AD

BRL R 1355

BRL

REPORT NO. 1355

EXPLICIT INCLUSION OF A MIXTURE OF VARIATES AS A SEPARATE CLASS IN PATTERN RECOGNITION PROBLEMS

by

W. J. Sacco

February 1967

U. S. ARMY MATERIEL COMMAND
BALLISTIC RESEARCH LABORATORIES
ABERDEEN PROVING GROUND, MARYLAND

B A L L I S T I C   R E S E A R C H   L A B O R A T O R I E S

REPORT NO. 1355

FEBRUARY 1967

EXPLICIT INCLUSION OF A MIXTURE OF VARIATES AS A SEPARATE CLASS IN PATTERN RECOGNITION PROBLEMS

W. J. Sacco

Computing Laboratory

RDT&E Project No. 1P014501A14B

A B E R D E E N   P R O V I N G   G R O U N D,   M A R Y L A N D

B A L L I S T I C   R E S E A R C H   L A B O R A T O R I E S

REPORT NO. 1355

WJSacco/bj
Aberdeen Proving Ground, Md.
February 1967

EXPLICIT INCLUSION OF A MIXTURE OF VARIATES AS
A SEPARATE CLASS IN PATTERN RECOGNITION PROBLEMS

ABSTRACT

Previous detection studies assumed that a "target" was either present in a region of interest with a probability p, or not, with a probability (1-p). Knowing the value of p, the "searcher" makes a measurement of a random variable that has a probability density function which depends on whether or not the target is present and attempts to make a decision regarding the presence of the target. In this paper a more general point of view is adopted in that we allow the possibility of both target and non-target elements to be present in any region of measurement. This provides an explicit consideration of a third possibility, that of a mixture of target and non-target elements.

The forms of the probability densities for the "mix" variate are obtained for a variety of assumptions regarding the relationships among the variates involved. A sequential decision problem is extended to include the three value case.

TABLE OF CONTENTS

INTRODUCTION

During the last two decades, a well developed theory of detection has been established. Detection theory is an adaptation of the statistical theory of hypothesis testing to problems arising in pattern recognition and radar, communication, and control engineering. A pattern recognition device is said to consist of two parts: a receptor, which generates a set of measurements of the physical sample to be recognized, and a categorizer, which assigns each set of measurements to one of a finite number of classes.

The purpose of the present paper is to discuss a new problem in detection theory. This problem arises, among other ways, when one attempts to detect a pattern in a two-dimensional optical display (hereafter called picture). The picture is to be scanned using a receptor of a fixed size and shape. For each positioning of the receptor, one obtains a set of measurements from the "subpicture" being sensed. Each measurement, x, can be thought of as a random variable, and is sometimes called a local property of the picture. For pattern recognition problems a discriminator (human or machine) may be assigned the problem of determining whether x is a sample from one or another class of inputs, say a "target scene" or perhaps a "non-target scene". The term "target scene" (or "target") is defined to be an object of primary interest to a decision-maker.

Previous detection studies [1], [2], [3], assumed that a target was either present in a region of interest with a probability, p, or not, with a probability 1-p. Knowing the value of p, the searcher makes a measurement of a random variable that has a probability density function which depends on whether or not the target is present, and attempts to make a decision regarding the presence of the target. In this paper a more general point of view is adopted in that we allow the possibility of both target and non-target elements to be present in any subpicture. We shall therefore be concerned with classifying a measurement as being a response from one of three sources, namely,

* References in brackets may be found on page 27.

1. a subpicture containing target elements only,

2. a subpicture containing non-target elements only,

3. a subpicture containing both target and non-target elements.

> The same problem arises [4] in certain analyses of radar search systems. The procedure is to continually sample a given range "bin" until a yes-no decision concerning the presence of a target is made. One often wishes to provide for the possibility of a target emerging into the bin during the sampling process.

