

10-1-1973

The JSC Clustering Program ISOCLS and its Applications

E. P. Kan

Lockheed Electronics Company, Inc.

W. A. Holley

Lockheed Electronics Company, Inc.

H. D. Parker Jr.

Lockheed Electronics Company, Inc.

Follow this and additional works at: http://docs.lib.purdue.edu/lars_symp

Kan, E. P.; Holley, W. A.; and Parker, H. D. Jr., "The JSC Clustering Program ISOCLS and its Applications" (1973). *LARS Symposia*. Paper 32.

http://docs.lib.purdue.edu/lars_symp/32

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Conference on
Machine Processing of
Remotely Sensed Data

October 16 - 18, 1973

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

Copyright © 1973
Purdue Research Foundation

This paper is provided for personal educational use only,
under permission from Purdue Research Foundation.

THE JSC CLUSTERING PROGRAM ISOCLS AND ITS APPLICATIONS*

E.P. Kan
W.A. Holley
H.D. Parker, Jr.

Lockheed Electronics Company, Inc.
Houston Aerospace Systems Division
Houston, Texas 77058

I. ABSTRACT

The clustering program ISOCLS developed at the Johnson Space Center, Houston, Texas, has been extensively used in the pattern analysis and classification of remote sensor data collected by aircraft and by the Earth Resources Technology Satellite ERTS-1. This paper discusses the theory behind this clustering algorithm. Several new ideas that have been incorporated in ISOCLS are discussed. Among these are the novel philosophy of operation behind the procedure, which assumes that a population (i.e. a class or a cluster) can be treated as the union of an appropriate number of subpopulations, and the termination of the clustering program by a 'chaining algorithm'. Finally, this paper reports the results of the application of ISOCLS to an investigation on rangeland vegetation mapping using ERTS-1 data.

* The material of this paper was developed under NASA Contract #NAS9-12200 and prepared for the Earth Observation Division, NASA Johnson Space Center, Houston, Texas 77058.

II. INTRODUCTION

Data clustering is a process in which similar data points are grouped into the same clusters while dissimilar data points are associated with different clusters. Clustering, i.e. data grouping, has played important roles in the study of statistical data, in numerical taxonomy, in the investigation of social as well as biological data, and recently also in the analysis and classification of remote sensor data. In the context of remote sensing, data clustering has emerged as one of the most important steps towards the pattern recognition of features or objects of interest, as observed by remotely located sensors. In particular, the areas in which clustering has been extensively applied are: (i) checking of homogeneity of data; (ii) aiding in detection of boundaries between fields, e.g. agricultural fields; and (iii) nonsupervised classification of data.

The clustering problem has been approached in many ways, as is apparent from the literature. There are basically two main categories of clustering techniques: (i) the iterative procedure, and (ii) the one-pass procedure. These two types of procedures are actually complementary to each other, since each type possesses certain advantages over the other and conversely, suffers specific disadvantages absent in the other. Basically, this is a trade-off between the speed and the dependability of the two types of procedures. One well known example of the former type is ISODATA, designed by Ball and Hall (Ball, 1966). An equally well recognized procedure of the latter type is the one-pass algorithm, designed by Nagy and Tolaba (Nagy, 1972).

The clustering program ISOCLS, developed at the Johnson Space Center (JSC), Houston, Texas, is of the iterative type, and is basically a derivative of ISODATA. However, several new ideas have been incorporated into ISOCLS that makes it operate quite differently from ISODATA. Actually, the philosophy of operation behind ISOCLS is novel. The new ideas involved are discussed in Section III and Section IV. Extensive application of the program to investigate ERTS-1 and non-ERTS data has been carried out at JSC (Kan, 1973), particularly in the area of nonsupervised classification. A recent rangeland investigation using ERTS-1 data to map vegetation types is reported in Section VI. The general consensus is that the ISOCLS clustering program proves to be a very useful and dependable tool for the pattern analysis and classification of statistical data such as those collected by remote sensors.

