General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

NASA CR- 147404

A Case Study Using LARSYS for Analysis of LANDSAT Data

BY

Tina K. Cary
John C. Lindenlaub

The Laboratory for Applications of Remote Sensing

Purdue University, West Lafayette, Indiana

1975

A Case Study Using
LARSYS for Analysis
of LANDSAT Data

by

Tina K. Cary

and

John C. Lindenlaub

TABLE OF CONTENTS

PREFACE TO THE STUDENT

Prerequisites

This Case Study is the last component of the LARSYS Educational Package. The presentation of material is based on the assumption that you have mastered the instructional objectives of the first five units of the sequence.

Instructional Objectives

The analysis of a set of multispectral scanner data can be broken down into a sequence of steps. By the time you have finished studying this volume and the recommended references, and you have carried out a detailed analysis yourself, you should be able to list the steps of the analysis sequence in the proper order. Furthermore, for each step in the analysis you should be able to do three things:

1) give a brief explanation of the significance of the analysis step with respect to the whole analysis sequence,

2) name and briefly describe any software tools available to carry out the analysis step, and

3) apply the analysis principles to a specific problem by writing control cards, running programs, and interpreting the results of the LARSYS functions used in the analysis sequence.

References

Throughout this case study references will be made to other written materials. The two most commonly referenced sources are considered part of this unit of instruction: LARSYS User's Manual, edited by T. L. Phillips, and Pattern Recognition: A Basis for Remote Sensing Data Analysis, by P. H. Swain (LARS Information Note 111572). These references should be in your site library. Be sure you have them available before you begin the case study. See your instructor if you need help locating them.

Student-Instructor Interaction	While this case study attempts to summarize the experiences of a great many multispectral data analysts, there is no real substitute for talking to someone who is already familiar with the use of the LARSYS programs. This is especially true when you begin carrying out the case study analysis. You are encouraged to discuss your progress periodically with your instructor.
Format	Each section of the case study follows this format: the instructional objectives are stated in italics, followed by a discussion of the purpose, philosophy, and analysis techniques associated with that step in the data analysis sequence; then there is an example showing control cards, computer output, and an interpretation of the results; exercises are provided to test your mastery of the section's instructional objectives; information and directions are provided to guide in the case study analysis.

The material is presented in this format so that a person wishing to become adept in the analysis of multispectral data (in particular, LANDSAT data) using LARSYS can proceed through this case study and learn what the analysis steps are, why each step is important, how each step is carried out, and gain practice in using these analysis techniques.

INTRODUCTION

Launch of the first Earth Resources Technology Satellite (ERTS), now called LANDSAT, in July, 1972 provided a new perspective for remote sensing research. Scientists interested in computer-aided analysis of multispectral scanner data had to devise new techniques appropriate to the new data characteristics and analysis problems.

The techniques described in this manual make use of the LARSYS data processing system. LARSYS has been used for a number of years as a tool in the analysis of multispectral scanner data collected from aircraft altitudes. Analysts experienced with LARSYS in that context have brought their experience to bear on the development of techniques suitable for LANDSAT data.

The procedures used for analysis of multispectral data fall into two categories: "supervised" approaches and "unsupervised" approaches. In a "supervised" approach, data points known to contain specific cover types are used to "train" the classification algorithm. If the cover types of interest (call them information classes) are spectrally distinct (have significantly different reflectance characteristics), the classifier will perform well. If the information classes are spectrally similar, the classifier will have difficulty distinguishing between them. For more information about a "supervised" approach, see Guide to Multispectral Data Analysis Using LARSYS by J. C. Lindenlaub.

On the other hand, in an "unsupervised" approach, spectrally distinct classes are determined without reference to the cover types present on the ground. The spectrally distinct classes are found by a clustering algorithm and are called cluster classes. The cluster classes are then used to train the classification algorithm. After the data has been classified, the analyst tries to associate the cluster classes with cover types of interest. If the relationship between cluster classes and information classes is one-to-one (Figure 1a), or if several cluster classes are associated with the same information class (Figure 1b), the procedure will be considered successful. If one cluster class is associated with two or more cover types of interest (Figure 1c), it is likely that these cover types are spectrally similar and cannot be differentiated using this data set.

The procedure described in this manual combines some techniques from both the supervised and unsupervised approaches. First, the analyst chooses the geographic area to be analyzed and determines the cover types he wishes to classify. Data from areas known to contain these cover types are input to the clustering algorithm.

a)

Information classes   Cluster Classes

    corn ←————————→ 1
    pasture ←————→ 2
    woods ←————→ 3
    water ←————→ 4

b)

Information classes   Cluster Classes

    commercial ←————→ 1
    older housing ←————→ 2
    newer housing ←————→ 3
    agriculture ←————→ 4
    → 5
    → 6

c)

Information classes   Cluster Classes

    corn ←————→ 1
    soybeans ←————→ 2
    bare soil ←————→ 3
    industrial ←————→ 4
    older housing ←
    newer housing ←

Figure 1.   Examples of relationships between information classes
and cluster classes.

As is the case in the unsupervised approach, cluster classes are
formed, but, in this modified procedure, the cluster classes are
then immediately associated with information classes (cover types
of interest) before classification. The spectral similarity of
the cluster classes is calculated in preparation for the next
step, in which spectrally similar cluster classes belonging to
the same information class are combined and labelled with their
information class identification. At this point this approach
resembles a supervised approach in its usage of the data points
known to contain specific cover types for "training" the class-
ifier.

This discussion of kinds of approaches points up the fact
that the procedure to be described here is presented as an approach,
not necessarily the approach.

For any of these procedures, the same LARSYS processors are
used, although they may be used in different sequences or for
somewhat different purposes in the various approaches. The way
in which an analyst uses a processor depends on the characteris-
tics of the data being analyzed, the analysis objectives, the
analyst's understanding of the processor, and his experience and
ingenuity.

The following steps comprise the procedure to be described
in this manual:

> examination of data quality
>
> coordination of multispectral scanner data with available
>   reference data
>
> selection of candidate training areas
>
> refinement of training set
>
> calculation of statistical characteristics of training
>   classes
>
> classification, results display, and evaluation
>
> information extraction and interpretation

At any step in the sequence, interpretation of the results
of that step can lead you to conclude that you need to go back
to a previous step and revise a decision made there. That is,
the procedure is not strictly sequential, but rather tends to be
iterative with feedback at various steps causing an analyst to
loop back and repeat a previous step. A flow chart of the analysis
sequence with dashed lines showing potential iteration loops is
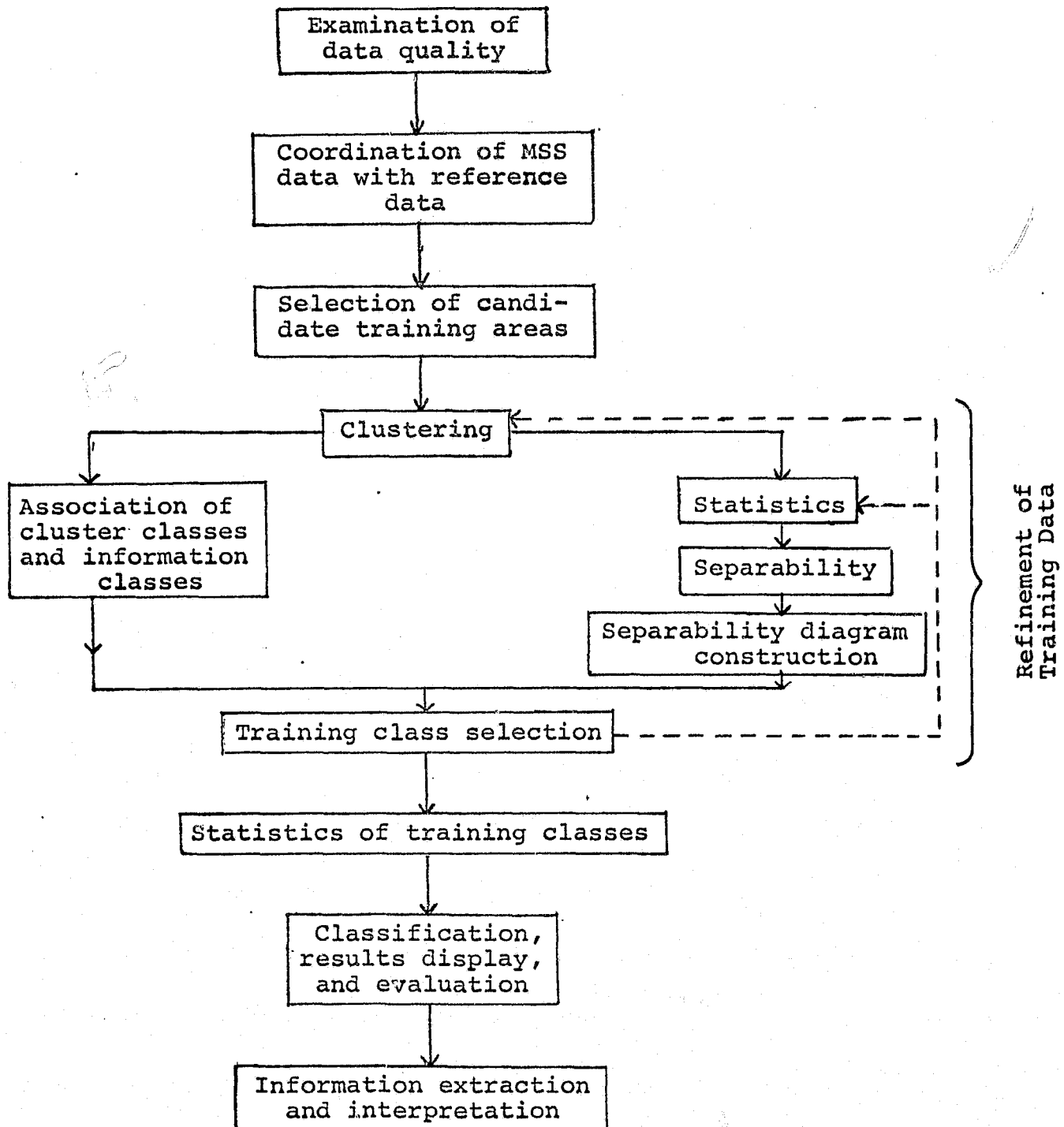shown in Figure 2.

ANALYSIS FLOWCHART

Figure 2. Flow chart indicating the sequence of steps undertaken in the analysis procedure described in this manual.

Section 1.   EXAMINATION OF DATA QUALITY

The instructional objectives for each section will be printed in italics at the beginning of the section.

State at least one reason why the quality of the data set being considered for analysis must be evaluated.

Name at least two sources of data quality information.

Name at least three data idiosyncrasies which might hinder analysis.

Use LARSYS processing functions to obtain identification information about a run and to obtain gray scale printouts of multispectral data.

Formulation of a problem or a hypothesis always precedes any analysis undertaking. To carry out the analysis, two components of the problem statement must be clearly specified: the geographic area of interest, and the particular cover types of interest. These two factors identify precisely what the analysis is supposed to accomplish, that is, the analysis objectives. An example of an analysis objective is "classify Fayette County, Illinois into corn, soybeans, and other cover types." Another example is "determine the percent of the San Juan National Forest in each of these cover types: ponderosa pine, spruce-fir, aspen, and other."

After the analysis objectives are stated, a data set must be selected. LANDSAT satellites with eighteen day cycles provide a wealth of data over any area. From this data, an analyst will choose a set at the time of year suitable for the cover types of interest. In many cases, data sets with much cloud cover or snow cover will not be as desirable as cloud-free or snow-free data sets.

A preliminary evaluation of data can be made by inspecting photographic imagery created from the digital data. This kind of imagery can be obtained from the data distribution centers

which supply digital data tapes. Gross data characteristics, including cloud cover and snow cover, will be apparent in these products. Figure 1-1 shows an example of this kind of imagery, with cloud cover on the right side of the scene.
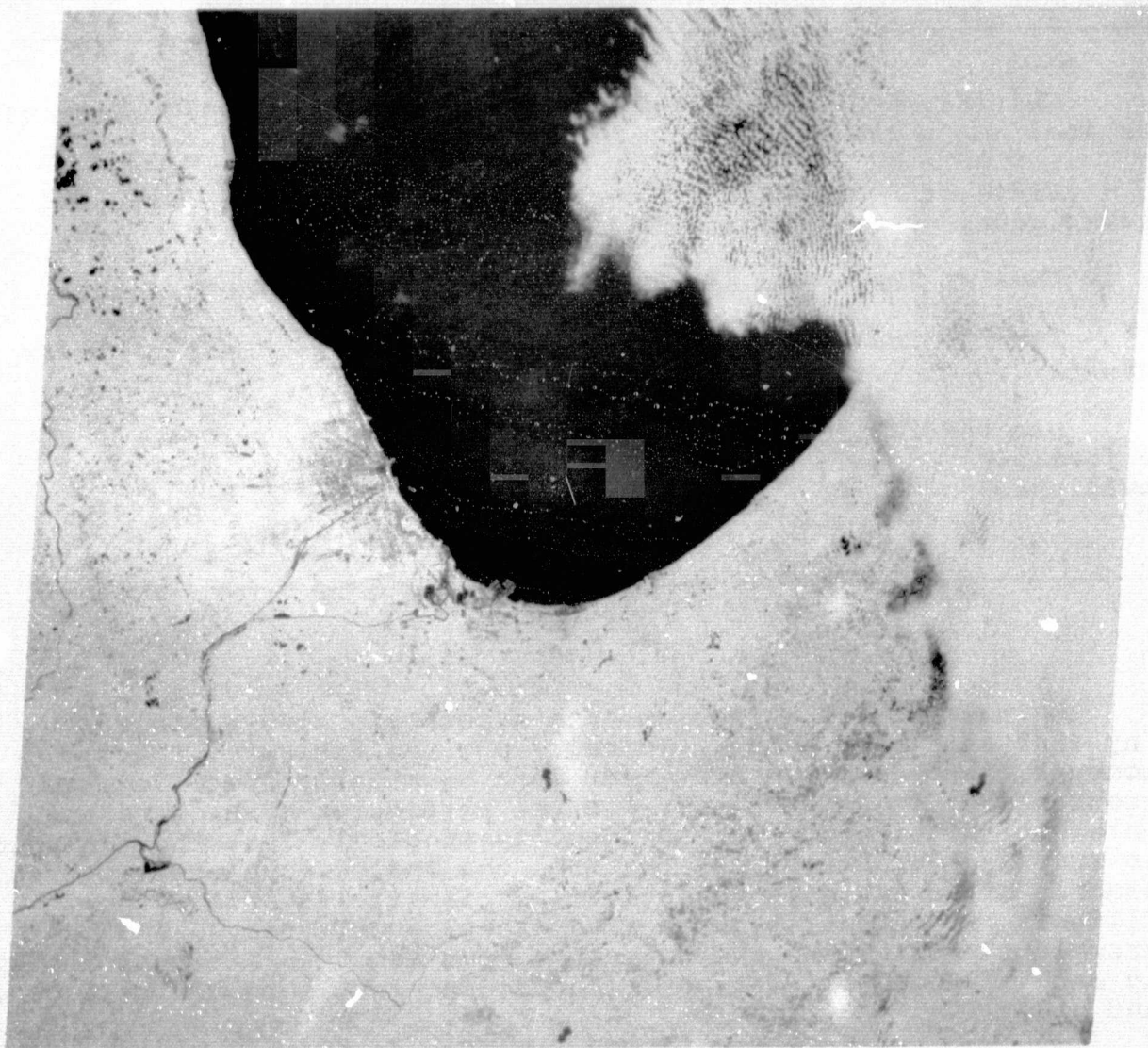

Figure 1-1. Scene number 1070-16041 over Chicago and the surrounding area has clouds obscuring the east side of the scene.

After a data set is selected, the analyst requests that the data preprocessing and reformatting group at LARS convert the LANDSAT data tapes into Multispectral Image Storage Tapes, the format required by LARSYS programs. When the data is reformatted, data log sheets are generated. A master file of these log sheets is maintained in the LARS Computer Center, and a copy is sent to the individual who requested that the data be reformatted. These log sheets are another source of information about data quality. An example is shown in Figure 1-2. Note that the section at the bottom of the form for comments includes the information that there are some bad data lines in that run.

Further examination of data quality can be done by use of LARSYS processors. Examples of various kinds of idiosyncrasies found in LANDSAT data will be given. The LARSYS run number for each data set is included, so that you can obtain gray scale printouts to observe these characteristics firsthand.

The examples were generated by use of the IMAGEDISPLAY processing function, which creates an image on a television screen by the same process PICTUREPRINT uses to create gray scale printouts.

The data in Figure 1-3 shows a phenomenon which appears in LANDSAT data due to the earth's rotation. A rectangular image on the ground appears as a skewed parallelogram, the top edge of the image being shifted to the right with respect to the bottom edge by approximately 5% of the height of the image. In addition the LANDSAT orbit is not oriented exactly over the north pole. This results in a rotation of the imagery which varies with latitude (this rotation is about 12° at 40° north latitude). Figure 1-3a shows non-rectangular fields in LANDSAT data, while Figure 1-3b shows the same data after it has been processed to remove these effects. The run number for Figure 1-3a is 72053602 and for Figure 1-3b is 72053609.
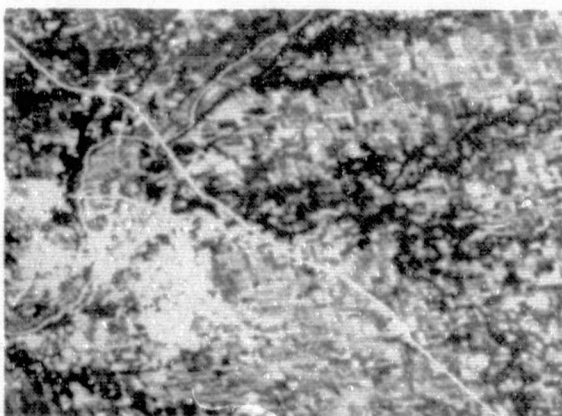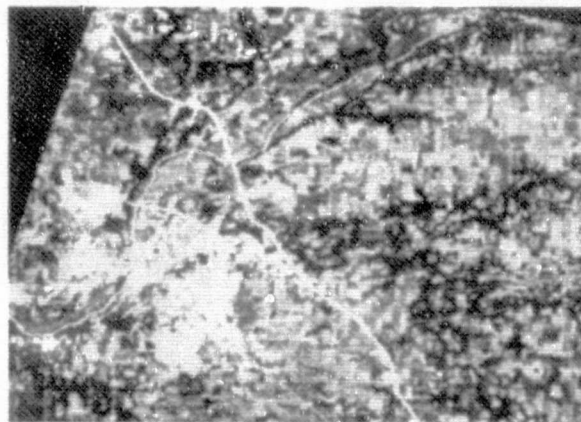
a) uncorrected   b) geometrically corrected

Figure 1-3.   LANDSAT data before and after processing to remove effects of the earth's rotation.

DATA STORAGE TAPE FILE

RUN NUMBER..................... 73052002	FLIGHTLINE ID.......... 134116111 ILL
DATE TAPE GENERATED..... MAR 16,1974	DATE DATA TAKEN................ 6/29/73
TAPE NUMBER.......................... 1353	TIME DATA TAKEN............... 1011 HOURS
FILE NUMBER.......................... 19	PLATFORM ALTITUDE......... 3062000 FEET
LINES OF DATA..................... 536	GROUND HEADING............. 190 DEGREES
SECONDS OF DATA................. 6.56	FIELD OF VIEW........... 0.027 RADIANS
MILES OF DATA.................. 23.41	DATA SAMPLES PER CHANNEL PER LINE 328
LINE RATE............ 81.68 LINES/SEC	SAMPLE RATE........ 0.09 MILLIRADIANS

SPECTRAL BANDWIDTH IN MICROMETERS..

CHAN	LOWER	UPPER	CHAN	LOWER	UPPER	CHAN	LOWER	UPPER
(1)	0.50	0.60	(2)	0.60	0.70	(3)	0.70	0.80
(4)	0.80	1.10	(5)	-----	-----	(6)	-----	-----
(7)	-----	-----	(8)	-----	-----	(9)	-----	-----
(10)	-----	-----	(11)	-----	-----	(12)	-----	-----
(13)	-----	-----	(14)	-----	-----	(15)	-----	-----
(16)	-----	-----	(17)	-----	-----	(18)	-----	-----
(19)	-----	-----	(20)	-----	-----	(21)	-----	-----
(22)	-----	-----	(23)	-----	-----	(24)	-----	-----
(25)	-----	-----	(26)	-----	-----	(27)	-----	-----
(28)	-----	-----	(29)	-----	-----	(30)	-----	-----

DATA RUN CONDITIONS..

DATA TAPE COMMENTS..

GEOMETRIC CORRECTION OF FAYETTE COUNTY, ILLINOIS - CITARS TIME II.
THIS RUN IS A GEOMETRIC CORRECTION OF RUN 73052001 (6/29/73) AND IS REGISTERED
TO THE COORDINATES OF RUN 73039101 (6/11/73).
CORRECTED TO LINEPRINTER ASPECT RATIO.
CHANNELS 1,3,AND 4 HAVE 1,7, AND 11 BAD LINES RESPECTIVELY.

Figure 1-2. Data Log Sheet for LANDSAT Data.

While the "uncorrected" data is adequate for some analysis tasks, "corrected" data simplifies the analyst's job of locating features, since it can more easily be compared with reference data (maps, aerial photography).

As previously mentioned, clouds can significantly decrease the usefulness of a data set. The example shown in Figure 1-4 is from run 72033000. A more subtle situation is the presence of haze. Figure 1-5 shows data collected on September 12 and 13, 1972, over southern Indiana.*

On the 12th a thin haze was present, but on the 13th the sky was clear. Images from two spectral bands on each date are shown. Notice that the channel two image (.6-.7 μm) on the 12th shows the haze, while the channel four image (.8-1.1 μm) from the same date does not. This is due to the fact that there is less scattering of the longer wavelengths.
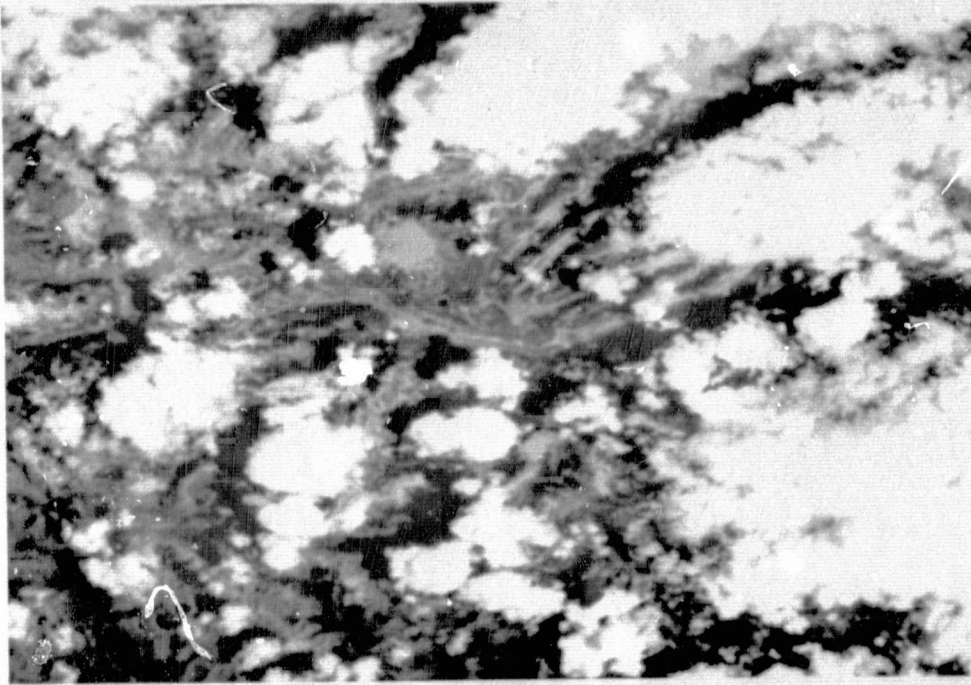
Figure 1-4. An example of clouds and their shadows.

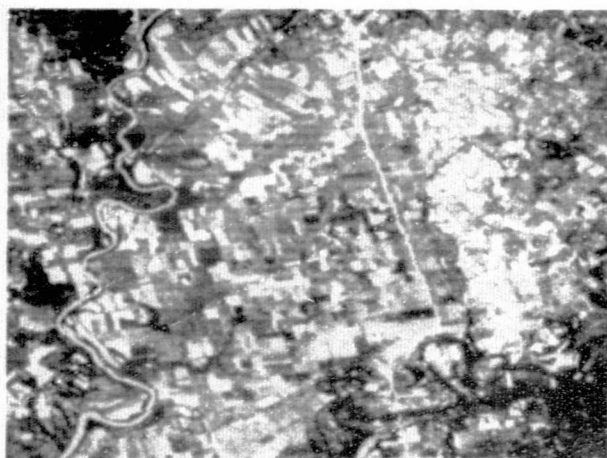
*The coverage pattern provides sidelap on succeeding passes ranging from 14% at the equator to more than 85% at the poles. The sidelap is approximately 30% for southern Indiana.

- 6 -
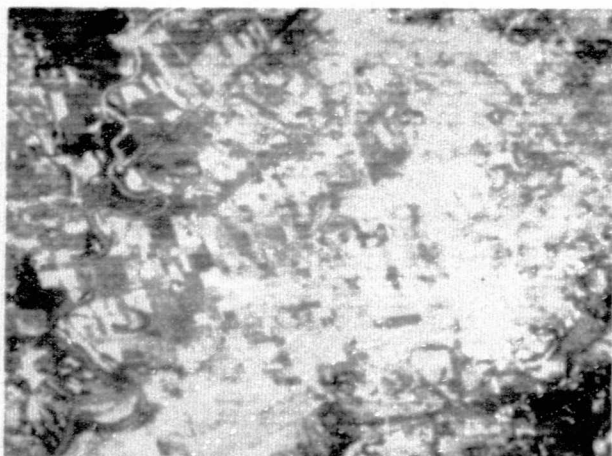
a)   September 12, channel 2
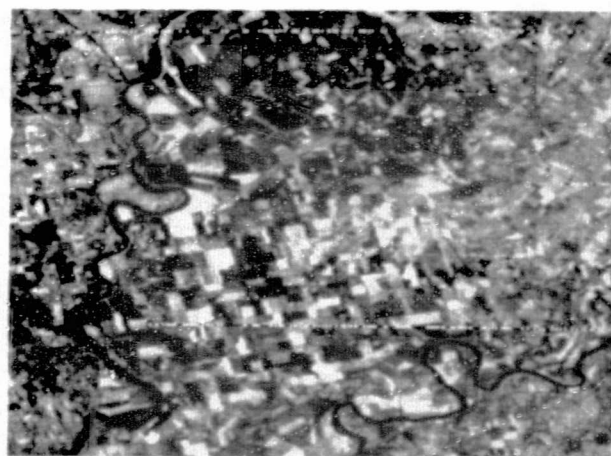
b)   September 12, channel 4

c)   September 13, channel 2
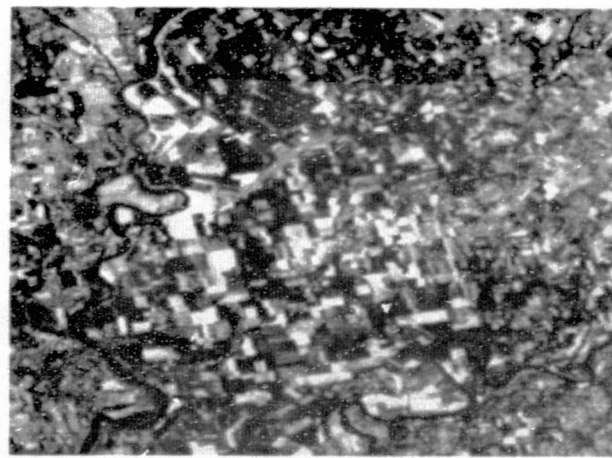
d)   September 13, channel 4

Figure 1-5.   Haze was present on September 12, but September 13 was clear.

The presence of snow can be a limitation in data analysis
if the cover types of interest are the vegetative cover types
under the snow. The data shown in Figure 1-6, run 73034300,
has quite a bit of snow cover. If the purpose of the analysis
is to determine the areal extent of snow cover, the presence of
snow would be desirable, but the presence of both clouds and
snow in the same data set would be undesirable, since they are
spectrally similar. An example of such a situation is shown in
Figure 1-7, from run 72051400. Of course, if the purpose of
the analysis is to compare the responses of clouds and snow,
this data set would be quite useful. This example points up the
necessity of clearly formulating analysis objectives and keeping
the objectives in mind.

Another idiosycrasy which can occur in LANDSAT data appears
as stripes in the image. In the LANDSAT scanner system, six scan
lines are swept out in each wavelength band each time the mirror
oscillates. A separate set of detectors is used for each of
these scan lines. If these detectors and their associated elec-
tronics are not properly matched or calibrated, a striping effect
may be noticeable in the imagery. A dramatic example can be
seen in Figure 1-8, channel 1 of run 72044401, a LANDSAT scene
which includes Lafayette, Indiana. The STATISTICS processing
function was used to obtain quantitative information about this

Figure 1-6. Snow covers the higher elevations
in the mountains of Colorado.

Figure 1-7.   The Rocky Mountains are on the
left and the Great Plains on the
right. There is snow on the
plains and haze over the mountains.

Figure 1-8.   Striping effect in imagery.

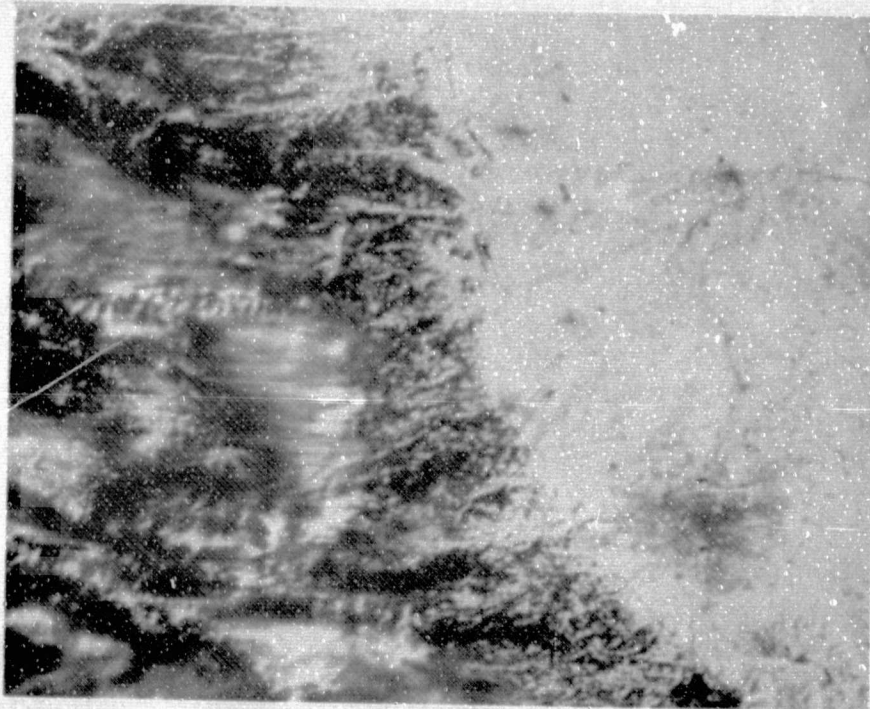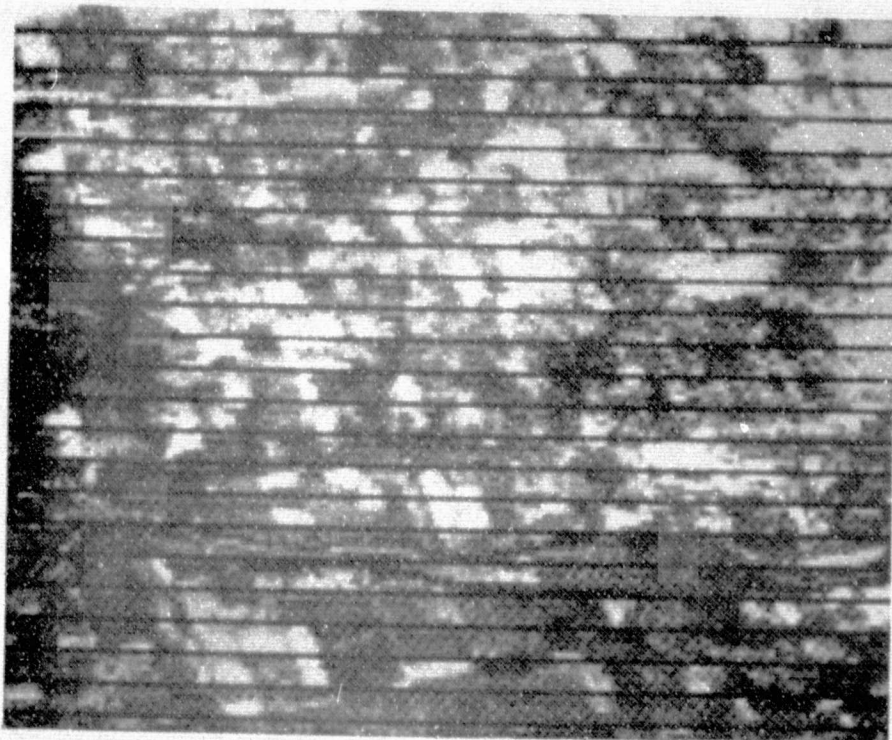striping. The table below shows the mean and standard deviation
for the output of each of the channel 1 detectors over the whole
frame. To obtain this information from the STATISTICS processor,
a line interval of six was used, with successive starting lines
of 1, 2, 3, 4, 5, and 6.

