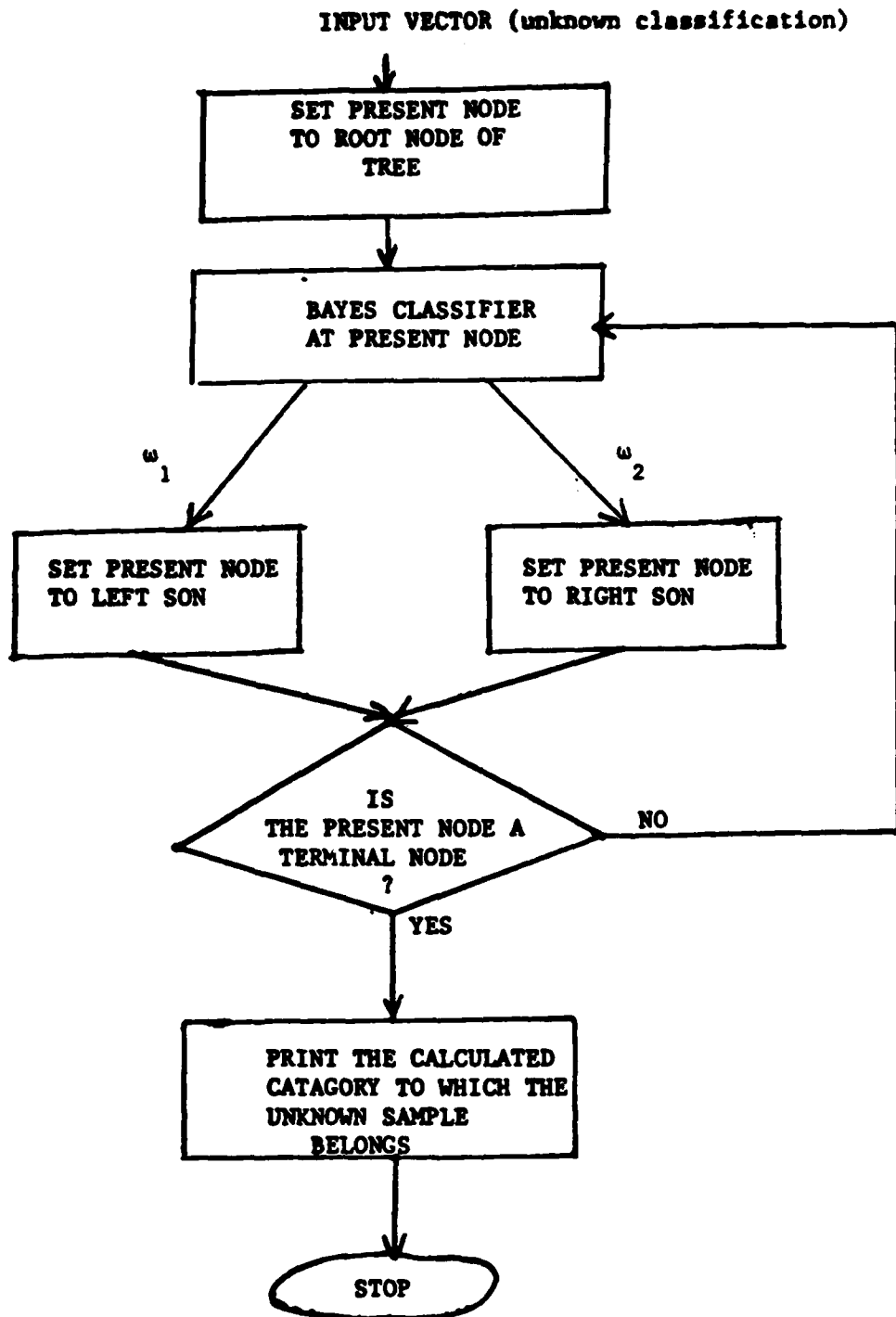


FIGURE 2. Flowchart of classifier.

The program was automated to execute all feature selection. The selection was done on the basis of how well classes were separated after clustering. The formula used is

$$\frac{I}{I+J} \sum_{n=1}^k I_n \log I_n + \frac{J}{I+J} \sum_{n=1}^k J_n \log J_n$$

where

- $I_n$  = number of vectors at left son from category n
- $J_n$  = number of vectors at right son from category n
- $I$  = number of vectors at left son
- $J$  = number of vectors at right son
- $k$  = number of categories in training set

A good separation is indicated by a large number.

The four categories of samples used for testing the tree classifier are

1. forest.
2. field.
3. water.
4. city.

The algorithms used in creating the tree and the flowchart of the tree classifier are shown in figures 1 and 2, respectively.

### IMAGERY USED FOR TEST

The proposed binary tree classifier was applied to two selected sets of high resolution synthetic aperture radar imagery taken over the Huntsville, Alabama, and the Elizabeth City, North Carolina, areas with the APD-10 and the UPD-4 radar systems, respectively. Sections of the radar imagery were digitized and stored on a disk unit. A Lexidata System 3400 display processor was used to display the images on a cathode ray tube. Software was written to take 100 samples for each of four terrain classes from imagery displayed on the screen. Each image sample consisted of 32 by 32 pixels that were located within a section of one particular terrain class. The four classes considered were (1) cities (combination of commercial and residential structures, DLMS category #504 FIC 301 and #505 FIC 401), (2) fields (agriculture used primarily for crop and pasture land, DLMS category #510 FIC 950), (3) water (rivers with smooth fresh water, DLMS category #510 FIC 940 and fresh water subject to ice, lakes and reservoirs, DLMS category #510 FIC 943), and (4) forests (mixed trees, deciduous and evergreens, DLMS category #510 FIC 954). A feature vector consisting of 15 components was computed for each sample. These 15 components were made up of the first- and second-order gray level histogram statistics computed from each sample. These components were then used as the input for the binary tree classifier.

## RESULTS

Two sets of radar imagery (Huntsville, Alabama and Elizabeth City, North Carolina) were used to create two binary decision trees. The resulting trees are shown in figures 3 and 4, and the results of classifying the training sets are shown in figures 5 and 6.

Often, only one or two feature vector components were used at a node, which illustrates the advantage of using a hierarchical approach rather than the traditional single-stage classifier. For example, although thresholding was an excellent feature vector component when separating water from field, forest, and city, it was a poor feature with respect to the separation of city and forest.

The overall classification accuracy for the Huntsville image samples was 99.25 percent and for the Elizabeth City samples, 98.50 percent.

## CONCLUSIONS

1. A binary tree classifier can be used to classify a selected set of radar images with high accuracy as illustrated. However, a new tree hierarchy will be needed if a new set of images from a different geographical area is to be classified optimally with a reasonably high accuracy.

2. Only the most pertinent feature vector components should be used at each node. This reduces processing time and improves classification accuracy.

3. An automated classifier is practical if a small set of feature vector components is used. For vectors with a large number of features, some human interaction in feature selection may be necessary to avoid lengthy processing time. The create-tree program took approximately 11 minutes to finish a tree using five features.

4. This work represents an application of a recently developed statistical classification method. A possible limitation is the requirement to create a new tree hierarchy for training, whenever a new set of images is to be classified with a reasonably high accuracy.

## REFERENCES

1. Jack K. Mui and King-Sun Fu, "Automated Classification of Nucleated Blood Cells Using a Binary Tree Classifier," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 5, Sep, pp. 429-442, 1980.
2. J. T. Tou and R. C. Gonzales, *Pattern Recognition Principles*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1974.

