UNIQUE DOCUMENT #SAC200122590000 **Enclosure 3**

UNCLASSI[redacted]

RS 0310/85/268
November 1, 1985

Systems Research Report

A Method for Automated Recognition of Satellite Radiometric Waveforms (U)

Charles V. Jakowatz, Jr.
Paul A. Thompson
Sandia National Laboratories
Albuquerque, New Mexico

November 1, 1985

DEPARTMENT OF ENERGY DECLASSIFICATION REVIEW	
1ST REVIEW-DATE: 6/16/97	DETERMINATION (CIRCLE NUMBER(S))
AUTHORITY: ☒ DC ☐ DD	1. CLASSIFICATION RETAINED
NAME: Zan Hohgen	2. CLASSIFICATION CHANGED TO:
	3. CONTAINS NO DOE CLASSIFIED INFO
2ND REVIEW-DATE:	4. ~~COORDINATE WITH: DTRA~~
AUTHORITY: DD	5. CLASSIFICATION CANCELED
NAME: Ronald L. Williams	6. CLASSIFIED INFO BRACKETED
	7. OTHER (SPECIFY): DTRA CONCURS W/ DOE BRACKETS 6/19/00

Albuquerque Operations Office memo dtd 1/30/98

Sandia National Laboratories

1221·01

RP

THIS DOCUMENT CONSISTS OF 30 PAGES.
COPY NO. 00018 OF .
SERIES A

V1B / DRP
0094321 01

UNCLASSIFIED

[redacted]

SECRET

98SA20C000017

[redacted] UNCLASSIFIED

RS 0310/85/268
November 1, 1985

Approved by: [signature]
P. A. Stokes, 0310

A Method for Automated Recognition of Satellite Radiometric Waveforms (U)

[redacted]

Charles V. Jakowatz, Jr.
Paul A. Thompson

Sandia National Laboratories
Albuquerque, New Mexico
November 1, 1985

[redacted] UNCLASSIFIED

[redacted] UNCLASSIFIED
RS 0310/85/268
November 1, 1985

Abstract

DOE
b(1)

The results of this work were presented at the 1982 Satellite Data Working Group meeting. This paper describes the application of similar techniques to RADEC data, pursuant to the implementation of automated event classification algorithms for AFTAC.

(U) There are several major processing steps that comprise the operational event classification procedure. The first of these is establishing a standard, satellite–independent representation for bhangmeter time histories, which will allow data from various systems to be processed with a common set of decision algorithms. A resampling scheme selected with this in mind serves as the first stage of the signal processing package to be described here. The second element is a simple Euclidean distance thresholding process that allows 'obviously uninteresting' events to be easily eliminated. Next, unsupervised learning (clustering) methods are applied to the data base so that the underlying substructure of the class of all non-nuclear events may be understood. Following this, a means for constructing a feature space in a relatively small number of dimensions is required. Here, this is accomplished by the method of Fisher for optimum linear feature selection. Finally, a Bayes minimum risk decision theoretic scheme is designed for final classification of data that have been transformed into the lower–dimensional feature space.

[redacted] UNCLASSIFIED

Introduction

(S) In this report a method is described for the automated recognition of satellite radiometric waveforms.

(U) The methods employed in this report are derived from the field of statistical pattern recognition. In this section, the basic idea of approach is explained, since it is different in several respects from past approaches to the same problem. It is then shown how the radiometric signature identification problem fits into this framework. In subsequent sections, the details of the implementation for actual satellite data are presented, along with an analysis of the performance of the resulting algorithms.

(U) The basic framework for problems that are solved by statistical pattern recognition is shown in Figure 2. It is imagined that there are a finite number of possibilities that the source can produce on any given trial of the basic experiment. These possibilities are usually called *hypotheses*, and are denoted $H_1, H_2, \ldots H_N$.