Detector	Mean	Standard Deviation
1	21.9	3.21
2	21.8	3.07
3	7.0	1.52
4	21.5	3.13
5	20.9	3.11
6	21.9	3.03

Notice that the mean value for detector 3 is very low compared
to that of the other detectors. Apparently a malfunction occurred
in the detector electronics, resulting in the striping illustrated
in Figure 1-8.

Another data idiosyncrasy which is sometimes present in LANDSAT
data is called a bad data line. In Figure 1-5, the example of haze,
look at the channel four image from September 12th. In the top half
of the picture there is one bad data line all the way across, and
near the middle of the picture one bad data line goes about two
thirds of the way across.

The material presented in this section has indicated that
analysis objectives include specification of geographic area and
cover types. After these are determined, a data set is chosen.
Next, the quality of the data must be examined to determine
whether it will be adequate for meeting the analysis objectives.

Several kinds of data characteristics were discussed and
illustrated, including cloud cover, snow cover, haze, striping,
and bad data lines.

The example and case study that follow will give you an
opportunity to use LARSYS processing functions to examine data
quality. In preparation for the material presented there, you
should read the following material in the LARSYS User's Manual:

a) Section 4 (Volume 1) of the LARSYS User's Manual, pages 4-1
to 4-3, gives a general description of LARSYS Control Commands.
The remaining pages in Section 4 describe the individual Control
Commands in detail. In particular, review the REFERENCE RUNTABLE
Control Command.

b) Section 6 (Volume 2), pages 6-1 to 6-3, gives a general
description of LARSYS Processing Functions. The remaining pages
in Section 6 describe the individual Processing Functions in
detail. Review the DUPLICATERUN, IDPRINT, and PICTUREPRINT
Processing Functions at this time. In particular note the last
paragraph on page PIC-7, concerning the BLOCK control card.

The examples given in conjunction with each step of the analysis include representative control card listings, computer printouts, and interpretations drawn from an analysis of LANDSAT data from the Kenosha Pass area of Colorado, run 73057902.*

The objectives of the Kenosha Pass analysis were the following:

1) to classify and inventory the area into these cover types: snow, grassland, deciduous forest, coniferous forest, barren (bare rock and bare soil);

2) to produce a classification map of these cover types;

3) to evaluate the classification accuracy.

In order to determine if the run was in the system runtable, the analyst typed in the following information at the terminal:

reference runtable 73057902

and the computer indicated that the run was in the system runtable by responding:

RUN NO.	TAPE	FILE	LINES	CHAN	SAMP	FLIGHTLINE ID
73057902	1087	3	1535	4	2200	138817134 COL

To obtain more information about the run, the analyst used the IDPRINT processing function. The following control card deck was set up:

```
*IDPRINT
PRINT RUN (73057902)
END
```

The output is shown in Figure 1-9. Note the wavelength bands of the LANDSAT scanner system. The first two bands are in the visible portion of the spectrum, and the last two are in the near infrared.

The next step was to make a working copy of the run. The DUPLICATERUN processing function was used. A working copy can serve several purposes, as you read on page DUP-1 in Section 6 of the LARSYS User's Manual. The analyst set up the following control cards to copy the data:

```
-COMMENT COPY OF RUN 73057902 FROM 1087 TO 253
*DUPLICATERUN
FROM RUN(73057902)
TO TAPE(253), FILE (1)
END
```

*The original analysis of this data was done for the U.S. Forest Service under USDA Contract 21-292.

- 10 -

```
   TECTRA                      LABORATORY FOR APPLICATIONS OF REMOTE SENSING            MAY   7,1975
   WILSSN                                  PURDUE UNIVERSITY                            1C 35 14 AM
                                                                                       LARSYS VERSIUN 3


   TAPE NUMBER.............. 1087        FILE NUMBER...............   3          RUN NUMBER............ 73C57902
   CONTINUATION CODE.......... 0        NUMBER OF DATA CHANNELS....  4          NUMBER OF DATA SAMPLES... 2200
   FLIGHT LINE.. 138817134    COL       DATE DATA TAKEN......  8/15/73          TIME DATA TAKEN.... 0913 HOURS
   PLATFORM ALTITUDE.3062000 FEET       GROUND HEADING.... 18J DEGREES          REFORMATTING DATE.APR   8,1974
                                        NUMBER OF LINES.......... 1535


                              SPECTRAL                                  CALIBRATION
                               BANC                                     PULSE VALUES

             CHANNEL     LOWER         UPPER                  CO            C1             C2

                1        0.5C          0.6C                   0.C          2.48C          0.C
                2        0.60          C.70                   0.C          2.CCC          C.C
                3        0.70          J.80                   C.9          1.760          0.C
                4        0.8C          1.1C                   0.C          4.6CC          J.0
```

Figure 1-9.   Output from the IDPRINT processing function.

For the remainder of the analysis, the analyst used a personal runtable so that he gained access to run 73057902 in the first file of tape 253.

The analyst then wanted to combine the task of investigating data quality with the tasks of (1) locating the line and column coordinates of the area of interest, and (2) producing gray scale printouts of the area of interest to use for selection of training areas.

From experience combined with knowledge of how Earth surface features interact with the sun's electromagnetic energy, the analyst knew that one of the two channels from the visible portion of the electromagnetic spectrum together with one of the two channels from the near infrared portion would provide sufficient information for locating areas. He therefore chose channels 1 and 3 for the task of locating the line and column coordinates of the area of interest. To generate the desired gray scale printouts, the following cards were used:

```
-COMMENT GRAY SCALE FOR LOCATION OF KENOSHA PASS COORDINATES
-RUNTABLE
DATA
RUN(73057902), TAPE(253), FILE(1)
END
*PICTUREPRINT
DISPLAY RUN (73057902), LINE (1, 1535, 4), COLUMN (1, 2194,4)
CHANNELS 1, 3
END
```

The analyst used a comment card so that the output would be readily identifiable. Using these printouts in conjunction with reference data, the analyst located the coordinates of the area of interest. To look at the area in more detail (with line interval of one and column interval of one) for subsequent steps in the analysis, the analyst used the following cards:

```
-COMMENT PICTUREPRINT-KENOSHA PASS TRAINING AREA SELECTION
-RUNTABLE
DATA
TAPE(235), FILE(1)
END
*PICTUREPRINT
DISPLAY RUN (73057902), LINE (197, 531, 1), COL (401, 803, 1)
CHANNELS 2, 4
BLOCK RUN (73057902), LINE (197, 531, 2), COL (401, 803, 2)
END
```

Two points should be noted in this control card setup. First, note that different channels were used than for the previous job. The analyst has now looked at all four channels to investigate data quality. Second, note the use of the BLOCK card. In the first PICTUREPRINT, no block card was used because the entire run was to be displayed, and histogramming every tenth line and column of the run (the default parameters) seemed reasonable. In this

second PICTUREPRINT, a subset of the run was being displayed,
and the analyst chose a line and column interval of two because
he wanted his histogram to be based on more data points than the
default interval of ten would provide. (See page PIC-7 in Volume 2
of the LARSYS User's Manual for an explanation of the block card.)

The analyst observed a couple of clouds and their shadows,
but judged the overall data quality to be adequate.

<div align="right">EXERCISES</div>

___

1. Explain in your own words why examination of data quality
should precede any extensive analysis.

2. Name at least two ways in which the analyst of remote
sensing can examine data quality.

3. Name at least three types of data idiosyncrasies an
analyst might find in LANDSAT data.

<div align="right">CASE STUDY</div>

___

As you progress through this guide, you will be asked to
carry out an analysis of a portion of LANDSAT scene 1321 - 15595,
collected June 9, 1973. (Figure 1-10). The run number is
73033802*. The scene is in southern Indiana. Take time now to
look at Figure 1-10 and state analysis objectives you would like
to pursue. Discuss them with your instructor.

Students who have analyzed this data set have used these
objectives: a) classify the area into cover type classes of
urban, forest, agriculture, and water; b) produce a cover type
map of these four classes; c) assess the accuracy of the class-
ification.

The data set you will be working with has been processed
for geometric correction, and it has a scale of 1:24,000, which
matches the scale of the U.S. Geological Survey 7.5 minute topo-
graphic series.

Obtain gray scale printouts for the channels corresponding
to the .6 - .7 μm and .7 - .8 μm wavelength bands, for lines 30
to 430 and columns 112 to 333. These printouts will be used in
the next step of the analysis.

___

*The DUPLICATERUN processing function has been used to make a copy
of this run for your terminal site. Consult your instructor for
the proper tape and file number.

PRECEDING PAGE BLANK NOT FILMED
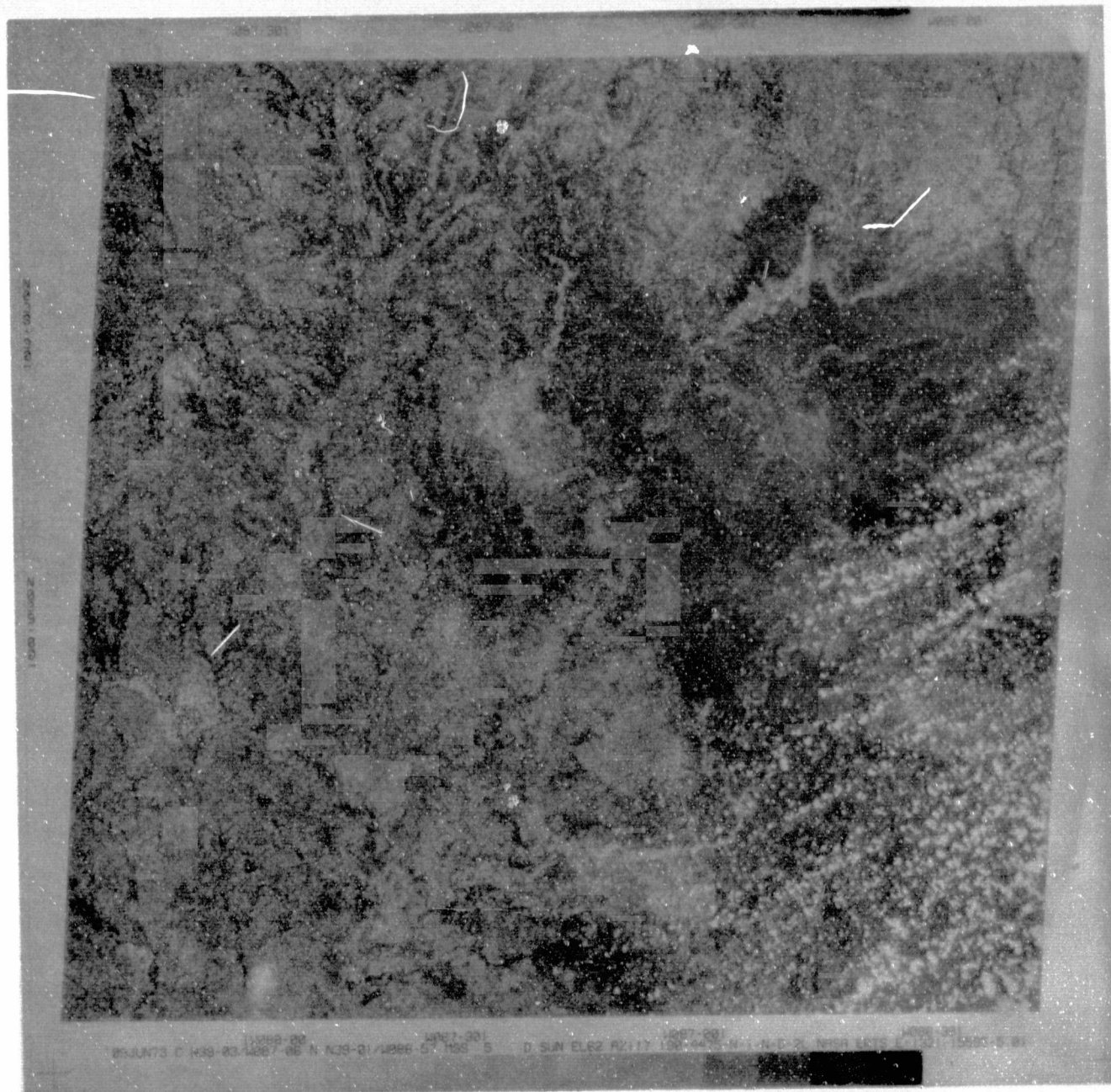
This is a print of ERTS scene 1321-15595, channel 5 (.6-.7μm),
collected June 9, 1973 at 9:59 a.m. Note that north is displaced
13° from vertical.

Figure 1-10

FOLDOUT FRAME

This print shows channel 7 (.8-1.1μm) of the same ERTS scene. The area outlined corresponds to the frame of aerial photography on the next page. This area includes Monroe Reservoir and Bloomington, Indiana.

FOLDOUT FRAME

This print was made from a 9 x 9 color infrared photograph
collected at an altitude of 60,000 feet at 11:40 a.m. the same
day.

Section 2. COORDINATION OF MULTISPECTRAL SCANNER DATA WITH
AVAILABLE REFERENCE DATA

---

*Upon completion of this section, you should be able to do
the following:*

State one reason for the necessity of reference data.
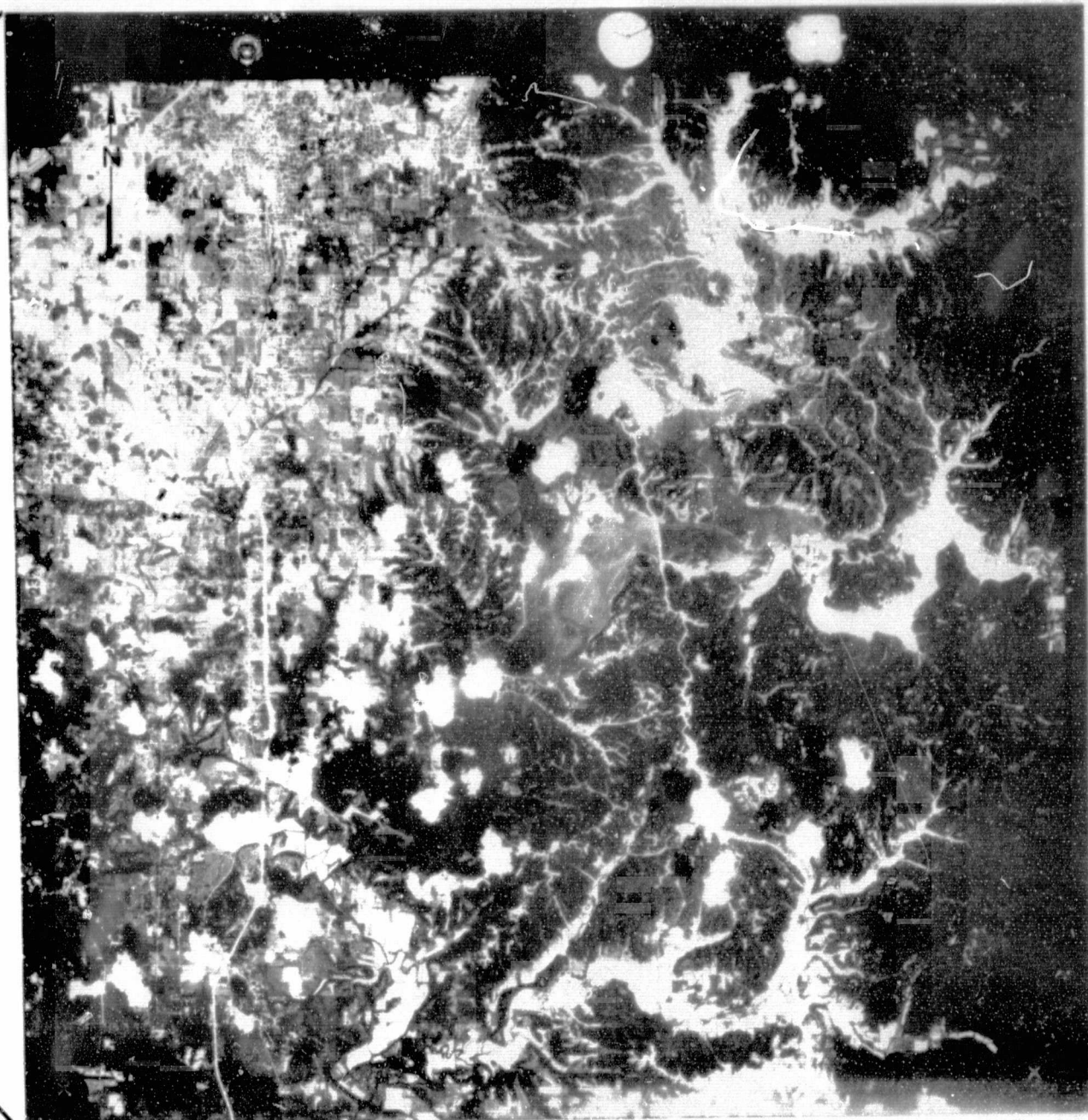
*State one reason for correlating multispectral scanner data
with reference data.*

List at least four kinds of reference data.

*Correlate the location of ground features apparent on multi-
spectral scanner data with the location of those features on an
aerial photograph.*

Coordination of multispectral scanner imagery with known
features on the ground is necessary in order to determine the
line and column coordinates of candidate training areas. The
importance of accurate information about the actual ground scene
is discussed in detail in LARS Information Note 120371, The
Importance of "Ground Truth" Data in Remote Sensing, by Roger
M. Hoffer. You should read this information note at this time.

What can an analyst use to obtain information about the
ground scene? Aerial photography can be a source of information.
Photography can be collected at various altitudes, resulting in
reference data over a range of scales. In general, as a plane
flies higher, each photograph will cover a larger area, but less
detail will be discernible.

Another variable in aerial photography besides altitude is
film type. Black and white film, color film, and color infrared
film can all record various kinds of information about a ground
scene, and serve as reference data for an analyst who understands
how to interpret photographic film.

Aircraft multispectral scanner data can also serve as reference
data for an analyst working with satellite data, by providing more

detailed information about the spectral characteristics of portions of a scene.

Maps (county highway maps or U.S. Geological Survey maps, for example) and historical records (past crop yields or weather patterns, for instance) can be useful to an analyst by helping him visualize an area and its characteristics.

Another source of information includes observations "at the scene" by the analyst, or other personnel. These observations can provide the key to successfully relating the spectral responses in the data to the cover types on the ground.

As a part of the case study, you will see examples of some of these kinds of reference data.

_____ EXAMPLE

In the example analysis of the Kenosha Pass area of Colorado, the analyst had small-scale (high altitude) photography available for reference data. By looking at the photography and gray scale printouts of the data, the analyst was able to correlate the location of ground features in the two images.

_____ EXERCISES

1. State one reason for obtaining reference data.

2. State one reason for correlating multispectral scanner data with reference data.

3. List at least four kinds of reference data.

_____ CASE STUDY

Obtain the following materials from your instructor: U.S. Geological Survey 7.5 minute topographic quadrangle sheets for six quads,* a 35 mm slide of a color-infrared aerial photograph, and a Monroe County map. Using these materials and Figure 1-10 in conjunction with the printouts generated during the previous analysis step, mark on the printouts with a felt tip pen the boundaries of as many features as possible (Lake Monroe, Highway 37, Bloomington, for example). Note that some features are more apparent on one printout than the other.

_____

*Oolitic, Bartlettsville, Clear Creek, Allen's Creek, Bloomington and Unionville, Indiana.

Section 3. SELECTION OF CANDIDATE TRAINING AREAS

Upon completion of this section, you should be able to do the following:

State in your own words why training areas must be selected.

Name at least two considerations that should go into the selection of candidate training areas.

Select candidate training areas, and specify their coordinates by means of Field Description Cards.

The next step in the analysis of multispectral scanner data is the selection of candidate training areas. This section will begin with a discussion of what training samples are and why they are needed, followed by a discussion of how candidate training areas are chosen.

To explain what training samples are and why they are needed, some pattern recognition concepts should be introduced. Pattern recognition provides the theoretical framework for LARSYS (Swain, 1972). The pattern recognition algorithms require that examples of typical data from each class of interest be supplied to the computer programs. These data, called training samples, are used to set certain parameters for the pattern recognition algorithms, in effect "training" the computer to recognize the training classes. When the classification operation is being carried out by the pattern recognition algorithms, each data point is "compared" to the training sample for each class, and the point is assigned to the "most likely" or most similar class. Further discussion of these concepts can be found in LARS Information Note 110474, An Introduction to Quantitative Remote Sensing by John Lindenlaub and James Russell, and in LARS Information Note 111572, Pattern Recognition: A Basis for Remote Sensing Data Analysis by Philip H. Swain. As you increase your understanding of the material presented in those readings, you will bring more insight to the interpretation of your analysis results.

To obtain training samples for this procedure, the first step is selection of candidate training areas. Experience gained during the development and evolution of this step in the analysis has indicated that a good starting point is to select candidate training areas from 40 to 100 lines by 40 to 100 columns in size, and containing from three to five cover types.

To select these candidate training areas, an analyst begins by reviewing the analysis objectives. In stating the objectives, the cover types of interest are listed. These cover types are called information classes. Candidate training areas are selected in such a way that every information class is represented in at least one of the areas. When possible, each information class is included in more than one candidate training area. This increases the likelihood that the training data will be representative of all of the variations in cover types in the scene being analyzed. When representative training data is available to the classifier, assignment of a data point to the most likely training class has a higher probability of being a correct assignment.

A common procedure for selecting the candidate training areas is to identify in the available reference data some general areas that contain the information classes. These areas are also located on gray scale printouts of the multispectral scanner data. From these areas, candidate training areas are selected, following the guidelines indicated previously: each area is from 40 to 100 lines by 40 to 100 columns, each area includes more than one cover type, and every cover type is included in at least one (preferably two or more) candidate training area. To help assure obtaining representative training data, the candidate training areas should be distributed uniformly throughout the area to be classified, but this may not be possible if adequate reference data is not available. Usually, representative training data for all information classes can be obtained by selecting from four to eight candidate training areas.

After the candidate training areas are selected, their coordinates must be specified in terms of lines and columns, so that the areas can be submitted to LARSYS processing functions in subsequent analysis steps. The formats of the Field Description Cards used to accomplish this are described in Volume 1 of the LARSYS User's Manual, pages 2 - 27 and 2 - 28. A coding sheet set up for the format is shown in Figure 3-1.

EXAMPLE

In the Kenosha Pass example, the analyst used his available reference data - color infrared aerial photography - to select candidate training areas. Since the photography was available only for a limited area, he restricted his choice of candidate training areas to fall within regions covered by photography. Review of the analysis objectives indicated that the cover types of interest were snow, grassland, deciduous forest, coniferous forest, and barren (bare rock and bare soil). With these cover

Run Number (1-8)	Field Designation (11-18)	First Line (21-25)	Last Line (26-30)	Line Interval (31-35)	First Column (36-40)	Last Column (41-45)	Column Interval (46-50)	Field Type (51-58)	Additional Information (59-72)

Figure 3-1. Field Description Card format.

types in mind, the analyst selected four candidate training areas. Each of the areas contained more than one cover type. Figure 3-2 shows a gray-level image (.6 - .7 μm band) of the Kenosha Pass area with the four candidate training areas outlined.

In terms of number of lines by number of columns, the four candidate training areas were of the following size: 89x73, 97x70, 44x103, and 56x66. Notice how these areas compare to the guideline of 40 to 100 lines by 40 to 100 columns.

By interpretation of his reference data, the analyst selected the areas so that one area included grassland, coniferous forest, and deciduous forest; a second area included those three cover types again and also included water; a third area included bare soil in addition to grassland, coniferous forest, and deciduous forest; and the fourth area included snow, grassland, bare rock, and conifers.
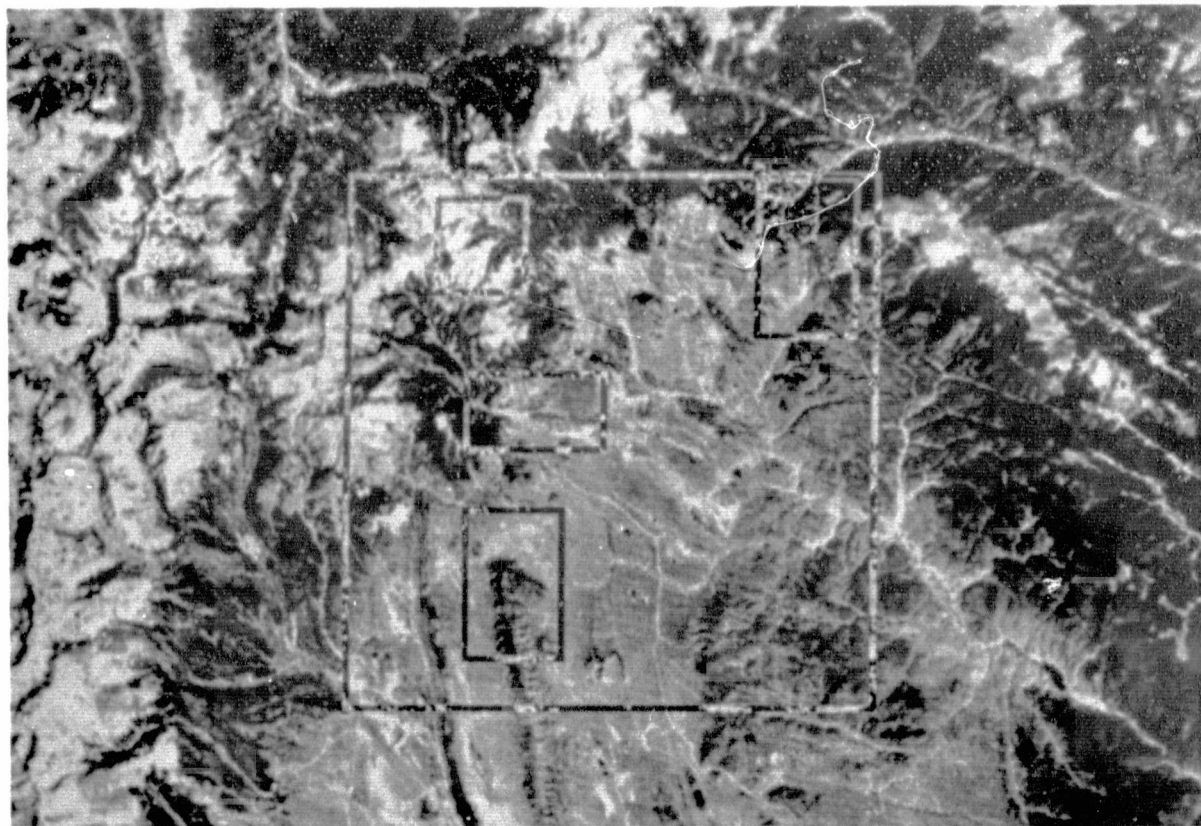

Figure 3-2.   Kenosha Pass area and the four candidate training
             areas outlined.

Then the analyst completed this step of the analysis by specifying the coordinates on Field Description Cards.

73057902	411	499	1	493	565	1
73057902	199	295	1	719	788	1
73057902	324	367	1	494	596	1
73057902	212	267	1	468	533	1

One point which should be discussed is the percent of total area used for training. In the Kenosha Pass analysis, the areas covered slightly more than 15% of the area. However, this is a larger proportion than usual. As the size of the area to be classified increases, the percent of the area used for training generally decreases. This trend is an expression of another idea: the amount of reference data available is an upper limit on the amount of training data that can be used, and as the area being considered increases in size, the logistics and expense of collecting reference data will limit the amount collected.

Where there has been a scarcity of reference data over a large area, an analyst has used as little as one tenth of one percent of the area for training. A more common proportion ranges from 1% to 10%.

EXERCISES

---

1. State why training areas must be selected.

2. Name two considerations that should go into the selection of candidate training areas.

CASE STUDY

Using the available reference data and the gray scale print-outs, select candidate training areas. Make sure that every cover type specified in your analysis objective is included in at least one of the candidate training areas. Specify your candidate training areas on Field Description Cards. Discuss the areas you selected with your instructor.

Section 4. CLUSTERING CANDIDATE TRAINING AREAS

Upon completion of this section, you should be able to do the following:

Describe at least two tasks the CLUSTER processing function can accomplish for you.

State the rule-of-thumb used to determine the number of clusters to request, and the reason behind it.

Given punched and printed output from the CLUSTER processor, explain the effect of choosing different MINPOINTS values.

Use the CLUSTER processing function to find spectrally distinct classes, given Field Description Cards of the areas to be clustered.

In the previous section, the concept of "training samples" was discussed, and candidate training areas were chosen. The process of getting from candidate training areas to training samples is a complex process as well as a crucial one. This section and the next five sections will all deal with aspects of this refinement process. Although the material is written in a linear or "straight-through" fashion, the process is somewhat circuitous, as indicated in Figure 4-1.

The portion of the refinement process to be discussed in this section involves use of the CLUSTER processor. LARS Information Note 111572, Pattern Recognition: A Basis for Remote Sensing Data Analysis by Philip H. Swain contains a section on clustering. Pages 27 through 36 are recommended for your reading.

The CLUSTER processor uses information from more than one channel or wavelength band (four channels in the Case Study) to produce a single image. Since more information is used, boundaries of ground features or cover types tend to be more distinct on cluster maps than on a single-channel gray scale printout. Thus, one task the CLUSTER processing function can accomplish is boundary enhancement.