III. PHILOSOPHY BEHIND ISOCLS

As in ISODATA (Ball, 1966), ISOCLS assigns data points to the current cluster centers by the minimum distance criterion. Cluster structures are also modified by appropriate 'splitting' or 'combining' of the existing clusters. However, a few ideas are new in the operation of ISOCLS. In particular, the program treats a population (i.e. a class or a cluster) as the union of an appropriate number of subpopulations of nominal sizes. In this way, even clusters of odd shapes, e.g. donut-shaped, are acceptable. As a result, assignment of points becomes less sensitive to the choice of distances. The distance between clusters is defined such that the magnitude of the distance (and thus the separation) relates to the probability of misclassification, i.e. the

degree of overlap of the two clusters. Of further interest is that this procedure terminates by chaining those clusters that are close to one another by a 'chaining algorithm'. This permits the discovery of those subpopulations, the union of which constitutes the parent population. Another highlight of the program is that it can be made by means of an artifice to operate in the same mode as Purdue's NSCLAS clustering program (Wacker, 1970).

To further clarify the philosophy of ISOCLS's operation, take the example of the two sets of points in Figure 1. Set (a) has the shape of a boomerang. Set (b) has the shape of a donut. Now, it is well known (Kan, 1972) that various clustering criteria with various measures of similarity will normally give rise to different groupings. As a matter of fact, in the existing clustering procedures, the final grouping is usually very sensitive to these criteria and measures. A most popular criterion is the minimum variance criterion which tends to produce eye-pleasing spherical (or ellipsoidal) groupings. This effect is most easily appreciated in vector spaces of three dimensions or less. However, operating under this particular criterion, clusters such as those in Figure 1 are unacceptable. The boomerang shape and the donut shape are odd shapes according to the standard of spherically shaped clusters. Furthermore, they can no longer be approximated to any satisfactory degree by Gaussian structures. But yet, such groupings are pleasing to the eye.

In order that data points are grouped into eye-pleasing structures, and also in order to circumvent the unnecessarily strict conformity to the standard of nice spherical clusters, the following philosophy of clustering is adopted in ISOCLS. Clusters are treated as unions of subclusters of nominal sizes. The intermediate steps in the clustering procedure try to assign points to subclusters of nominal sizes. In this way, the two data sets in Figure 1 would be grouped into subclusters 1-7 as shown in Figure 2. At the final step of the procedure, subclusters are chained into one group if each is close to at least one other in the group. Thus, subclusters 1, 2 and 3 are chained together because subclusters 1 and 2 are close to each other, and subclusters 2 and 3 are close to each other, even though subclusters 1 and 3 are by no means close to each other. Likewise, subclusters 4, 5, 6 and 7 are chained together to form a big cluster.

The chaining of neighboring clusters is performed by a 'chaining algorithm'. As a result of the chaining operation, the output classification map from ISOCLS represents chained clusters by the same symbol, signifying that the chained clusters basically belong to the same parent cluster. However, the statistics of these chained subclusters are not merged together. The program, as a matter of fact, outputs the statistics of the individual subclusters. This is precisely because not all the eye-pleasing clusters can adequately be approximated by Gaussian structures. Thus, if ISOCLS statistics are used in supervised classifications, as in the pattern recognition system LARSYS, the statistics of the subclasses should be used, unless there is reason to believe that the subclasses can be merged together. If merging is appropriate, a simple calculation would give the composite statistics of the parent class.

The distance measure used in ISOCLS for the assignment of points is the l_1 distance. That is, the distance $d(.,.)$ between point $X=(x_1, x_2, \dots, x_n)$ and $Y=(y_1, y_2, \dots, y_n)$ is defined as

$$d(X,Y) = \sum_{i=1}^n |x_i - y_i| \quad (1)$$

The measure of similarity $D(\cdot, \cdot)$ between clusters C_1 and C_2 , where C_1 is characterized by mean $\mu^{(1)} = (\mu_1^{(1)}, \dots, \mu_n^{(1)})$ and standard deviation $\sigma^{(1)} = (\sigma_1^{(1)}, \dots, \sigma_n^{(1)})$ and C_2 by $\mu^{(2)}$ and $\sigma^{(2)}$, is defined as

$$D(C_1, C_2) = \left[\sum_{i=1}^n \frac{(\mu_i^{(1)} - \mu_i^{(2)})^2}{\sigma_i^{(1)} \sigma_i^{(2)}} \right]^{\frac{1}{2}} \quad (2)$$