RS 0310/85/268
November 1, 1985

DOE
b( )

RS 0310/85/268
November 1, 1985

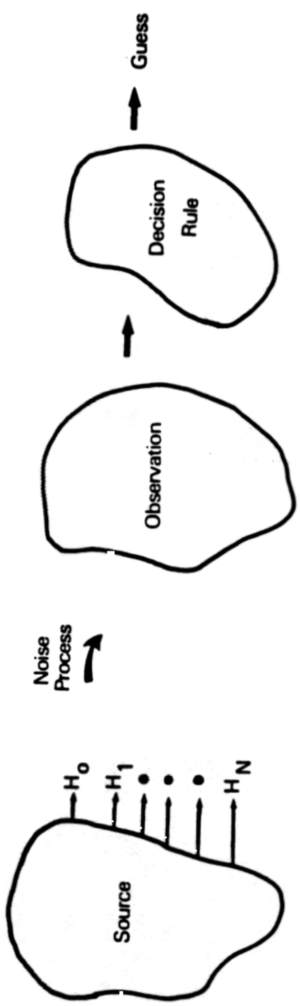

Figure 2 (U)

The observer is not allowed to see the output of the source directly. Instead, he obtains an observation which is related to the source output, but which has a random, or stochastic, component to it. This randomness is introduced by a *chance mechanism*, or *noise process*. The goal of pattern recognition is to design a decision rule that will map any given observation into a *guess* as to which of the possible source hypotheses actually produced the given observation. The *best* decision rule will usually be that which produces the fewest number of incorrect guesses, averaged over all possible observations. The design of such an optimum rule must take into account a description of the noise process. Sometimes this will be known from basic physical principles, but often it must be empirically derived from a controlled set of observations.

(U) A simple example of a familiar problem that fits easily into this framework is that of the communication system depicted in Figure 3. In this case, the source is the sender of a message to a receiver, which is the observer. There are a variety of messages that the sender can choose to convey to the receiver, and these messages constitute the source hypotheses. A different shaped *signal* that will be transmitted over a communications channel is used to encode each of the messages. In the real world, all such channels are *noisy*, in the sense that they will introduce some amount of random distortion into the signal as it is being transmitted. This distortion, then, comprises the noise process of Figure 2. The decision rule in this case would be a signal processing scheme that uses samples of the received waveform as input, and produces as output the best guess as to which message was intended by the sender. Again, the exact design of the best (minimum probability of error) rule would necessitate knowing, either from theoretical considerations or from experimental data, a mathematical description of the noise process.

(U) The problem of automated (machine) recognition of satellite radiometric waveforms can readily be addressed in terms of the above framework. The various possible source outputs (hypotheses) are the different kinds of physical events that can produce radiometric signatures. These events include, to name only a few, lightning bolts, sun glints from lakes, charged particles impacting on the radiometers, and detonations of nuclear devices in the atmosphere. The observation, of course, is the set of digitized samples from the from the collected

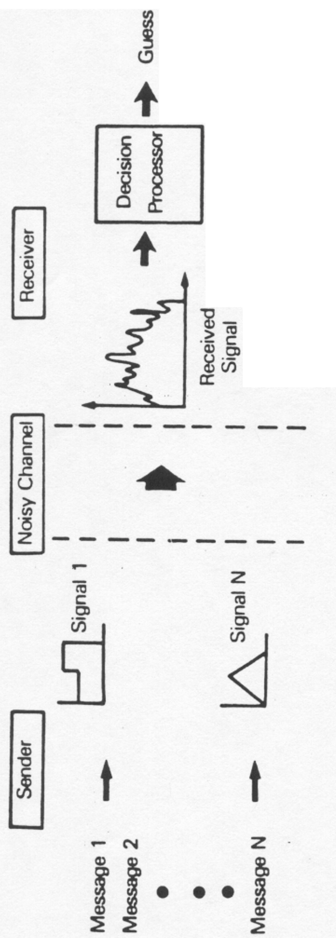
Sender
Noisy Channel
Receiver
Message 1
Message 2
Message N
Signal 1
Signal N
Received Signal
Decision Processor
Guess

Figure 3 (U)