- 25 -

Refinement of Training Data

```
                           ⋮
                    ┌──────────────┐
        ┌───────────│  Clustering  │◄─────────────────────────────┐
        │           └──────────────┘                              │
        │                                 ┌────────────┐          │
        │                                 │ Statistics │◄────┐    │
        │                                 └────────────┘     │    │
        ▼                                        │           │    │
┌────────────────┐                               ▼           │    │
│  Association of │                       ┌──────────────┐    │    │
│ cluster classes│                       │ Separability │    │    │
│ and information│                       └──────────────┘    │    │
│     classes    │                              │            │    │
└────────────────┘                              ▼            │    │
        │                          ┌─────────────────────────┐    │
        │                          │  Separability diagram   │    │
        │                          │      construction       │    │
        │                          └─────────────────────────┘    │
        │                                        │                │
        └──────────┐        ┌─────────────────────┘               │
                   ▼        ▼                                      │
            ┌──────────────────────────┐                          │
            │ Training class selection │- - - - - - - - - - - - - ┘
            └──────────────────────────┘
                           ⋮
```
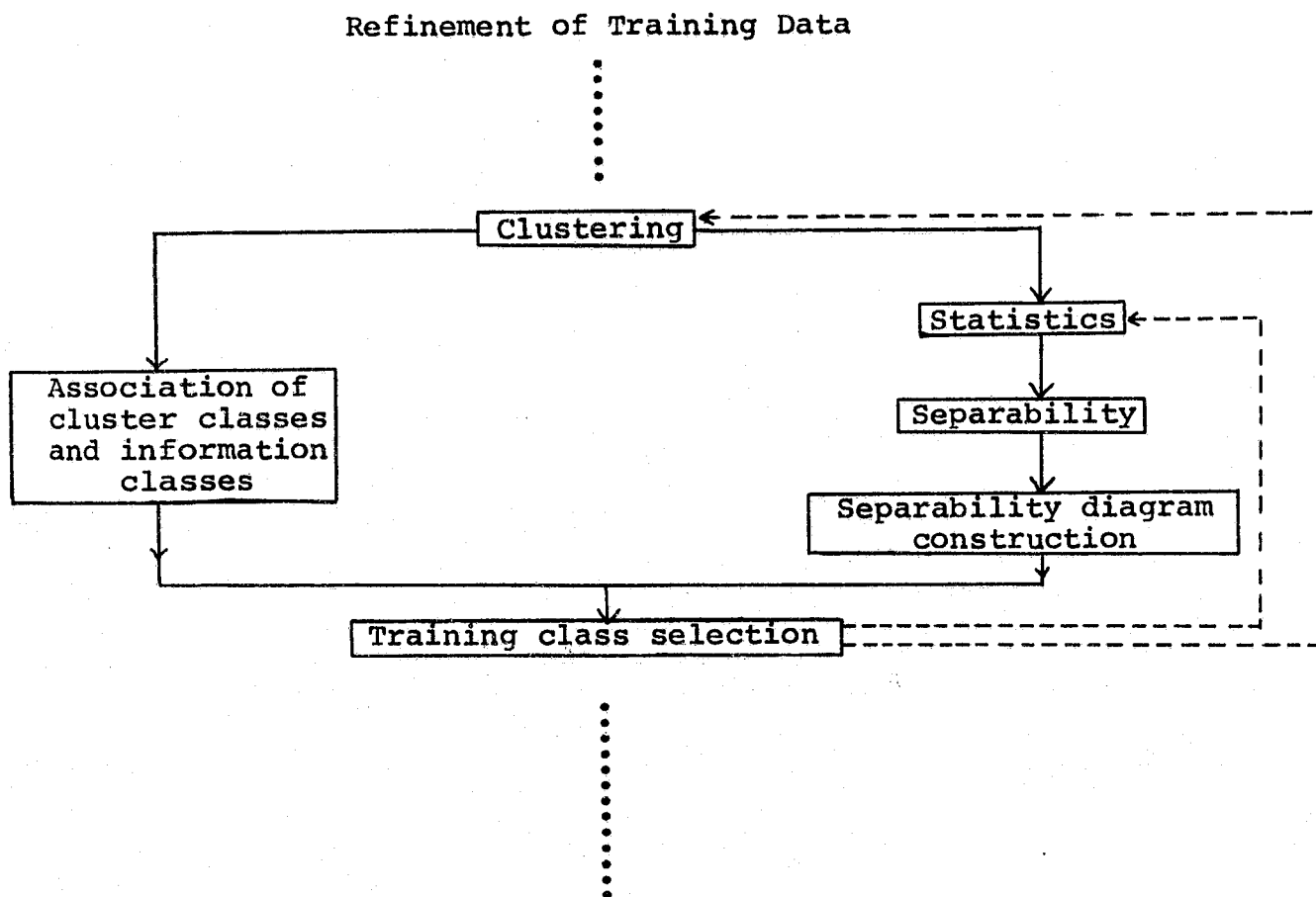
Figure 4-1. Flow chart indicating the steps involved in refinement of candidate training areas. Dashed lines indicate potential iteration loops. This is a portion of the flow chart shown in the Introduction, Figure 2.

The clustering algorithm is called an unsupervised class-
ifier, because it finds natural groupings in multispectral
scanner (MSS) data strictly on the basis of inherent properties
within the data. These natural groupings in the data are called
cluster classes. Thus, another task the CLUSTER processing
function can accomplish is to determine cluster classes within
a data set.

When data from a natural scene is clustered there is a
tendency for the data points within each cluster class to be
distributed in a Gaussian fashion.

Figure 4-2a shows a typical Gaussian function in one dimen-
sion -- commonly called a "normal curve." Figure 4-2b shows a
two-dimensional Gaussian density function. The fact that clusters
in remotely sensed data tend to be Gaussian is important because
the classification algorithm to be used is based upon a Gaussian
assumption, i.e., that the data to be classified can be approx-
imated by a set of Gaussian density functions.

The distribution of the data associated with an information
class is likely to be non-Gaussian such as that shown in Figure
4-3a. As an example, an agricultural crop might exhibit a multi-
modal distribution (more than one peak) due to different soils,
moisture content, planting dates, crop density, seed varieties,
or a combination of these factors. The multimodal non-Gaussian
density function in Figure 4-3a could be decomposed into two
Gaussian components by clustering, as shown in Figure 4-3b.
These components are commonly referred to as subclasses. The
subclass concept is an important one as it allows the analyst to
use a classification algorithm based upon a Gaussian assumption
even though the information class distributions may be non-Gaussian.

In this step of the analysis the CLUSTER processing function
is used to determine cluster classes in the training areas. The
enhanced boundaries on the cluster maps will be used in Section 5
to help establish associations between the cluster classes and
information classes.

The clustering algorithm implemented in LARSYS requires that
the analyst specify the number of clusters to be found. Experi-
ence has indicated that most cover types have multimodal distri-
butions, and a rule-of-thumb is to request twice the number of
expected information classes, except in areas of great topographic
relief (such as the Kenosha Pass example), where three times the
number of expected information classes seems to be a better
guideline.

If an analyst requests an insufficient number of clusters,
some of the clusters will be multimodal, and further clustering
will still be necessary. If "too many" clusters are requested,
they can be grouped back together without much trouble.

A good understanding of the interactions between solar
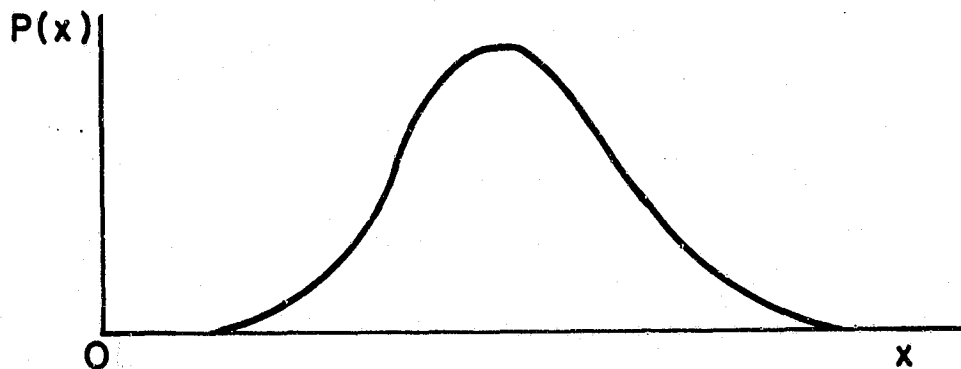energy and matter will help you know how many cluster classes a

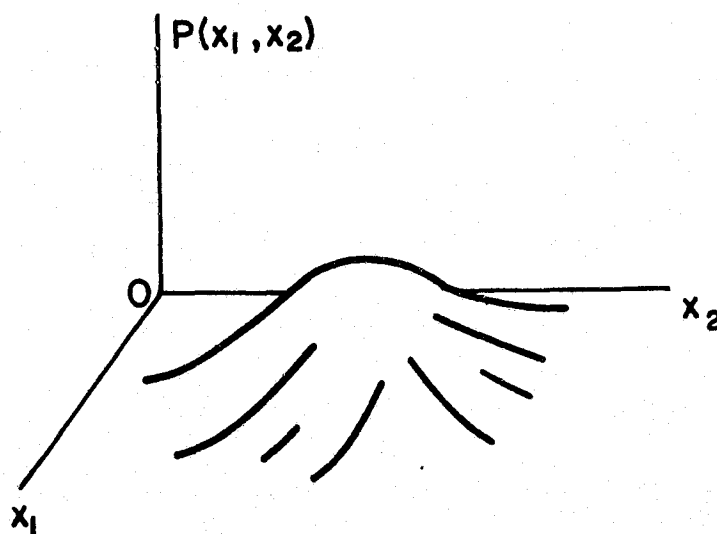Figure 4-2a.   Gaussian density function in one dimension.

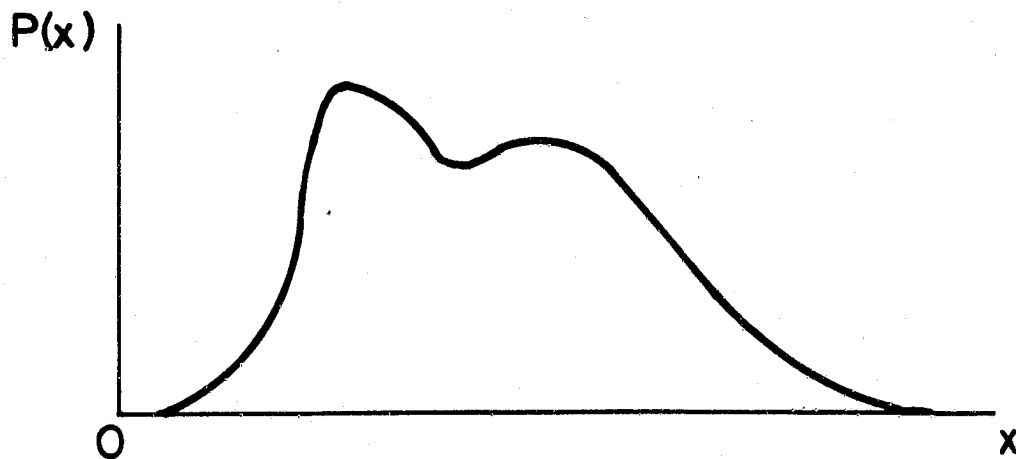Figure 4-2b.   Gaussian density function in two dimensions.

P(x)

Figure 4-3a.   Multimodal non-Gaussian density function.
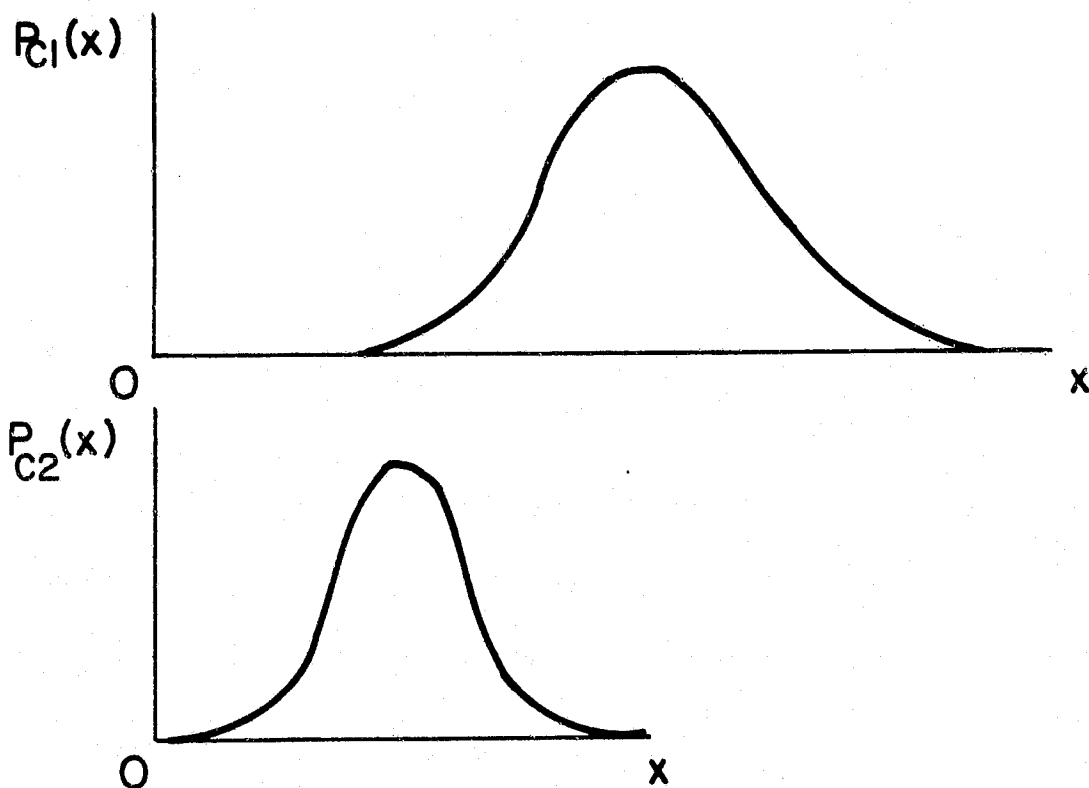
$P_{C1}(x)$

$P_{C2}(x)$

Figure 4-3b.   Multimodal function decomposed into two
Gaussian components.

given scene could be expected to have. Basic information of
interest to the analyst of MSS data would be the reflectance
properties of bare soil, green vegetation, and water, shown in
Figure 4-4. For more specific information, refer to LARS
Information Note 011069, Ecological Potentials in Spectral Signa-
ture Analysis, by R. M. Hoffer and C. J. Johannsen, and also
LARS Information Note 072473, Emission and Reflectance from
Natural Targets, by R. Kumar and L. Silva.

The analyst wanted to cluster each of his candidate training
areas into distinct cluster classes, so he set up his deck in
the following way:

```
-COMMENT KENOSHA PASS CANDIDATE TRAINING AREA 1
-RUNTABLE
DATA
RUN (73057902), TAPE (253), FILE (1)
END
*CLUSTER
OPTIONS MAXCLAS (15)
PUNCH FIELD, MINPOINTS (3)
CHANNELS 1, 2, 3, 4
DATA
73057902              411  499  1    494  565  1
END
-COMMENT KENOSHA PASS CANDIDATE TRAINING AREA 2
*CLUSTER
OPTIONS MAXCLAS (15)
PUNCH FIELD, MINPOINTS (3)
CHANNELS 1, 2, 3, 4
DATA
73057902              199  295  1    719  788  1
END
-COMMENT KENOSHA PASS CANDIDATE TRAINING AREA 3
*CLUSTER
OPTIONS MAXCLAS (15)
PUNCH FIELD, MINPOINTS (3)
CHANNELS 1, 2, 3, 4
DATA
73057902              212  267  1    468  533  1
END
```

Why did the analyst choose a MAXCLAS of 15? The Kenosha
Pass study area includes some very rugged terrain. The analyst
was looking for five cover types - snow, grassland, deciduous
forest, coniferous forest, and barren. He expected an average
of three subclasses per cover type due to the complexity of topo-
graphic relief. If he had wanted to, the analyst could have
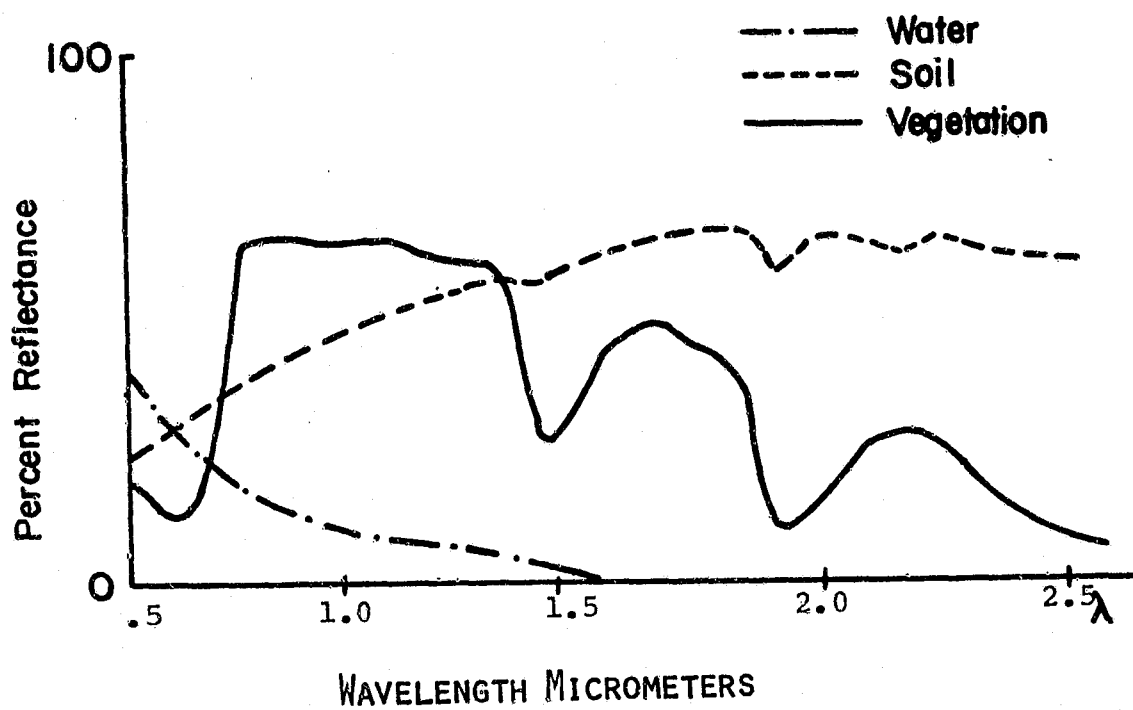chosen a different MAXCLAS parameter for each area clustered.

Figure 4-4.   Typical reflectance properties of bare soil, green vegetation and water.

Each of these CLUSTER jobs produced a punched deck of Field Description Cards to be used in a later step in the refinement of candidate training areas (Section 6).

1. Describe two tasks the CLUSTER processing function can accomplish for you.

2. State the rule-of-thumb used to determine the number of cluster classes to request, and the reason behind it.

3. The following control cards were used to generate the cluster map in Figure 4-5 and the punched output listed in Figure 4-6:

```
-COMMENT CLUSTER EXAMPLE, MINPOINTS EXERCISE
*CLUSTER
OPTIONS MAXCLAS (6)
PUNCH FIELD, MINPOINTS (2)
CHANNELS 1, 2, 3, 4
DATA
RUN (72072302), LINE (500, 515, 1), COL (445, 465, 1)
END
```

The format of the punched cards should be familiar to you from Figure 3-1. Look at the first punched card, find its coordinates on the cluster map, and mark the map with a colored pencil or felt tip marker. Follow the same procedure for all the punched cards describing cluster class 1.

Now, prepare the control cards shown above, BUT use a MINPOINTS parameter of 3. Then, looking at your cluster map, which should be the same as the one in Figure 4-5, and a listing of your punched output, again locate the data points corresponding to the Field Description Cards for the first cluster class. Observe the differences in the two punched decks. Discuss with your instructor the factors you should keep in mind when choosing a MINPOINTS parameter.

CASE STUDY

Set up the control cards to cluster separately each of the candidate training areas you chose in the last section. Remember, you can ask for a different number of cluster classes from each candidate training area if you have reason to do so.

The printed cluster maps will be used in the next section, Section 5, and your punched deck will be used in Section 6.

SBLARSYS
RUSSELL/TKC
                    LABORATORY FOR APPLICATIONS OF REMOTE SENSING
                               PURDUE UNIVERSITY

CLUSTER EXAMPLE, MINPOINTS EXERCISE

FIELD INFORMATION

FIELD		TYPE		LINES	500- 515
RUN NO. 72072302		NO. OF SAMPLES 336		COLUMNS	445- 465
OTHER INFORMATION					

```
                        4444444444444444444444
                        4444455555555555666666
                        5678901234567890123456
                        56789012345678901234 5

            500    XXSS SSSSSS SNNXMNMXX
            501    XS  SS    SS SXXNMNXXX
            502    XS  SSS      XXNXXXXXX
            503    XS S   S     XNM)XNXNXX
            504    SXS      S SXMN)XXXXXX
            505    SSSS     S SNN)))XNXNN
            506     SS    S)  S SNM))NXXNX
            507     S     S)SSSNMMX))XXNN
            508            SSSXMMX)XNNNNN
            509         SS5 XXNMNNNNXNNN
            510        SSSS SXXXMNNXXXNSS
            511    SSSSS  MNXXMMNNNNSSSX
            512    SS  NSMMNXNXMNXNXXXSN
            513       SMMMMNXXNNMMNNNSSSN
            514    SSMXMMMMMMMMNNNNNNSNM
            515    SNX)))))XNMNXXNNNNNMM
```

NUMBER OF POINTS PER CLUSTER

CLUSTER	1	2	3	4	5	6
SYMBOL		)	S	X	N	M
POINTS	67	17	75	70	70	37

Figure 4-5. An example of a printed cluster map.

```
CLASS NS- 1/ 6
72072302        3    501  501   1   447  448    1NS-  1/  6
72072302        4    501  501   1   451  453    1NS-  1/  6
72072302        6    502  502   1   447  448    1NS-  1/  6
72072302        7    502  502   1   452  456    1NS-  1/  6
72072302        9    503  503   1   449  450    1NS-  1/  6
72072302       10    503  503   1   452  455    1NS-  1/  6
72072302       11    504  504   1   448  452    1NS-  1/  6
72072302       13    505  505   1   449  452    1NS-  1/  6
72072302       16    506  506   1   448  450    1NS-  1/  6
72072302       20    507  507   1   447  450    1NS-  1/  6
72072302       21    508  508   1   445  451    1NS-  1/  6
72072302       22    509  509   1   445  449    1NS-  1/  6
72072302       24    510  510   1   445  447    1NS-  1/  6
72072302       26    511  511   1   450  451    1NS-  1/  6
72072302       27    512  512   1   447  448    1NS-  1/  6
72072302       28    513  513   1   445  446    1NS-  1/  6
CLASS NS- 2/ 6
72072302        3    505  505   1   458  460    1NS-  2/  6
72072302        5    506  506   1   459  460    1NS-  2/  6
72072302        7    507  507   1   460  461    1NS-  2/  6
72072302        9    515  515   1   448  452    1NS-  2/  6
CLASS NS- 3/ 6
72072302        1    500  500   1   447  448    1NS-  3/  6
72072302        2    500  500   1   450  455    1NS-  3/  6
72072302        5    501  501   1   449  450    1NS-  3/  6
72072302        6    501  501   1   454  455    1NS-  3/  6
72072302        9    502  502   1   449  451    1NS-  3/  6
72072302       17    505  505   1   445  448    1NS-  3/  6
72072302       20    506  506   1   446  447    1NS-  3/  6
72072302       26    507  507   1   453  455    1NS-  3/  6
72072302       27    508  508   1   452  454    1NS-  3/  6
72072302       28    509  509   1   450  452    1NS-  3/  6
72072302       29    510  510   1   448  451    1NS-  3/  6
72072302       31    510  510   1   464  465    1NS-  3/  6
72072302       32    511  511   1   445  449    1NS-  3/  6
72072302       33    511  511   1   462  464    1NS-  3/  6
72072302       34    512  512   1   445  446    1NS-  3/  6
72072302       38    513  513   1   462  464    1NS-  3/  6
72072302       39    514  514   1   445  446    1NS-  3/  6
CLASS NS- 4/ 6
72072302        1    500  500   1   445  446    1NS-  4/  6
72072302        3    500  500   1   464  465    1NS-  4/  6
72072302        5    501  501   1   458  459    1NS-  4/  6
72072302        6    501  501   1   463  465    1NS-  4/  6
72072302        8    502  502   1   457  458    1NS-  4/  6
72072302        9    502  502   1   460  465    1NS-  4/  6
72072302       14    503  503   1   464  465    1NS-  4/  6
72072302       17    504  504   1   460  465    1NS-  4/  6
72072302       20    506  506   1   462  463    1NS-  4/  6
72072302       23    507  507   1   462  463    1NS-  4/  6
72072302       27    509  509   1   454  455    1NS-  4/  6
72072302       29    510  510   1   454  456    1NS-  4/  6
72072302       30    510  510   1   460  462    1NS-  4/  6
72072302       31    511  511   1   454  455    1NS-  4/  6
72072302       36    512  512   1   461  463    1NS-  4/  6
72072302       37    513  513   1   453  454    1NS-  4/  6
72072302       41    515  515   1   457  458    1NS-  4/  6
CLASS NS- 5/ 6
72072302        1    500  500   1   458  459    1NS-  5/  6
72072302       10    505  505   1   456  457    1NS-  5/  6
72072302       12    505  505   1   464  465    1NS-  5/  6
72072302       17    507  507   1   464  465    1NS-  5/  6
72072302       18    508  508   1   461  465    1NS-  5/  6
72072302       20    509  509   1   458  461    1NS-  5/  6
72072302       21    509  509   1   463  465    1NS-  5/  6
72072302       22    510  510   1   458  459    1NS-  5/  6
72072302       25    511  511   1   458  461    1NS-  5/  6
72072302       33    513  513   1   455  456    1NS-  5/  6
72072302       34    513  513   1   459  461    1NS-  5/  6
72072302       36    514  514   1   457  462    1NS-  5/  6
72072302       41    515  515   1   459  463    1NS-  5/  6
CLASS NS- 6/ 6
72072302        7    507  507   1   457  458    1NS-  6/  6
72072302        8    508  508   1   456  457    1NS-  6/  6
72072302       12    511  511   1   456  457    1NS-  6/  6
72072302       13    512  512   1   451  452    1NS-  6/  6
72072302       15    513  513   1   448  451    1NS-  6/  6
72072302       16    513  513   1   457  458    1NS-  6/  6
72072302       18    514  514   1   449  456    1NS-  6/  6
72072302       21    515  515   1   464  465    1NS-  6/  6
```

Figure 4-6. Listing of punched Field Description Cards generated by CLUSTER.

Warning! You'll get a punched deck for each cluster area.
Be sure to label the cards and keep them in order.

Since clustering is both complex, and crucial to the
success of your analysis, be sure to ask your instructor any
questions that come up.

Section 5. ASSOCIATION OF CLUSTER CLASSES AND INFORMATION CLASSES

Upon completion of this section, you should be able to do the following:

Describe why cluster classes are associated with information classes.

Given printed cluster maps and available reference data, associate cluster classes with information classes.

Up to this point in the process of meeting analysis objectives, the LANDSAT imagery has been examined for data quality, the imagery has been correlated with reference data, candidate training areas have been selected and the data within each candidate training area has been clustered. Previous sections have introduced the concepts of information classes and cluster classes. Recall that information classes were defined during the process of stating the analysis objectives. Cluster classes for each candidate training area were determined by means of the CLUSTER processing function. The objective of this step of the analysis is to associate each cluster class with one of the information classes in order to obtain information class (or subclass) training data.

To carry out this step of the analysis, maximum use is made of all available reference data, so that the cluster classes can be reliably identified. If errors occur in this step of the analysis the training data supplied to the classifier will not be representative of the information classes. The association of cluster classes and information classes is difficult and time consuming, but this step is most important for insuring that the classifier is correctly trained.

EXAMPLE

The reference data available for the Kenosha Pass area included color infrared 9" x 9" transparencies. The analyst used

- 37 -

an overhead projector to superimpose the photography on the printed cluster maps. By varying the projector-to-wall distance, he was able to project the transparency to the scale of the printout, and since the data had been geometrically corrected, a good match could be obtained.

U.S. Geological Survey quadrangle sheets of the area were useful for general location, and for indicating which areas would fall in topographic shadow at the time of the satellite overpass.

The analyst identified the cover types as accurately as he could, making use of his knowledge of the area and his skill as a photointerpreter. The results of his efforts at identification are shown in Table 5-1.

There are several points to observe about Table 5-1. One point is that if the analyst felt a single cluster corresponded to more than one cover type, then he identified it that way. For example, cluster 13 in candidate training area 1 is associated with both deciduous and coniferous cover types.

Another point to observe is that the list includes grassland 1, grassland 2, and 3 and 4. There were four distinctly different kinds of grassland that the analyst could distinguish in the color infrared photography: dry grass, mountain bunch grass, wet meadows, and tundra. Since they were clearly distinct in the photography, the analyst wanted to keep that information available. Grassland 1 is dry grass. Grassland 2 is mountain bunch grass. Grassland 3 is wet meadow, and grassland 4 is tundra.

Another point to observe is that cluster class numbers from different training areas do not necessarily correspond to the same information classes. For instance, cluster 12 in candidate training area 1 was deciduous, in area 2 it was coniferous, in area 3 it was grassland 3, and in area 4 it was again coniferous. This occurred because each candidate training area was clustered separately, and the results of clustering always depend on the data being clustered. In this example, each candidate training area contained different information classes.

<div style="text-align: right">EXERCISES</div>

1. State why cluster classes are associated with information classes.

<div style="text-align: right">CASE STUDY</div>

To carry out the next step in your case study analysis, assemble the following materials: the printed maps you got from clustering your candidate training areas, the 35 mm slide, the Monroe County map, and the topographic maps. Then identify as completely as possible each cluster class in every candidate

Table 5-1. Kenosha Pass cluster identification

| | Candidate Training Area | | | |
Cluster Number	1	2	3	4
1	grassland 1	grassland 2	bare soil	snow
2	grassland 2	grassland 3	grassland 2	snow and edge
3	grassland 2	grassland 2	grassland 3	grassland 4
4	grassland 2	grassland 3	grassland 3	grassland 4
5	grassland 2	grassland 3	grassland 2	bare rock
6	grassland 2	grassland 3	grassland 3	grassland 4
7	grassland 2	grassland 3	grassland 3	bare rock?
8	grassland 3	deciduous and grassland 3	grassland 3	coniferous and grassland
9	grassland 3	edge	grassland 3	?
10	coniferous	deciduous	deciduous	bare rock
11	deciduous	water +	grassland 3	coniferous and grassland
12	deciduous	coniferous	grassland 3	coniferous
13	deciduous and coniferous	coniferous	deciduous	bare rock in shadow
14	deciduous	coniferous	coniferous	coniferous
15	coniferous	coniferous and water	coniferous	coniferous

training area. Since the topographic maps are the same scale as the MSS data, you will probably want to start with them. They will be most useful in the vicinity of the reservoir. The Monroe County map shows the Highway 37 By-Pass well.*

The 35mm slide will be helpful for making more detailed identifications, such as distinguishing fields of bare soil from green vegetation in agricultural areas or shopping centers from residential neighborhoods in the urban area.

One way to use the 35mm slide is to use a regular slide pro-jector, and instead of projecting the slide onto a screen, pro-ject it onto the cluster map (taped to a wall). The first time you do this, the scale of the slide image will probably not match the scale of the cluster map. If the slide image is bigger than the corresponding part of the cluster map, move the projector closer and re-focus. This procedure is somewhat tedious, but it does work.

After you have identified each cluster class in every candi-date training area, make a table of your results. This informa-tion will be used in Section 9.

---

*If you have access to a zoom transfer scope, you could use it to match the Monroe County map and the cluster maps to the same scale.

Section 6. CALCULATION OF STATISTICAL CHARACTERISTICS OF CLUSTER CLASSES

Upon completion of this section, you should be able to do the following:

Name the two statistical parameters which define a Gaussian distribution.

Use the LARSYS processing function to obtain statistics for classes described by a set of Field Description Cards.

This section describes another step in the process of getting from candidate training areas to training samples. So far, candidate training areas have been selected to contain representative data from every information class, the candidate training areas have been clustered, and the cluster classes have been associated with specific information classes.

In this section the STATISTICS processing function will be run, and the resulting statistics file will be punched out on cards for use in the next analysis step.

Input to the STATISTICS processing function includes a deck of Field Description Cards. In this step, the input will be all of the cards punched out by CLUSTER in the step described in Section 4. From the data points specified on the Field Description Cards, the STATISTICS processing function will calculate the mean vector and covariance matrix for each cluster class. The mean vector is the average of all the data vectors in the class, and the covariance matrix is a measure of the "spread" of the data.