Notice the similarity between the distance measure (2) and the divergence measure between C_1 and C_2 (Marill, 1963). The divergence measure is defined as, in the case of univariate normal distributions,

$$J = \frac{1}{2} \left[(\sigma^{(1)})^2 - (\sigma^{(2)})^2 \right] \left[\frac{1}{(\sigma^{(2)})^2} - \frac{1}{(\sigma^{(1)})^2} \right] + \frac{1}{2} \left[\mu^{(1)} - \mu^{(2)} \right] \left[\frac{1}{(\sigma^{(1)})^2} + \frac{1}{(\sigma^{(2)})^2} \right] \left[\mu^{(1)} - \mu^{(2)} \right] \quad (3)$$

When $\sigma^{(1)}$ and $\sigma^{(2)}$ are almost equal, i.e.

$$\alpha < 1, \quad \text{where} \quad \frac{\sigma^{(1)}}{\sigma^{(2)}} \equiv 1 + \alpha \quad (4)$$

expression (3) can be approximated by

$$J \approx 2\alpha^2 + \frac{[\mu^{(1)} - \mu^{(2)}]^2}{\sigma^{(1)} \sigma^{(2)}} [1 + \frac{1}{2}\alpha^2] \quad (5)$$

which can in turn be approximated by the same form as (2). Since the program ISOCLS treats clusters as the union of subclusters, and since subclusters have nominal sizes, the assumption of $\alpha < 1$ is indeed valid. Thus, the distance measure (2) is well justified.

Consider the case of a univariate distribution which can be approximated by the sum of two equally probable, overlapping, normal $\mathcal{N}(0,1)$ distributions. Suppose this distribution is treated as two disjoint clusters. Notice that neither of these two clusters is $\mathcal{N}(0,1)$. Table 1 and Figure 3 summarize the relation between the distance measure D between the two clusters and the probability of misclassification P_e between the two component distributions. A distance of 3.2 or larger between two clusters indicates a P_e of less than 10%. A distance of less than 3.2 indicates a more extensive overlapping and thus a higher degree of confusion. The following sections will discuss more on the effects of the distance threshold on the grouping of data points by the ISOCLS clustering program.

IV. THE ISOCLS PROGRAM

The operation of the ISOCLS program is most easily explained by the functional sequence of its ISTOP iterations, where ISTOP is a preassigned input number. The operational sequence can be coded as

$$\underbrace{SS\dots S}_m \underbrace{(CS)(CS)\dots(CS)}_{2n} C_h,$$

initial intermediate final
mode mode mode

where S, C and C_h stand for a 'split', 'combine' and 'chain' iteration respectively. The functions of each of these iterations are explained below in detail.

The first m iterations serve as the automatic initialization of cluster centers in case they are not input. The number of iterations, m , is determined such that p percent of all the existing clusters is of nominal sizes. After this initial start of m iterations, the program enters its intermediate stage. In this stage, cluster structures are modified by appropriately combining clusters that are close to each other, and by appropriately splitting large sized clusters into nominal sized subclusters. The value of n is such that $m+2n+1 = \text{ISTOP}$ or $m+2n = \text{ISTOP}$. The final C_h iteration searches for clusters that should be linked together, and prepares for the symbol assignment in the output of the cluster map.

While complete program documentation can be found in (Minter, 1972), an informal presentation of the ISOCLS clustering procedure is as follows:

Procedure

0. SETUP and DATAIN. If initial cluster centers are input, go to Step 2; otherwise, continue.
1. Automatic initialization of cluster centers:
 - 1.1 CLASSify, STATistics and DELETE.
 - 1.2 SPLIT.Repeat Steps 1.1 - 1.2 as necessary.
2. Intermediate cluster structure changes:
 - 2.1 CLASSify, STATistics and DELETE.
 - 2.2 Alternately COMBINE and SPLIT.
3. Final iteration: CHAIN.
4. OUTPUT.

Descriptions of the subprograms are as follows:

SETUP and DATAIN : The values of the various input parameters are read in and default values are set up. Data is read into core (and into other memory units as necessary).

CLASS and STAT : Data points are assigned to the cluster centers retained from the previous iteration. Assignment of points is according to the minimum distance criterion, using a l_1 distance as defined in (1). Also, the mean and standard deviation of the newly formed clusters are iteratively computed.

DELETE : When a cluster contains less than NMIN data points, NMIN being an input parameter, the cluster center is removed.