radiometric signature. The noise process is a complicated collection of processes that, for example, make different individual lightning bolts have different signatures at the satellite, and that result in every nuclear detonation having a different signature. Thus, some of what makes up this noise process really represents the observer's lack of precise knowledge as to the nature of each physical event, e.g., the exact design parameters of a nuclear device that was detonated, or the exact structure of the cloud masses that produced a lightning bolt. Also contributing to the noise process would be such items as random distortions introduced into the radiometric waveform by fluctuating atmospheric conditions. A partial list of still other sources of randomness would include electronic noise in the sensor circuits and possible errors introduced into the digital data as it is transmitted back to the earth station. *The central idea of this analysis is to characterize all of these components of the noise process not from any theoretical physical considerations, but rather simply from the study of a large amount of collected satellite radiometric data.* This should be contrasted with most, if not all, past approaches to this problem, which have attempted to see how theoretical models of nuclear detonations would fit, or fail to fit, observed radiometric collections. The real disadvantage of such past analyses is that they treated only one type of event as a valid hypothesis. That is, no other classes of events except NUDETS were used as possible models for an unknown event. The weakness in this kind of one–hypothesis approach is that it leads to statements concerning the observer's *confidence* in his assessments that may not be totally meaningful. This is because one can only state that a given event appears to be consistent or not consistent, at some confidence level, with the hypothesis that the event was generated by a NUDET. But one cannot calculate the *a posteriori* proability (i.e., the conditional probability given the observation) that the event was generated from a NUDET, *unless one has other specific hypotheses which can also be evaluated.* That is to say, the second of the following two statements, in the authors' opinion, is to be greatly preferred: 1) The unknown event is consistent at the 80 per cent confidence level with the hypothesis that it was generated by a NUDET (but, in fact, any number of other hypotheses could also be consistent at this or an even greater level of confidence); and 2) The unknown event is identified as NUDET with 80 per cent probability of correct classification, when the classes of NUDETS, lightning bolts, and zoo events are considered to be *a priori* equally likely.

(S) The construction of theoretical models for all possible kinds of triggers would, of course, be an extremely difficult, if not impossible, task. It is chiefly for this reason that the authors of this work have chosen the approach of using empirical data to construct statistical models for various event classes.

DOE

b\

RS 0310/85/268
November 1, 1985

Statistical Decision Theory Applied to Radiometric Waveform Classification

A Sensor–Independent Sampling Scheme

(U) The first step in designing a decision theoretic classification scheme for radiometric waveforms is to design a sampling scheme that will place data from any one of the several types of bhangmeters that are currently deployed (or future types that will be deployed) on some common basis. Various bhangmeters in existing and planned satellite systems differ considerably in sampling rate, trigger criteria, and quantization levels. The resampling scheme described below is logarithmic in both amplitude and time and is scaled relative to a common trigger level.

DOE
b(1)

(U) The resampling is performed at the constant logarithmic rate of eight samples per decade over five decades of time. Since the exact beginning of the optical time history is unknown, the samples are arbitrarily shifted in time so that the level-eight threshold crossing occurs at the 100-microsecond point on the time axis. Two pre-trigger samples are combined with 40 post-triggger samples to form a 42-sample representation of the time history. These 42 samples may be thought of as the components of a 42-dimensional vector representing the time history.

[redacted] UNCLASSIFIED

RS 0310/85/268
November 1, 1985

DOE
b(1)

[redacted] UNCLASSIFIED

1.2

[redacted] UNCLASSIFIED

RS 0310/85/268
November 1, 1985

DOE
b(1)

[redacted]

UNCLASSIFIED

3

Pre–Processing to Eliminate Uninteresting Events

(U) The total number of radiometric waveforms collected by the VELA and RADEC sensors over the past decade is enormous, and numbers in the hundreds of thousands. A significant portion of these waveforms are rather *uninteresting*, in the sense that they were caused by well–known phenomena and have essentially no similarity to waveforms caused by nuclear detonations. An example of such a class of waveforms is that of *particle triggers.* These events are generated when certain types of cosmic ray particles impact on the radiometric sensor element and trigger the system. There is a relatively easy way to cull out these as well as certain other types of events so that the job of discriminating between nuclear detonation waveforms and interesting non–nuclear events may be more readily addressed. This amounts to simply computing the Euclidean distance from a given event to the class of known nuclear events, i.e., the *training set* of nuclear events, and then rejecting those events that are sufficiently far away by this measure. This simple idea is depicted in Figure 5. Thresholding the Euclidean distance at level R amounts to retaining only those events which fall *inside a hypersphere of radius R.* The figure only shows three dimensions, but should be thought of in the number of dimensions of the signal vector space. In this case, the dimension would be 42 (or 84 for two channel concatenated data) because of the resampling scheme discussed in the previous section.