These two statistical parameters define a Gaussian probability density function. A Gaussian distribution is assumed in processing functions to be used later (Sections 7 and 11).

In addition to the mean vectors and covariance matrices, the STATISTICS processing function can produce several kinds of printed output. The input and output are described in Volume 2 of the

---

After clustering his candidate training areas, and associating each cluster class with a ground cover type, the analyst was ready to obtain the statistical characteristics of all the cluster classes. In order to have the statistics available for the next step, he requested a punched statistics deck.

The control card deck listed below was used.

```
-COMMENT STATISTICS OF 60 CLUSTER CLASSES
-RUNTABLE
DATA
RUN(73057902), TAPE(253), FILE(1)
END
*STATISTICS
PUNCH
CHANNELS 1,2,3,4
SCALE SPCINT (1)
DATA
  .
  .
  .
 all field description cards punched out by CLUSTER for
  cluster area 1
  .
  .
  .
 all field description cards punched out by CLUSTER for
  cluster area 2
  .
  .
  .
 all field description cards punched out by CLUSTER for
  cluster area 3
  .
  .
  .
 all field description cards punched out by CLUSTER for
  cluster area 4
  .
  .
  .
END
```

Notice that the analyst did not request histograms. When the Field Description Cards used as input for STATISTICS come from CLUSTER, individual *fields* generally have so few points that histograms of fields are not meaningful. The analyst could have requested histograms of *classes*, but they would have added considerable bulk to the printed output. Since the STATISTICS processor will be run again at a later stage in the refinement process (with a smaller number of classes), the analyst chose to wait until then to look at histograms.

Why was the SCALE control card used? One of the printed outputs from STATISTICS is a coincident spectral plot. The analyst had learned in his first few analyses of LANDSAT data

that data values range up to 128 in the first three channels, and to 64 in the fourth channel, so using an interval of 1 (rather than the default interval of three) would show more detail in the plot while still fitting on the page.

The analyst inspected the coincident spectral plot to gain insight into relationships between his various classes. He kept the Field Description Cards for use at a later stage in the analysis, and he kept the punched statistics deck for use in the next step.

EXERCISES

1. Name the two statistical parameters which define a Gaussian distribution.

2. Explain why statistics are needed at this point in the analysis.

CASE STUDY

Set up the control cards for the STATISTICS processing function. Use the punched output from CLUSTER as your input data. Request a punched statistics deck so that you will have the mean vectors and covariance matrices of each cluster available for the next analysis step.

Section 7. CALCULATION OF DISTANCES BETWEEN CLUSTER CLASSES

Upon completion of this section, you should be able to do the following:

Given two pairs of one-dimensional density functions, identify the pair which is separated by the larger statistical distance.

Name two measures of statistical distance calculated in LARSYS.

Name the two characteristics of Gaussian probability density functions which determine the statistical distance between the density functions.

Set up the control cards and run the SEPARABILITY processing function when given a punched statistics deck.

At this point in the analysis sequence, candidate training areas have been chosen and clustered, the cluster classes have been associated with information classes, and the statistical characteristics of the cluster classes have been calculated. It would be possible at this point to use all of these cluster classes to train the classifier, but that is usually not done for a couple of reasons. First, the number of clusters available at this point is normally greater than the number of classes needed to adequately train the classifier. For instance, one of the cluster classes in area 1 identified as forest may be spectrally similar to a cluster class of forest in area 3. An analyst would like to reduce the number of training classes in such cases, because a smaller number of classes saves computer time and simplifies interpretation of results. Also, some of the clusters may have too few data points to get good estimates of the mean vector and covariance matrix. By combining spectrally similar clusters, the number of data points used to calculate the mean vector and covariance matrix for a training class will be greater, and this will generally lead to a better representation of the cover type.

– 45 –

The second major reason a classification is not performed at this point is that the analyst would like to have some indication of the probability of correct classification in advance of doing the classification. If there appears to be confusion between classes, an analyst can do more clustering on the areas already used, asking for a different number of clusters, or perhaps the analyst would choose to go back and select additional candidate training areas in an effort to get good distinction between classes.

The SEPARABILITY processing function in LARSYS can help an analyst determine which cluster classes are similar, and it can serve as an indicator of probability of correct classification.

To explain how this can be accomplished, the concept of statistical distance must first be discussed. Figure 7-1 shows two cases of one-dimensional density functions. Intuitively you know that the "distance" between the density functions is greater in case b than in case a. There are a number of ways of measuring *statistical distance*. (Section 2.4 of LARS Information Note 100771, The Minimum Distance Approach to Classification, by Wacker and Landgrebe). The distance measure implemented in SEPARABILITY, transformed divergence, assumes that the density functions are Gaussian. The distance between two Gaussian probability density functions depends not only on the ordinary (Euclidean) distance between the mean values but also on the "spread" of the data. Figure 7-2 illustrates this point. The Euclidean distances between the mean values are equal in both of the cases shown, but the smaller variances (smaller "spread") in part b result in a larger statistical distance between the two density function.

At this point another distance measure, the one that is calculated in CLUSTER, will be discussed. The Swain-Fu distance is printed out in the column labelled QUOT on the page showing separability information for the clusters. Figure 7-3 shows an example of this information. The Swain-Fu distance is discussed in detail on pages 30 to 33 of Pattern Recognition: A Basis for Remote Sensing Data Analysis by Philip H. Swain. Analysts have observed that if the Swain-Fu distance for two clusters is less than .75, the two clusters can probably be combined back into a single class without creating a multimodal distribution. From this experience, an algorithm has been programmed that indicates which pairs of clusters could be combined together on the basis of having a distance less than .75 (or some user-specified threshold). An example of a grouping table output from CLUSTER is shown in Figure 7-4. A detailed explanation of the algorithm used to get the suggested groupings can be found on page CLU-20 of the LARSYS User's Manual.

Now, a natural question to ask is "Why don't we just use the Swain-Fu distance and the suggested grouping table?" There are two reasons why transformed divergence, calculated by SEPARABILITY, is used instead.
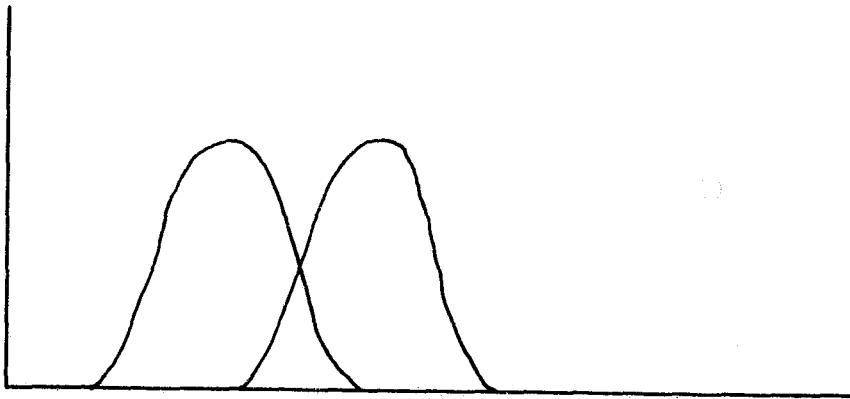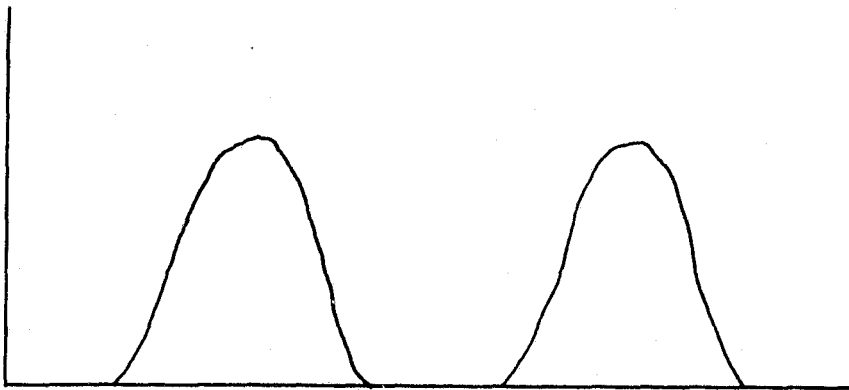
Figure 7-1a

Figure 7-1b

Figure 7-1. Two pairs of one-dimensional density functions.
The statistical distance between the density func-
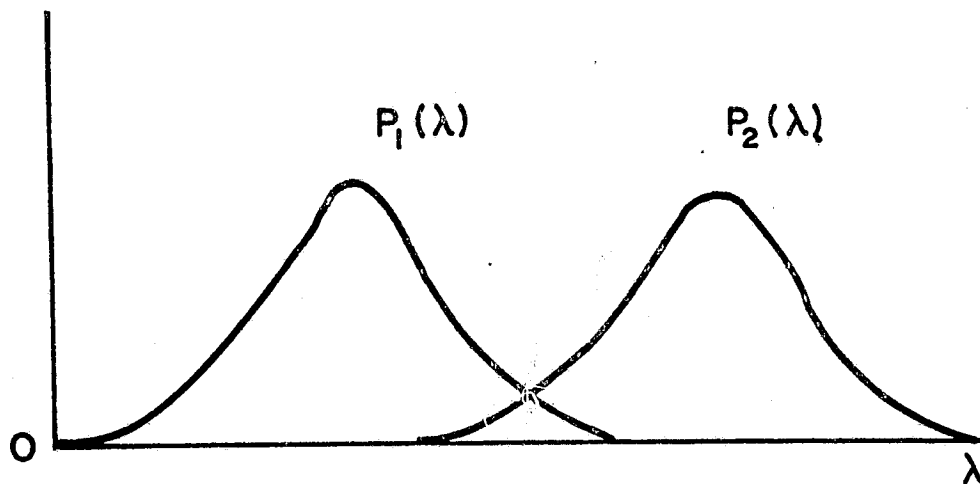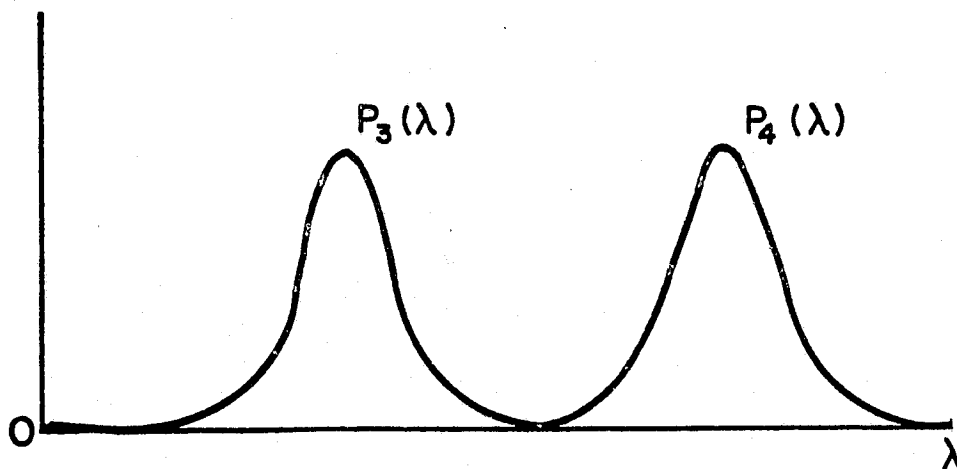tions in part b is greater than in part a.

Figure 7-2a

Figure 7-2b

Figure 7-2. Each pair of distribution functions shown above has equidistant means, but the smaller variance in $P_3(\lambda)$ and $P_4(\lambda)$ cause them to have a larger statistical distance.

```
                    SEPARABILITY INFORMATION
                    ----------------------------

    I   J    D(I,J)      D(I)      D(J)     D(I)+D(J)    QUOT

    1   2    21.467      9.426     5.219     14.645      1.466
    1   3     9.860      8.931     6.479     15.409      0.640
    1   4    19.002      9.232     5.754     14.986      1.268
    1   5    19.568      9.054     6.094     15.148      1.292
    1   6    25.497      8.797     5.934     14.730      1.731
    2   3    16.691      6.335     7.020     13.354      1.250
    2   4    11.819      7.060     5.290     12.350      0.957
    2   5    17.396      7.436     4.510     11.946      1.456
    2   6    25.314      7.246     4.351     11.597      2.183
    3   4     9.860      6.087     5.817     11.904      0.828
    3   5     9.716      6.476     6.092     12.568      0.773
    3   6    16.247      7.155     5.726     12.881      1.261
    4   5     5.621      5.423     4.473      9.895      0.568
    4   6    13.542      5.230     4.419      9.649      1.403
    5   6     8.255      4.662     4.655      9.317      0.886

    AVERAGE QUOTIENT           1.197
```

Figure 7-3. Separability information calculated in CLUSTER.

```
              RESULTS OF CLUSTER GROUPING
              ----------------------------------

                 THRESHOLD = 0.750

        GROUP      CLUSTERS      NO. PTS.

          1
                      1            67
                      3            75

          2
                      2            17

          3
                      4            70
                      5            70

          4
                      6            37
```

Figure 7-4. A CLUSTER grouping table.

First, in order to have a Swain-Fu distance calculated
between all the cluster pairs in all of the candidate training
areas, all of the areas would have to be clustered together in
one job. This would increase the computer time significantly
(a factor of 7 was observed in one example).

The second reason transformed divergence is used is that
transformed divergence seems to correlate better with probability
of correct classification than Swain-Fu distance does.

How does transformed divergence relate to probability of
correct classification? You might expect that a greater sta-
tistical distance between density functions would be accompanied
by greater classification accuracy. In general, that does
happen, although the relationship is not linear.

For a detailed discussion of the relationship between sta-
tistical distance and probability of correct classification,
see LARS Information Note 042673, Two Effective Feature Selection
Criteria for Multispectral Remote Sensing, by Swain and King.
In particular, experimental results of plotting probability of
correct classification versus transformed divergence for training
data are shown in the graph in Figure 7-5. A transformed diver-
gence value of 1.5 on that graph corresponds to a value of 1500
in SEPARABILITY output. See Figure 7-6 for an example of SEPARA-
BILITY output. Notice that the distances are given for *pairs* of
classes, and that the largest value appearing in the table is
2000.

Before proceeding to the exercises and case study you will
want to familiarize yourself with the material on the SEPARABILITY
processing function in Section 6 (Volume 2) of the LARSYS User's
Manual.
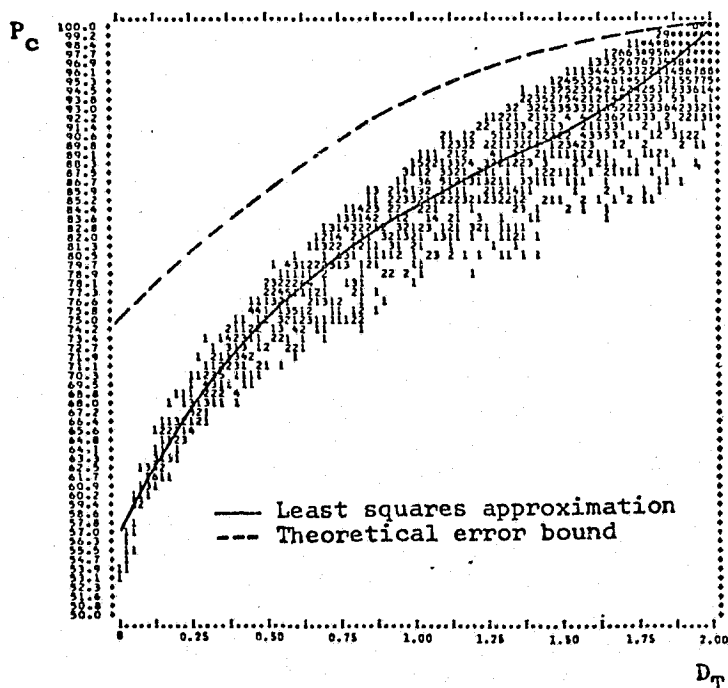
Figure 7-5. Observed
values of probability
of correct classifica-
tion versus transformed
divergence. To get
transformed divergence
values as they are printed
by SEPARABILITY, multiply
the x-axis by 1000.

D_T

CHANNELS	DIJ(MIN)	D(AVE)	WEIGHTED INTERCLASS DIVERGENCE					
			AB (10)	AC (10)	AD (10)	AE (10)	AF (10)	AG (10)
1. 1 2 3 4	1253.	1976.	1626	1632	1986	200^	1948	2000

CHANNELS	WEIGHTED INTERCLASS DIVERGENCE (DIJ)										
	AL (10)	AM (10)	AN (10)	AO (10)	AP (10)	AQ (10)	AR (10)	AS (10)	BC (10)	BD (10)	BE (10)
1. 1 2 3 4	2000	2000	2000	1994	2000	2000	2000	2000	1994	2000	2000

CHANNELS	WEIGHTED INTERCLASS DIVERGENCE (DIJ)										
	BJ (10)	BK (10)	BL (10)	BM (10)	BN (10)	BO (10)	BP (10)	BQ (10)	BR (10)	BS (10)	CD (10)
1. 1 2 3 4	2000	2000	2000	2000	2000	1899	2000	2000	200^	2000	1688.

CHANNELS	WEIGHTED INTERCLASS DIVERGENCE (DIJ)										
	CI (10)	CJ (10)	CK (10)	CL (10)	CM (10)	CN (10)	CO (10)	CP (10)	CQ (10)	CR (10)	CS (10)
1. 1 2 3 4	1998	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000

CHANNELS	WEIGHTED INTERCLASS DIVERGENCE (DIJ)										
	DI (10)	DJ (10)	DK (10)	DL (10)	DM (10)	DN (10)	DO (10)	DP (10)	DQ (10)	DR (10)	DS (10)
1. 1 2 3 4	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000

Figure 7-6.   Output from the SEPARABILITY processing function.

- 51 -

The analyst working on the Kenosha Pass area ran SEPARABILITY to determine which clusters could be combined. To do this, he used the following control cards:

```
-COMMENT SEPARABILITY FOR 60 CLUSTER CLASSES
-RUNTABLE
DATA
RUN(73057902), TAPE(253), FILE(1)
END
*SEPARABILITY
COMBINATIONS 4
SYMBOLS A,B,C,D,E,F,G,H,I,J,K,L,M,N,O,P,Q,R,S,T
SYMBOLS U,V,W,X,Y,Z,$,+,=,/,A,B,C,D,E,F,G,H,I,J
SYMBOLS K,L,M,N,O,P,Q,R,S,T,U,V,W,X,Y,Z,$,+,=,/
CARDS READSTATS
PRINT DIV(1000)
DATA
     .
     .
     .

    Punched statistics deck from previous STATISTICS
    run - 60 classes, 4 channels
     .
     .
     .

END
```

Since the analyst intends to use all four of the available channels (rather than selecting a subset of features), he used the parameter 4 on the required COMBINATIONS card. A (one-character) symbol had to be assigned to every class. The DIV(1000) on the PRINT card caused a summary listing of all class pairs whose pairwise distance was less than or equal to 1000 to be printed out. The analyst knew that this condensed listing would be useful when he constructed his "separability diagram" (Section 8), because he had chosen 1000 as his first threshold for combining classes. This part of the procedure frequently involves two or more iterations, depending on how simple or complex the analysis problem is, and the threshold may change from one iteration to the next. 1000 has been a useful starting value for combining clusters in many problems.

Notice in Figure 7-7 the legend accompanying the SEPARABILITY output. Since there were more classes than printer symbols, the analyst had to use symbols twice, requiring extra care in interpreting the results.

Figure 7-8 shows the condensed list at the end of the printout that resulted from the PRINT DIV(1000) card. The analyst annotated the list as shown in Figure 7-9 to eliminate any confusion due to duplicate symbols.

CLASSES CONSIDERED

SYMBOL   CLASS

A   NS- 1/15
B   NS- 2/15
C   NS- 3/15
D   NS- 4/15
E   NS- 5/15
F   NS- 6/15
G   NS- 7/15
H   NS- 8/15
I   NS- 9/15
J   NS-10/15
K   NS-11/15
L   NS-12/15
M   NS-13/15
N   NS-14/15
O   NS-15/15
P   NS- 1/15
Q   NS- 2/15
R   NS- 3/15
S   NS- 4/15
T   NS- 5/15
U   NS- 6/15
V   NS- 7/15
W   NS- 8/15
X   NS- 9/15
Y   NS-10/15
Z   NS-11/15
$ NS-12/15
+   NS-13/15
=   NS-14/15
/   NS-15/15
A   NS- 1/15
B   NS- 2/15
C   NS- 3/15
D   NS- 4/15
E   NS- 5/15
F   NS- 6/15
G   NS- 7/15
H   NS- 8/15
I   NS- 9/15
J   NS-10/15
K   NS-11/15
L   NS-12/15
M   NS-13/15
N   NS-14/15
O   NS-15/15
P   NS- 1/15
Q   NS- 2/15
R   NS- 3/15
S   NS- 4/15
T   NS- 5/15
U   NS- 6/15
V   NS- 7/15
W   NS- 8/15
X   NS- 9/15
Y   NS-10/15
Z   NS-11/15
$ NS-12/15
+   NS-13/15
=   NS-14/15
/   NS-15/15

Figure 7-7.   The legend accompanying SEPARABILITY output for the
              Kenosha Pass analysis.

BB	945.		PB	152.
BU	838.		QD	599.
FQ	950.		QY	413.
FD	489.		RD	996.
FY	740.		RE	814.
GR	284.		SG	754.
GE	874.		SZ	931.
HQ	932.		TH	847.
HC	394.		TX	340.
HW	887.		UI	493.
IS	671.		UJ	873.
IY	988.		VF	468.
JE	634.		V$	658.
J$	797.		YM	597.
KU	690.		Z=	421.
KX	995.		$N	271.
KJ	541.		$/	959.
KZ	544.		+/	507.
LV	343.		=O	203.
LF	462.		BU	850.
MZ	518.		DY	238.
M=	834.		F$	957.
NY	866.		F=	753.
NM	593.		HX	581.
O+	473.		N/	646.
OO	704.		O/	666.
O/	397.			

Figure 7-8. SEPARABILITY output resulting from the PRINT DIV(1000)
card.

BB	945.	B1-B2		PB	152.	P1-B2
BU	838.	B1-U2		QD	599.	Q1-D2
FQ	950.	F1-Q1		QY	413.	Q1-Y2
FD	489.	F1-D2		RD	996.	R1-D2
FY	740.	F1-Y2		RE	814.	R1-E2
GR	284.	G1-R1		SG	754.	S1-G2
GE	874.	G1-E2		SZ	931.	S1-Z2
HQ	932.	H1-Q1		TH	847.	T1-H2
HC	394.	H1-C2		TX	340.	T1-X2
HW	887.	H1-W2		UI	493.	U1-I2
IS	671.	I1-S1		UJ	873.	U1-J2
IY	988.	I1-Y2		VF	468.	V1-F2
JE	634.	J1-E2		V$	658.	V1-$2
J$	797.	J1-$2		YM	597.	Y1-M2
KU	690.	K1-U1		Z=	421.	Z1-=2
KX	995.	K1-X1		$N	271.	$1-N2
KJ	541.	K1-J2		$/	959.	$1-/2
KZ	544.	K1-Z2		+/	507.	+1-/2
LV	343.	L1-V1		=O	203.	=1-O2
LF	462.	L1-F2		BU	850.	B2-U2
MZ	518.	M1-Z1		DY	238.	D2-Y2
M=	834.	M1-=2		F$	957.	F2-$2
NY	866.	N1-Y1		F=	753.	F2-=2
NM	593.	N1-M2		HX	581.	H2-X2
O+	473.	O1-+1		N/	646.	N2-/2
OO	704.	O1-O2		O/	666.	O2-/2
O/	397.	O1-/2				

Figure 7-9. List of class pairs from Figure 7-8 annotated to
eliminate confusion due to duplicate symbols.

1. Look at the two pairs of one-dimensional density functions shown in Figure 7-10. For which pair is the statistical distance between density functions the largest?

2. Name two statistical distance measures calculated in LARSYS.

3. Name the two characteristics of Gaussian probability density functions which determine the statistical distance between the density functions.

CASE STUDY

Set up the control cards to run the SEPARABILITY processing function. Use the statistics deck you punched out in the previous step as input data. You should note that when you ran the STATISTICS processing function, you put the punched decks from CLUSTER in a certain order, and the punched STATISTICS deck kept the classes in the same order. When the punched statistics deck is used as input to SEPARABILITY, the same class order is still preserved. You may want to make use of the PRINT control card as the analyst did in the example.
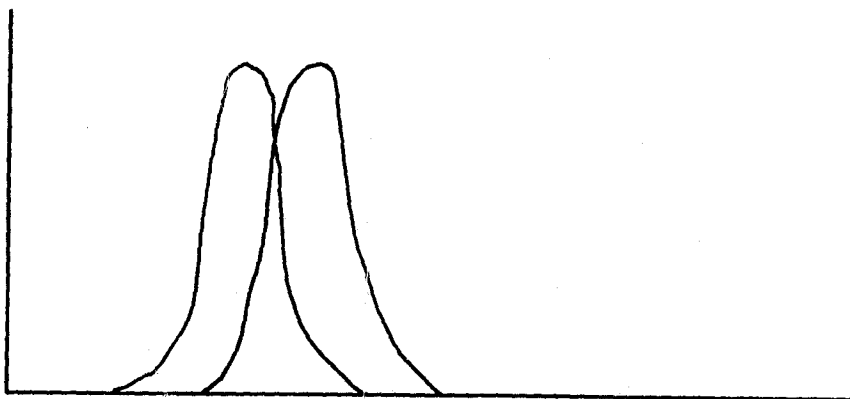
Figure 7-10a
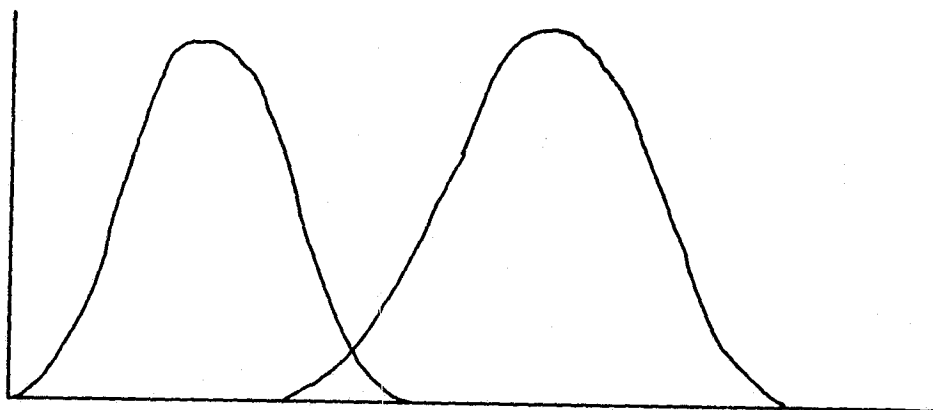
Figure 7-10b

Figure 7-10. Two pairs of one-dimensional density functions.

Section 8.   CONSTRUCTION OF SEPARABILITY DIAGRAM

Upon completion of this section, you should be able to construct a "separability diagram," when given SEPARABILITY output.

The SEPARABILITY processing function was run to determine how to combine cluster classes from different cluster areas to form training samples for information classes. In this section, a technique for graphically portraying the information from SEPARABILITY will be demonstrated. In Section 9, this separability diagram will be interpreted in conjunction with the information on cluster class identification obtained in Section 5. Review the analysis flow chart, Figure 2 in the Introduction, to see how these steps fit into the analysis sequence.

There are several ways in which the information can be diagramed. Two of the possibilities are demonstrated here.

To work with a fairly simple case first, consider this example: two areas were clustered, and in each case eight clusters were requested. Statistics were calculated, and SEPARABILITY was run with the option to print out class pairs whose interclass divergence was less than 1000. The sixteen classes were assigned the symbols A,B,C,D,E,F,G,H,1,2,3,4,5,6,7, and 8. The following list was printed out:

A1	815.
C2	823.
C3	760.
D4	70.
E6	382.
F8	732.
G7	194.