SPLIT : A cluster that is larger than the nominal size is split into two subclusters, a cluster being considered to be of nominal size when the standard deviation in every channel is less than the threshold STDMAX. The two new cluster centers just created will be at $(\mu_1, \mu_2, \dots, \mu_k + \sigma_k, \dots, \mu_n)$ and $(\mu_1, \mu_2, \dots, \mu_k - \sigma_k, \dots, \mu_n)$, where $(\mu_1, \mu_2, \dots, \mu_n)$ and $(\sigma_1, \sigma_2, \dots, \sigma_n)$ are the mean and the standard deviation of the original cluster, and where the k^{th} channel is that in which the original cluster has the largest standard deviation.

COMBINE : If the distance $D(.,.)$, as defined in (2), between two clusters is less than the threshold DLMIN, the two clusters are combined into one. That is, the original two cluster centers are removed, and replaced by a single new cluster center at a weighted mean of the two original clusters.

CHAIN : In the last iteration, all clusters that are close to each other are linked together, as explained in Section III and Figure 2. The subprogram that performs this operation is included in the Appendix.

OUTPUT : After the last iteration, i.e. the CHAIN iteration, the program outputs the statistics of all the subclusters that exist PRIOR to the chaining operation. The statistics of the subclusters are the number of data points assigned to them, their mean vectors and covariance matrices. However, in the display map, those points belonging to subclusters that have been chained together are represented by the same symbol.

V. EXERCISING THE ISOCLS PROGRAM

The following discusses two special cases of ISOCLS, how to interpret the cluster results and how best to set the values of the input parameters.

A Special Case: NSCLAS The ISOCLS program has a user-input parameter, MAXCLS, which governs the maximum number of clusters permitted to be generated. During a 'split' iteration, if the MAXCLS limit has been reached, a cluster of size larger than the nominal size would not be allowed to be split. The program simply proceeds to reassign the points in this particular iteration. Recall the form of the iteration sequence:

$$\underbrace{SS \dots S}_m \underbrace{(CS)(CS) \dots (CS)}_{2n} C_h \quad .$$

If, in the initial splitting stage, MAXCLS is reached before the program switches to the CS intermediate operation mode, the program would actually take the following sequence of actions:

$$\underbrace{SS \dots S}_m C_h \quad .$$

ISTOP

In this way, the program essentially attempts to best assign the data points to MAXCLS clusters. This is basically the same philosophy taken by the clustering program NSCLAS at Purdue University (Wacker, 1970).

To achieve this goal, the user can assign to the program the desired value of MAXCLS, and should input a very small value to STDMAX, e.g. zero. This assures that the MAXCLS limit will be reached, and the program will remain in the initial splitting mode. The same effect also be attained by inputting MAXCLS initial cluster centers and setting DLMIN to a small value, e.g. zero. In this way, a 'C' iteration reduces to a mere reassignment of points.

Another Special Case: ISODATA By skipping the initial mode, and by inputting an initial set of cluster centers, the program operates in the following sequence:

CSCS...CS C_n

This is the mode in which ISODATA (Ball, 1966) operates.

Interpretation of Cluster Results As mentioned earlier, ISOCLS outputs the statistics of the individual subclusters, even if some of these subclusters are chained together. These subclusters would have nominal sizes if the MAXCLS limit is not reached. Recall that the nominal size of a cluster depends on the threshold STDMAX, which the user can input. Hence, the clustering results obtained from the ISOCLS program should be interpreted with this particular understanding. That is, even if a 'homogeneous' field (e.g. agricultural field) is clustered into several clusters that have not been chained together, the user should interpret that, indeed, that field is the union of several subclusters. He should by no means be dissatisfied with the outcome of the clustering results. As a matter of fact, if the field is indeed spectrally homogeneous, and is representable by some compact distribution, it is likely that the subclusters so obtained are chained together if the proper DLMIN threshold is used. In other words, if the user fully comprehends how he is exercising the program, he should be able to interpret his results, and satisfactorily too.

Setting of Parameter Values Three input parameters are of particular interest, namely DLMIN, STDMAX and MAXCLS.

The role of DLMIN has been thoroughly discussed above. A value of 3.2 for DLMIN has been suggested in connection with the probability of misclassification. Values of DLMIN lower than 3.2 give finer groupings, while higher values of DLMIN give coarser groupings. In general, values of DLMIN outside the range 2.9 -3.6 are discouraged for normal applications.