The three categories of subpictures will be called pattern classes, and denoted by R_1, R_2, and R_3, respectively. To treat the problem at hand let $p(x|R_i)$ be the probability of occurrence of the measurement, x, given that it belongs to class R_i. We shall assume that $p(x|R_1)$ and $p(x|R_2)$ are known and set, for our first task, the determination of $p(x|R_3)$. To facilitate the following discussion, we shall denote the measurement x by r, s, or u depending on whether x is a measurement from R_1, R_2, or R_3, respectively. Furthermore we let $f_1(r) \equiv p(r|R_1)$, $f_2(s) \equiv p(s|R_2)$, and $h(u) = p(u|R_3)$. The next section shall be devoted to determining h(u) under various assumptions about the form of u.

1. DETERMINATION OF THE DENSITY FUNCTION, h(u)

1.1 Relationships Among the Random Variables.

We shall consider the determination of the density function h(u) for the following proposed expressions relating the random variables r, s, and u.

I. $u = wr(w,S_1) + (1-w)s(1-w,S_2)$

II. $u = wr(w) + (1-w)s(w)$

III. $u = wr + (1-w)s$

IV. $u = r(w,S_1) + s(w,S_2)$

V. $u = r(w) + s(w)$.

The first three expressions result when one conceives of x as some average value taken over the subpicture. In picture processing x could be an average value of some textural property, say average intensity. One can think of the subpicture S as partitioned into two parts, P_1 and P_2, where P_1 contains target elements only and P_2 contains non-target elements only. Let S_1 and S_2 be vector parameters which somehow characterize the geometric shapes of P_1 and P_2. Let $w(= \frac{\text{area }(P_1)}{\text{area }(S)})$ be the fraction of the subpicture containing target elements only. The first equation expresses u as a weighted average of r and s. The measurement r is taken over P_1; the measurement s is taken over P_2. In general the measurement r will be a function of both w and S_1. Similarly s is a function of w and S_2. As P_2 is the complement of P_1 with respect to S, a specification of S_1 also yields S_2. Therefore we can rewrite expressions I and IV in the form

$$\text{I}' \quad u = wr(w,S_1) + (1-w)s(w,S_1) \ ,$$

$$\text{IV}' \quad u = r(w,S_1) + s(w,S_1) \ .$$

In expression II, r and s are deemed to be independent of S_1 and S_2. In expression III, r and s are assumed to be independent of w and S_1. Expressions IV and V arise naturally when one considers the measurement x to be a count of the number of occurrences of a certain event in a subpicture (such as the number of edges, or perhaps, the number of closed contours). The order in which we shall discuss the expressions will be II, III, V, I', and IV'.

1.2 Case 1. $u = wr(w) + (1-w)s(w)$.

In this case we have assumed that r and s are independent of S_1 and S_2. The quantities r, s, and w are random variables with densities $f_1(r)$, $f_2(s)$, and $k(w)$ respectively. Our objective is to derive an expression for the density function, h(u), of the variate u. The variate u is a monotonic function of s when r(w) and w are held constant.

To obtain the distribution of u we imagine that the two variates $r(w)$ and w are held fixed at some arbitrary value while s varies over its possible range according to its distribution. The probability distribution of u which then results is the conditional distribution of u given $r(w)$ and w. Moreover, u becomes a monotonic function of the variable s although the equation relating u and s will contain the parameters w and $r(w)$. Hence the conditional density $\varphi(u|w, r(w))$ can be derived from $f_2(s|w)$ by a univariate method giving

$$\varphi(u|w,\ r(w)) = \frac{1}{1-w}\ f_2\left(\frac{u-wr(w)}{1-w}\Big|\ w\right) .$$

Furthermore one can then produce the joint density function, $g(u,r(w)|w)$, of u and $r(w)$ conditioned upon w by the expression

$$g(u,r(w)|w) = \frac{1}{1-w}\, f_1(r(w)|w)\ f_2\left(\frac{u-wr(w)}{1-w}\Big|\ w\right) .$$

One obtains the density of u (conditioned on w) by integrating the function $g(u,r(w)|w)$ with respect to $r(w)$, <u>taking pains</u> to keep the argument of the function f_2 in the range of the variable s. One obtains