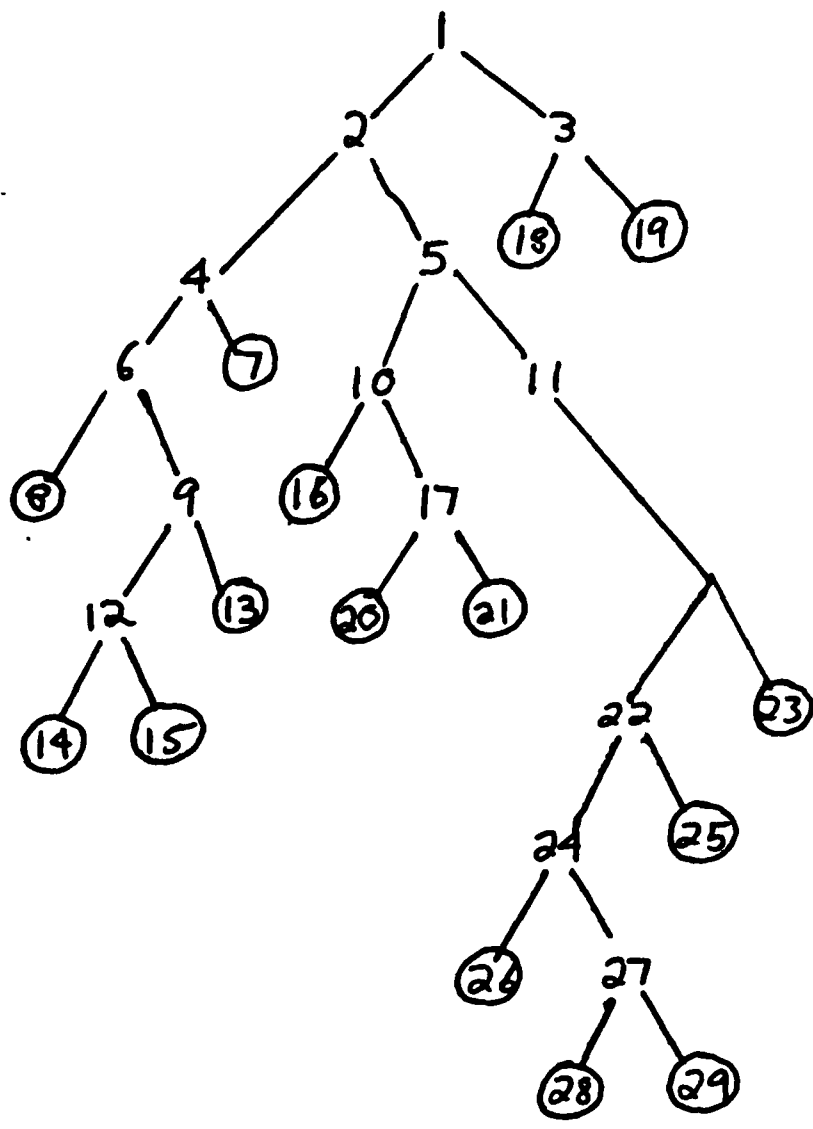


FIGURE 3. Decision tree for image samples from Huntsville, AL.

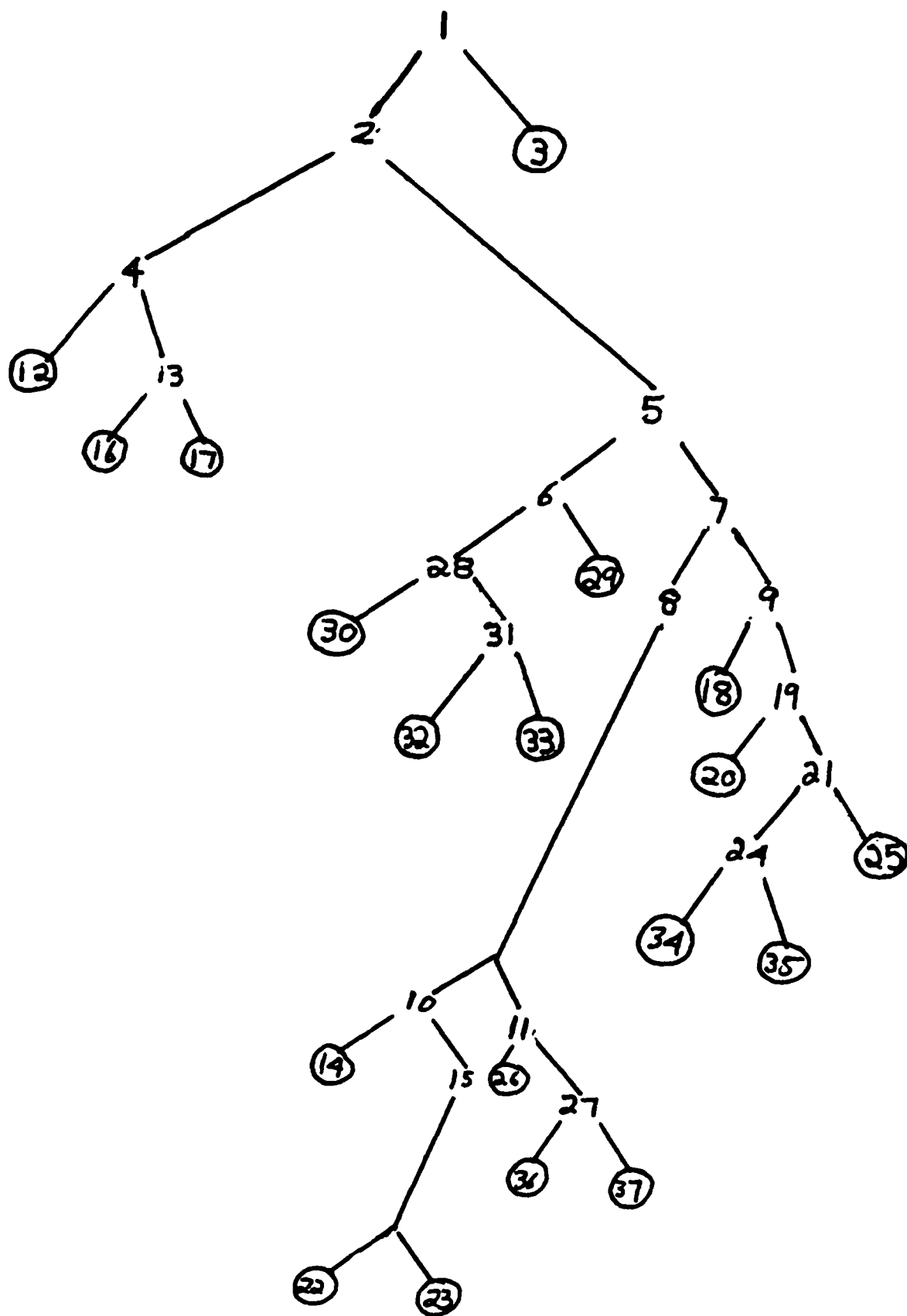


FIGURE 4. Decision tree for image samples from Elizabeth City, NC.

TRUE CAT \ REC CAT	FIE	WAT	FOR	CIT
FIE	98	0	1	0
WAT	2	100	0	0
FOR	0	0	99	0
CIT	0	0	0	100

FIGURE 5. Classified results for image samples from Huntsville, AL.

TRUE CAT \ REC CAT	FIE	WAT	FOR	CIT
FIE	100	0	1	0
WAT	0	100	0	0
FOR	0	0	96	2
CIT	0	0	3	98

FIGURE 6. Classified results for image samples from Elizabeth City, NC.



## Appendix A. Data for Creation of Decision Tree

### Huntsville, Alabama

Node #	Features Used in Separation of Node
1	THRSH, HOU, COV
2	THRSH, HOU
3	HOU
4	AVV, THRSH, HOU
5	AVV, SKW, THRSH
6	SKW
9	SKW
10	SKW
11	HOU
12	THRSH
17	HOU
22	THRSH, HOU
24	AVV
27	AVV

Key: HOU — Hough Transform  
SKW — Skewness  
COV — Covariance  
THRSH — Threshold  
AVV — Average Pixel Value.