One way to approach the diagram construction would be to set

all 16 symbols down in a circle, as shown in Figure 8-1. Then
start at the top of the printed list with the pair Al and draw
a line connecting A and 1. Along the line, write the divergence
value, 815, as shown in Figure 8-2. Then proceed to draw in
lines indicating the remaining class pairs, as shown in Figure 8-3.
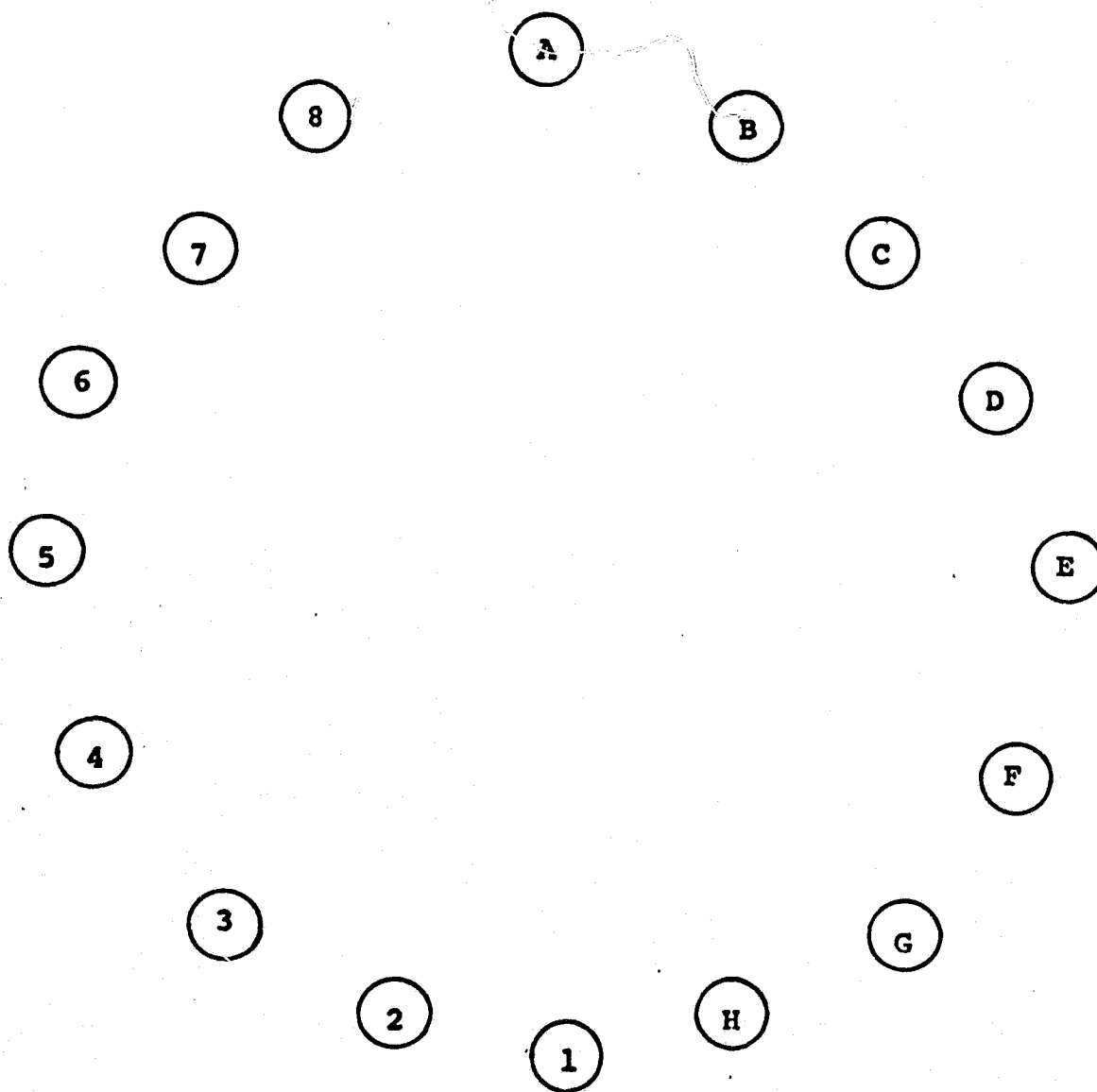
Figure 8-1.   First step in one method of constructing a
              separability diagram.
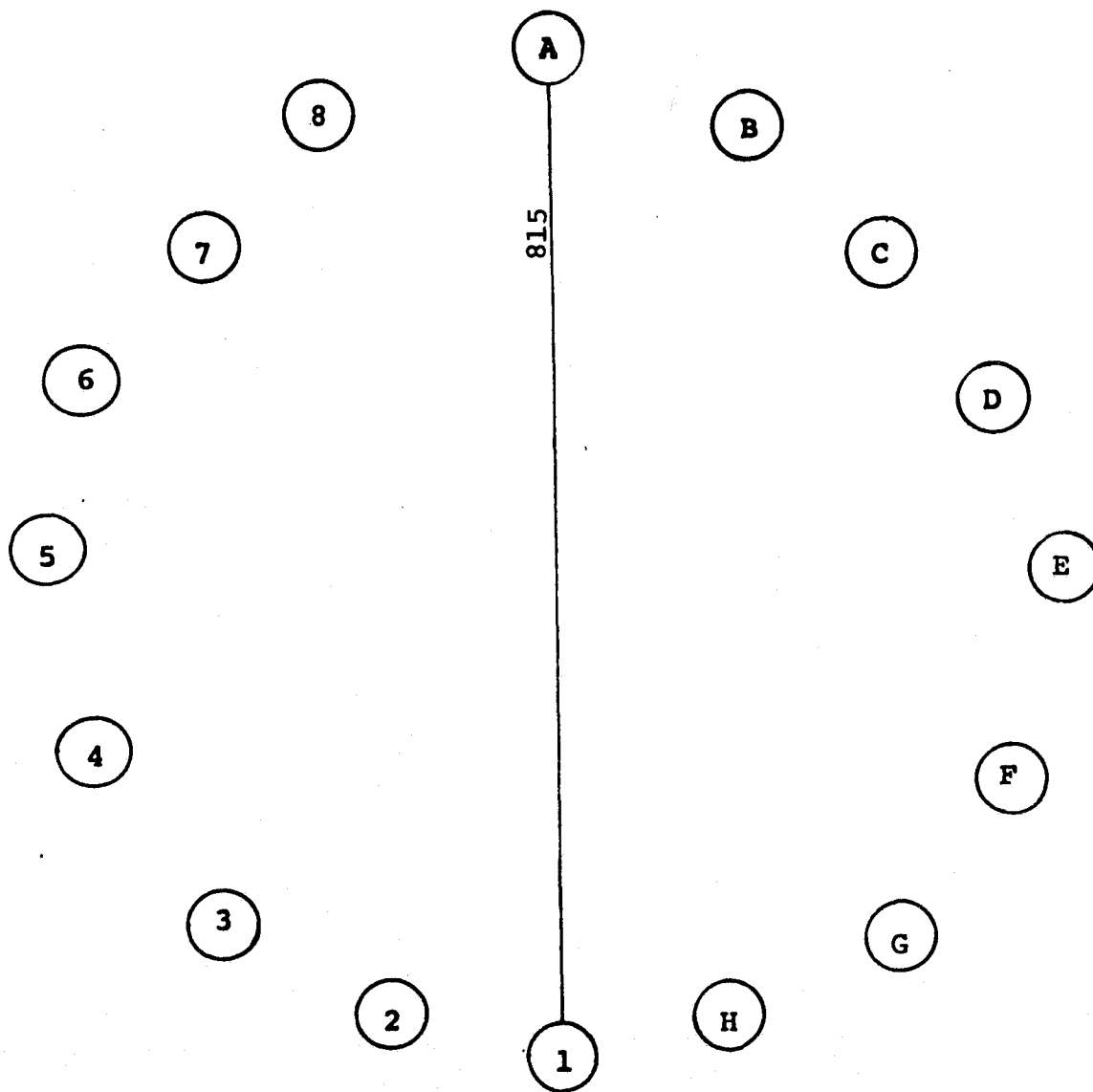
Figure 8-2. Second step in one method of constructing a separability diagram.
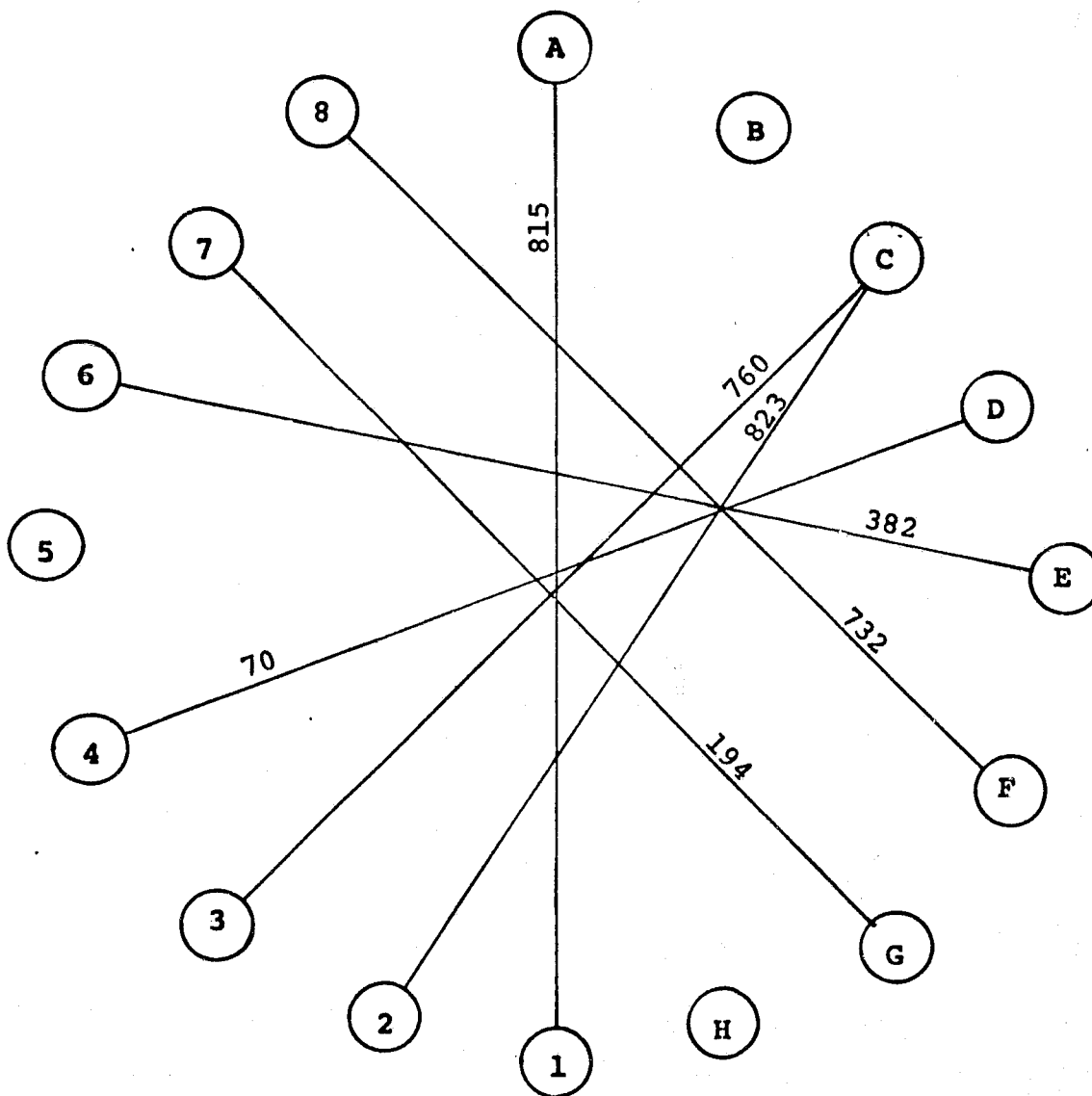
Figure 8-3.   A completed separability diagram.

An alternative approach is to start at the top of the list of pairs and draw the pair relationships. First, make a list of all of the class symbols. Then, noting that the first pair is A1, draw the A, the 1, cross them off the symbol list, and draw the connecting line, as shown in Figure 8-4. The second pair is C2, and neither C nor 2 is represented yet, so draw the C, the 2, cross them off the symbol list, and draw the connecting line as shown in Figure 8-5. The next pair is C3. A check of the symbol list shows that the C has already been drawn, so draw the 3 only, cross it off the symbol list, and connect C3, as shown in Figure 8-6.

X B C D E F G H

X 2 3 4 5 6 7 8

A —— 815 —— 1
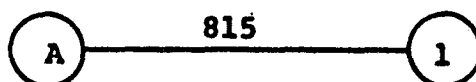
Figure 8-4.
Step one of a
separability
diagram.

X B Ȼ D E F G H

X X 3 4 5 6 7 8

A —— 815 —— 1

C —— 823 —— 2
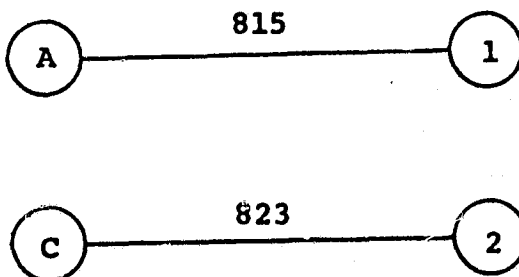
Figure 8-5.
Step two.

X B Ȼ D E F G H

X X X 4 5 6 7 8

A —— 815 —— 1

C —— 823 —— 2

760

3
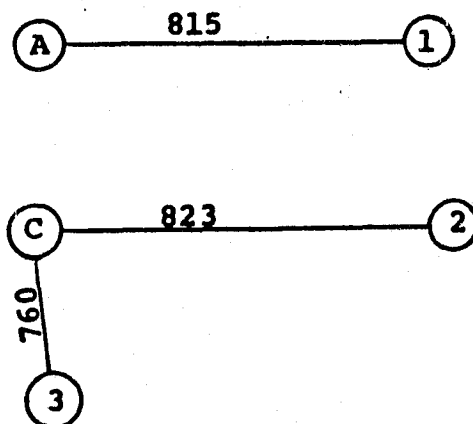
Figure 8-6.
Third step.

- 63 -

When the list of class pairs has been exhausted, the diagram shown in Figure 8-7 results. Note that the symbol list shows that 3 classes (B, H, and 5) had a statistical distance greater than 1000 from every other class.

Figures 8-3 and 8-7 contain the same information, and you could come up with still other diagrams equivalent to these two. When carrying out this step of the analysis you will have to choose some way of representing the separability information so that the class relationships are apparent to you.

A B C D E F G H

1 2 3 4 5 6 7 8
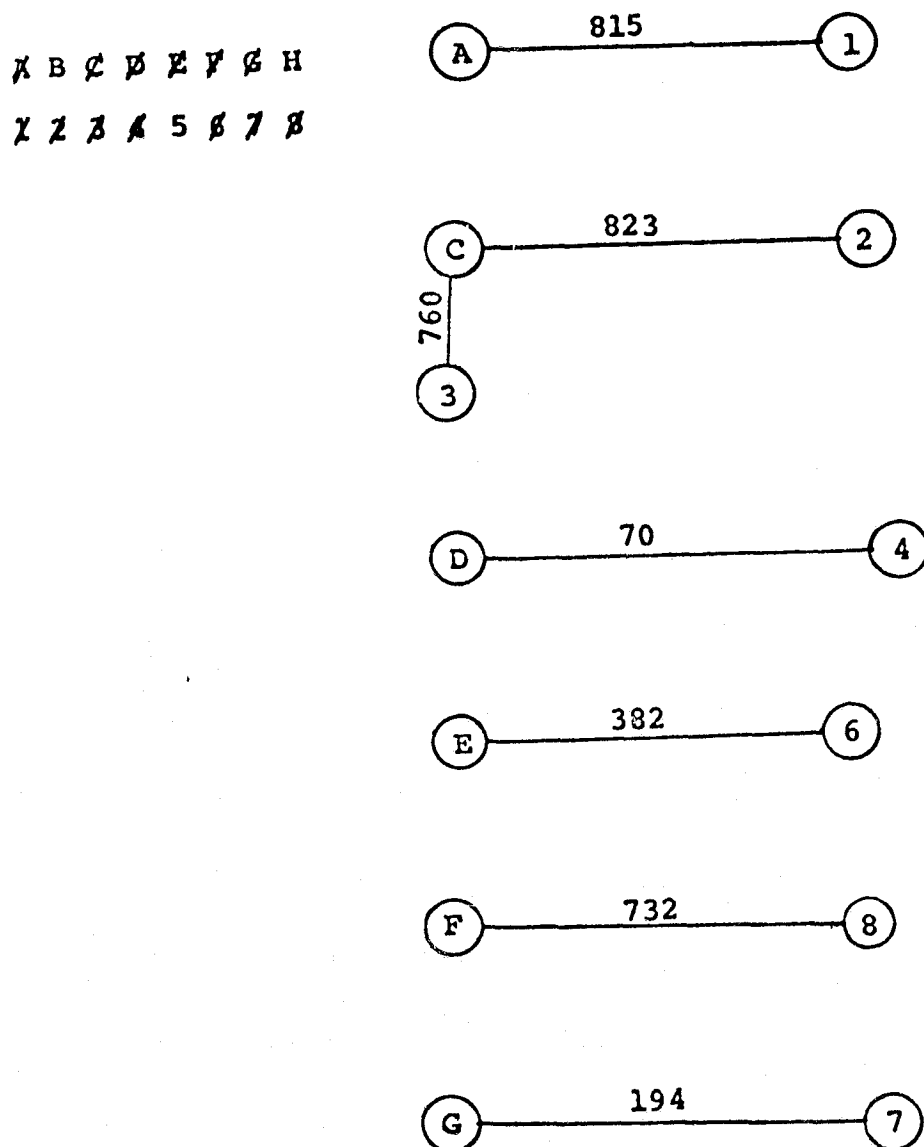
Figure 8-7. A completed separability diagram.

In the next section, you will use the separability diagram constructed here along with the cluster class identification you found in Section 5 to select training samples for the information classes you wish to classify.

The analyst working on the Kenosha Pass area used his list of class pairs having interclass divergence values less than or equal to 1000 to construct a separability diagram. He had already annotated the list, as discussed in the last section, to prevent confusion about symbols. The annotated list is reproduced as Figure 8-8 for your convenience.

If the analyst had chosen to diagram the information in a circle, Figure 8-9 would have resulted. However, he chose instead to construct the diagram shown in Figure 8-10. (Don't be surprised if your first few diagrams do not look so orderly -- you can redraw them if they look too confusing.)

CASE STUDY

Using the SEPARABILITY output you generated in the last section, construct a separability diagram. Discuss the results of this operation with your instructor.

BB	945.	B1-B2		PB	152.	P1-B2
BU	838.	B1-U2		QD	599.	Q1-D2
FQ	950.	F1-Q1		QY	413.	Q1-Y2
FD	489.	F1-D2		RD	996.	R1-D2
FY	740.	F1-Y2		RE	814.	R1-E2
GR	284.	G1-R1		SG	754.	S1-G2
GE	874.	G1-E2		SZ	931.	S1-Z2
HQ	932.	H1-Q1		TH	847.	T1-H2
HC	394.	H1-C2		TX	340.	T1-X2
HW	887.	H1-W2		UI	493.	U1-I2
IS	671.	I1-S1		UJ	873.	U1-J2
IY	988.	I1-Y2		VF	468.	V1-F2
JE	634.	J1-E2		V$	658.	V1-$2
J$	797.	J1-$2		YM	597.	Y1-M2
KU	690.	K1-U1		Z=	421.	Z1-=2
KX	995.	K1-X1		$N	271.	$1-N2
KJ	541.	K1-J2		$/	959.	$1-/2
KZ	544.	K1-Z2		+/	507.	+1-/2
LV	343.	L1-V1		=O	203.	=1-O2
LF	462.	L1-F2		BU	850.	B2-U2
MZ	518.	M1-Z1		DY	238.	D2-Y2
M=	834.	M1-=2		F$	957.	F2-$2
NY	866.	N1-Y1		F=	753.	F2-=2
NM	593.	N1-M2		HX	581.	H2-X2
O+	473.	O1-+1		N/	646.	N2-/2
OO	704.	O1-O2		O/	666.	O2-/2
O/	397.	O1-/2				

Figure 8-8
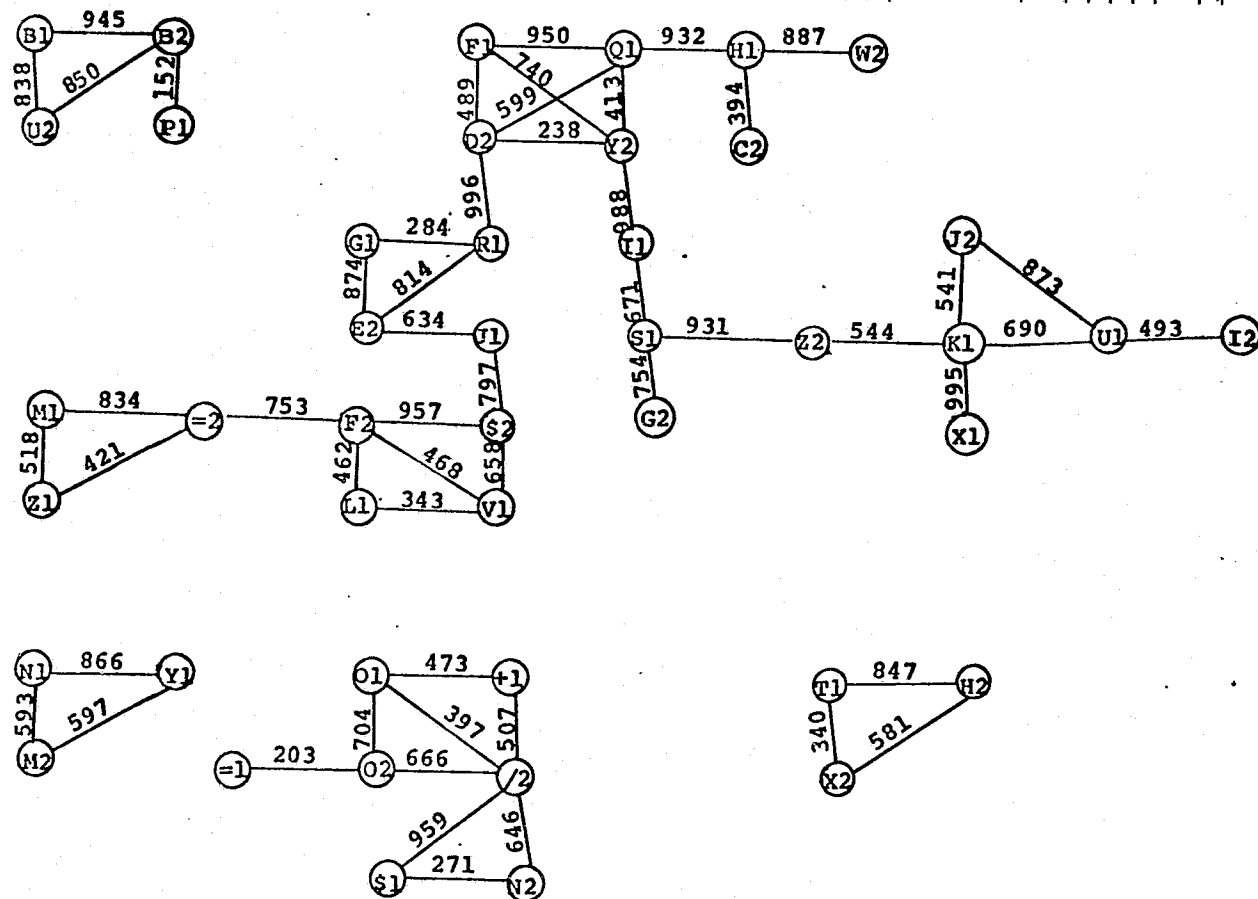
Figure 8-9

Figure 8-10. Separability diagram for Kenosha Pass example.

Section 9. SELECTION OF TRAINING CLASSES

Upon completion of this section, you should be able to select training classes for use in classification, given a separability diagram and a list of cluster-class/information-class associations (and an analysis objective).

This section describes the last step in the process of refining the candidate training areas to obtain training samples. So far, refinement of candidate training areas has included clustering the areas, associating cluster classes with information classes (cover types), calculating statistics of the clusters, running SEPARABILITY to get a measure of the distance between clusters, and constructing a separability diagram. In this section, training classes will be selected from the cluster classes. To select training classes, the separability information diagramed in Section 8 and the cluster-class/information-class associations determined in Section 5 will be used.

First, analysis objectives should be reviewed, to bring clearly to mind the cover types of interest.

There are a number of ways the problem of selecting training classes from cluster classes could be approached. A few possibilities will be suggested in general terms in the discussion, and then one of the possibilities will be pursued in detail in the example.

One possibility would be to begin by transferring the cluster class identification information onto the separability diagram. Then, group together cluster classes from the same cover type whose interclass statistical distances are less than the chosen threshold (e.g., a transformed divergence < 1000). When cluster classes from the same cover type have a pairwise statistical distance greater than the chosen threshold, the conclusion would be that the cluster classes are spectrally distinct subclasses of that cover type. When cluster classes from DIFFERENT cover types are spectrally similar, as shown by a small statistical distance, there is a problem. Possible courses of action include the following: the classes could be clustered again to refine them further; the identification of cluster classes could be

checked to verify that the cluster classes are indeed from different cover types; the confusion between classes could be accepted and no action taken. Another possible course of action is more easily described by an example: assume that an analyst is interested in classifying an urban area, and has discovered that a cluster class identified as urban is similar to a cluster class identified as agriculture. He could decide that for his purposes, the error of classifying some agriculture data points as urban would not be too troublesome, while the error of classifying some urban points into agriculture would be disastrous. In that case, the analyst could choose to eliminate the cluster class identified as agriculture from any subsequent processing.

Another possible way to select training classes would again begin with the transfer of cluster class identification information onto the separability diagram. Then, when cluster classes from the same cover type have interclass statistical distances less than the chosen threshold, just one of those clusters would be chosen to represent that cover type, and the rest would not be used. The philosophy behind this approach is that the single cluster class will have a smaller variance than would a group of cluster classes, and therefore the single class would be less likely to be confused with other cover type classes. However, a criticism of this approach is that the number of data points is not as large as it would be if all clusters were used, and in general a larger number of data points in a training class is more representative.

Again if cluster classes from DIFFERENT cover types have a small interclass distance, the same problem arises as before, and one of the previously mentioned ways of dealing with the situation can be chosen.

A third possible approach to the selection of training samples would begin with the separability diagram. Tentative class groupings would be determined on the basis of the statistical distances without reference to the cover type identifications of the clusters. After cluster groupings have been determined on the basis of statistical distances, then the cluster class identification information is transferred onto the diagram. Then the groupings are inspected to see if the spectrally similar clusters are from the same cover type. The difference between this approach and the first one is rather subtle and can best be understood by trying both ways.

One point which should be apparent by now is that there is no single correct way to progress through an analysis sequence. As you increase your understanding of the pattern recognition concepts used in LARSYS and gain experience in analysis, you may even develop new procedures yourself.

This section has described ways in which training samples can be selected from the cluster classes generated earlier. The techniques described make use of the separability diagram constructed in the last section and the cluster class-information class associations determined earlier.

The techniques described here are still subject to a great deal of investigation. The construction of the separability diagram can be expressed in terms of a programmable algorithm, and the numerical criteria for grouping clusters together could be programmed. Experimental work along these lines is in progress at this writing.

<div align="right">EXAMPLE</div>

In the Kenosha Pass example, remember that the analyst constructed the separability diagram shown in Figure 9-1. We will now go through a step-by-step interpretation of that diagram to convey the kinds of thinking an analyst does.
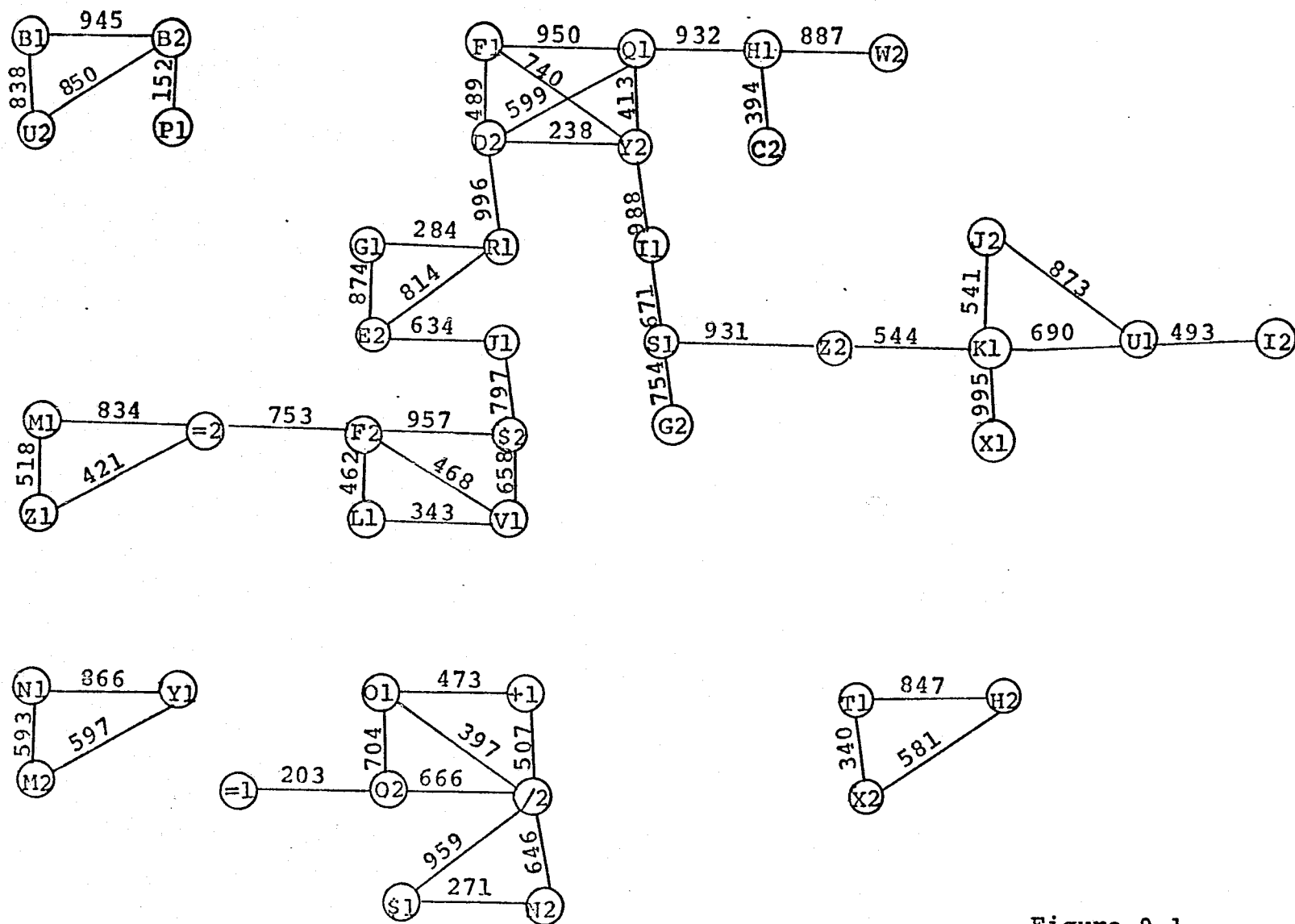
Figure 9-1

The analyst who did this work chose to begin by looking at the statistical distances first without reference to the cover type identification of the various clusters. Two of the cluster groups were straightforward. Look in the lower left-hand corner at the N1-Y1-M2 group. It was not connected to any other group and every pair of clusters in the group has transformed divergence less than 1000. Similarly, the T1, X2 and H2 clusters were close to each other while being separate from any other clusters. The analyst then interpreted those by circling those two groups as shown in Figure 9-2.

B1 —945— B2
838 850 152
U2 P1

F1 —950— Q1 —932— H1 —887— W2
740
489 599 413 394
D2 —238— Y2 C2

996
G1 —284— R1
874 814 988
E2 —634— J1
671
797
S1 —931— Z2 —544— X1 —690— U1 —493— I2
754 541
M1 —834— =2 —753— F2 —957— S2 995
518 421 462 658 J2 —273—
Z1 L1 —343— V1 X1
468
G2

N1 —866— Y1
593 597
M2

=1 —203— O2 —666— /2
O1 —473— +1
704 397 507
/2
959 646
S1 —271— N2

T1 —847— H2
340 581
X2
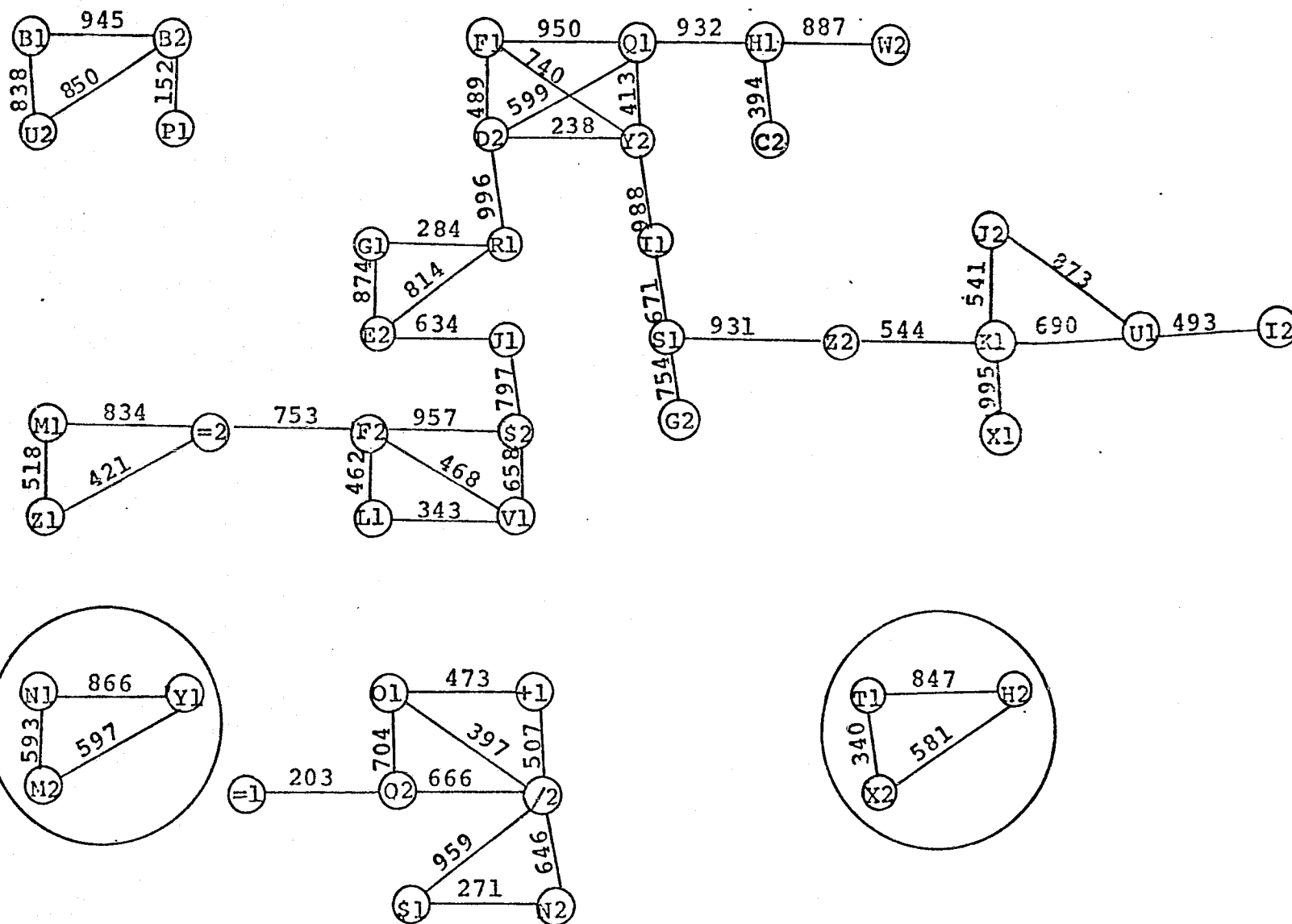
Figure 9-2

Then what? Look at the B1-B2-U2-P1 group in the upper left-hand corner. Since the transformed divergence values between B1 and P1 and between U2 and P1 were greater than   1000, how did this group get handled? The analyst who did this work set a second criterion for cases such as this. This criterion will be expressed in terms of an example. Suppose you have three clusters a, b, and c. The a-b distance is less than 1000 and the a-c distance is less than 1000, but the b-c distance is greater than 1000. If this b-c distance is less than 1500 go ahead and group all three together. If the b-c distance is greater than 1500 only group together a-b or a-c, whichever pair has the smaller transformed divergence.