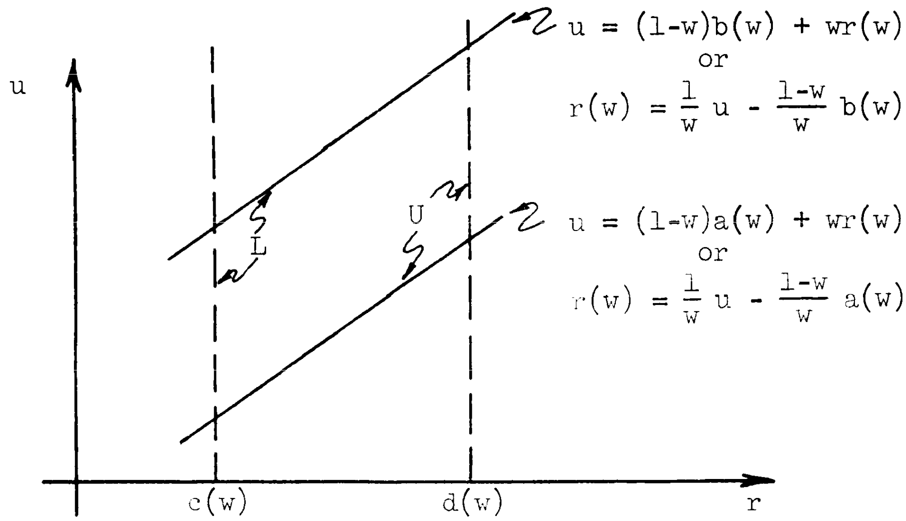
$$h(u|w) = \frac{1}{1-w}\left[\int_{T(u,w)} f_1(r(w)|w)\ f_2\left(\frac{u-wr(w)}{1-w}\Big|\ w\right) dr\right] .$$

In the above, the set $T(u,w) = \left\{r(w)\,|\,r(w) \in R(w) \text{ and } \frac{u-wr(w)}{1-w} \in S(w)\right\}$, where $R(w)$ is the range of $r(w)$ and $S(w)$ is the range of $s(w)$. If we let $R(w) = [c(w), d(w)]$, and $S(w) = [a(w), b(w)]$, then

$$T(u,w) = \left\{r\,|\,c(w) \le r \le d(w) \text{ and } a(w) \le \frac{u-wr}{1-w} \le b(w)\right\} .$$

From the inequalities $a(w) \le \frac{u-wr}{1-w} \le b(w)$, we obtain $(1-w)a(w) + wr(w) \le u \le (1-w)b(w) + wr(w)$. From Figure 1 we can visualize the dependence of the limits of integration upon u and w.

FIGURE 1



Hence

$$h(u|w) = \frac{1}{1-w} \int_{L(u,w)}^{U(u,w)} f_1(r(w)|w)\ f_2\left(\frac{u-wr(w)}{1-w}\middle|w\right) dr,$$

with

$$L = \max(c(w);\ \frac{1}{w} u - \frac{(1-w)}{w} b(w)) \ ,$$

and

$$U = \min\ (\frac{1}{w} u - \frac{1-w}{w} a(w);\ d(w))$$

There are two cases to distinguish. If $(1-w)a(w) + wd(w) \leq (1-w)b(w) + wc(w)$, then $T(u,w)$ is

$$\begin{cases} [c(w),\ (1-w)a(w) + u] \text{ for } (1-w)a(w) + wc(w) \leq u \leq (1-w)a(w) + wd(w) \\ [c(w),\ d(w)] \text{ for } (1-w)a(w) + wd(w) \leq u \leq (1-w)b(w) + wc(w) \\ [(1-w)b(w) + wu,\ d(w)] \text{ for } (1-w)b(w) + wc(w) \leq u \leq (1-w)b(w) + wd(w). \end{cases}$$

If $(1-w)a(w) + wd(w) \geq (1-w)b(w) + wc(w)$, then $T(u,w)$ is

$$\begin{cases} [c(w),\ (1-w)a(w) + wu] \text{ for } (1-w)a(w) + wc(w) \leq u \leq (1-w)b(w) + wc(w) \\ [(1-w)b(w) + wu,\ (1-w)a(w) + wu] \text{ for } (1-w)b(w) + wc(w) \leq u \leq (1-w)a(w) + wd(w) \\ [(1-w)b(w) + wu,\ d(w)] \text{ for } (1-w)a(w) + wd(w) \leq u \leq (1-w)b(w) + wd(w). \end{cases}$$