DOE
b

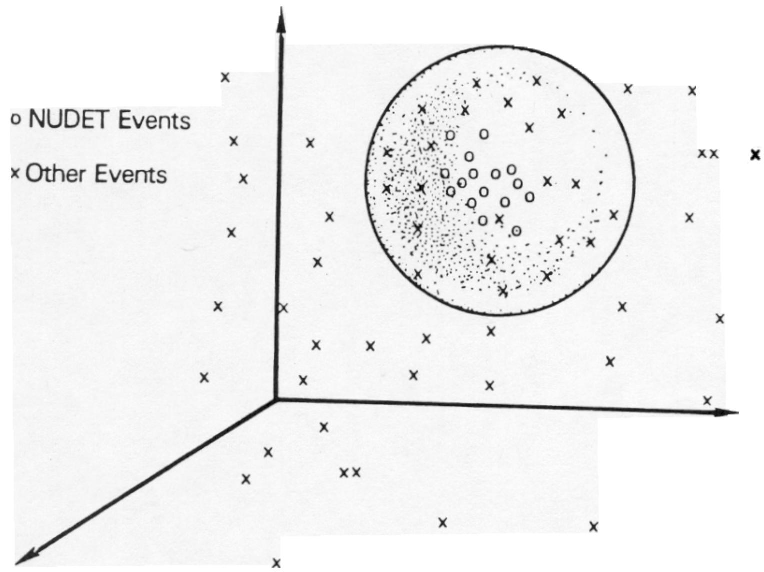

Figure 5 (U)

DISTANCE HISTOGRAMS

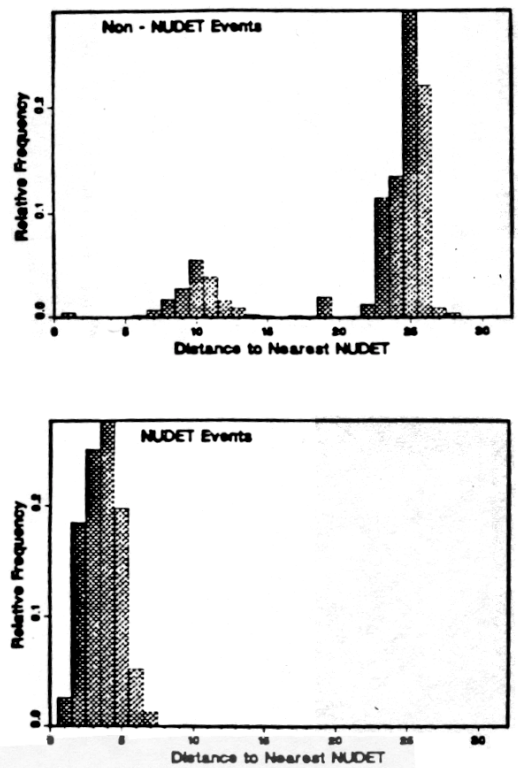

Figure 6 (U)

Clustering to Depict Underlying Substructure of Non-Nuclear Class

(U) Once a data base of non-nuclear events that are reasonably close to the class of known nuclear events has been established, it is important to attempt to understand the substructure these events may possess. The reason this is important is because it is usually more difficult to construct a statistical description for a class of events if the distribution of the events in that class is multimodal, as opposed to unimodal. For example, a very useful distribution that often fits real–world problems well is the Gaussian, or normal, distribution, which is unimodal. When a multimodal distribution can be approximately broken up into component unimodal functions, a set of Gaussian distributions is often useful in describing the class.

(U) One method for discovering the modes of a multimodal distribution in many dimensions (e.g., 84 dimensions for samples of a pair of RADEC waveforms) is that of hierarchical clustering. The fundamental idea of this technique is depicted in Figure 7. This shows how data in two dimensions can be grouped by clustering into three classes, which would seem to be the 'correct' substructure for this particular data set. The clustering algorithm starts by computing the distance between all possible pairs of data points. It then orders this list of distances from smallest to largest. Starting with the two samples for which the inter–sample distance is smallest, the routine *links* these two samples into the same *cluster*. Each such link is denoted in the figure by a solid line drawn between the corresponding two points. This kind of linking could proceed until all points are linked into precisely one large cluster. Clearly, one large cluster would reveal nothing about data substructure, and so the linking process must be terminated at some point before this happens. The usual method for determining when to cease linking in this type of clustering is to examine the sequence of link–distances as the algorithm proceeds. When a 'natural' substructure is approached, the link distance should exhibit a substantial jump. This is evident from examinination of Figure 7. Note that the links corresponding to the two dotted lines, which are merges that should not be performed, represent substantially larger distances than do the earlier links.