The value of STDMAX governs the nominal size of a cluster. As a rule of thumb, if it pleases the human eye to consider a statistical class to be a good compact class if the class distributes itself over n units of measurement, STDMAX should be set at roughly n/3. This rule of thumb is suggested because, with a probability of 0.9, a Gaussian distribution takes values within ± 1.5 standard deviation from the statistical mean of the distribution. Experience with aircraft agricultural data, like the C1 Flightline data over Indiana, indicates that the data values of a statistically homogeneous field varies over 10-15 units (in a full range of 256 units), and thus suggests a choice of value 4.0-4.5 for STDMAX.

MAXCLS limits the maximum number of clusters to be generated by the program. For normal applications, this value is set free; i.e. the program is allowed to generate as many subclusters as is necessary within the system's limit. For special applications, MAXCLS and the other two parameters DLMIN and STDMAX can be manipulated to induce the desired results. This will be further discussed in Section VI.

VI. APPLICATIONS: INVESTIGATION ON ERTS-1 DATA FOR RANGE VEGETATION MAPPING

In the Summer of 1972, the ERTS-1 (Earth Resources Technology Satellite) Application project was initiated for the investigation of the utility of ERTS satellite data, which are collected by various sensors. Among these sensors is the multispectral scanner (MSS), which collects four channels of data, with a spectral range from 0.5 micron to 1.1 micron. Several analysis teams were formed to study specific features on Earth, as observed by ERTS-1. One such team was the Range Analysis Team*. The objective of this team (Parker, 1973) was to estimate the utility of ERTS-1 MSS data for mapping the various range vegetation types in the Houston Area Test Site of Texas. In broad generality, range vegetation types include grassland, shrubland, pasture, wooded and noncommercial forested areas, wetland etc.

In particular, the Range Analysis Team chose several study sites to investigate in the ERTS-1 project, one of which was the 'Snook site'. This is north-east of Somerville in Burleson County, Texas. This site is approximately 300 km². It lies astride the ecotonal boundary between the Texas blackland prairie region and the Texas claypan savannah. The study site is characterized by fairly extensive stands of post-oak, bottomland hardwood, mesquite, native grassland and manipulated pastures of bermuda grasses, the totality of which makes up the rangeland category in the study site, as distinguished from other features such as agricultural fields and water, e.g. Lake Somerville. The scope of the investigation was to perform vegetation type mappings into the first hierarchical as well as the second hierarchical levels.

For this purpose, ISOCLS was exercised in various ways to achieve classification of the Snook study site. In fact, there are four approaches to classification:

(a) Normal mode: The program ISOCLS is allowed to generate as many subclusters of nominal sizes as is necessary. DLMIN is set at 3.2 as recommended earlier. STDMAX is set between 2.5 and 3.5. (ERTS-1 data take discrete values from 0 to 127.) This gives rise to a second order classification.

(b) For a coarser second order classification or even a first order classification, the user inputs the number of classes he wishes to see by setting MAXCLS to that number. Also, STDMAX is set to a larger value than that which gives rise to nominal sized clusters. In addition, since a mere assignment of data points to MAXCLS classes is desired, DLMIN is set to a small value, e.g. zero.

(c) A first order classification based on spectral similarity can be obtained, by manipulating the results obtained in (b). That is, clusters that are of distance less than 3.2 are grouped together either automatically in the chaining iteration of the program, or manually during the display of the output results.

(d) Another first order classification based on biological and ecological similarity can be obtained as follows. First, the results of the clustering runs as obtained in (a) or (b) are interpreted. Then, clusters that correspond to classes of the same or similar vegetation types (or eco-systems) are grouped together to give a first order classification.

* The team members of the Range Analysis Team were J.M. Disler, E.P. Kan, H.D. Parker, Jr., A. Simon of Lockheed Electronics Co., Inc., HASD, Houston, Texas, and D. Pendleton of the Soil Conservation Service, USDA.

One such cluster map is shown in Figure 4. This is the result of applying ISOCLS to the August 30, 1972 ERTS-1 MSS data on the Snook study site. The GSFC (Goddard Space Flight Center) identification number of the raw data tape is 1038-16303-MB-3. The study area corresponds to scan line 651 to 1000, resolution element 1841 to 2190. Out of the 25 clusters obtained in a second order classification following approach (a), 10 clusters were considered major (by population) and were colored as shown. Interpretation of these clusters is found in Table 2. More cluster maps and their interpretation can be found in (Kan, 1973).