Finally, to determine $h(u)$ we must integrate with respect to w; the result is

$$h(u) = \int_W k(w)h(u|w)\, dw$$

1.3 Case 2. $u = wr + (1-w)s$.

In this case we have assumed that r and s are independent of w and the shape vectors S_1 and S_2. As in Section 1.2 the objective is to determine the density function, $h(u)$, for the variate u. Before proceeding to the general formulation we shall consider several special cases.

1. Suppose

   a. r and s are constants,
   b. w is uniformly distributed on $[o,m]$; $0 < m \leq 1$,
   c. $r > s$.

Then u is uniformly distributed on $[s, s + m(r-s]$.

2. Suppose

   a. r is a constant,
   b. s and w are statistically independent,
   c. s has a uniform distribution, $f_2(s)$, on $[a,b]$,
   d. w is uniformly distributed on $[0,1]$.

We obtain, using rationale similar to that of Section 1.2,

$$h(u) = \int_{T(u)} \left(\frac{1}{b-a}\right)\left(\frac{1}{1-w}\right) f_2\left(\frac{u-wr}{1-w}\right) dw ,$$

or

$$(b-a)h(u) = \int_{T(u)} \frac{1}{1-w}\, dw ,$$

where

$$T(u) = \left\{ w \,\middle|\, 0 \le w \le 1 \text{ and } a \le \frac{u-wr}{1-w} \le b \right\} .$$

We consider three subcases:

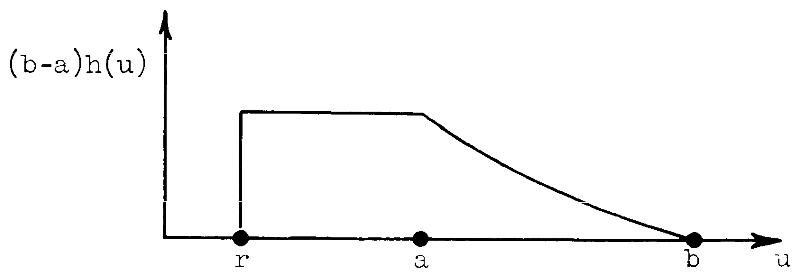
i. $r < a$. Then

$$T(u) = \begin{cases} \left[\dfrac{a-u}{a-r}, \dfrac{b-u}{b-r}\right] & \text{for } r < u < a \\ \\ \left[0, \dfrac{b-u}{b-r}\right] & \text{for } a < u < b \end{cases} ,$$

and

$$(b-a)h(u) = \begin{cases} \ell n\left(\dfrac{b-r}{a-r}\right) & \text{for } r \le u < a \\ \\ \ell n\left(\dfrac{b-r}{u-r}\right) & \text{for } a < u < b \end{cases} .$$

The graph of this function has the form



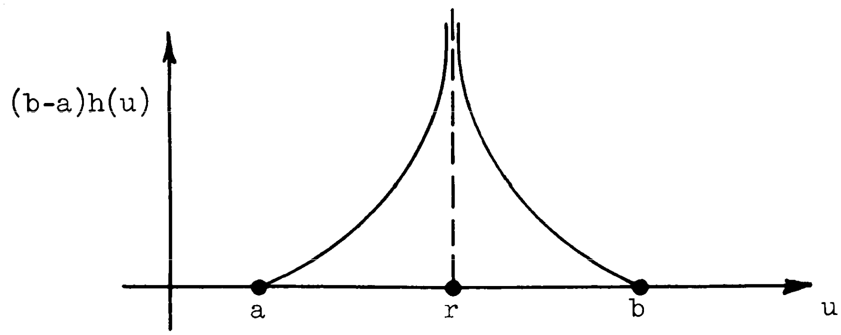
ii. $a < r < b$. Then

$$T(u) = \begin{cases} \left(0, \dfrac{u-a}{r-a}\right) & \text{for } a \le u < r \\ \\ \left(0, \dfrac{b-u}{b-r}\right) & \text{for } r < u \le b \end{cases} ,$$