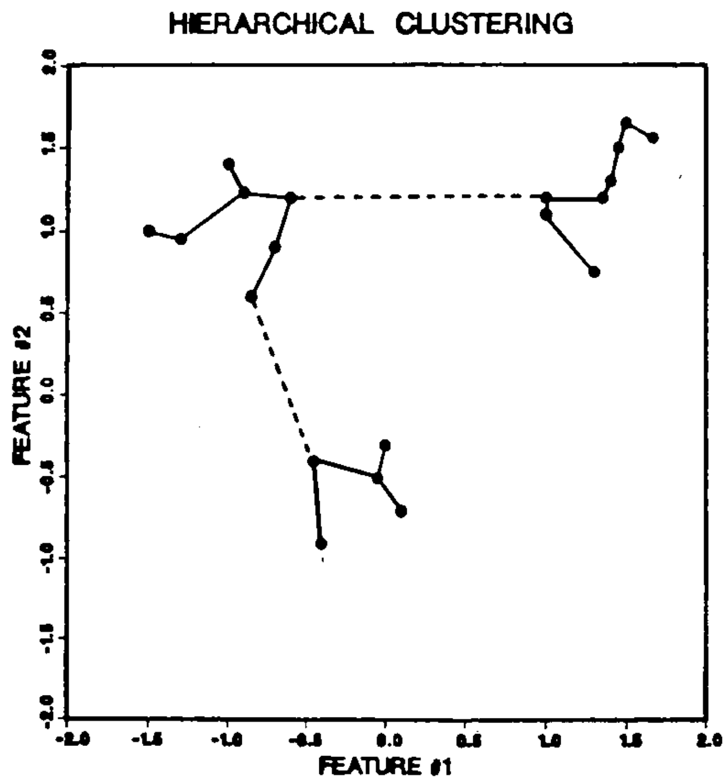
HIERARCHICAL CLUSTERING
FEATURE #2
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
FEATURE #1

Figure 7 (U)

[redacted] SS. ED

be

b( )

[redacted] JNC ASSIF FD

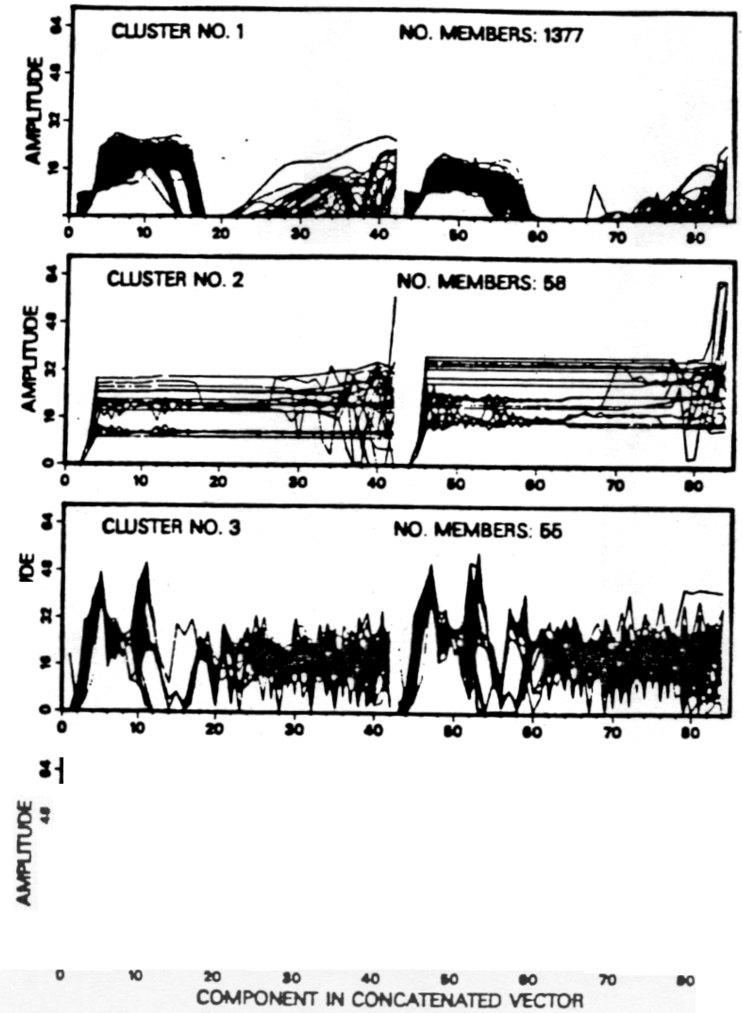
CLUSTER NO. 1
NO. MEMBERS: 1377
CLUSTER NO. 2
NO. MEMBERS: 58
CLUSTER NO. 3
NO. MEMBERS: 55
AMPLITUDE
COMPONENT IN CONCATENATED VECTOR