A comparison of these interpreted classification results with ground-truth information on the Snook study site revealed that the classification results are indeed satisfactory. No attempt was made to derive a quantitative measure of the effectiveness and accuracy of the classification results because of the lack of an adequate set of compatible ground-truth measurements and the shortage of time in the analysis team effort.

VII. CONCLUSION

The JSC clustering program ISOCLS is thoroughly discussed in this paper. In particular, the philosophy of operation of this clustering procedure is emphasized in that classes are treated as the unions of appropriate numbers of subclasses. Also, chaining of neighboring subclasses is performed at the end of the clustering procedure in order to discover the eye-pleasing cluster structures that exist in the set of statistical data. Furthermore, the usage of the inter-cluster distance measure is carefully discussed which relates to the similarity and thus probability of misclassification between clusters.

Applications of the ISOCLS clustering program to achieve classification of ERTS-1 and non-ERTS data have been discussed in this paper and elsewhere. A strong feel for the effectiveness of the procedure has evolved from these investigations. The general consensus is arrived at that the ISOCLS clustering program is a very useful and dependable tool for the pattern analysis and classification of statistical data such as those collected by remote sensors.

VIII. REFERENCES

1. G.H. Ball and D.J. Hall, "ISODATA, an Iterative Method of Multivariate Analysis of Pattern Classification," Proceedings of the International Communication Conference, Philadelphia, Pa., Jun. 1966.
2. E.P. Kan, "Data Clustering: An Overview," Lockheed Electronics Co., Inc., HASD, Houston, Texas, Tech. Rep. 640-TR-080, Mar. 1972.
3. E.P. Kan, "The JSC Clustering Program and Its Applications," Lockheed Electronics Co., Inc., HASD, Houston, Texas, Tech. Rep. 640-TR- , Jul. 1973.
4. T. Marill and D.M. Green, "On the Effectiveness of Receptors in Recognition Systems," IEEE Transactions on Information Theory, Vol. IT-9, pp. 11-17, Jan. 1963.
5. R. Minter, "ISOGLS, an Iterative Self-Organizing Clustering Program: Computer Program Documentation," Lockheed Electronics Co., Inc., HASD, Houston, Texas, Computer Program Documentation CPD 202, Oct. 1972.
6. G. Nagy and J. Tolaba, "Nonsupervised Crop Classification Through Airborne MSS Observations," IBM Journal of R&D, Mar. 1972.
7. H.D. Parker, Jr., J. Disler and E.P. Kan, "Automatic and Manual Vegetation Mapping with ERTS-1," Lockheed Electronics Co., Inc., HASD, Houston, Texas, Tech. Rep. 640-TR- , Jun. 1973.
8. A.G. Wacker and D.A. Langrebe, "Boundaries in Multispectral Imagery by Clustering," 1970 IEEE Symposium on Adaptive Processing (9th), Proc., Decision and Control, pp. X14.1-X14.8, Dec. 1970.

IX. APPENDIX

CHAIN: This subprogram chains into one cluster all those subclusters that are close to at least one other subcluster in the group. Two subclusters are said to be close if they are within DLMIN distance apart, as discussed in COMBINE. This subroutine creates an index array P(I), I=1,2,...,NCLUSTR from the table D(I,J) of intercluster distances, NCLUSTR being the current number of existing subclusters. The end product of this subroutine will be a converged index array. Subclusters that are to be chained together (and thus considered to belong to one composite cluster) will have the same index value:

0. Form table D(I,J) and set P(I)=I, i=1,...,NCLUSTR. Then, set I=1.
1. Find all J's, J>I, such that D(I,J)<DLMIN. If no such J exists, go to Step 3.
2. For all these J's, set P(I) and P(J) to min[P(I),P(J)].
3. Increment I by one (1). If I<NCLUSTR, go to Step 1; otherwise, continue.
4. If there is a change in P(I), for some I, in this pass through the table D(I,J), set I=1 and go to Step 1; otherwise, return.

For an illustration of this procedure, it is suggested that the interested reader execute the chaining procedure for the example shown in Figure 5.