Now to return to the B1-B2-P1-U2 group, the analyst went back to his SEPARABILITY output to find the transformed divergences for B1-P1 and U2-P1. The B1-P1 distance was 1312 and the U2-P1 distance was 1117. Both of these were less than 1500 so the analyst grouped all four clusters together and his separability diagram looked like Figure 9-3.
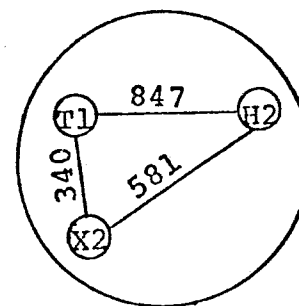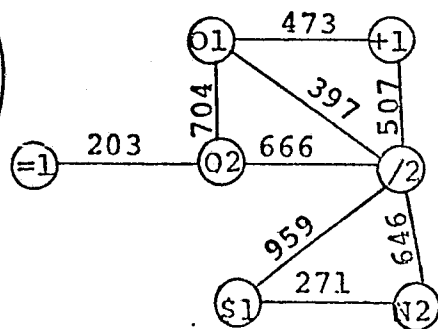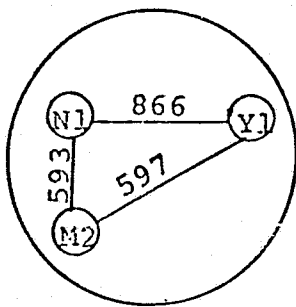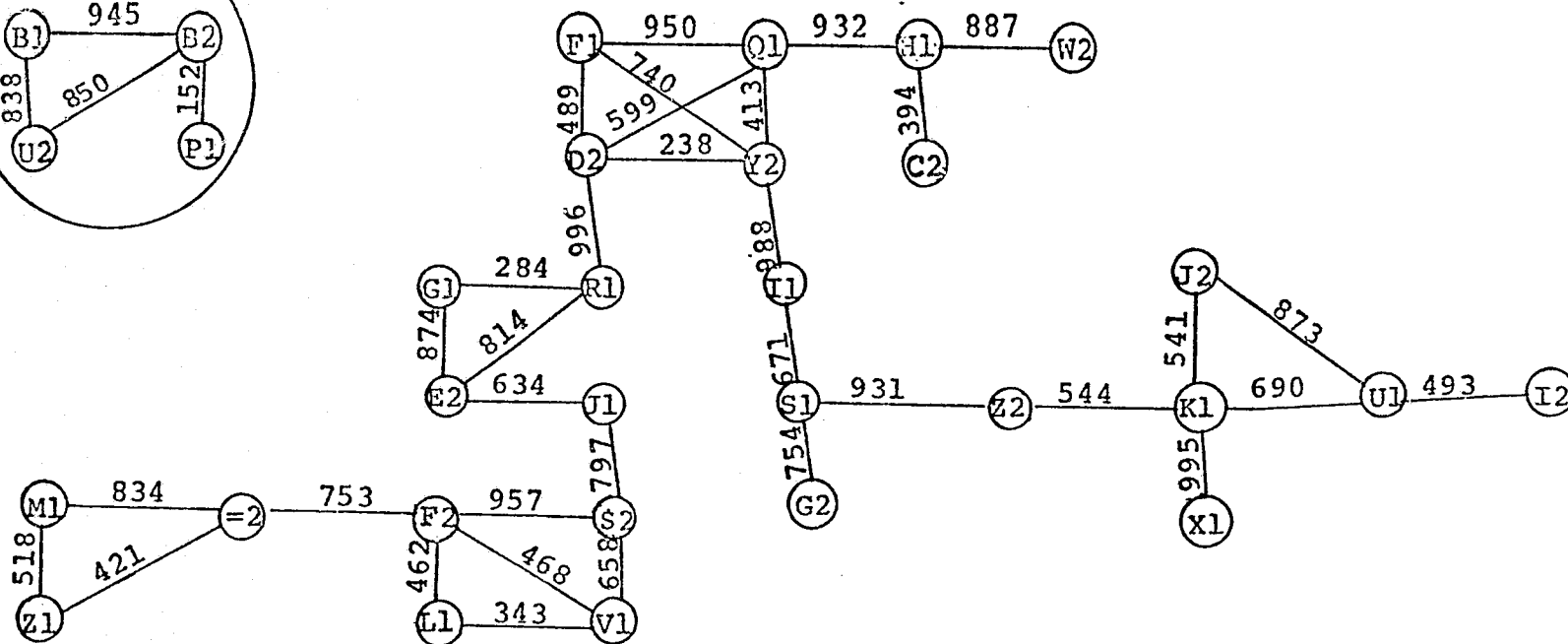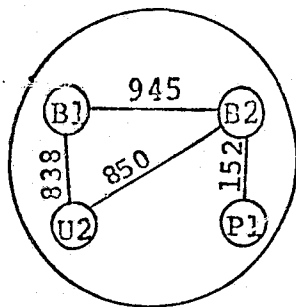
Figure 9-3

Next go the the bottom of the page,the 01,+1,=1,02,/2,$1 and
N2 group. In order to decide how to handle this the analyst again
had to go back to his SEPARABILITY output to find the transformed
divergences associated with some of the other class pairs. To help
in following the discussion, draw in the indicated distances with
dashed lines on Figure 9-3. The 01-=1 distance was 1170. The 02-+1
distance was 1072. The 02-N2 distance was 1579. The 02-$1 distance
was 1747. The =1-$1 distance was 1850. Since the =1-$1 distance,
02-$1 distance and the 02-N2 distance were all greater than 1500 the
analyst decided that the clusters represented by $1 and N2 belonged
together in a group separate from the other clusters. For thorough-
ness the analyst also inspected the =1-+1 distance, 1470, and the
=1-/2 distance, 1131. Both were less than 1500 so the 01,+1,=1,02,/2
clusters were all grouped together. At this point the separability
diagram looked like Figure 9-4.

Figure 9-4
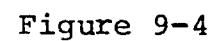
Now how did he interpret this string that is left? Start at
the left end with the M1-Z1-=2 group. Should it be combined with
other clusters or not? Going back to his **SEPARABILITY** output,
the analyst found that the M1-F2 distance, 1073, and the Z1-F2 dis-
tance, 1266, were both less than 1500. However, the M1-L1 distance,
1538, and the Z1-L1 distance, 1731, both exceeded 1500 so these
three clusters became a separate group. The separability diagram
looked like Figure 9-5.

B1 — 945 — B2
838
850
152
U2
P1

F1 — 950 — Q1 — 932 — H1 — 887 — W2
489
740
599
413
394
D2 — 238 — Y2
C2
996
988
G1 — 284 — R1
874
814
J2
873
E2 — 634 — J1
541
T1
671
797
S1 — 931 — Z2 — 544 — K1 — 690 — U1 — 493 — I2
995
M1 — 834 — =2 — 753 — F2 — 957 — S2
518
421
462
468
658
690
X1
Z1
754
G2
L1 — 343 — V1

N1 — 866 — Y1
593
597
M2

O1 — 473 — +1
704
397
507
=1 — 203 — O2 — 666 — /2
959
646
271
S1 — 271 — I2

T1 — 847 — H2
340
581
X2
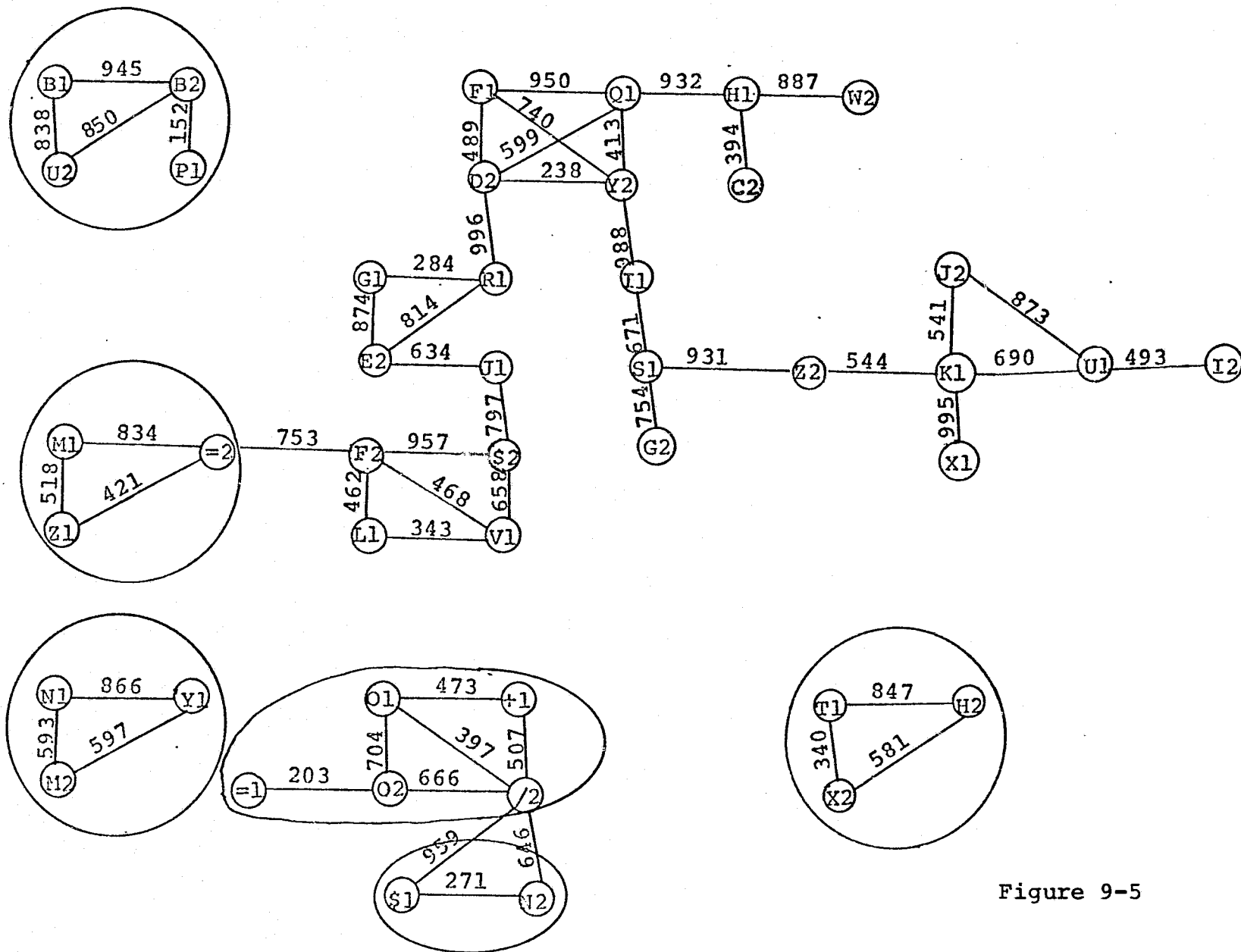
Figure 9-5

To proceed with the interpretation, the F2-L1-V1-$2 group is next. Again the analyst had to go back to his SEPARABILITY output to find out some more distances. The L1-$2 distance was 1197, the V1-J1 distance was 1611, and the F2-J1 distance was 1683. The V1-J1 distance and the F2-J1 distance were both greater than 1500 so the F2-$2-L1-V1 group was separated from the rest of the string. However the analyst wasn't yet ready to circle this group and finish with it. He looked at the mean and standard deviation of the clusters as plotted on the coincident spectral plot which he got as part of his STATISTICS output. He decided that the cluster represented by $2 would cause confusion so he chose to delete that cluster. His separability diagram then looked like Figure 9-6.
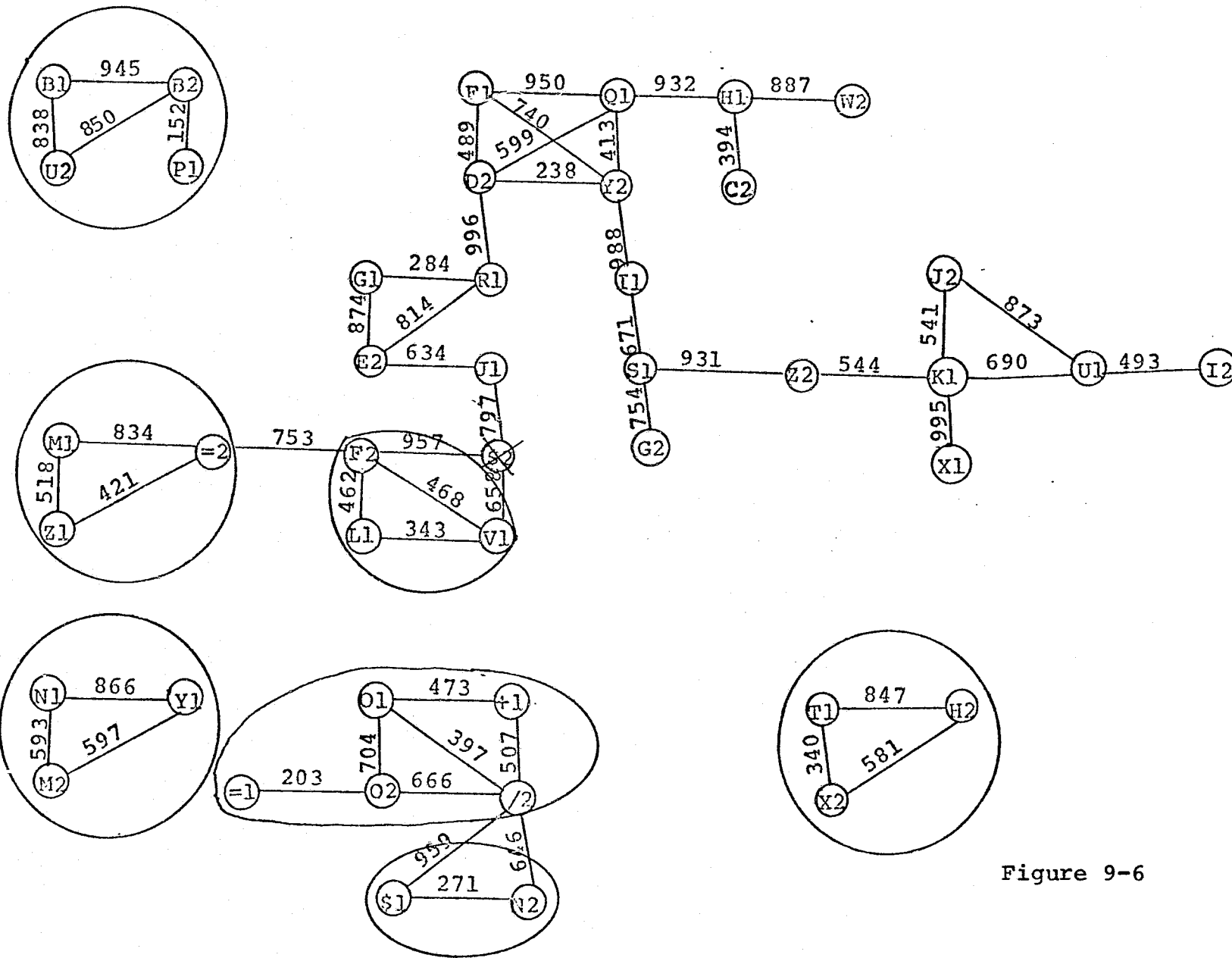
Figure 9-6

The next part of this string is the J1-E2-G1-R1 group. Other distances of interest include the G1-J1 distance, 1703, the R1-J1 distance, 1464, and G1-D2, 1612. Since the G1-D2 distance and the G1-J1 distance were both greater than 1500, the analyst grouped E2 and J1 together in one group, and made a separate group of G1 and R1. The separability diagram at this point looked like Figure 9-7.

ABCDEF¢G₩‡JKL₩NΦPФRS₮ФѴW¥¥₹$‡‡/AB¢ÐE₮¢G₩‡JKL₩NΦPQRST₩V₩¥¥₹$‡‡Ӿ
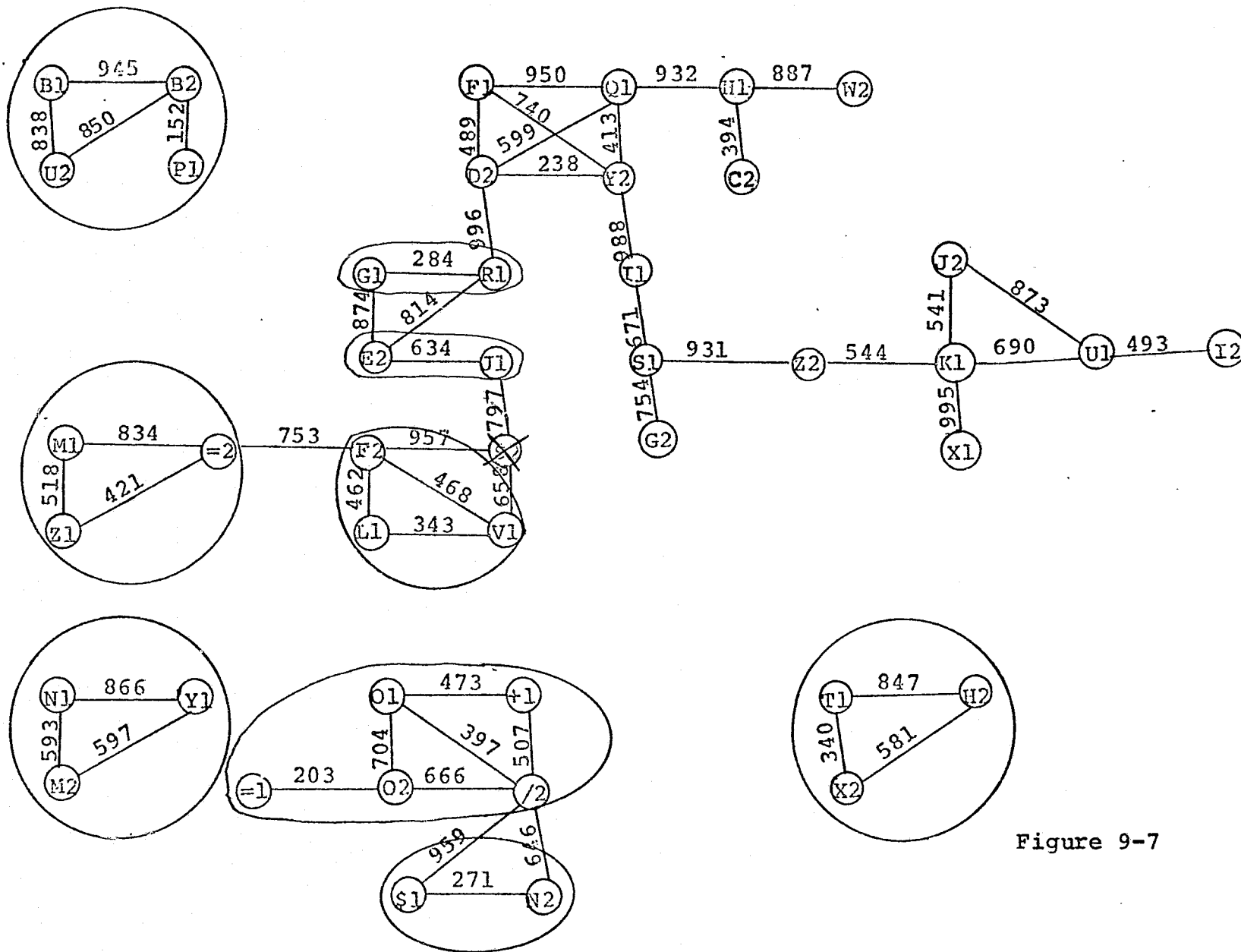
Figure 9-7

Next, move to the right-hand side of the diagram, the part that
has the I2, U1, J2, K1, and X1 clusters. The analyst again went back
to his SEPARABILITY output to find out some more distances. The J2-I2
distance was 1675, the X1-U1 distance was 1797, and the X1-I2 distance
was 1970. Because the J2-I2, the X1-I2, and the X1-U1 distances were
all greater than 1500, only U1 and I2, the pair with the smallest trans-
formed divergence, got grouped together. The diagram then looked like
Figure 9-8.

ABCDEFGHIJKLMNOPQRSTUVWXYZ$+‡/ABCDEFGHIJKLMNOPQRSTUVWXYZ$+‡/
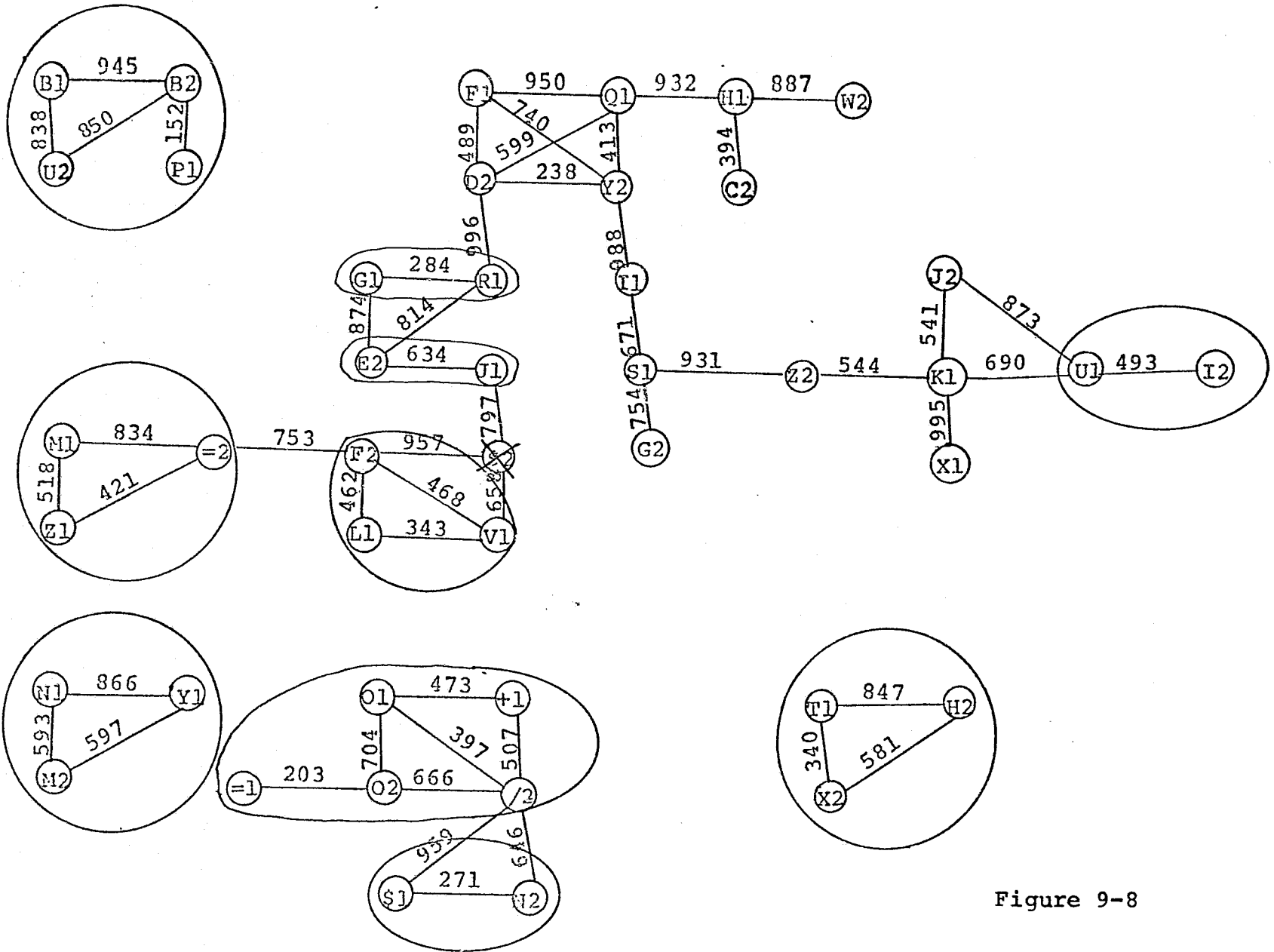
Figure 9-8

Moving to the left of the Ul-I2 group, what about the J2, Kl, Xl clusters? Well, the J2-Xl distance was 1404 so those three clusters at least could be put together. But how did Z2 fit in? Well the Z2-J2 distance was 1228, the Z2-Xl distance was 1079. The Z2-Il distance was 1586 and the Z2-G2 distance was 1119. Z2 was closer to the J2-Kl-Xl group than to the other clusters and all pairwise distances for Z2, J2, Kl and Xl were less than 1500 so the analyst grouped those four clusters together. The separability diagram in Figure 9-9 shows the interpretation to this point.
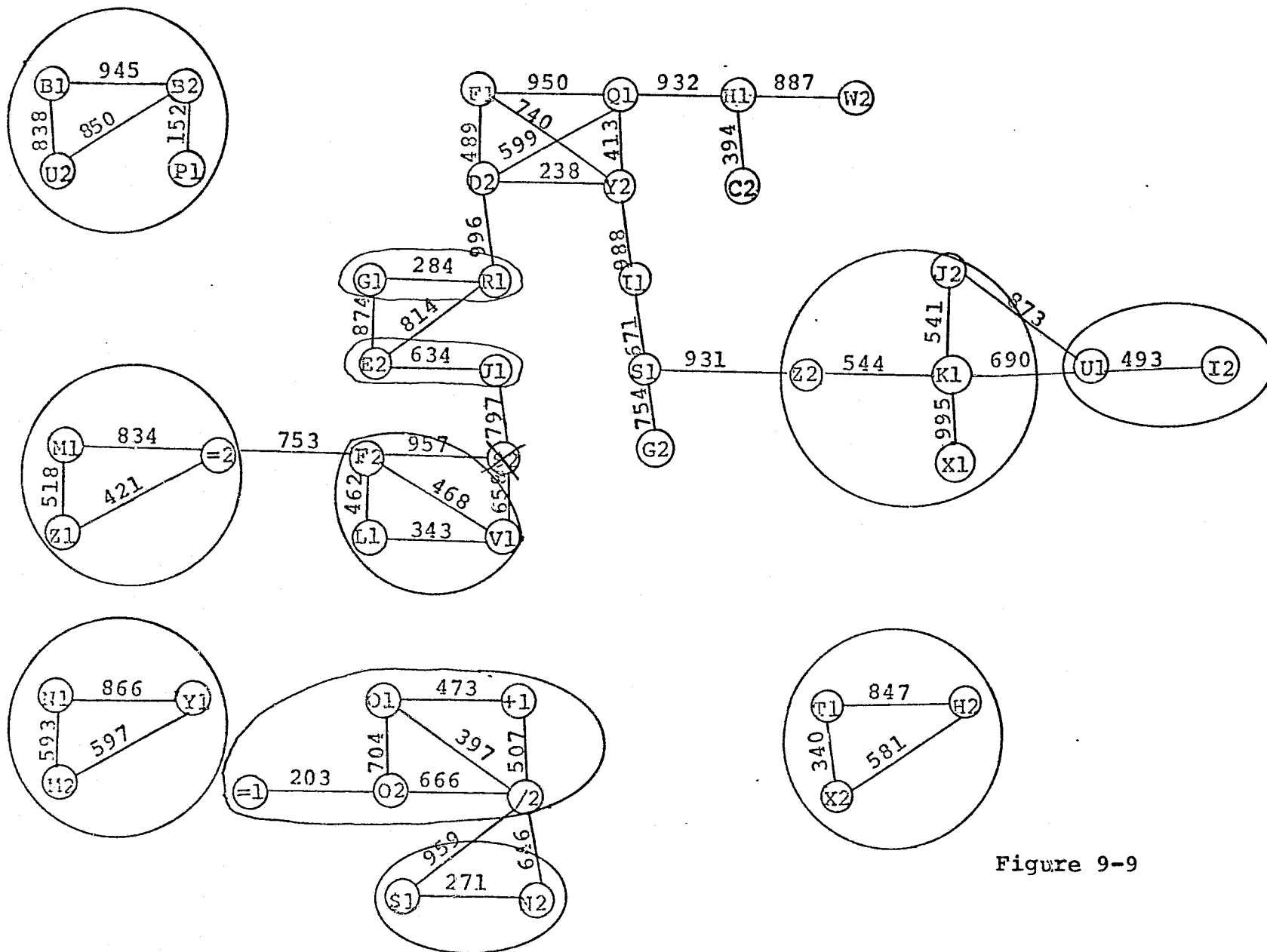
Figure 9-9

Next look at the G2-Il distance. The analyst found on his
SEPARABILITY output that that was 1121. Since it was less than
1500, G2, Sl and Il all went into one group. What about Y2?
How did it fit with this group? The Y2-Sl distance, 1578, was
greater than 1500 so Y2 did not go with this group. The separa-
bility diagram looked like Figure 9-10.