Figure 8 (U)

Feature Selection and Classification

(S) With the set of non-NUDET events broken down into its natural substructure, the final phase of constructing an automated processing algorithm is to design a pattern recognizer that will categorize an unknown collection as either NUDET, or one of the non-NUDET subclasses. This is accomplished via the formalism of statistical decision theory. Briefly, a minimum probability of error classifier is constructed as follows. For each of the N hypotheses, the *a posteriori* probability $P(H_j \mid \Re)$ is computed. This number is the conditional probability that the j^{th} hypothesis is correct, given that the observation $\Re$ was made. The hypothesis for which the *a posteriori* probability is greatest is the one chosen. This maximum *a posteriori* processor can be shown to be that which produces the minimum probability of mis-classification [4]. The chief problem in attempting to compute this probability from an empirical data set is that the number of training samples (events) that are required to establish the function grows with the number of dimensions of the observation vector. For example, for a multi-dimensional Gaussian statistical model, the mean vector and covariance matrix must be calculated to specify the *a posteriori* probability function. If the number of dimensions in the data is M, then the minimum number of training samples (events) required to estimate the covariance matrix is roughly 10 times M.

(U) When the problem of dimensionality reduction is considered, it must be understood that any such process, by its very nature, *destroys information*. That is, the theoretical performance of an optimum classifier in *any lower dimensional space* can be no better than that of the best classifier in the original data space. Therefore, the feature selection process must be viewed as *undesireable* from the standpoint of loss of possible discriminatory information, but *necessary* from the standpoint of producing data that can be processed in a meaningful way. The critical question, then, is exactly how should the features be chosen so that the amount of discriminatory information that is thrown away is minimized. The problem is extremely difficult if all possible forms of maps from M dimensions down to L dimensions ($L < M$) are considered. However, R. A. Fisher [2] demonstrated that if the maps are constrained to be linear, then the feature selection problem can be solved in closed form. Fisher's solution involves computing a ratio of a *between-class distance* to a *within-class* distance in the lower-dimensional space. The particular linear transformation of the data that *maximizes* this ratio is the desired one. The idea is illustrated in Figure 9 for a case in which the original data dimensionality is 2, and the desired feature space is one-dimensional. The linear transformation that would map the data onto line L_2 is not very effective since the two-class data are greatly overlapped. The linear transformation that projects the original samples onto line L_1, however, is a very effective one, as the resulting one-dimensional data are well-separated.

DOE
b(1)

[redacted]