and

$$(b-a)h(u) = \begin{cases} \ell n\left(\dfrac{r-a}{r-u}\right) & \text{for } a < u < r \\ \infty & \text{for } u = r \\ \ell n\left(\dfrac{b-r}{u-r}\right) & \text{for } r < u < b \end{cases} ,$$

whose graph has the form



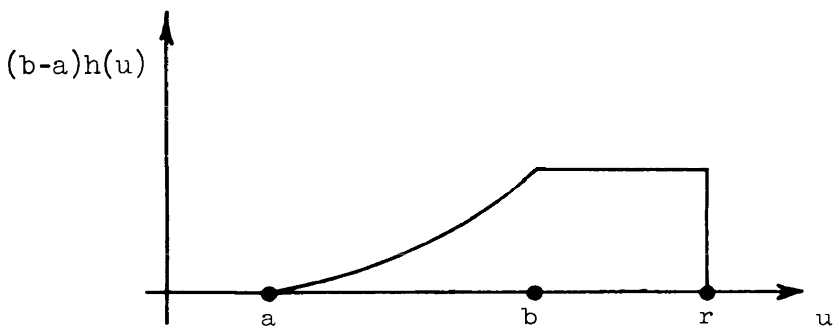
iii. $r > b$. Then

$$T(u) = [a, \min(u,b)] ,$$

and

$$(b-a)h(u) = \begin{cases} \ell n(\frac{r-a}{r-u}) & \text{for} \quad a \le u < b \\ \ell n(\frac{r-a}{r-b}) & \text{for} \quad b \le u < r \end{cases}$$

The graph of $(b-a)h(u)$ has the form



In deriving the general form for $h(u)$, we choose to fix the variables r and s at constant values initially. Again we see that u behaves as a monotonic function of w when r and s are held constant. Therefore the conditional density $\ell(u|r,s)$ can be derived from the probability density of w, $k(w)$, by a univariate method, yielding

$$\ell(u|r,s) = k(\frac{u-s}{r-s}) \frac{1}{|r-s|} \cdot$$

Moreover one can derive the joint density function $g(u,r,s)$ from the expression

$$g(u,r,s) = f(r,s)\ell(u|r,s)$$

where $f(r,s)$ is the joint density function of r and s. Finally the density of u is obtained by integrating $g(u,r,s)$ with respect to r and s. Thus we have

$$h(u) = \iint_{T(u)} f(r,s)\ k\left(\frac{u-s}{r-s}\right) \left|\frac{1}{r-s}\right|\ dr\ ds \tag{1.3.1}$$

where

$$T(u) = \left\{(r,s)|r \in R,\ s \in S, \text{ and } \frac{u-s}{r-s} \in W\right\},$$

$$R = \text{range of } r,$$

$$S = \text{range of } s,$$

and

$$W = \text{range of } w.$$

Letting $R = [c,d]$, $S = [a,b]$, $W = [0,1]$, and assuming that $r,s,$ and w are independent, we obtain the expression

$$\begin{aligned} h(u) = &\int_{\gamma(u)}^{d} \int_{a}^{\beta(u)} f_1(r)f_2(s)k\left(\frac{u-s}{r-s}\right) \left(\frac{1}{r-s}\right) ds\ dr\ + \\ &\int_{\gamma'(u)}^{b} \int_{c}^{\beta'(u)} f_1(r)f_2(s)k\left(\frac{u-s}{r-s}\right) \left(\frac{1}{s-r}\right) dr\ ds, \end{aligned} \tag{1.3.2}$$

where

$$\beta(u) = \max\{a, \min(u,b)\},\quad \gamma(u) = \min\{d, \max(c,u)\},$$

$$\beta'(u) = \max\{c, \min(u,b)\},\quad \gamma'(u) = \min\{b, \max(u,a)\}.$$

The two integrals evolve as a result of partitioning $T(u)$ into two sets $T_1(u)$ and $T_2(u)$, where