Figure 9-10

945
838
850
152
B1 B2 U2 P1

950   932   887
F1 Q1 H1 W2
489   740   599   413   238   394
D2 Y2 C2
896   288   671   754
R1 284 G1
874   814
E2 634 J1
797
I1 S1 G2
931
J2 873
541
Z2 544 K1 690 U1 493 I2
995
X1

834   753   957
M1 =2 F2
518   421   462   468   650
Z1 L1 343 V1

866   473
N1 Y1 O1 +1
593   597   704   397   507
M2 =1 203 O2 666 /2
959   646
S1 271 12

847
T1 H2
340   581
X2
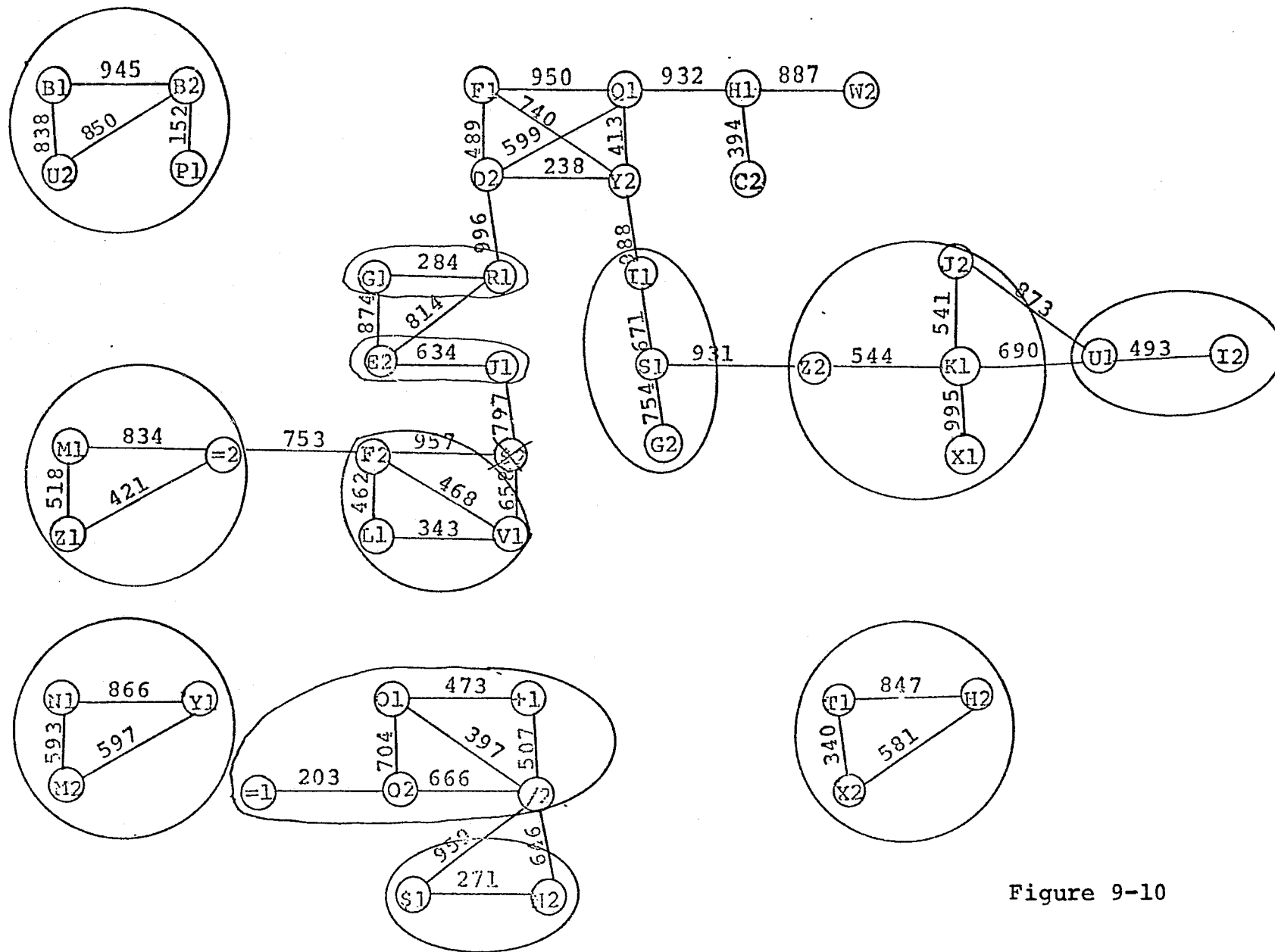
Now did W2 go with C2 and H1? The C2-W2 distance was 1230, but the analyst looked at the means and standard deviations of the clusters and decided that the W2 cluster would cause too much confusion so the W2 cluster was deleted. Now, did H1 and C2 go with the F1-D2-Q1-Y2 group? The H1-Y2 distance, 1459, and the Q1-C2 distance, 1031, were less than 1500. The Y2-C2 distance was 1500. The F1-H1 distance, 1860, and the D2-C2 distance, 1745, were so large as to prevent C2 and H1 from being included in that group. Therefore H1 and C2 became a separate cluster group and the diagram looked like Figure 9-11.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z $ + / A B C D E F G H I J K L M N O P Q R S T U V W X Y Z $ + /
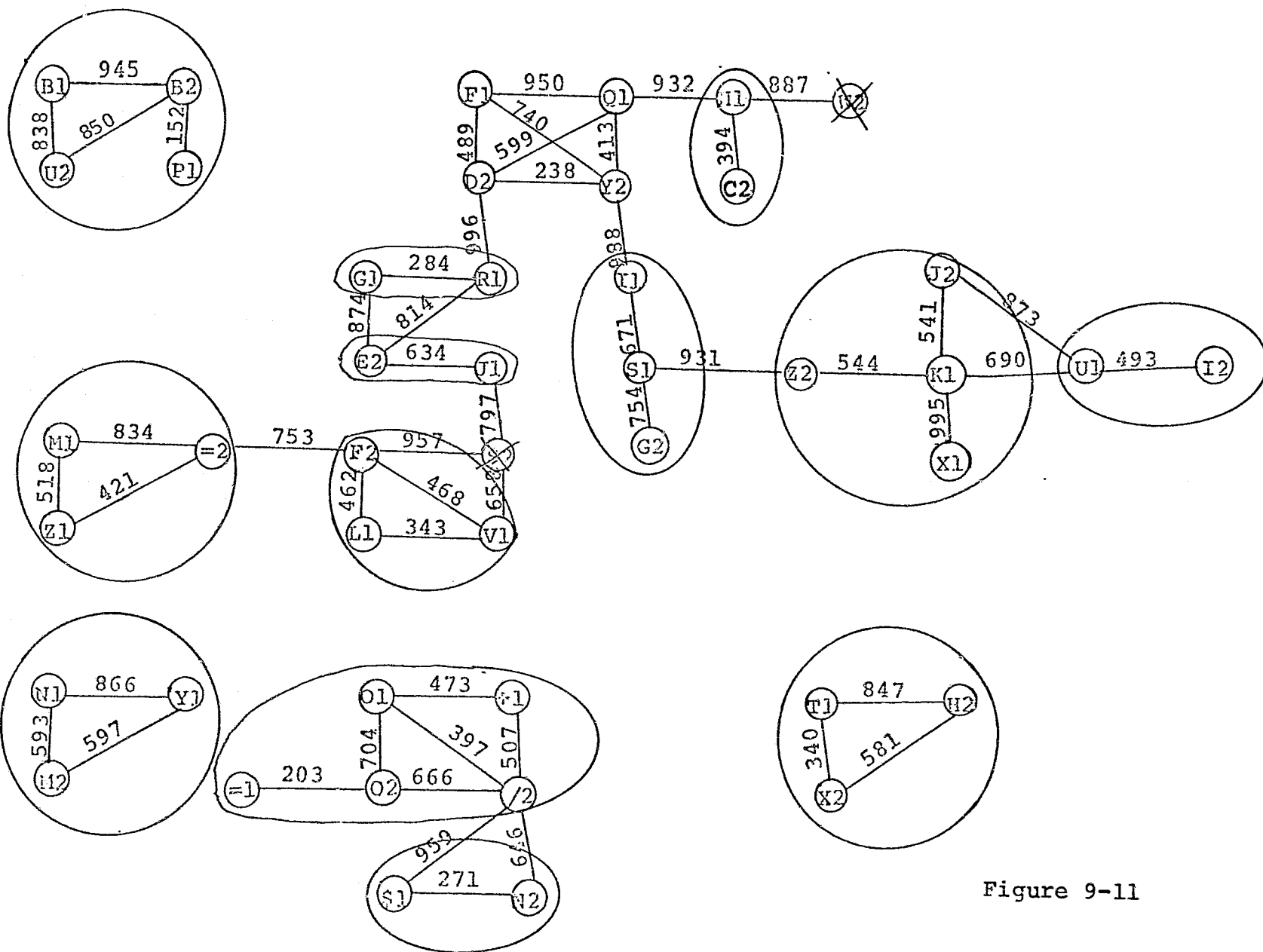
Figure 9-11

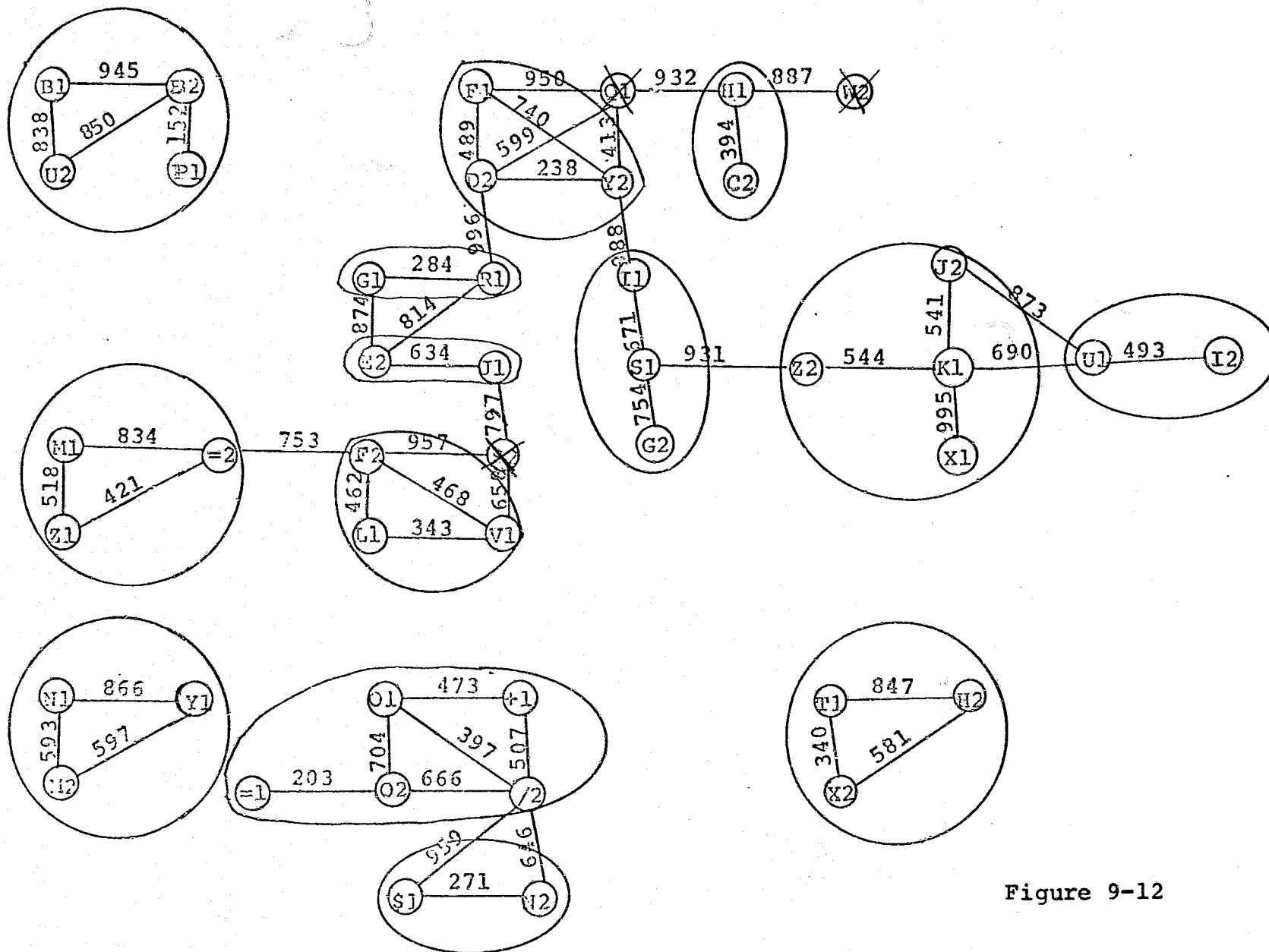Before circling the F1, Q1, D2, Y2 clusters and calling them a cluster group the analyst reviewed the cluster means and standard deviations and found that the Q1 cluster was likely to cause confusion with the H1-C2 group, so that cluster was also deleted. Figure 9-12 was the result of the interpretation to this point.

Figure 9-12

The line across the top of Figure 9-12 shows that several symbols were not crossed off, indicating that the clusters they represented were still classes by themselves. The rest of Figure 9-12 shows how the analyst had decided to combine clusters. This iteration reduced the number of classes from 60 to 30, but 30 is still a large number of classes for an analysis with an objective of differentiating five cover types.

The analyst combined clusters, assigned new symbols to all classes, and ran STATISTICS and SEPARABILITY over again.

Table 9-1 shows how the 60 classes were combined into 30, and the new symbols corresponding to the 30 classes are shown.

Table 9-1.

symbols in first iteration	new symbols	symbols in first iteration	new symbols
A1	A	B1,B2,P1,U2	Q
C1	B	F1,D2,Y2	R
D1	C	H1,C2	S
E1	D	G1,R1	T
W1	E	J1,E2	U
/1	F	I1,S1,G2	V
A2	G	K1,X1,J2,Z2	W
K2	H	U1,I2	X
L2	I	M1,Z1,=2	Y
P2	J	L1,V1,F2	Z
Q2	K	N1,Y1,M2	$
R2	L	O1,+1,=1,O2,/2	+
S2	M	T1,H2,X2	=
T2	N	$1,N2	/
V2	O		
+2	P		

Figure 9-13 shows the information from Table 9-1 put on the separability diagram. The analyst combined the Field Description Cards from CLUSTER in the manner indicated, and added to those cards a set of Field Description Cards describing clouds, cloud shadows, and water, chosen from outside the Kenosha Pass area boundaries. Then he took all these Field Description Cards, ran STATISTICS again, ran SEPARABILITY again, and constructed another separability diagram, the one shown in Figure 9-14.
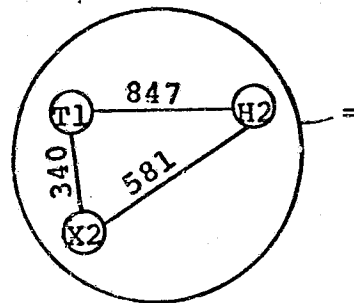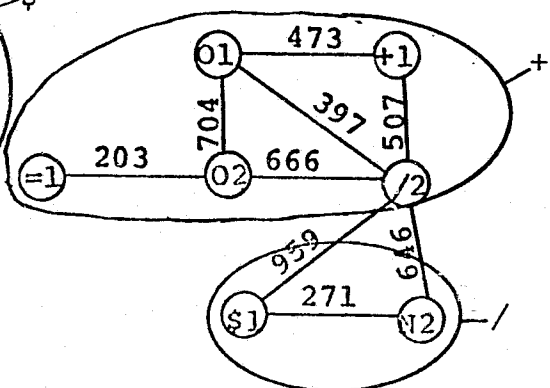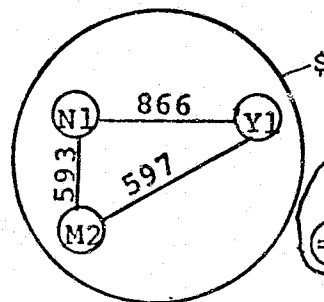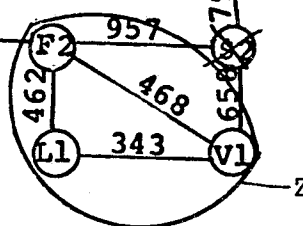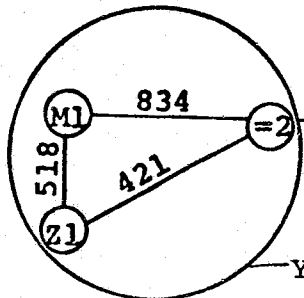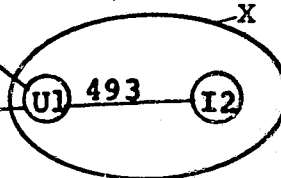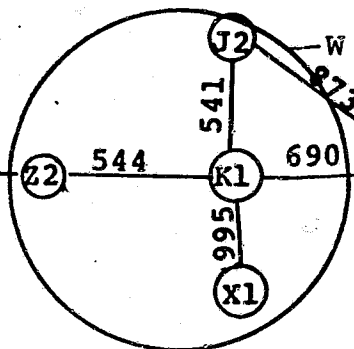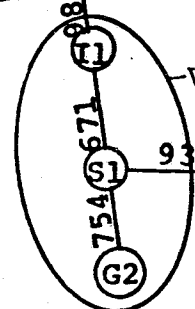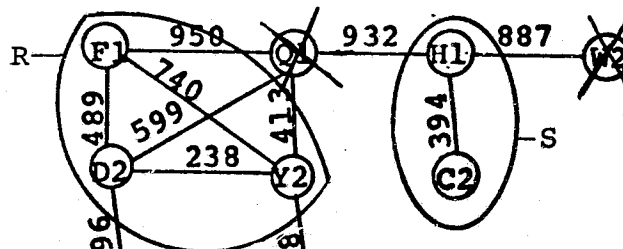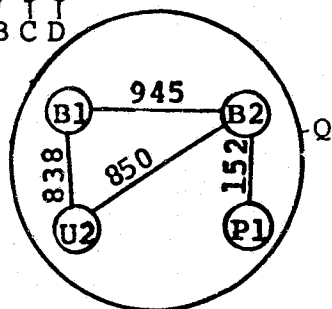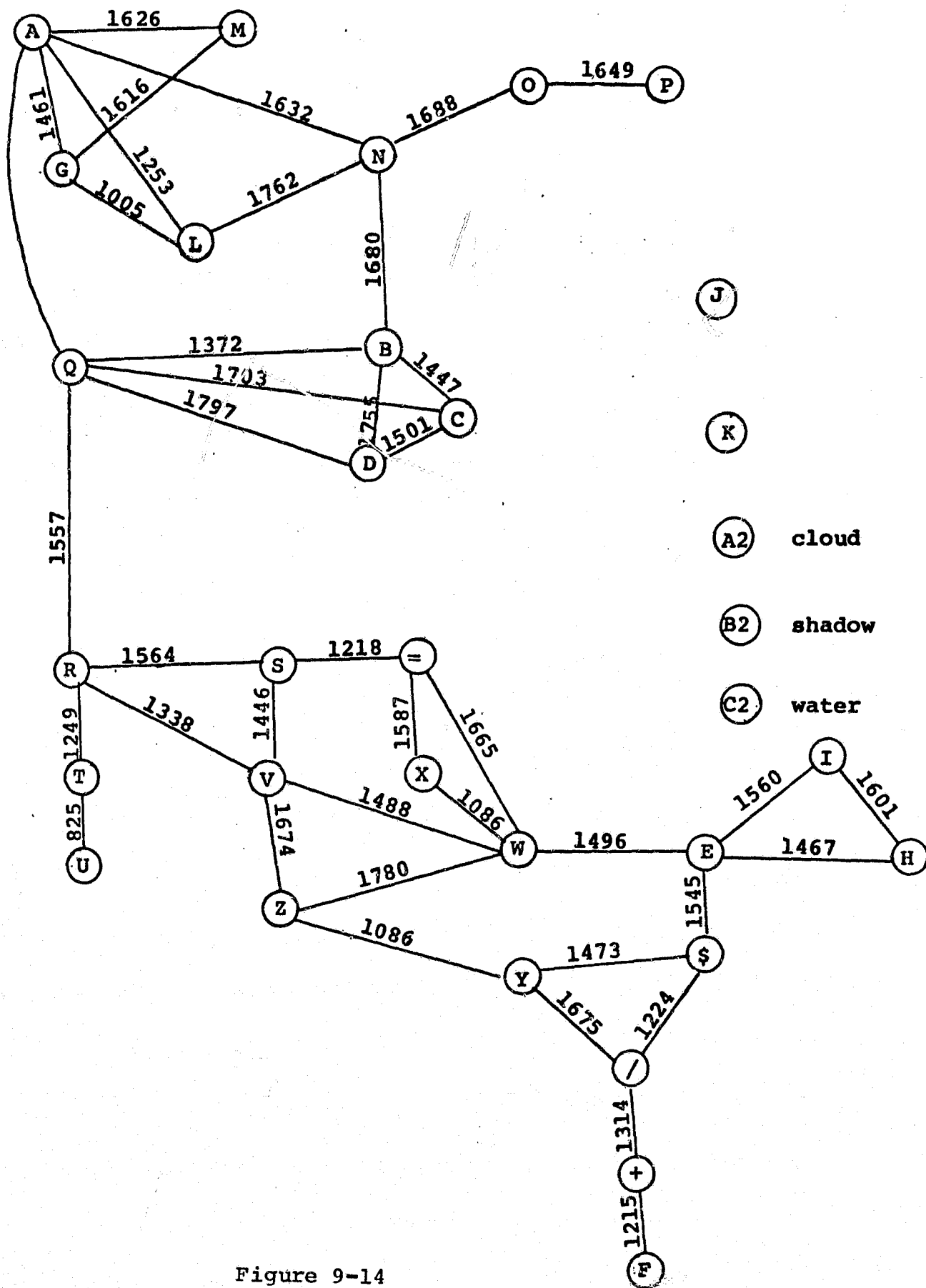
Figure 9-13

Figure 9-14

To interpret this diagram the analyst started by putting the cluster class identification information on the diagram. The result is shown in Figure 9-15.
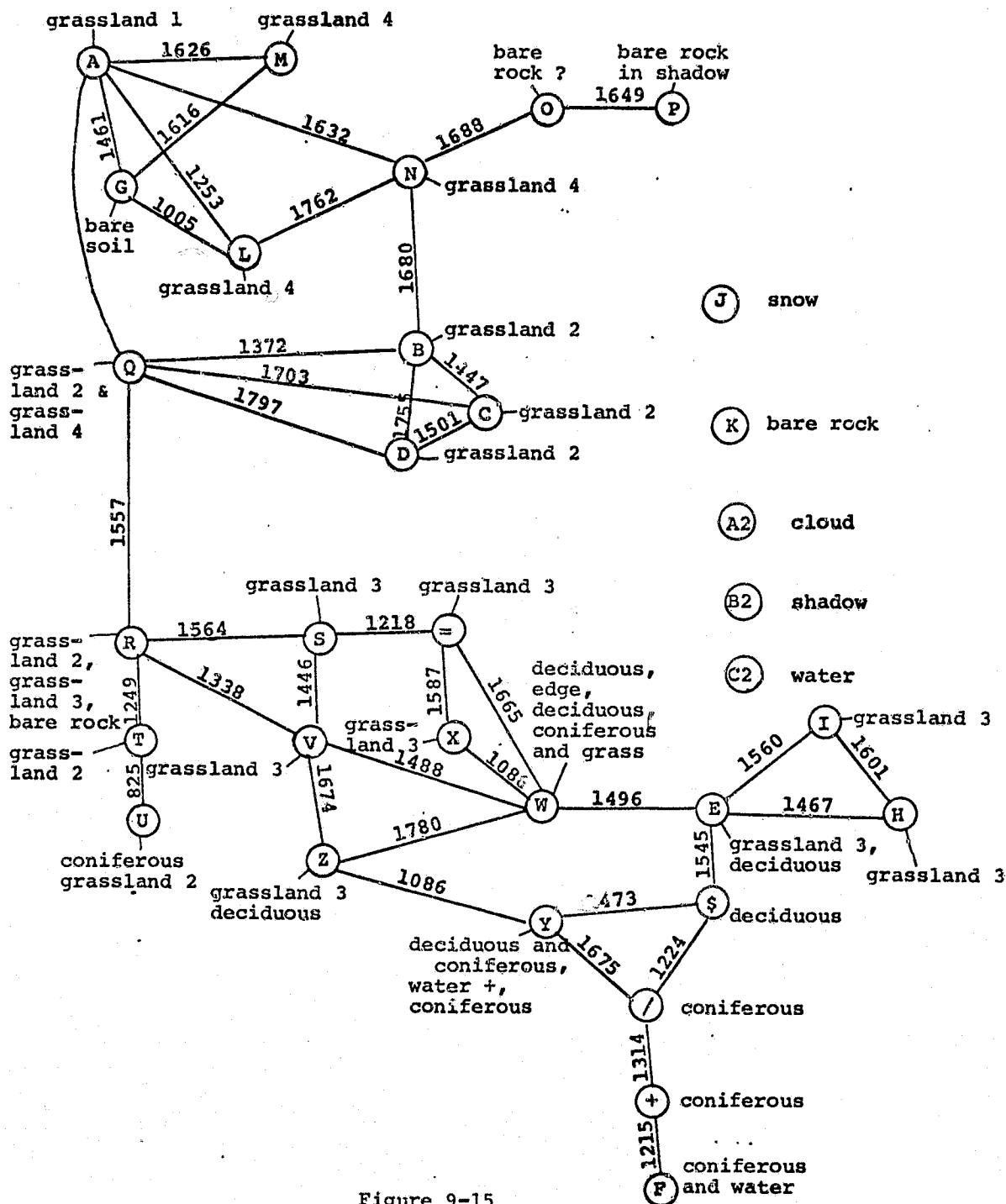
Figure 9-15

To interpret the diagram, classes J, K, A2, B2 and C2 were
rather straight-forward. They were not connected to any other
clusters, so they could each be circled individually and maintained
as clusters.

The next thing the analyst did was to delete any classes
composed of more than one cover type. For instance, R represented
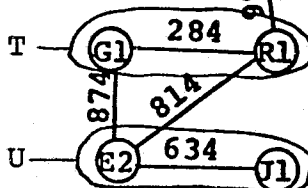grassland 2, grassland 3, and bare rock, so it was deleted. U
represented both coniferous and grassland 2, so it was deleted.
Z was deleted, and W, E, Q and Y were deleted. The diagram then
looked like Figure 9-16.

Figure 9-16

- 104 -
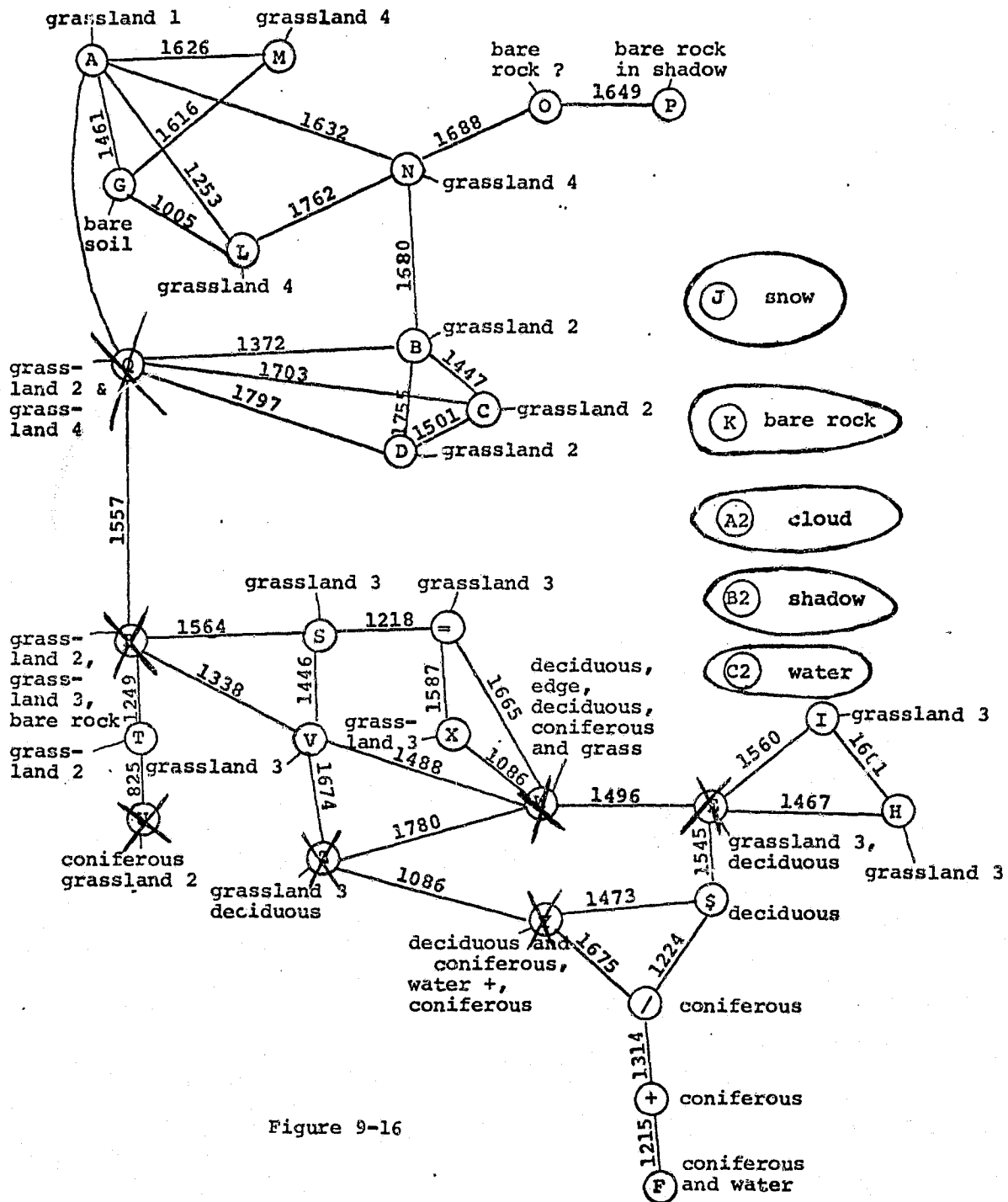
In this iteration, when classes of the same cover type had transformed divergence values less than 1750, the analyst grouped them together. As the interpretation of the rest of the diagram is discussed, you may wish to annotate Figure 9-16 to indicate the interpretation. Let's start at the bottom. Cluster F was coniferous and water. The next cluster up was coniferous. F was kept separate and by itself. Then the + and / were both coniferous, so they were grouped together. The $ was deciduous, so that became a class by itself. The I and H were both grassland 3, so they got grouped together. The S,=,V, and X classes were all grassland 3 and they were all connected to each other, so the analyst grouped them together into a single class. The X was also grassland 3 but it was not close to V or S, so it became another subclass of grassland 3. T, grassland 2, became a class by itself. Moving up to the B and D and C classes, the B-D distance was rather large and by inspection of the coincident spectral plot from STATISTICS the analyst decided that C was between B and D and therefore would cause confusion. He interpreted these classes by deleting class C, having B a separate class, and D as a separate class.

Look next at cluster A in the corner. It was identified as grassland 1 so it became a class by itself. The class G, bare soil, looked rather suspicious because it was close to class A, grassland 1, and would cause confusion. Also there is a class M, of grassland 4, and a class N of grassland 4, so the analyst felt that grassland 4 could be adequately represented without L, and it was deleted also. Them M became one grassland 4 subclass and N was the second grassland 4 subclass. O was a class of bare rock by itself and P was another class of bare rock. The interpretation of the diagram is shown completed in Figure 9-17.
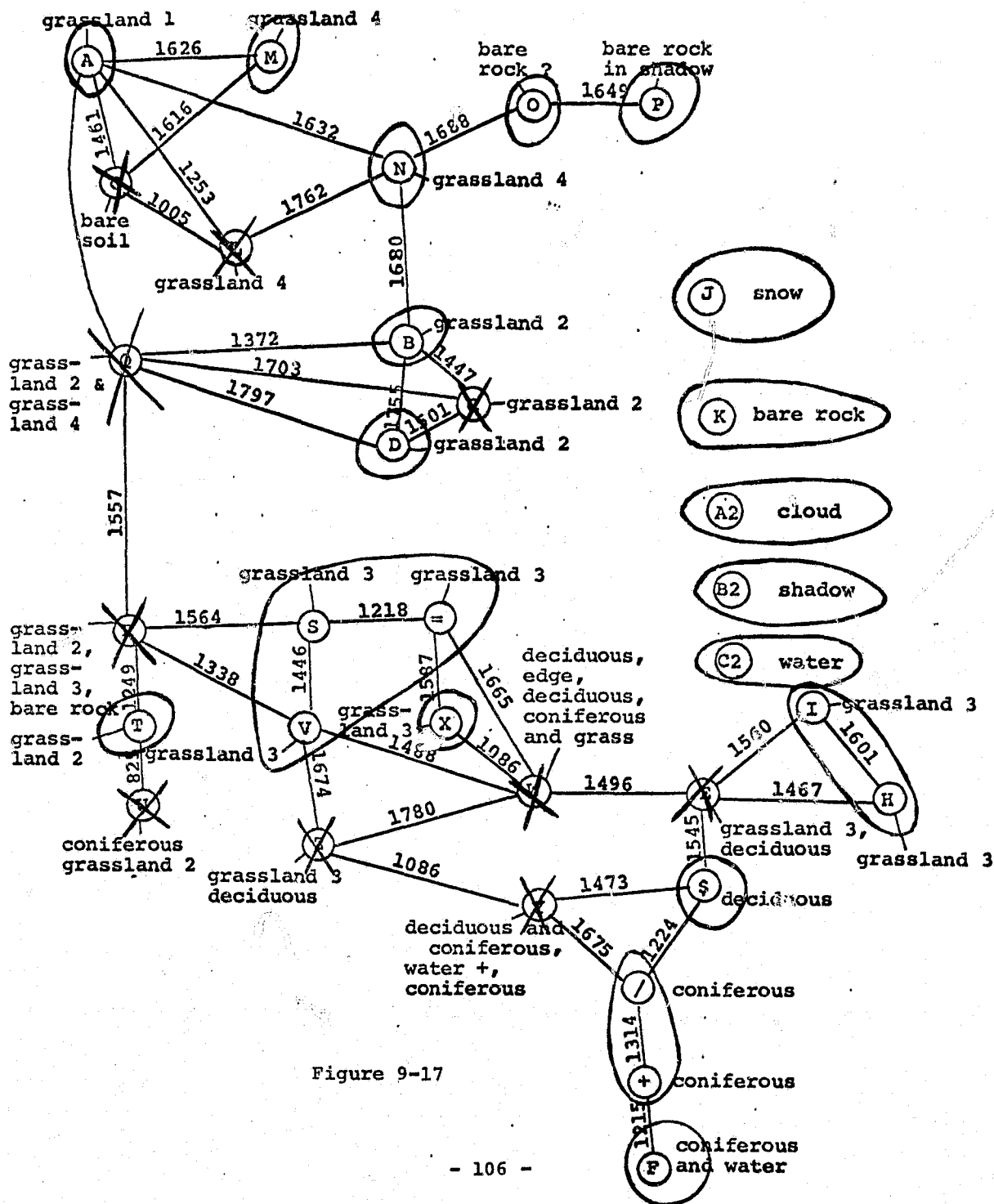
Figure 9-17

Using this diagram the analyst went back to his Field Description Cards again and did the combining and deleting indicated in Figure 9-17 to get together the description of all the training classes to be used to classify the Kenosha Pass test site.