Table 1

Quantities Related to the Two Clusters which are Separated from
a Composite of Two Normal Distributions

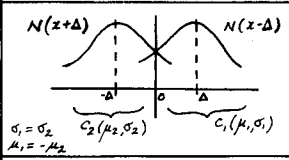
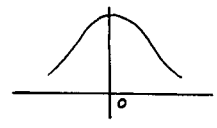
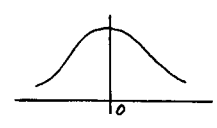
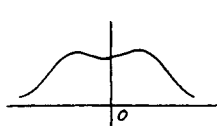
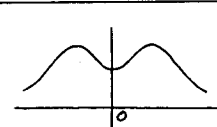
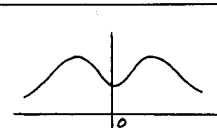
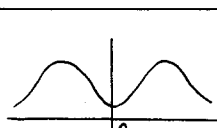
2Δ	$p(x) = N(x+\Delta) + N(x-\Delta)$ $N(x) \equiv \frac{1}{\sigma} \pi_{\sigma,1}(x)$	μ_1	σ_1	$D(C_1, C_2)$	P_e
		$\int_0^{\infty} x N(x+\Delta) dx$ $+ \int_0^{\infty} x N(x-\Delta) dx$	$\left[\int_0^{\infty} x^2 N(x+\Delta) dx \right.$ $\left. + \int_0^{\infty} x^2 N(x-\Delta) dx \right. - \mu_1^2 \left. \right]^{1/2}$	$\frac{\mu_1 \mu_2}{\sqrt{\sigma_1 \sigma_2}}$ $= 2\mu_1 / \sigma_1$	$\int_0^{\infty} N(x+\Delta) dx$ $+ \int_{-\infty}^0 N(x-\Delta) dx$
1.5		1.09	0.74	2.78	0.23
2.0		1.17	0.80	2.90	0.16
2.5		1.35	0.86	3.13	0.11
3.0		1.56	0.91	3.42	0.07
3.5		1.78	0.94	3.80	0.04
4.0		2.02	0.97	4.14	0.02

Table 2

Interpretation and Statistics of the Clusters in Figure 4

Cluster Color	light green	dark green	dark brown	medium green	turquoise	red	blue	yellow	light brown	off-white
Statistics	mean st. dev.	mean st. dev.	mean st. dev.	mean st. dev.	mean st. dev.	mean st. dev.	mean st. dev.	mean st. dev.	mean st. dev.	mean st. dev.
Channel 1	32.4 1.5	35.6 1.6	29.7 1.5	32.8 1.9	33.9 2.2	32.1 1.3	27.5 1.8	32.4 1.9	38.0 1.8	39.4 2.2
Channel 2	25.4 1.9	30.4 2.1	21.3 2.0	25.1 2.3	29.0 3.1	21.2 2.7	17.3 1.9	23.5 2.5	33.2 2.7	38.7 2.8
Channel 3	41.3 2.3	44.0 2.3	36.8 2.8	48.0 2.4	33.2 3.5	79.7 4.2	13.6 4.2	56.8 3.3	50.4 2.5	44.3 4.2
Channel 4	22.2 1.5	23.2 1.6	20.5 1.5	26.4 1.6	17.5 3.3	46.1 2.1	4.2 2.6	32.0 2.3	26.3 1.6	22.5 2.6
Population	30149	16998	31096	13838	7742	1319	2115	5784	4726	2990
Interpretation	open range, grassland	abandoned cropland; open range, sparsely vegetated	woods; post-oak	open grassland; bermuda fields, greener than the light green cluster	wet bottomland (non-wooded)	healthy agricultural fields, very vigorous	water (Lake Somerville)	agricultural fields, not as vigorous as the red cluster	grassland, sparse and dry	bare soil; asphalt