**Appendix A. (Cont'd)**

**Elizabeth City, North Carolina**

<b>Node #</b>	<b>Features Used in Separation of Node</b>
1	SKW
2	THRSH
4	HOU
5	AVV, THRSH
6	AVV, COV
7	AVV, HOU
8	AVV
9	SKW, HOU
10	AVV, SKW
11	SKW, HOU, COV
13	THRSH
15	AVV
19	AVV
20	COV
21	SKW
24	SKW
27	SKW
28	AVV, THRSH
31	AVV, SKW

Appendix A. (Con't)

Elizabeth City, North Carolina

NODE #	WATER	FOREST	CITY	FIELD
1	100	100	100	100
2	0	100	100	100
3	100	0	0	0
4	0	12	0	100
5	0	88	100	0
6	0	63	2	0
7	0	25	98	0
8	0	19	24	0
9	0	6	74	0
10	0	9	21	0
11	0	10	3	0
12	0	5	0	0
13	0	7	0	100
14	0	1	17	0
15	0	8	4	0
16	0	1	0	100
17	0	6	0	0
18	0	0	67	0
19	0	6	7	0
20	0	2	2	0
21	0	4	5	0
22	0	8	1	0
23	0	0	3	0
24	0	2	5	0
25	0	2	0	0
26	0	8	0	0
27	0	2	3	0
28	0	45	2	0
29	0	18	0	0
30	0	17	0	0
31	0	28	2	0
32	0	27	0	0
33	0	1	2	0
34	0	0	5	0
35	0	2	0	0
36	0	2	0	0
37	0	0	3	0
38	0	0	2	0
39	0	2	0	0

Vector Distribution at Each Node of Decision Tree

Appendix A. (Con't)

Huntsville, Alabama

<b>NODE #</b>	<b>WATER</b>	<b>FOREST</b>	<b>CITY</b>	<b>FIELD</b>
1	100	100	100	100
2	5	100	100	97
3	95	0	0	3
4	5	10	88	95
5	0	90	12	2
6	0	10	88	95
7	5	0	0	0
8	0	0	0	95
9	0	10	88	0
10	0	7	12	0
11	0	83	0	2
12	0	10	25	0
13	0	0	63	0
14	0	10	0	0
15	0	0	25	0
16	0	4	0	0
17	0	3	12	0
18	0	0	0	2
19	95	0	0	1
20	0	3	0	0
21	0	0	12	0
22	0	62	0	2
23	0	21	0	0
24	0	13	0	2
25	0	49	0	0
26	0	2	0	0
27	0	11	0	2
28	0	10	0	0
29	0	1	0	2

Vector Distribution at Each Node of Decision Tree

Appendix B. Software Listing

TREC3 T=00004 IS ON CR00024 USING 00030 BLKS R=0000

```

0001 FTN4X
0002 PROGRAM TREC3
0003 C
0004 C THIS PROGRAM READS DATA OFF THE DISK WHICH WAS CREATED BY
0005 C PROGRAM "TREEF". THIS DATA IS USED TO IMPLEMENT A DECISION
0006 C TREE.
0007 C
0008 DIMENSION COV(5,5,61),MEAN(5,61),IDCB(144),ITREE(61,2)
0009 DIMENSION XIN(400,5),IDCB4(144),LABS(4),INP2(400)
0010 DIMENSION LUOT(5),LL(4,61),LABL(4),WV(5,61),XS(5)
0011 DIMENSION INAME(3),X(5),INDEX(61)
0012 DIMENSION NODE(61),MEAN1(5),COV1(5,5),COV2(25)
0013 DIMENSION MEANS(5),NAME(3)
0014 REAL MEAN,MEAN1,MEANS
0015 C
0016 C VARIABLE ARRAYS
0017 C -----
0018 C
0019 C COV - COVARIANCE MATRICES OF ALL NODES
0020 C COV1 - COVARIANCE MATRIX FOR A NODE BEFORE MASK
0021 C COV2 - COVARIANCE MATRIX FOR A NODE AFTER MASK
0022 C INDEX - NUMBER OF VECTORS AT EACH NODE
0023 C INP2 - LABELS OF VECTORS AT A NODE
0024 C ITREE - TREE STRUCTURE READ FROM TREE FILE
0025 C LABL - THE CLASS LABELS READ FROM TREE FILE
0026 C LABS - THE CLASS LABELS READ FROM PRP FILE (NOT USED)
0027 C LL - NUMBER OF VECTORS IN A CLASS AT A NODE
0028 C MEAN - MEANS OF ALL NODES (REAL)
0029 C MEAN1 - MEAN OF A NODE BEFORE MASK (REAL)
0030 C MEANS - MEAN OF A NODE AFTER MASK (REAL)
0031 C NODE - THE NODE NUMBERS OF THE TREE (NOT USED)
0032 C WV - THE MASKS FOR ALL NODES READ FROM TREE FILE
0033 C X - THE VECTOR BEING CLASSIFIED BEFORE MASK
0034 C XIN - THE VECTORS TO BE CLASSIFIED READ FROM PRP FILE
0035 C XS - THE VECTOR BEING CLASSIFIED AFTER MASK
0036 C
0037 C VARIABLES
0038 C -----
0039 C
0040 C IANS - USER INPUT: ANSWER
0041 C INDX1 - INDEX USED IN PACKING VECTOR VIA THE MASK
0042 C INDX2 - INDEX USED IN PACKING VECTOR VIA THE MASK
0043 C IRGT - THE RIGHT NODE OF NODEP
0044 C KLAC - THE NUMBER OF CLASSES (READ FROM PROPERTY FILE)
0045 C KLAS - THE NUMBER OF CLASSES (READ FROM TREE FILE)
0046 C LEFT - THE LEFT NODE OF NODEP
0047 C LVEC - THE NUMBER OF VECTORS IN THE PROPERTY FILE
0048 C MAX - THE INDEX NUMBER OF THE DOMINANT CLASS OF A NODE
0049 C MAXN - THE NUMBER OF VECTORS IN THE DUMINANT CLASS OF A NODE
0050 C NELE - NUMBER OF ELEMENTS BEFORE PACKING (USING MASK)
0051 C NODEP - THE NODE AT WHICH THE VECTOR IS BEING TESTED
0052 C NUMND - THE NUMBER OF NODES IN THE TREE STRUCTURE
0053 C
0054 C
0055 CALL RMPAR(LUOT)
0056 CALL ERLU(LUOT)
0057 C
0058 OPEN FILES AND READ TREE STRUCTURE