$$T_1(u) = \left\{(r,s)|r \in R,\ s \in S,\ r > s, \text{ and } 0 \le \frac{u-s}{r-s} \le 1\right\},$$

and

$$T_2(u) = \left\{(r,s)|r \in R,\ s \in S,\ r < s, \text{ and } 0 \le \frac{u-s}{r-s} \le 1\right\},$$

or equivalently,

$$T_1(u) = \{(r,s) | r \in R,\ s \in S, \text{ and } s \le u < r\} ,$$

and

$$T_2(u) = \{(r,s) | r \in R,\ s \in S, \text{ and } r \le u \le s\} .$$

1.4 Numerical Example for Case 2.

Suppose

a. r, s, and w are statistically independent, uniformly distributed random variables,

b. $R = [c,d]$, $S = [a,b]$, and $W = [0,1]$,

c. for definiteness, $0 < c < a < d < b$.

Then

$$(d-c)(b-a)h(u) = I_1(u) + I_2(u) ,$$

where

$$I_1(u) = \begin{cases} 0 & \text{for } u \le a \\ (d-a)\ln(d-a) + (\gamma(u) - \beta(u))\ln(\gamma(u) - \beta(u)) & \\ -(d-\beta(u))\ln(d-\beta(u)) - (\gamma(u)-a)\ln(\gamma(u)-a) & \text{for } u > a \end{cases} ,$$

and

$$\begin{aligned} I_2(u) = {} & (b-c)\ln(b-c) + (\gamma'(u) - \beta'(u))\ln(\gamma'(u) - \beta'(u)) \\ & - (b-\beta'(u))\ln(b-\beta'(u)) - (\gamma'(u)-c)\ln(\gamma'(u)-c) . \end{aligned}$$

Furthermore

$$(d-c)(b-a)h(u) = \begin{cases} k_1 + h_1(u) & \text{for } c \le u < a \\ k_2 - h_2(u) & \text{for } a < u \le d \\ k_3 + h_3(u) & \text{for } d < u \le b \end{cases} ,$$

where,

$$k_1 = (b-c)\ell n(b-c) - (a-c)\ell n(a-c),$$
$$k_2 = (d-a)\ell n(d-a) + (b-c)\ell n(b-c),$$
$$k_3 = (b-c)\ell n(b-c) - (b-d)\ell n(b-d),$$
$$h_1(u) = (a-u)\ell n(a-u) - (b-u)\ell n(b-u),$$
$$h_2(u) = (d-u)\ell n(d-u) + (u-a)\ell n(u-a) + (b-u)\ell n(b-u) + (u-c)\ell n(u-c),$$

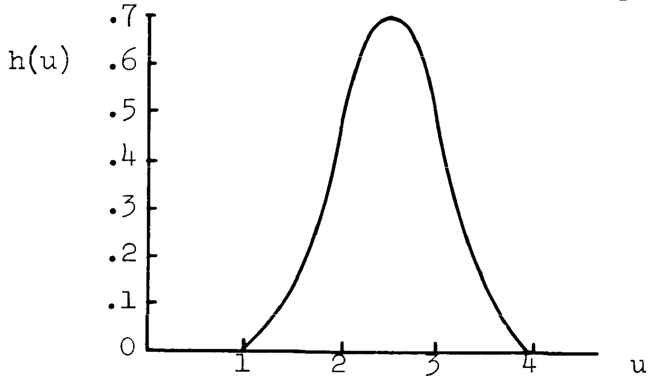
and

$$h_3(u) = (u-d)\ell n(u-d) - (u-c)\ell n(u-c).$$

A very lengthy and tedious computation verifies that $\int_c^b h(u)\, du = 1.$

Let c=1, a=2, d=3, and b=4. Then $h(u) = \dfrac{I_1(u) + I_2(u)}{4}$, and $k_1 = k_2 = k_3 = 3\ \ell n\ 3.$

The graph of h(u) is given in the following figure.



We shall make use of this example again in Section 2.1.

1.5 Plausible Density Functions for the variable, w.

In this section we shall derive several plausible density functions, k(w), for the random variable w. We shall describe the circumstances which give rise to these functions and discuss the importance of obtaining a catalog of such functions.