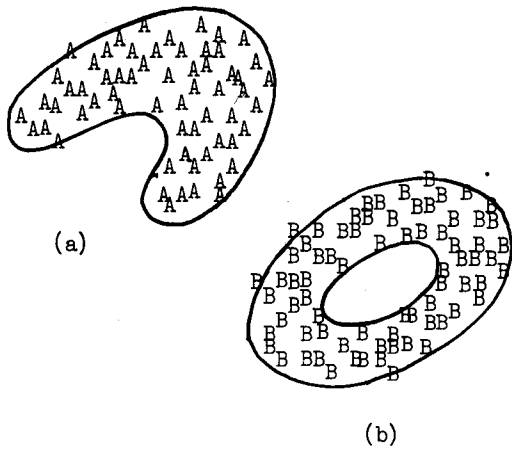


Figure 1
 (a) The Boomerang-shaped Cluster
 (b) The Donut-shaped Cluster

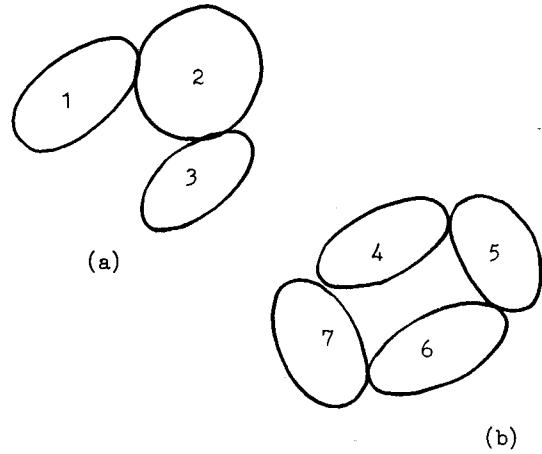


Figure 2
 Breaking Up of the Clusters (a) and (b)
 of Figure 1 into Subclusters

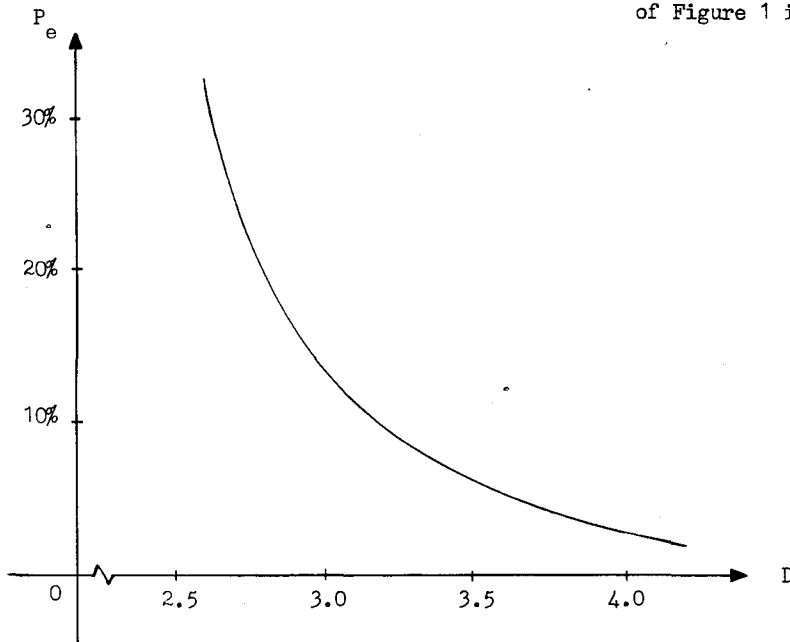


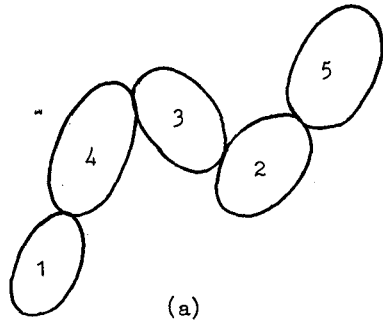
Figure 3
 Relation Between the Probability
 of Misclassification P_e and the
 Intercluster Distance e Measure D

NASA
S-73-28027



Figure 4

ISOCLS Classification Map of 8/30/72 ERTS-1
MSS Data on Snook Study Site (Image ID 1038-
16303-MB-3, Scan 651-1000, Pixel 1841-2190):
10 major clusters



(a)

	I					
	J	1	2	3	4	5
1	0.0	7.5	6.2	3.2	11.8	
2	7.5	0.0	3.1	5.6	3.0	
3	6.2	3.1	0.0	3.1	6.3	
4	3.2	5.6	3.1	0.0	9.7	
5	11.8	3.0	6.3	9.7	0.0	

D_{MIN} = 3.2

(b)

Figure 5

Example for Chaining Procedure

(a) Cluster Structure

(b) Intercluster Distance Table D(I,J)