```

## Appendix B. (Con't)

```

0059 C
0060 WRITE(LUOT,774)
0061 774 FORMAT(" ENTER NAME OF TREE-FILE")
0062 READ(LUOT,776)INAME
0063 776 FORMAT(3A2)
0064 WRITE(LUOT,778)
0065 778 FORMAT(" ENTER PROPERTY FILE TO BE CLASSIFIED")
0066 READ(LUOT,776)NAME
0067 WRITE(LUOT,150)
0068 150 FORMAT(" DO YOU WANT A HARD COPY ?")
0069 READ(LUOT,250)IANS
0070 250 FORMAT(1A1)
0071 IF(IANS.EQ.1HY) LUOT=38
0072 CALL OPEN (IDCB,IERR,INAME)
0073 CALL READF(IDCB,IERR,KLAS)
0074 CALL READF(IDCB,IERR,NUMND)
0075 CALL READF(IDCB,IERR,LABL)
0076 CALL READF(IDCB,IERR,LL)
0077 CALL READF(IDCB,IERR,ITREE)
0078 CALL READF(IDCB,IERR,WV)
0079 CALL READF(IDCB,IERR,INDEX)
0080 DO 1 N=1,NUMND
0081 CALL READF(IDCB,IERR,NODE(N))
0082 CALL READF(IDCB,IERR,NELE)
0083 CALL READF(IDCB,IERR,MEAN1)
0084 CALL READF(IDCB,IERR,COV1)
0085 DO 2 I=1,NELE
0086 DO 2 J=1,NELE
0087 2 COV(I,J,N)=COV1(I,J)
0088 DO 3 I=1,NELE
0089 3 MEAN(I,N)=MEAN1(I)
0090 1 CONTINUE
0091 C
0092 CALL CLOSE(IDCB)
0093 C
0094 C
0095 C INPUT THE VECTORS TO BE CLASSIFIED
0096 C
0097 CALL OPEN(IDCB4,IERR,NAME)
0098 CALL READF(IDCB4,IERR,LVEC)
0099 CALL READF(IDCB4,IERR,LELE)
0100 CALL READF(IDCB4,IERR,KLAC)
0101 CALL READF(IDCB4,IERR,LABS)
0102 CALL READF(IDCB4,IERR,XIN)
0103 CALL READF(IDCB4,IERR,INP2)
0104 CALL CLOSE(IDCB4)
0105 5 DO 1223 IVEC=1,LVEC
0106 DO 99 IP=1,NELE
0107 99 X(IP)=XIN(IVEC,IP)
0108 C
0109 C PRESENT NODE OPERATED ON IS SET TO "1"
0110 C
0111 NODEP=1
0112 C
0113 C THE LEFT AND RIGHT CHILD OF EACH NODE IS FOUND
0114 C
0115 10 LEFT=ITREE(NODEP,1)
0116 IRGT=ITREE(NODEP,2)
0117 C
0118 C SEE IF NODE IS TERMINAL

```

## Appendix B. (Con't)

```

0119 C
0120 IF( (LEFT.EQ.0).OR.(IRGT.EQ.0) ) GOTO 200
0121 C
0122 C FIND DISTANCE FOR LEFT AND RIGHT NODES
0123 C
0124 DO 50 I=1,NELE
0125 DO 50 J=1,NELE
0126 50 COV1(I,J)=COV(I,J,LEFT)
0127 DO 55 I=1,NELE
0128 55 MEAN1(I)=MEAN(I,LEFT)
0129 INDX1=0
0130 INDX2=0
0131 DO 20 ITEL=1,NELE
0132 IF(WV(ITEL,NODEP).NE.0.0) THEN
0133 INDX1=INDX1+1
0134 MEANS(INDX1)=MEAN1(ITEL)
0135 XS(INDX1)=X(ITEL)
0136 ENDIF
0137 DO 20 JTEL=1,NELE
0138 IF(WV(ITEL,NODEP).NE.0.AND.WV(JTEL,NODEP).NE.0) THEN
0139 INDX2=INDX2+1
0140 COV2(INDX2)=COV1(ITEL,JTEL)
0141 ENDIF
0142 20 CONTINUE
0143 RLFT=INDEX(ITREE(NODEP,1))
0144 ALLV=INDEX(NODEP)
0145 CALL MDST(COV2,NELE,MEANS,XS,W1,INDX1,NELE**2,ALLV,RLFT)
0146 DO 60 I=1,NELE
0147 DO 60 J=1,NELE
0148 60 COV1(I,J)=COV(I,J,IRGT)
0149 DO 65 I=1,NELE
0150 65 MEAN1(I)=MEAN(I,IRGT)
0151 INDX1=0
0152 INDX2=0
0153 DO 2000 ITEL=1,NELE
0154 IF(WV(ITEL,NODEP).NE.0) THEN
0155 INDX1=INDX1+1
0156 MEANS(INDX1)=MEAN1(ITEL)
0157 XS(INDX1)=X(ITEL)
0158 ENDIF
0159 DO 2000 JTEL=1,NELE
0160 IF(WV(ITEL,NODEP).NE.0.AND.WV(JTEL,NODEP).NE.0) THEN
0161 INDX2=INDX2+1
0162 COV2(INDX2)=COV1(ITEL,JTEL)
0163 ENDIF
0164 2000 CONTINUE
0165 C
0166 RRGT=INDEX(ITREE(NODEP,2))
0167 CALL MDST(COV2,NELE,MEANS,XS,W2,INDX1,NELE**2,ALLV,RRGT)
0168 C
0169 C MAKE THE DECISION OF WHETHER TO MAKE TO LEFT OR RIGHT NUDE THE
0170 C PRESENT NODE
0171 C
0172 IF(W1.LE.W2) THEN
0173 NODEP=LEFT
0174 ELSE
0175 NODEP=IRGT
0176 ENDIF
0177 C
0178 C CLEAR COV2 AND MEANS

```

## Appendix B. (Con't)

```

0179 C
0180 DO 9000 I=1,NELE**2
0181 9000 COV2(I)=0.0
( :2 DO 9001 I=1,NELE
0183 9001 MEANS(I)=0.0
0184 C
0185 C GO TO NEXT NODE
0186 C
0187 C GOTO 10
0188 C
0189 C WRITE THE FINAL DESTINATION OF THE VECTOR
0190 C
0191 200 MAX=1
0192 MAXN=LL(1,NODEP)
0193 DO 220 I=1,KLAS
0194 IF (MAXN.LT.LL(I,NODEP)) THEN
0195 MAX=I
0196 MAXN=LL(I,NODEP)
0197 ENDIF
0198 220 CONTINUE
0199 WRITE(LUOT,101) LABL(MAX),NODEP
0200 101 FORMAT(" CLASS: ",1A2,3X,I3)
0201 C
0202 C REPEAT BY ASKING FOR ANOTHER VECTOR
0203 C
0204 1223 CONTINUE
0205 C
0206 C STOP
0207 C END
0 3 C
0209 C SUBROUTINE MDST(COV,NELE,MEAN,X,W,INDX1,NELEQ,ALLV,RORL)
0210 C
0211 C THIS IMPLEMENTS THE MINIMUM DISTANCE CLASSIFIER
0212 C
0213 C DIMENSION COV(NELEQ),MEAN(NELE),X(NELE)
0214 C DIMENSION R(5),TR(5),R1(5)
0215 C DIMENSION L(5),M(5)
0216 C REAL MEAN
0217 C
0218 C
0219 C CALL INV(COV,INDX1,DET,L,M)
0220 C CALL SUMAT(X,MEAN,R,INDX1,1)
0221 C CALL TRMAT(R,TR,INDX1,1,0)
0222 C CALL PRMAT(TR,COV,R1,1,INDX1,INDX1)
0223 C CALL PRMAT(R1,R,W,1,INDX1,1)
0224 C
0225 C CALCULATE PROBABILITY
0226 C
0227 C PROB=RORL/ALLV
0228 C IF(DET.NE.0.0) THEN
0229 C W=PROB-(ALOG(ABS(DET)))/2.0-W/2.0
0230 C W=-W
0231 C ELSE
0232 C W=PROB+20E20/2.0-W/2.0
0 3 C W=-W
0234 C ENDIF
0235 C RETURN
0236 C END
0237 C ENDS

```



Appendix B. (Con't)

TREEF T=00004 IS ON CR00024 USING 00072 BLKS R=0000

001 FTN4X

PROGRAM TREEF

003 C  
 004 C THIS PROGRAM CREATES A BINARY TREE STRUCTURE AND  
 005 C STORES IT ON DISK. AT EACH NODE A K-MEANS ALGORITHM  
 006 C IS CALLED TO SEPARATE THE VECTORS INTO TWO NEW NODES.  
 007 C  
 008 C THIS PROGRAM READS IN A PROPERTY FILE CREATED BY PROGRAM  
 009 C 'MKPRP'. THE FORMAT OF THIS FILE IS:

010 C	011 C	VARIABLE	012 C	SIZE	013 C	DESCRIPTION
014 C	015 C	-----	016 C	----	017 C	-----
014 C	014 C	NVEC	014 C	1	014 C	✦ OF VECTORS
016 C	016 C	NELE	016 C	1	016 C	✦ OF ELEMENTS
018 C	018 C	KLAS	018 C	1	018 C	✦ OF CLASSES
020 C	020 C	LABL	020 C	KLAS	020 C	NAMES OF CLASSES
022 C	022 C	XX	022 C	NVEC X NELE X 2	022 C	VECTORS
024 C	024 C	IN1	024 C	NVEC	024 C	VECTOR LABELS / THEY IN THE SAME ORDER AS THE VECTORS IN XX.