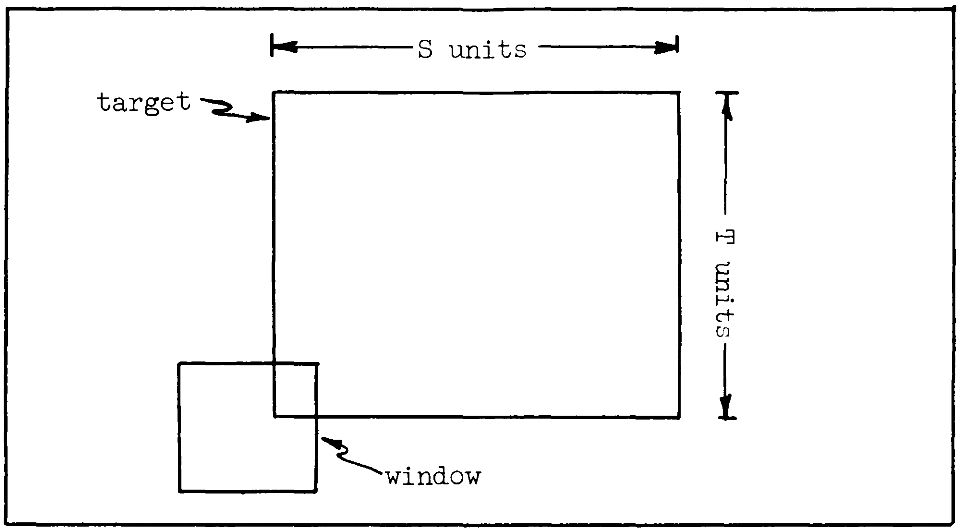
Consider the problem of scanning, horizontally, a two-dimensional optical display, such as a photograph or a television image with a receptor (window).

Suppose,

a. the target in the display is rectangular-shaped

b. the window is square-shaped

c. the edges of both the window and target are parallel to the corresponding edges of the optical display as illustrated,



d. each possible position of the window is equally likely.

Let the window be a unit square and let the target be $S(> 1)$ units long and $T(> 1)$ units wide. Let w be the ratio, $\frac{\text{area of overlap of target and window}}{\text{area of window}}$. We wish to find the probability density function, $k(w)$, given there exists an overlap of window and target. Assuming that each admissible position of the window with respect to the target is equally likely we can derive the expression

$$k(w) = \frac{S + T - 2 - 2 \ell n\, w}{S+T}\ ;\ 0 < w \leq 1.$$

For $S=T$, we obtain,

$$k(w) = \frac{S - 1 - \ell n\, w}{S}\ ;\ 0 < w \leq 1.$$

As either S or T approaches infinity, $k(w)$ approaches a uniform distribution on $[0,1]$.

Suppose

a. the target and window are circular-shaped with radii of $R(> 1)$ units and one unit, respectively,

b. each possible position of the window is equally likely,

c. $w(s)$ defines the ratio,

$$\frac{\text{area of overlap of the target and window}}{\text{area of window}} .$$

Then

$$w(s) = \frac{S_1(s) + S_2(s)}{\pi}$$

where

$$S_1(s) = \begin{cases} \frac{\pi}{2} - \left[y(s)\sqrt{1-y^2(s)} + \text{Sin}^{-1}\, y(s)\right] & \text{for} \quad 0 \leq y(s) \leq 1 \\ \\ -\frac{\pi}{2} + \left[|y(s)|\sqrt{1-y^2(s)} + \text{Sin}^{-1}|y(s)|\right] & \text{for} \quad -1 \leq y(s) \leq 0 \end{cases} ,$$

$$S_2(s) = \frac{\pi R^2}{2} - \left[(t(s) - y(s))\sqrt{R^2 - (t(s) - y(s))^2} + R^2\text{Sin}^{-1}\left(\frac{t(s) - y(s)}{R}\right)\right],$$