Using the procedure illustrated in the example, interpret the separability diagram you created at the end of the preceding section, using the cluster class identification information you derived earlier, and keeping in mind your analysis objectives. You will have to decide how many iterations are necessary.

You should arrange to discuss your progress with your instructor at this time.

Section 10. CALCULATION OF STATISTICAL CHARACTERISTICS OF TRAINING CLASSES

Upon completion of this section, you should be able to do the following:

Explain why statistics are needed at this point in the analysis.

Use the LARSYS processing function to obtain statistics of training classes.

The analysis sequence so far has resulted in a set of Field Description Cards indicating the coordinates of data to be used for training. Before the classification can be performed, the STATISTICS processing function must be run on this final training set. The classifier available in LARSYS for classifying data on a point-by-point basis (CLASSIFYPOINTS) is based on the assumption that the training classes can be represented by multivariate Gaussian probability density functions, defined by mean vectors and covariance matrices. The STATISTICS processing function calculates the mean vectors and covariance matrices for the data described on Field Description Cards.

A detailed description of the input and output of the STATISTICS processor can be found on pages STA-1 to STA-22 in Volume 2 of the LARSYS User's Manual.

EXAMPLE

After the analyst had defined training patterns as discussed in the last section, he was ready to generate statistics for training. The control cards listed on the next page gave him the output he wanted:

```
-COMMENT KENOSHA PASS FINAL TRAINING STATISTICS
-RUNTABLE
DATA
RUN(73057902), TAPE(253), FILE(1)
END
*STATISTICS
CHANNELS 1,2,3,4
PUNCH
PRINT HIST(C), CORRE(C)
SCALE SPCINT(1)
DATA
CLASS GRASLND1
    .
    .
   field description cards for Dry Grass
    .
    .
CLASS GRASLND4
    .
    .
   field description cards for first Tundra subclass
    .
CLASS GRASLND4
    .
    .
   field description cards for second Tundra subclass
    .
    .
CLASS BAREROK1
    .
    .
   field description cards for first subclass of Bare Rock
    .
    .
CLASS BARESHAD
    .
    .
   field description cards for Bare Rock in Shadow
    .
    .
CLASS GRASLND2
    .
    .
   field description cards for first subclass of
   Mountain Bunch Grass
    .
    .
CLASS GRASLND2
    .
    .
   field description cards for second subclass of
   Mountain Bunch Grass
    .
    .
CLASS GRASLND2
    .
    .
   field description cards for third subclass of
   Mountain Bunch Grass
    .
    .
```

```
CLASS GRASLND3
  .
  .
  field description cards for first subclass of Wet Meadow
  .
  .
CLASS GRASLND3
  .
  .
  field description cards for second subclass of Wet Meadow
  .
  .
CLASS GRASLND3
  .
  .
  field description cards for third subclass of Wet Meadow
  .
  .
CLASS DECIDOUS
  .
  .
  field description cards for Deciduous
  .
  .
CLASS CONIFER1
  .
  .
  field description cards for first subclass of
  Conifer
  .
  .
CLASS CONIFER2
  .
  .
  field description cards for second subclass of
  Conifer
  .
  .
CLASS BAREROK2
  .
  .
  field description cards for second subclass of Bare Rock
  .
  .
CLASS SNOW
  .
  .
  field description cards for Snow
  .
  .
CLASS CLOUD
  .
  .
  field description cards for Clouds
  .
  .
CLASS C SHADOW
  .
  .
  field description cards for Cloud Shadows
  .
  .
CLASS WATER
  .
  .
  field description cards for Water
  .
  .
END                        - 111 -
```

This time the analyst wanted to see the histograms of the classes and the correlation matrices of the classes. Again, he used the SCALE card so that the coincident spectral plot would have a scale appropriate to LANDSAT data.

After running STATISTICS, the analyst ran SEPARABILITY one more time for the purpose of checking on probability of correct classification. The transformed divergence between all class pairs was greater than 1750, which generally corresponds to better than 90% probability of correct classification for training data. The analyst considered this acceptable, and decided that he was ready to classify the data, so the punched statistics deck from this job was saved for use in the next analysis step.

EXERCISES

---

1. Explain why statistics are needed at this point in the analysis.

CASE STUDY

Set up the control cards to run the STATISTICS processing function with your Field Description Cards from CLUSTER combined in the way determined in Section 9. Remember to request punched output.

Section 11. CLASSIFICATION, RESULTS DISPLAY, AND EVALUATION

Upon completion of this section, you should be able to do the following:

Name and briefly describe the decision rule implemented in the CLASSIFYPOINTS processing function.

Identify the kind of distribution training classes are assumed to have, and name the two parameters needed to define such distributions.

Given a statistics deck and the coordinates of an area to be classified, set up the control cards and run the CLASSIFYPOINTS processing function.

Set up control cards for PRINTRESULTS to display the classification, and run the job.

Given an example of a class performance matrix, indicate points correctly classified, errors of omission, and errors of commission for a specified class.

The CLASSIFYPOINTS processing function classifies multi-spectral (and multitemporal) data one point at a time into classes defined by the training statistics. This is the last major step in the process of deriving useful information from remote sensing data. Of course, an analyst may decide that the first classification produced is not satisfactory for the objectives to be met. In that case, decisions made in previous steps would have to be revised. The first few times an analyst deals with the spectral properties of cover types by using the pattern recognition techniques available in LARSYS, it may be necessary to go all the way back to the step where candidate training areas were selected. However, with experience and increased analysis skill, such drastic revisions can be avoided.

The decision rule implemented in LARSYS is called a maximum likelihood classification rule. Each data point to be classified is compared to all of the training classes, and is assigned

to the most likely class. To express the concept of a classifier
in a quantitative way so that the computer can do the work, we
would like to have a set of functions corresponding to the train-
ing classes. These functions would have the property that when
a data vector to be classified is substituted into all of them, **the**
function having the largest value corresponds to the class to
which the data vector belongs. This would provide a quantitative
way of discriminating between classes. Such functions are called
discriminant functions. One way to get such a set of functions
(the way LARSYS gets discriminant functions) is to start with
training classes, and assume that they have multivariate Gaussian
(or multivariate normal) probability density functions. Then
the mean vector and covariance matrix define the distribution of
the training data. The discriminant functions are expressed in
terms of the mean vector and covariance matrix, which were cal-
culated for all training classes by use of the STATISTICS process-
ing function in the last section.

For more detailed information about the classification
algorithm, see pages CLA-25 through CLA-29 of Volume 2 of the
LARSYS User's Manual and also LARS Information Note 111572,
Pattern Recognition: A Basis for Remote Sensing Data Analysis
by Philip H. Swain.

The classification which is produced is stored on disk or
tape (whichever you specify). In order to access and evaluate
classification results, another LARSYS processing function is
used. The PRINTRESULTS processing function can provide an alpha-
numeric printout, and it has a capability for providing quantita-
tive information about a classification in the form of tables.
The analyst can specify the coordinates of areas of interest,
called "test fields." The computer then examines and tabulates
the classification decision for each data point, and prints out
a summary by fields, or classes, or both, as specified by the
analyst. An example of tabular results for classes is shown in
Figure 11-1. Such a table can be called a test class performance
matrix.

	NO OF SAMPS	PCT. CORCT	OATS	CORN	WHEAT	SOYB	GRASS
OATS	66	98.5	65	0	0	1	0
CORN	93	93.5	0	87	0	6	0
WHEAT	69	100.0	0	0	69	0	0
SOYB	57	93.0	2	0	0	53	2
GRASS	31	90.3	0	3	0	0	28
TOTAL	316		67	90	69	60	30

Figure 11-1. Test class performance matrix

What do the numbers in the performance matrix tell you about
the classification? Look first at the 66 samples of OATS. The
table indicates that 65 of those points, or 98.5%, were correctly
identified. Looking across that row, the table also indicates
that one data point which the analyst knows to be oats was in-
correctly classified as soybeans. That is, there was one error
of omission for the 66 oats samples. Looking down the column
labelled OATS, there were two errors of commission for the class
oats. That is, two samples were called oats that should not
have been.

The diagonal elements of the matrix can be summed, and that
total divided by the total number of samples. The result is
called overall performance. For Table 11-1, the overall per-
formance is (65+87+69+53+28) ÷ 316 = 95.5%.

Another way of evaluating the classification is to sum the
percent correct for each class, and divide by the number of
classes. This result is called average performance by class.
The average performance by class in Table 11-1 is (98.5+93.5+
100.0+93.0+90.3) ÷ 5 = 95.1%.

The capability to get tabular results for a specified
area can be used to obtain area estimates for the cover types
in a classification. To do this, the coordinates of the en-
tire area can be put in as a test field. Then the output,
instead of being the usual performance matrix, will be a
one-line table indicating the number of data points classified
into each cover type. Given the total number of data points
and the total area, the area per data point can be calculated.
Then this area per data point multiplied by number of data
points per cover type will give area per cover type.

Furthermore, the information in the performance matrix
about the error rates associated with a classification can
be used to adjust areal estimates derived from the classifi-
cation, so that they more nearly estimate the true amounts
of each cover type. More detailed information about this
use of the performance matrix can be found in Appendix A,
material extracted from the final report for Contract NAS5-
21773: A Study of the Utilization of ERTS-1 Data from the
Wabash River Basin by Marvin E. Bauer, titled Identification
and Area Estimation of Agricultural Crops by Computer Class-
ification of ERTS-1 MSS Data.

For the Kenosha pass example, the analyst used the following control cards to classify the data:

```
-COMMENT KENOSHA PASS CLASSIFICATION
-RUNTABLE
DATA
RUN(73057902), TAPE(253), FILE(1)
END
*CLASSIFYPOINTS
RESULTS TAPE(TTT), FILE(F)
CARDS READSTATS
CHANNELS 1,2,3,4
DATA
    .
    .
    .
    punched statistics file from STATISTICS processing function
    .
    .
DATA
RUN(73057902), LINE(197,531,1), COL(401,803,1)
END
```

Note that the analyst specified that his results were to go onto tape, where they remain until he writes over them. When you classify the case study data, you should put your results on disk. However, this means that you must run your PRINTRESULTS job in the same terminal session (or batch job) as the CLASSIFY-POINTS, because disk results are cleared at logout.

When you do need to keep results on tape, check with your supervisor to learn the numbers of any tapes assigned to you.

The next control card, CARDS READSTATS, indicates that the statistics (mean vectors and covariance matrices) of the training classes are to be read from punched cards.

The CHANNELS card of course specifies which channels are to be used for classification. The analyst chose to classify with all four channels. If a subset of channels had been desired (for instance, if one of the channels had striping so bad it couldn't be used), this is where the subset would be specified.

The first data deck contains the statistics of the training classes. The second data deck indicates the area to be classified.

To display the results, the analyst used the following control card setup:

```
-COMMENT PRINTOUT OF KENOSHA PASS
*PRINTRESULTS
RESULTS TAPE(TTT), FILE(F)
PRINT TRAIN(C)
SYMBOLS +,+,+,-,-,+,+,+,+,+,+,/,0,0,-, , ,$,W
GROUP GRASSLND(1/1,2,3,6,7,8,9,10,11/)
GROUP BARREN(2/4,5,15/),DECIDOUS(3/12/)
GROUP CONIFER(4/12,14/),SNOW(5/16/),CLOUD(6/17/)
GROUP C SHADOW(7/18/),WATER(8/19/)
END
```

The analyst specified that training class performances were
to be printed out. Since CLASSIFYPOINTS stores the coordinates
of the training data at the beginning of the results file, the
analyst did not have to include a data deck to get training per-
formance as he would have for test performance.

On the SYMBOLS card, alphanumeric symbols were assigned to
the classes in the order that the classes appeared in the statis-
tics deck. Using the same symbol for all subclasses of a single
cover type simplified the printed map. By the way, a blank can
be used as the symbol for a class (for example, classes 16 and
17 above).

The next four cards the analyst used were GROUP cards.
Notice that the conventions for continuing function control
cards include repeating the key word, not ending with a comma,
and not punching past column 72. These cards provide special
instructions to the computer about how results are to be tab-
ulated. The first GROUP card, for instance, indicates that
if a data point specified to the computer as grassland was
classified into class 1,2,3,6,7,8,9,10 or 11, it was correctly
classified. Similarly, if a point specified to the computer
as barren was classified into class 4 or class 5 or class 15,
it was correctly classified.

For this Kenosha Pass study, the analyst did not have ade-
quate reference data to specify the coordinates of test fields.
The training class performance provided one indication of accur-
acy. The training class performance matrix is shown in Figure 11-2.

Notice that the headings across the top of Figure 11-2 in-
clude water, cloud, and cloud shadow, but the column on the side
doesn't. The explanation is that the column on the side lists
only those groups of classes for which there are training fields
within the area classified. Remember that the analyst selected
training data for water, clouds, and cloud shadows from outside
the Kenosha Pass area.

GROUP	NO OF SAMPS	PCT. CORCT	WATER	CLOUD	C SHADOW	SNOW	GRASSLND	DECIDOUS	CONIFER	BARREN
SNOW	8	100.0	0	0	0	8	0	0	0	0
GRASSLND	3185	99.7	0	0	0	0	3175	0	0	10
DECIDOUS	552	93.3	0	0	0	0	9	515	28	0
CONIFER	823	95.7	0	0	2	0	0	33	788	0
BARREN	306	98.7	0	0	0	1	3	0	0	312
TOTAL	4874		0	0	2	9	3187	548	812	312

OVERALL PERFORMANCE (4788/4874) = 98.2

AVERAGE PERFORMANCE BY CLASS (487.4/5) = 97.5

Figure 11-2.   Training class performance, Kenosha Pass example.

Further evaluation of the results was not available for the Kenosha Pass area at this writing. However, the Forest Service, Rocky Mountain Forest and Range Experiment Station had provided evaluation information for two other areas. Since the two available areas, Manitou and Eleven Mile, are similar to Kenosha Pass, the evaluation procedure and results for those two areas will be discussed.

Test fields were found by the following procedure: In each area, the part corresponding to a 1:50,000 color infrared aerial photograph was used for evaluation. A grid of cells, each the same size as a 4 x 4 block of data points, was superimposed on the photography, and the grid was sampled systematically. Sample cells which were completely within one cover type became test fields. To minimize possible location errors, only the results of the interior 2 x 2 block of data points were tabulated. For the Manitou area, the hand-calculated results table is shown in Figure 11-3. The other area which was evaluated by this procedure, the Eleven Mile area, had the results (or test class performance matrix) shown in Figure 11-4.

You have followed the Kenosha Pass example through classification, results display, and evaluation, the topics of this section. However, since you will be instructed to use test fields to evaluate your classification, another example will be shown.

An example of the PRINTRESULTS processing function where test fields have been included for evaluation purposes will be presented. The analysis was done on MSS data collected from aircraft altitude over an agricultural scene. The analyst has already examined data quality, coordinated MSS data with reference data, selected candidate training areas, refined the candidate training areas, and calculated statistics for training classes. We will just look over his shoulder to see how he set up his card deck to classify the data and display the results.

Computer classification category

Photointer-pretation Category	Number of Samples	Percent Correct	Grassland	Deciduous Forest	Coniferous Forest	Water
Grassland	257	84.4	217	10	30	0
Deciduous Forest	21	23.8	8	5	8	0
Coniferous Forest	1239	86.1	146	26	1067	0
Water	3	66.7	0	0	1	2
Total	1520	84.9*	371	41	1106	2

*overall performance=number correct/total number of samples

Figure 11-3. Hand-tabulated test class performance matrix for Manitou area.

Photointer-pretation Category	Number of Samples	Percent Correct	Grassland	Deciduous Forest	Coniferous Forest	Water	Barren
Grassland	433	80.1	347	7	74	0	5
Deciduous Forest	94	51.1	35	48	11	0	0
Coniferous Forest	739	66.0	76	175	488	0	0
Water	60	95.0	0	0	3	57	0
Barren	54	1.9	34	1	18	0	1
Total	1380	68.2*	492	231	594	57	6

*overall performance = number correct/total number of samples.

Figure 11-4. Hand-tabulated test class performance matrix for Eleven Mile area.

He used the following card deck:

```
-COMMENT CLASSIFY, DISPLAY, AND EVALUATE 66000652
-RUNTABLE
DATA
RUN(66000652), TAPE(TTT), FILE(F)
END
*CLASSIFYPOINTS
RESULTS DISK
CARDS READSTATS
CHANNELS 1,6,8,12
DATA
  •
  •
  stat deck from previous STATISTICS run
  •
  •
DATA
RUN(66000652), LINES(1,950,2), COL(1,222,2)
END
*PRINTRESULTS
RESULTS DISK
PRINT OUTLINE (TRAIN,TEST), TRAIN(F,C),TEST(F,C,P)
SYMBOLS O,O,O,C,C,C,C,W,W,W,S,S,S,G,G,G,G
GROUP OATS(1/1,2,3/), CORN(2/4,5,6,7/), WHEAT(3/8,9,10/)
GROUP SOYBEANS(4/11,12,13/), GRASS(5/14,15,16,17/)
DATA
TEST 1
  •
  •
  Field Description Cards for OATS test fields
  ~
  TEST cards and Field Description Cards for corn,wheat,and soybeans
  •
  •
TEST 5
  •
  •
  Field Description Cards for Grass test fields
  •
  •
END
```

There are several points to be observed about this deck.
The CLASSIFYPOINTS and PRINTRESULTS jobs are being run "back-to-
back". The classification results file is to be stored on disk.
Training and test fields are to be outlined on the PRINTRESULTS
display map. If your training fields are small, three- or four-
point fields, outlining them may not be desirable. Training field
and class performances are to be printed out (again, if your train-
ing fields are small, you would probably print performances for
classes, not fields). Test field, class, and percentage tables
are to be printed out.

Also observe that the group number on the GROUP cards deter-
mines the order in which the test fields appear in the data deck.

1. Name and briefly describe the kind of decision rule implemented in the CLASSIFYPOINTS processing function.

2. Identify the kind of distribution training classes are assumed to have, and name the two parameters needed to define such distributions.

3. In Figure 11-1 indicate the points correctly classified, errors of omission, and errors of commission for soybeans.

CASE STUDY

Using the punched STATISTICS deck you generated in the last section, set up the control cards to run the CLASSIFYPOINTS processing function, putting your results on disk. The area to be classified is from line 30 to 430, and from column 112 to 333.

Also set up the control cards to run the PRINTRESULTS processing function. Your instructor has a set of test fields which you can use to evaluate your classification. Ask for any output products you think would be useful.

Run the CLASSIFYPOINTS and PRINTRESULTS jobs. Consider the possibility of running more than one PRINTRESULTS job. For instance, you might wish to assign every class and subclass a different symbol on one map, while on another map you could assign the same symbol to all subclasses of a class. If one cover type is of particular interest to you, you could assign a symbol to it, and blanks to all other classes.

If your overall test performance is not satisfactory, consider what you could do to improve the classification. Discuss the possibilities with your instructor.

Section 12. INFORMATION EXTRACTION AND INTERPRETATION

Upon completion of this section, you should be able to do the following:

Name at least two kinds of information that can be extracted from a classification.

Give an example of useful information extracted from multi-spectral classifications in your discipline.

The final and most important step is interpretation of results. The classification results themselves are not usually the product of interest. Instead, the objective of an analysis is usually to gain information for use in such things as forest management or land use planning. For instance, the objective usually involves learning where specific cover types are located, or what proportion of the area belongs to each cover type.

To complete the analysis, the original objectives must be reviewed, and the desired information extracted.

Examples of results analysis and the extraction of useful information from multispectral data classifications may be found in several journals, including those listed here:

Remote Sensing of the Environment
IEEE Transactions on Geoscience Electronics
Remote Sensing in Ecology
Journal of Soil and Water Conservation
Photogrammetric Engineering and Remote Sensing
Agronomy Journal
Applied Optics

Samples can also be found in a number of LARS Information Notes, published proceedings of remote sensing conferences, etc.

The objectives of the Kenosha Pass analysis were 1) to classify and inventory the area into these cover types: water, snow, grassland, deciduous forest, coniferous forest, barren (bare rock and bare soil); 2) to produce a classification map of these cover types; 3) to evaluate the classification accuracy.

The third objective, evaluation of accuracy, was discussed in the previous section. The analyst requested training class performance from PRINTRESULTS, and overall performance was 98.2%. Test performance was determined manually, rather than by submitting test field coordinates to PRINTRESULTS.

The second objective was to produce a classification map of the cover types of interest. The control cards used to produce the map were listed and discussed in the previous section.

And now for the first objective--a classification and inventory of the cover types. In addition to the map showing how cover types were distributed, an inventory of the amount of each cover type was needed. To get this information, the analyst ran PRINT-RESULTS again, and input the coordinates of the entire area as a test field. Then the output, instead of being the usual performance matrix, was a one-line table indicating the number of data points classified into each cover type.

The control cards used are listed here:

```
-COMMENT KENOSHA PASS INVENTORY
*PRINTRESULTS
RESULTS TAPE(TTT), FILE(F)
PRINT TEST(C), MAPS(0)
GROUP GRASSLND(1/1,2,3,6,7,8,9,10,11/)
GROUP BARREN(2/4,5,15/), DECIDOUS(3/12/)
GROUP CONIFER(4/13,14/), SNOW(5/16/), CLOUD(6/17/)
GROUP C SHADOW(7/18/), WATER(8/19/)
DATA
TEST 1
RUN(73057902), LINE(197,531,1), COL(401,803,1)
END
```

The table produced by these cards contained the following information:

NUMBER OF SAMPLES CLASSIFIED INTO

GRASSLND	BARREN	DECIDOUS	CONIFER	SNOW	CLOUD	C SHADOW	WATER
97742	3597	13519	19640	38	196	253	20

Since the total number of data points (135,005) and the total area in acres (154,965) are known, the number of acres per data point could be determined (154965÷135005=1.14785), and from this the number of acres classified into each cover type was calculated.

The analyst has carried out an analysis sequence which met his objectives.

EXERCISES

1. Name at least two kinds of information that can be extracted from a classification.

2. Check with your instructor on the availability of the listed references, and skim through one or more of them. Then give an example of useful information extracted from multispectral classification.

CASE STUDY

Study your classification analysis results. Did you meet your analysis objectives? What information can you extract from the results? Based on your results, would you say that the cover type classes you initially selected were sufficiently distinct spectrally to provide adequate classification accuracy? Would you consider it worthwhile to use these classes as the basis for a "real life" application of remote sensing?

Discuss your results and conclusions with your instructor.

Appendix A

The following material has been reprinted from Section 2 of
Identification and Area Estimation of Agricultural Crops by Com-
puter Classification of ERTS-1 MSS Data, the final report for
Contract NAS5-21773: A Study of the Utilization of ERTS-1 Data
from the Wabash River Basin.

2.333 UNBIASING CLASSIFICATION RESULTS

Experience has shown that it is inevitable that some points
are incorrectly identified by the maximum likelihood classifier.
In this experiment, only about 80% of the test samples were cor-
rectly classified. The primary source of these errors is over-
lapping density functions for two or more classes. For example,
some corn "looks" like soybeans and some soybeans are spectrally
similar to corn. As described above, prior probability informa-
tion or class weights can be used to good advantage to at least
partially reduce the effects of such circumstances. A second pro-
cedure which can be used after the classification has been per-
formed is to unbias or adjust the results based on the correct
classification proportions and error rates. The latter procedure
was first used during the 1971 Corn Blight Watch Experiment.

The source of the correct classification proportions and
error rates are the matrices of test field classification per-
formance such as shown in Table 2.6. From such information, we
can determine the proportions, for instance, of corn classified
as corn and non-corn and the proportions of non-corn classified
as non-corn and corn. With this information, it is then possible
to unbias or adjust the classification results for a county or
several counties so that they more nearly estimate the true amounts
of each class present in the classified area.

Theoretically, if the true values of the error rates of
omission and commission were known, the classification results
could be adjusted so that in effect the area estimates based on
the classification closely approximated the true amounts of each
crop present. In practice, of course, this situation is seldom
found. The primary limitation is that the test samples are not
completely representative of the total area classified and only
provide estimates of the true error rates. Possible causes of
non-representative test samples are that samples come from only
a small part of the total area being classified and that many
cover types such as farmsteads, idle land, roads, and urban areas
generally have not been included in the set of test fields.

The method we have used for unbiasing classification results
involves multiplying the county classification results (Table 2.7)
by the inverse of the test field classification performance
matrix (Table 2.6) as follows:

$$A = CP^{-1}$$

Table 2.6 Classification of corn, soybean, and "other" test fields, DeKalb, Ogle, and Lee Counties, Illinois, with and without the use of prior probability information in the classification decision rule.

(a) No prior probability information used, equal class weights assumed.

| | NO. POINTS | NO. POINTS CLASSIFIED AS | | | PERCENT CORRECTLY CLASSIFIED |
CLASS		CORN	SOYBEANS	"OTHER"	
Corn	9290	7546	973	771	81.2
Soybeans	2235	244	1732	259	77.5
"Other"	1121	150	307	664	59.2
TOTAL	12646	7940	3012	1694	78.6

(b) Prior probability information used, unequal class weights.*

| | NO. POINTS | NO. POINTS CLASSIFIED AS | | | PERCENT CORRECTLY CLASSIFIED |
CLASS		CORN	SOYBEANS	"OTHER"	
Corn	9290	7983	382	925	85.9
Soybeans	2235	395	1556	284	69.6
"Other"	1121	206	220	695	62.0
TOTAL	12646	8584	2158	1904	80.9

*Class weights were 44, 16, and 40 for corn, soybeans, and "other", respectively.

Table 2.7   Number of samples classified into corn, soybeans, and "other" for DeKalb, Ogle, and Lee Counties, Illinois.

(a)   Equal Class Weights

| County | No. Points Classified As | | |
	Corn	Soybeans	"Other"
DeKalb	131,451	85,148	74,311
Ogle	146.108	112,385	135,058
Lee	150,992	122,101	120,266
TOTAL	428,551	319,634	329,635

(b)   Unequal Class Weights

| DeKalb County | No. Points Classified As | | |
	Corn	Soybeans	"Other"
DeKalb	152,920	54,948	83,042
Ogle	170,220	74,940	148,391
Lee	178,177	80,241	134,941
TOTAL	501,317	210,129	366,374

where, C is the classification vector with n crops or classes, P^{-1} is the inverse of the n x n classification performance matrix, and A is a 1 x n vector of the crop acreages.

The results of applying this correction procedure are presented in Table 2.8 and discussed in the next section along with further results on the use of prior probabilities to the classification function.

2.334 ACREAGE ESTIMATION

The classification performance indicated by 80% correct recognition of test fields is believed to be adequate for satisfactorily estimating crop acreages. To determine how well crop acreages could be estimated from the ERTS classification, the ERTS coordinates of the three counties were obtained, the counties were classified, and the number of pixels classified into each class tabulated (Table 2.7). In Table 2.8, four acreage estimates based on the ERTS classifications are compared to each other and to estimates made by the Illinois Cooperative Crop Reporting Service (SRS/USDA). The ERTS estimates are the four combinations of using prior probability information in the classification decision rule and unbiasing the classification results as discussed in the previous two sections.

The standard to which the ERTS classifications are compared is the acreage estimates (shown as percentage of total land area) made by SRS/USDA. The mean squared differences between the SRS/USDA estimates and the several ERTS estimates are shown as a means of comparing the overall goodness of each ERTS estimate.

One of the most difficult aspects of remote sensing technology is quantitatively evaluating classification results. It is physically impossible to collect sufficient ground data of crop identification and acreage over large areas, to determine how accurate area estimates made from the ERTS classification are. We have therefore used the USDA county estimates as the reference for comparison. However, the crop surveys conducted by the USDA are designed to achieve prescribed levels of accuracy at only the national and state levels. For this reason the USDA does not publish accuracy figures for their county estimates. However, in those states, including Illinois, in which an annual farm census is conducted, the acreage estimates are considered to be quite accurate. Their estimates are probably within three to five percent of the actual acreages.

The ERTS estimates, particularly those adjusted for classification bias, are very close to those made by the USDA. It seems clear that the USDA and ERTS estimates are of the same parameter. The estimates agree best for the total of the three counties. There is more variation between the two estimates for the individual counties. However, this is simply a result of having a larger sample and can be expected as long as there is not a consistent bias in one direction in the ERTS classification, eg., corn is always over-estimated.

Table 2.8. Comparison of crop acreage estimates by USDA and estimates based on ERTS classifications. The results of utilizing prior probability information in classification and bias correction of classifications are shown.

County	Class	SRS-USDA	ERTS			
			Uncorrected		Bias Corrected	
			Equal Wts.	Non-Eq. Wts.	Equal Wts.	Non-Eq. Wts.
			(Percent of Total Land Area)			
DeKalb	Corn	41.5	45.2	52.6	47.6	50.8
	Soybeans	21.3	29.3	18.9	19.9	14.3
	"Other"	37.2	25.5	28.5	32.5	34.9
	r.m.s.*		6.5	6.9	8.3	8.5
Ogle	Corn	41.3	37.1	43.3	35.5	37.0
	Soybeans	11.5	28.6	19.0	14.4	10.2
	"Other"	47.2	34.3	37.7	50.1	52.8
	r.m.s.		12.6	7.1	4.1	4.1
Lee	Corn	37.9	38.4	45.3	37.6	40.0
	Soybeans	21.9	31.0	20.4	19.9	14.0
	"Other"	40.2	30.6	34.3	42.5	46.0
	r.m.s.		7.6	5.5	1.8	5.8
Total for all Counties	Corn	40.2	39.8	46.5	39.6	41.8
	Soybeans	18.0	29.6	19.5	17.8	12.7
	"Other"	41.8	30.6	34.0	42.6	45.5
	r.m.s.		9.3	5.8	0.6	3.9

*r.m.s.-root mean square difference between USDA and ERTS estimates

Two additional things are clear from the results: (1) the use of class weights or prior probability information in classification gave substantially better estimates of the amounts of corn and soybeans present (reduction of the r.m.s. difference from 9.3 to 5.8) and (2) the application of the unbiasing procedure after classification further improved the ERTS estimates (reduction of r.m.s. difference to 3.9 and 0.6) for the classification with and without class weights, respectively. In conclusion, these two procedures should be used whenever possible in making crop acreage estimates from classifications of ERTS type data.

CREDITS

The analysis example was done by Michael D. Fleming, LARS. Information about evaluation procedures and results was provided by Dr. Richard S. Driscoll, Rocky Mountain Forest and Range Experiment Station.