026 C

027 C

028 C

029 C

030 C

031 C

032 C

033 C

034 C

035 C

036 C

037 C

038 C

039 C

040 C

041 C

042 C

043 C

044 C

045 C

046 C

047 C

048 C

049 C

050 C

051 C

052 C

053 C

054 C

055 C

056 C

057 C

058 C

VARIABLE ARRAYS

034 C C1 -- WORK ARRAY FOR MVCOV  
 035 C C2 -- WORK ARRAY FOR MVCOV  
 036 C COV - COVARIANCE MATRIX IN MVCOV  
 037 C IN1 - LABEL BUFFER FOR LEFT NODE  
 038 C IN2 - LABEL BUFFER FOR RIGHT NODE  
 039 C INB - LABEL BUFFER FOR INPUT TO CLUS  
 040 C INDEX - THE NUMBER OF VECTORS AT EACH NODE  
 041 C IREC - STORAGE FOR RECORD THAT CORRESPONDS TO A NODE NUMBER  
 042 C ITREE - THE TREE STRUCTURE  
 043 C LABL - THE NAMES (LABELS) OF THE CLASSES  
 044 C LL -- THE NUMBER OF VECTORS IN EACH CLASS AT EACH NODE  
 045 C MEAN - THE MEAN CALCULATED IN MVCOV. THIS IS A REAL ARRAY.  
 046 C MEAN1 - THE TWO MEANS CALCULATED FROM CLUS. THIS IS A REAL ARRAY.  
 047 C OLDM -- THE OLD MEAN VALUES IN CLUS  
 048 C W -- THE MASKS FOR ALL NODES  
 049 C WOLD -- THE PREVIOUS MASK VALUE WHEN IN "SQ" ROUTINE  
 050 C WORTH - STORES HOW WELL SEPARATED EACH NODE IS  
 051 C WPRIM - THE BEST MASK IS STORED HERE AFTER "SQ" ROUTINE  
 052 C X1 -- THE VECTOR BUFFER FOR THE LEFT NODE  
 053 C X2 -- THE VECTOR BUFFER FOR THE RIGHT NODE  
 054 C XX -- THE VECTOR BUFFER FOR THE NODE BEING CLUSTERED

VARIABLES

-----

Appendix B. (Con't)

```

0059 C      ICMND - INPUT FROM USER IN COMMAND LOOP
0060 C      IND1 - NUMBER OF VECTORS IN LEFT NODE
0061 C      IND2 - NUMBER OF VECTORS IN RIGHT NODE
0062 C      KLAS - TOTAL NUMBER OF CLASSES
0063 C      KNODE - MAXIMUM NUMBER OF NODES ALLOWED
0064 C      KREC -- PRESENT NUMBER OF RECORDS
0065 C      N1 -- LEFT NODE NUMBER
0066 C      N2 -- RIGHT NODE NUMBER
0067 C      NELE - MAXIMUM NUMBER OF ELEMENTS
0068 C      NODE - NUMBER OF NODE TO BE CLUSTERED OR OPERATED ON
0069 C      NVEC - NUMBER OF VECTORS IN PROPERTY FILE
0070 C      OLDVL - THE OLD VALUE OF THE VARIABLE "VALUE"
0071 C      PCLAS - NUMBER OF VECTORS IN DOMINANT CLASS AT A NODE
0072 C      SMSH1 - INTERMEDIATE VARIABLE FOR CALCULATING "VALUE"
0073 C      SMSH2 - INTERMEDIATE VARIABLE FOR CALCULATING "VALUE"
0074 C      VALUE - MEASURE OF THE CLUSTER. LARGE VALUES ARE GOOD.
0075 C      VIN1 - INTERMEDIATE VARIABLE FOR CALCULATING "VALUE"
0076 C      VIN2 - INTERMEDIATE VARIABLE FOR CALCULATING "VALUE"
0077 C      WMIN - THE MINIMUM VALUE OF WORTH(NODE). THIS IS THE BEST VALUE.
0078 C
0079          DIMENSION LUOT(5),XX(400, 5),OLDM( 2, 5),C1( 5, 5)
0080          DIMENSION IN1(400),C2(5,5),COV(5,5),MEAN1(2,5),IN2(400)
0081          DIMENSION LABL(4),INAME(3),MEAN(5),LL(4,61),INB(400),W(5,61)
0082          DIMENSION X1(400,5),X2(400,5),ITREE(61,2),INDEX(61),WPRIM(5)
0083          DIMENSION INAM1(3),WOLD(5),WORTH(61)
0084          REAL MEAN,MEAN1
0085          DIMENSION IDCB(144),IDCB1(144),IDCB2(144),IREC(100)
0086 C
0087 C      SET OLD VALUE TO 0
0088 C
0089          OLDVL=0.0
0090 C
0091 C      SET LEFT, RIGHT, AND STARTING NODE TO INITIAL VALUES
0092 C
0093          NODE=1
0094          N1=0
0095          N2=1
0096 C
0097 C      SET MAX. NODES
0098 C
0099          KNODE=61
0100 C
0101 C
0102 C      INIT. LU
0103 C
0104          CALL RMPAR(LUOT)
0105          LU=LUOT(1)
0106          CALL ERLU(LU)
0107 C
0108 C
0109 C      FORMAT(1A2)
0110 C
0111 C      READ FILES
0112 C
0113          WRITE(LUOT,333)
0114 333          FORMAT(" ENTER NAME OF PROPERTY FILE")
0115          READ(LUOT,334)INAME
0116          WRITE(LUOT,335)
0117 335          FORMAT(" ENTER THE FILE-NAME FOR THE TREE STRUCTURE")
0118          READ(LUOT,334)INAM1