r be

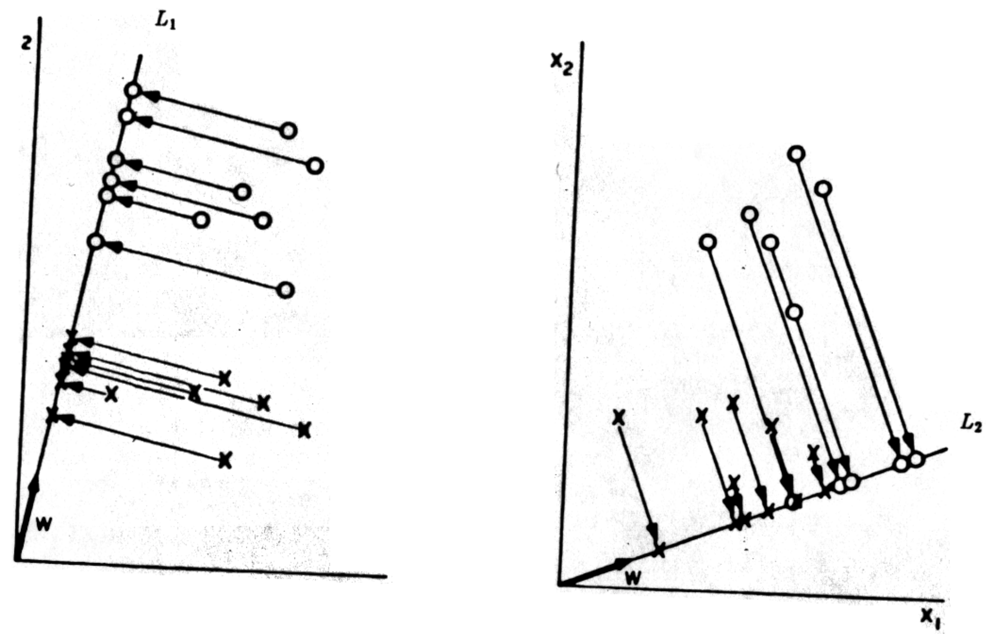

Figure U) (taken fr Duda and Hart[1 page 5)

RESULTS FROM 2-DIMENSIONAL FISHER ANALYSIS

DOE
b(1)

Figure 10 (U)

TABLE I (U)
Classification Error Matrix
Double Events – Two dimensions

Actual class:	Assigned class: 1	2	3	4	5*
1	490	782	0	105	0
2	11	43	0	4	0
3	0	0	55	0	0
4	38	34	2	106	3
5*	0	0	0	0	76

* NUDET Class

TABLE II (U)
Classification Error Matrix
Double Events – Four dimensions

Actual class:	Assigned class: 1	2	3	4	5*
1	1330	0	0	47	0
2	0	55	0	3	0
3	0	0	55	0	0
4	36	5	2	140	0
5*	0	0	0	0	76

* NUDET Class

TABLE III (U)
Classification Error Matrix
Single Events – Five dimensions

Actual class:	Assigned class: 1	2	3	4	5	6*
1	97	0	0	0	121	0
2	0	57	0	0	1	0
3	0	0	110	1	2	4
4	0	0	24	775	19	7
5	1	40	14	99	408	20
6*	0	1	0	0	0	94

* NUDET Class

[redacted] SS

RS

Nov mbe

[redacted] UNCLASS F FD

RS 0310/85/268
November 1, 1985

References

1. Duda and Hart, *Pattern Classification and Scene Analysis*, John Wiley ans Sons, Inc., 1968.
2. Fisher, R. A.,' The use of multiple measurements in taxonomic problems', *Ann. Eugenics*, 7, Part II, 179-188 (1936).
3. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, 1972.
4. Melsa and Cohn, *Decision and Estimation Theory*, McGraw-Hill, 1978.
5. Van Trees, *Detection and Estimation Theory, Part I*, John Wiley and Sons, Inc., 1968.

RS 0310/85/268
November 1, 1985

RS 0310/85/268, Series A, ~~28~~ 30 Pages, 37 Copies, SNSI, 11/1/85
Subject: A Method for Automated Recognition of Satellite Radiometric Waveforms (U)

Distribution:

1-6	–	M3283	Major R. Warren, AFTAC/TNT
7	–	M4697	F. Dickerson, DOE
8	–	M4697	R. Ewing, DOE
9	–	M4697	M. Koontz, DOE
10	–	M4697	C. Boykin, DOE
11	–	M1256	R. Alewine, DARPA
12	–	M3256	A. Conner, DIA
13	–	M4378	J. Leong, CIA
14	–	1	G. Dacey
15	–	1	I. Welber, 2
16	–	20	T. B. Cook
17	–	0300	R. L. Hagengruber
18	–	0310	P. A. Stokes
19	–	0310	J. M. Fletcher
20-24	–	0310	C. V. Jakowatz
25	–	0310	P. A. Thompson
26	–	0320	R. E. Spalding
27	–	1000	W. F. Brinkman
28	–	1600	R. G. Clem
29	–	1623	L. D. Hostetler
30	–	5000	O. E. Jones
31	–	5300	L. D. Smith
32	–	5310	H. M. Dumas
33	–	5320	G. H. Mauth
34	–	5321	D. A. Reynolds
35	–	5322	L. S. Walker
36	–	5323	J. C. Mitchell
37	–	5324	L. J. Ellis

THIS IS AN EXACT COPY:7447:1/15/98:cmg

Distribution:
MS: 0175 R. B. Craner, 7447

[redacted]
UNCLASSIFIED

UNCLASSIFIED
[redacted]