```

Appendix B. (Con't)

```

J119 334  FORMAT(3A2)
J120      CALL OPEN(IDC.B,IERR,INAME)
J121      CALL READF(IDC.B,IERR,NVEC)
J122      CALL READF(IDC.B,IERR,NELE)
J123      CALL READF(IDC.B,IERR,KLAS)
J124      CALL READF(IDC.B,IERR,LABL)
J125      CALL READF(IDC.B,IERR,XX)
J126      CALL READF(IDC.B,IERR,IN1)
J127      CALL CLOSE(IDC.B)
J128      C
J129      C   PURGE OLD AND CREATE NEW SCRATCH FILE
J130      C
J131      CALL PURGE(IDC.B1,IERR,5HXBUFR)
J132      CALL PURGE(IDC.B2,IERR,5HIBUFC)
J133      CALL CREAT(IDC.B1,IERR,5HXBUFR,500,3,0,-24)
J134      CALL CREAT(IDC.B2,IERR,5HIBUFC,500,3,0,-24)
J135      IF(IERR.LT.0) WRITE(LUOT,50)IERR
J136      C
J137      C   SET FIRST INDEX
J138      C
J139      INDEX(NODE)=NVEC
J140      C
J141      C   FILL LABEL ARRAY FOR NODE 1
J142      C
J143      DO 1192 I=1,KLAS
J144 1192  LL(I,1)=100
J145      C
J146      C   STORE VECTOR NAMES ON SCRATCH FILE
J147      C
J148      CALL WRITF(IDC.B2,IERR,IN1,NVEC)
J149      C
J150      C   INITIALIZE RECORD NUMBER
J151      C
J152      KREC=1
J153      C
J154      C   STORE DATA ON SCRATCH FILE
J155      C
J156      CALL WRITF(IDC.B1,IERR,XX,2*NVEC*NELE)
J157      IF(IERR.LT.0) WRITE(LUOT,50)IERR
J158      IREC(1)=KREC
J159      C
J160      C   ENTER COMMAND LOOP. THE POSSIBLE COMMANDS ARE :
J161      C
J162      C   LU -- CHANGE OUTPUT DEVICE
J163      C   TR -- TERMINATE PROGRAM
J164      C   SQ -- SEQUENCE W-MASK THROUGH ALL COMBINATIONS AND CREATE TREE
J165      C   CT -- CREATE TREE USING A W-MASK OF ALL 1'S (ALL FEATURES USED)
J166      C
J167 96    WRITE(LUOT,250)
J168 250  FORMAT(" ENTER COMMAND",/, "??")
J169      READ(LUOT,3)ICMND
J170      IF(ICMND.EQ.2HSQ) GOTO 252
J171      IF(ICMND.EQ.2HCT) THEN
J172          DO 3000 I=1,NELE
J173          DO 3000 J=1,KNODE
J174 3000  W(I,J)=1.0
J175          GOTO 252
J176      ENDIF
J177      IF(ICMND.EQ.2HTR) STOP
J178      C

```

Appendix B. (Con't)

```

0179      IF(ICMND.EQ.2HLU) THEN
0180 256    READ(LUOT,*)LU
) 1      IF(LU.LE.0.OR.LU.GE.39) GOTO 256
0182      ENDIF
0183      C
0184      C    IF NO RECOGNIZABLE COMMAND IS ENTERED, THEN ASK FOR ANOTHER
0185      C
0186      GOTO 96
0187      C
0188      C    SPLIT THE NODE SELECTED INTO TWO NEW NODES
0189      C
0190 252    NODE=1
0191 4444   IF(ICMND.EQ.2HSQ) THEN
0192      DO 7721 I=1,NELE
0193 7721   W(I,NODE)=0.0
0194      CALL SEQUN(WOLD,W,NODE,KNODE,NELE)
0195      ENDIF
0196 5994   FORMAT(/,10X,"W MASK ==) ",5F2.0)
0197      N1=N1+2
0198      N2=N2+2
0199      C
0200      C    STORE LEFT AND RIGHT CHILD IN ARRAY, "ITREE"
0201      C
0202      ITREE(NODE,1)=N1
0203      ITREE(NODE,2)=N2
0204      C
0205      C    FIND CORRECT RECORDS
0206      C
) 7      CALL POSNT(IDCBI,IERR,IREC(NODE),1)
0208      CALL POSNT(IDCBI,IERR,IREC(NODE),1)
0209      C
0210      IF(IERR.LT.0) WRITE(LUOT,50)IERR
0211 50     FORMAT(" FILE-ERROR #",I4)
0212      C
0213      CALL READF(IDCBI,IERR,XX)
0214      IF(IERR.LT.0) WRITE(LUOT,50)IERR
0215      C
0216      C    READ LABELS FOR VECTORS AT THIS NODE
0217      C
0218      CALL READF(IDCBI,IERR,INB)
0219      C
0220      C    CLUSTER THE VECTORS
0221      C
0222 770    CALL CLUS(XX,NELE,INDEX(NODE),X1,IND1,X2,IND2,NODE,N1,N2,MEAN1
0223      $,OLDM,KNODE,NVEC,INB,IN1,IN2,W)
0224      C
0225      IF(ICMND.EQ.2HNX) WRITE(LU,5994)(W(K,NODE),K=1,NELE)
0226      IF(ICMND.EQ.2HCT) WRITE(LU,5994)(W(K,NODE),K=1,NELE)
0227      C
0228      C    STORE LEFT AND RIGHT NODE NAMES IN LIST
0229      C
0230      INDEX(N1)=IND1
0231      INDEX(N2)=IND2
) 7      C
0233      C    PRINT THE RESULTS OF THE CLUSTERING AND ASK IF IT IS TO BE STORED
0234      C
0235      DO 27 I=N1,N2
0236      IF(ICMND.EQ.2HNX) WRITE(LU,29)I
0237      IF(ICMND.EQ.2HCT) WRITE(LU,29)I
0238 29     FORMAT(16X,"NODE #",I3)

```

Appendix B. (Con't)

```

0239      IF(ICMND.EQ.2HNX) WRITE(LU,26) (LABL(J),J=1,KLAS)
0240      IF(ICMND.EQ.2HCT) WRITE(LU,26) (LABL(J),J=1,KLAS)
0241      DO 30 NUMK=1,KLAS
0242      LL(NUMK,I)=0
0243      DO 30 J=1,INDEX(I)
0244      IF(I.EQ.N1) THEN
0245      IF(IN1(J).EQ.LABL(NUMK))LL(NUMK,I)=LL(NUMK,I)+1
0246      ELSE
0247      IF(IN2(J).EQ.LABL(NUMK))LL(NUMK,I)=LL(NUMK,I)+1
0248      ENDIF
0249      30 CONTINUE
0250      C
0251      IF(ICMND.EQ.2HNX) WRITE(LU,32)(LL(INUMB,I),INUMB=1,KLAS)
0252      27 IF(ICMND.EQ.2HCT) WRITE(LU,32)(LL(INUMB,I),INUMB=1,KLAS)
0253      32 FORMAT(12X,4(1X,I3),/)
0254      C
0255      IF(ICMND.EQ.2HSQ) THEN
0256      C
0257      C      HERE IS THE IMPLEMENTATION OF A FORMULA THAT WILL
0258      C      TELL HOW GOOD A SEPARATION IS
0259      C
0260      SMSH1=0.0
0261      SMSH2=0.0
0262      DO 7225 IJ=1,KLAS
0263      IF(LL(IJ,N1).NE.0) THEN
0264      SMSH1=FLOAT(LL(IJ,N1))*ALOGT(FLOAT(LL(IJ,N1)))+SMSH1
0265      ENDIF
0266      IF(LL(IJ,N2).NE.0) THEN
0267      SMSH2=FLOAT(LL(IJ,N2))*ALOGT(FLOAT(LL(IJ,N2)))+SMSH2
0268      ENDIF
0269      7225 CONTINUE
0270      C
0271      VIN1=FLOAT(IND1)/FLOAT(IND1+IND2)
0272      VIN2=FLOAT(IND2)/FLOAT(IND1+IND2)
0273      VALUE=VIN1*SMSH1+VIN2*SMSH2
0274      C
0275      C      SET VALUE TO 0 IF IND1 OR IND2 IS 0 OR 1
0276      C
0277      IF(IND1.EQ.0.OR.IND2.EQ.0.OR.IND1.EQ.1.OR.IND2.EQ.1) VALUE=0.0
0278      C
0279      C      IF THIS IS THE LARGEST VALUE YET, THEN SAVE THE CORRESPONDING
0280      C      W MASK
0281      C
0282      IF(VALUE.GT.OLDVL) THEN
0283      DO 6461 ISW=1,NELE
0284      6461 WPRIM(ISW)=W(ISW,NODE)
0285      OLDVL=VALUE
0286      ENDIF
0287      C
0288      CALL SEQUN(WOLD,W,NODE,KNODE,NELE)
0289      IQUIT=1
0290      DO 1055 I=1,NELE
0291      1055 IF(W(I,NODE).EQ.1) IQUIT=0
0292      C
0293      IF(IQUIT.EQ.1) THEN
0294      ICMND=2HNX
0295      VALUE=0.0
0296      OLDVL=0.0
0297      DO 9911 I=1,NELE
0298      9911 W(I,NODE)=WPRIM(I)

```

Appendix B. (Con't)

```

0299          ENDIF
0300          GOTO 770
0301          ENDIF
0302 C
0303 C
0304 C      STORE INFORMATION ON FILE
0305 C
0306 C      MOVE TO EOF
0307 C
0308          CALL POSNT(IDC B1, IERR, KREC+1, 1)
0309          CALL POSNT(IDC B2, IERR, KREC+1, 1)
0310 C
0311 C      INCREMENT RECORD NUMBER
0312 C
0313          KREC=KREC+1
0314 C
0315 C      STORE RECORD *
0316 C
0317          IREC(N1)=KREC
0318 C
0319          IF(IERR.LT.0) WRITE(LUOT,50)IERR
0320 C
0321 C      WRITE CLASS ONE VECTORS TO FILE (THESE ARE FROM THE LEFT NODE)
0322 C
0323          CALL WRITF(IDC B1, IERR, X1, 2*NVEC*NELE)
0324          CALL WRITF(IDC B2, IERR, IN1, NVEC)
0325 C
0326          IF(IERR.LT.0)WRITE(LUOT,50)IERR
0327 C
0328 C      INCREMENT RECORD *
0329 C
0330          KREC=KREC+1
0331 C
0332 C      STORE RECORD *
0333 C
0334          IREC(N2)=KREC
0335 C
0336 C      WRITE CLASS TWO VECTORS TO FILE (RIGHT NODE VECTORS)
0337 C
0338          CALL WRITF(IDC B1, IERR, X2, 2*NVEC*NELE)
0339          CALL WRITF(IDC B2, IERR, IN2, NVEC)
0340 C
0341          IF(IERR.LT.0) WRITE(LUOT,50)IERR
0342 C
0343 C      CALCULATE THE STATUS OF EACH NODE
0344 C
0345          DO 882 I=1,N2
0346          WORTH(I)=0.0
0347          PINDX=INDEX(I)
0348          IF(INDEX(I).NE.0) THEN
0349              PCLAS=0.0
0350              DO 881 J=1,KLAS
0351 881          IF(PCLAS.LT.LL(J,I)) PCLAS=LL(J,I)
0352          IF(PCLAS.NE.0) WORTH(I)=PCLAS*ALOGT(PCLAS/PINDX)
0353          IF(INDEX(I)-PCLAS.EQ.1) WORTH(I)=0.0
0354          ENDIF
0355 882          CONTINUE
0356 24          FORMAT(3X,20I3,/)
0357 26          FORMAT(13X,4(2X,1A2) )
0358 C

```

Appendix B. (Con't)

```

0359 C
0360 C THE INITIAL MINIMUM VALUE IS SET TO THE PROVIOUSLY CREATED LEFT
0 1 C NODE. THIS IS DONE TO INSURE THAT A NODE WITH ITREE ROOTS OF 0
0362 C IS SELECTED.(EG. NO NODES ALREADY DONE CAN BE DONE AGAIN.)
0363 C
0364 C WMIN=WORTH(N1)
0365 C
0366 C DO 9933 I=2,N2
0367 C IF(ITREE(I,1).EQ.0.AND.ITREE(I,2).EQ.0.AND.WORTH(I).LE.WMIN)THEN
0368 C WMIN=WORTH(I)
0369 C NODE=I
0370 C ENDIF
0371 9933 C CONTINUE
0372 C IF(ICMND.EQ.2HNX) ICMND=2HSG
0373 4923 C FORMAT(15X,"THE NEXT NODE IS",I3,/)
0374 C IF(N2.LT.KNODE.AND.WMIN.NE.0.0) WRITE(LU,4923) NODE
0375 C IF(N2.LT.KNODE.AND.WMIN.NE.0.0) GOTO 4444
0376 C
0377 C PURGE OLD FILES AND STORE NEW DATA IN THEM
0378 C
0379 100 C CALL PURGE(IDC.B,IERR,INAM1)
0380 C CALL CREAT(IDC.B,IERR,INAM1,500,3,0,-24)
0381 C CALL WRITF(IDC.B,IERR,KLAS,1)
0382 C CALL WRITF(IDC.B,IERR,N2,1)
0383 C CALL WRITF(IDC.B,IERR,LABL,KLAS)
0384 C CALL WRITF(IDC.B,IERR,LL,KLAS*KNODE)
0385 C CALL WRITF(IDC.B,IERR,ITREE,2*KNODE)
0386 C CALL WRITF(IDC.B,IERR,W,2*KNODE*NELE)
0 7 C CALL WRITF(IDC.B,IERR,INDEX,KNODE)
0388 C DO 91 NODE=1,N2
0389 C INND=INDEX(NODE)
0390 C
0391 C POSITION AND READ FILES
0392 C
0393 C CALL POSNT(IDC.B1,IERR,IREC(NODE),1)
0394 C CALL READF(IDC.B1,IERR,XX)
0395 C
0396 C WRITE NODE # AND # OF FEATURES AT THAT NODE
0397 C
0398 C CALL WRITF(IDC.B,IERR,NODE,1)
0399 C CALL WRITF(IDC.B,IERR,NELE,1)
0400 C
0401 C
0402 C CREATE MEAN VECTOR AND COVARIANCE MATRIX
0403 C CALL MUCOV(XX,MEAN,COV,C1,C2,NELE,INND,NVEC)
0404 C
0405 C CALL WRITF(IDC.B,IERR,MEAN,NELE*2)
0406 91 C CALL WRITF(IDC.B,IERR,COV,(NELE**2)*2)
0407 C CALL CLOSE(IDC.B1)
0408 C CALL CLOSE(IDC.B2)
0409 C CALL CLOSE(IDC.B)
0410 C WRITE(LU,6119)(LABL(I),I=1,KLAS)
0411 6119 C FORMAT(57X,4(2X,1A2))
0412 C DO 6117 I=1,N2
0413 6117 C WRITE(LU,6116) I,(ITREE(I,J),J=1,2),(LL(INUM,I),INUM=1,KLAS)
0414 6116 C FORMAT(10X,"NODE#",I3,5X,"LEFT-->",I3,5X,"RIGHT-->",I3,7X,
0415 C #4(1X,I3))
0416 C STOP
0417 C END
0418 C

```

Appendix B. (Con't)

```

0419 C          SUBROUTINE MVCOV
0420 C
0421 C
0422 C
0423 C          A ----- MEAN VECTOR
0424 C          NFC --- DIMENSION OF VECTORS
0425 C          C1,C2 - BUFFERS
0426 C          C ----- COVARIANCE MATRIX
0427 C          NDATA-- NUMBER OF DATA POINTS
0428 C          X ----- INPUT VECTORS
0429 C
0430 C          SUBROUTINE TO COMPUTE MEAN VECTOR AND COVARIANCE MATRIX
0431 C          BASED ON THE INCOMING TEXTURE MEASUREMENT VECTOR .X
0432 C
0433 C          SUBROUTINE MVCOV(X,A,C,C1,C2,NFC,NDATA,NVEC)
0434 C          DIMENSION C(NFC,NFC),C1(NFC,NFC),C2(NFC,NFC),A(NFC)
0435 C          DIMENSION X(NVEC,NFC)
0436 C          DO 100 I=1,NFC
0437 C          DO 100 J=1,NFC
0438 C          C1(I,J)=0.0
0439 C          C2(I,J)=0.0
0440 C          DO 10 I=1,NFC
0441 C          A(I)=0.0
0442 C          Z=NDATA
0443 C          DO 20 J=1,NDATA
0444 C          A(I)=A(I)+X(J,I)
0445 C          A(I)=A(I)/Z
0446 C          DO 50 I=1,NFC
0447 C          DO 40 J=1,NFC
0448 C          C(I,J)=0.0
0449 C          DO 30 K=1,NDATA
0450 C          C1(I,J)=C1(I,J)+X(K,I)*X(K,J)
0451 C          C1(I,J)=C1(I,J)/Z
0452 C          40 CONTINUE
0453 C          50 CONTINUE
0454 C          DO 70 I=1,NFC
0455 C          DO 60 J=1,NFC
0456 C          C2(I,J)=A(I)*A(J)
0457 C          60 CONTINUE
0458 C          70 CONTINUE
0459 C          DO 90 I=1,NFC
0460 C          DO 80 J=1,NFC
0461 C          C(I,J)=C1(I,J)-C2(I,J)
0462 C          80 CONTINUE
0463 C          90 CONTINUE
0464 C          RETURN
0465 C          END
0466 C
0467 C          THIS SUBROUTINE TAKES A GROUP OF VECTORS AND SEPARATES THEM
0468 C          INTO TWO CLASSES. THIS IS DONE WITH A K-MEANS ALGORITHM.
0469 C
0470 C          SUBROUTINE CLUS(X,NELE,NVEC,X1,IND1,X2,IND2,NODE,N1,N2,MEAN
0471 C          *,OLDM,KNODE,NVE2,INB,IN1,IN2,W)
0472 C
0473 C          THIS SUBROUTINE TAKES AN INPUT VECTOR SET." X ", AND
0474 C          CLASSIFIES IT'S MEMBERS INTO ONE OF TWO REGIONS
0475 C          DEFINED BY "X1" AND "X2" VIA A K-MEANS ALG. THIS ROUTINE
0476 C          ALSO RETURNS THE TWO MEAN VECTORS IN THE ARRAY "MEAN".
0477 C
0478 C          DIMENSION X(NVE2,NELE),X1(NVE2,NELE),X2(NVE2,NELE)

```



Appendix B. (Con't)

```

0479          DIMENSION MEAN(2,NELE),OLDM(2,NELE),INB(NVE2)
0480          DIMENSION IN1(NVE2),IN2(NVE2),W(NELE,KNODE)
0   1      REAL MEAN
0482 C
0483 C          INITIAL VALUES ARE CHOSEN FOR THE MEAN
0484 C
0485          DO 99 I=1,2
0486          DO 99 J=1,NELE
0487          MEAN(I,J)=0.0
0488 99      OLDM(I,J)=0.0
0489          DO 10 IELE=1,NELE
0490          MEAN(1,IELE)=X(1,IELE)
0491 10      MEAN(2,IELE)=X(2,IELE)
0492          OK=0.0
0493          DO 11 I=1,NELE
0494 11      IF(MEAN(1,I)*W(I,NODE).NE.MEAN(2,I)*W(I,NODE))OK=1.0
0495 C
0496 C          IF THE MEANS ARE THE SAME, INCREMENT THE ELEMENTS OF THE
0497 C          FIRST MEAN BY ONE.
0498 C
0499          IF(OK.EQ.0.0) THEN
0500          DO 1001 I=1,NELE
0501 1001      MEAN(1,I)=MEAN(1,I)+1.0
0502          ENDIF
0503 C
0504 C
0505 C          INITIALIZE INDICES FOR THE TWO NEW CLASSES
0506 C
0507 100      IND1=0
0508          IND2=0
0509          DO 20 J=1,NVEC
0510          SQ1=0.0
0511          SQ2=0.0
0512          DO 30 I=1,NELE
0513          SQ1=SQ1+(X(J,I)*W(I,NODE)-MEAN(1,I)*W(I,NODE))**2
0514 30      SQ2=SQ2+(X(J,I)*W(I,NODE)-MEAN(2,I)*W(I,NODE))**2
0515          SQ1=SQRT(SQ1)
0516          SQ2=SQRT(SQ2)
0517          IF ( SQ2 .LE. SQ1 ) GOTO 50
0518          IND1=IND1+1
0519          DO 40 I=1,NELE
0520          IN1(IND1)=INB(J)
0521 40      X1(IND1,I)=X(J,I)
0522          GOTO 20
0523 C
0524 C
0525 50      IND2=IND2+1
0526          DO 60 I=1,NELE
0527          IN2(IND2)=INB(J)
0528 60      X2(IND2,I)=X(J,I)
0529 C
0530 20      CONTINUE
0531 C
0532 C
0533 C          THE MEAN IS COPIED FOR COMPARISON
0534 C
0535 C
0536          DO 5 IELE=1,NELE
0537          OLDM(1,IELE)=MEAN(1,IELE)
0538 5      OLDM(2,IELE)=MEAN(2,IELE)

```

## Appendix B. (Con't)

```

0539 C
0540 C   A NEW MEAN IS CREATED FOR THE TWO NEW CLASSES
0   1 C
0542   DO 14 K=1,2
0543   DO 14 I=1,NELE
0544 14   MEAN(K,I)=0.0
0545   DO 15 I=1,IND1
0546   DO 15 J=1,NELE
0547 15   MEAN(1,J)=X1(I,J)/FLOAT(IND1)+MEAN(1,J)
0548   DO 16 I=1,IND2
0549   DO 16 J=1,NELE
0550 16   MEAN(2,J)=MEAN(2,J)+X2(I,J)/FLOAT(IND2)
0551 C
0552 C   IF THE LAST ITERATION IS THE SAME AS THE PRESENT VALUE
0553 C   THEN THE SOLUTION HAS BEEN REACHED AND THE SUBR. RETURNS.
0554 C
0555   FLAG=0.0
0556   DO 45 K=1,2
0557   DO 45 I=1,NELE
0558 45   IF(OLDM(K,I)*W(I,NODE).NE.MEAN(K,I)*W(I,NODE))FLAG=1.0
0559   IF (FLAG.EQ.1.0) GOTO 100
0560   RETURN
0561   END
0562 C
0563 C   THIS SUBROUTINE DOES A BINARY COUNT OF THE W-MASK. THE
0564 C   SEQUENCE IS:           00           (ON INITIAL ENTRY)
0565 C                       10
0566 C                       01
0   ' C                       11
0568 C                       00
0569 C   FOR A TWO BIT MASK.
0570 C
0571   SUBROUTINE SEQUN(WOLD,W,NODE,KNODE,NELE)
0572   DIMENSION W(NELE,KNODE),WOLD(NELE)
0573 C
0574 C   STORE OLD MASK
0575 C
0576   DO 10 I=1,NELE
0577 10   WOLD(I)=W(I,NODE)
0578 C
0579   IF(WOLD(1).EQ.0.0) THEN
0580     W(1,NODE)=1.0
0581   ELSE
0582     W(1,NODE)=0.0
0583   ENDIF
0584 C
0585   DO 20 J=2,NELE
0586     ITOG=1
0587     DO 30 I=1,J-1
0588 30   IF(WOLD(I).EQ.0) ITOG=0
0589     IF(ITOG.EQ.1) THEN
0590       IF(WOLD(J).EQ.0) THEN
0591         W(J,NODE)=1
0592       ELSE
0593         W(J,NODE)=0
0594     ENDIF
0595   ENDIF
0596 20   CONTINUE
0597   RETURN
0598   END

```

END

9-87

Dtic