$$t(s) = R - 1 + s ,$$

and

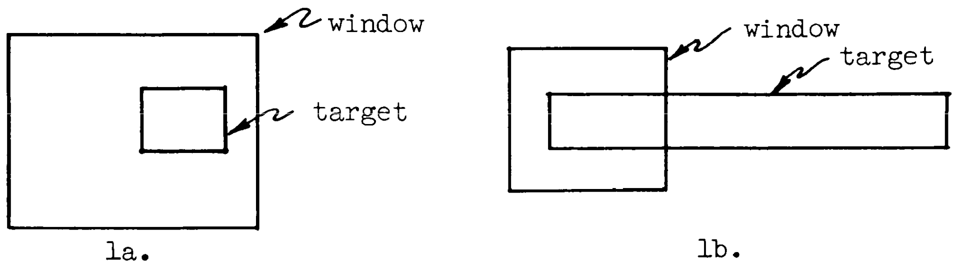
$$y(s) = \frac{s^2 + 2(R-1)(s-1)}{2(s + R-1)} .$$

The variable s is a random variable having a density function

$$\ell(s) = \frac{R + 1 - s}{2R} \qquad 0 \leq s \leq 2 .$$

Estimates of the probability density function for w may be obtained by computer simulation. One could sample s, many times, from a discrete version of its cumulative distribution, evaluate $w(s)$, and construct a discrete approximation to $k(w)$. The procedure could be repeated for a family of R values.

For the discussion above, we selected two examples where the variate w could range in the interval $[0,1]$. Another consideration is motivated by Figures 1a. and 1b.



In Figure 1a the target is smaller than the window. In Figure 1b we see an elongated target, so that the target never completely fills the window. These examples show that the values of w might be restricted to an interval, say, $[0,m]$, where $m < 1$.

The computations of density functions for w for such cases, and the implications with regard to Equation 1.3.1 will be given in another paper.

Our purpose is to provide a catalog of plausible density functions for the variate w. A function from the catalog could prove to be useful as an approximation in an application where a target shape is not regular. Moreover, knowledge of the properties of the functions might dictate the design of receptor shapes.

1.6 Use of Histogram Data in Conjunction with Equation (1.3.2).

In many pattern recognition problems, probability density functions for r and s are approximated by histograms obtained from sample sets, frequently called training sets.

With this in mind, let us define,

$$f_1(r) = \sum_{i=1}^{m} a_i H(r-r_i) , \tag{1.6.1}$$

and

$$f_2(s) = \sum_{j=1}^{n} b_j H(s-s_j) , \tag{1.6.2}$$

where

$$H(x) = \begin{cases} 1 \ ; \ x > 0 \\ 0 \ ; \ x \le 0 \end{cases}$$

We shall determine the function $h(u)$ with $f_1(r)$ and $f_2(s)$ defined by equations (1.6.1) and (1.6.2) and a uniform density function for w. It follows from equation (1.3.2) that

$$h(u) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j \left[I_1(i,j,u) + I_2(i,j,u) \right] ,$$

where

$$I_1(i,j,u) = \int_{\gamma_1(u)}^{d} \int_{s_j}^{\beta_1(u)} k\left(\frac{u-s}{r-s}\right) \frac{1}{r-s} \, ds \, dr \ ,$$

and

$$I_2(i,j,u) = \int_{\gamma_2(u)}^{b} \int_{c}^{\beta_2(u)} k\left(\frac{u-s}{r-s}\right) \frac{1}{s-r} \, dr \, ds \ ,$$

and

$$\beta_1(u) = \max \left[s_j, \ \min(u,b) \right] ,$$

$$\gamma_1(u) = \min \left[d, \ \max(r_i,u) \right] ,$$

$$\beta_2(u) = \max \left[r_i, \ \min(u,d) \right] ,$$

$$\gamma_2(u) = \min \left[b, \ \max(u,s_j) \right] .$$

For w uniformly distributed on [0,1], we obtain

$$I_1(i,j,u) = (d-s_j)\ell n(d-s_j) + (\gamma_1(u) - \beta_1(u))\ell n(\gamma_1(u)-\beta_1(u)) - (d-\beta_1(u))\ell n(d-\beta_1(u)) - (\gamma_1(u) - s_j)\ell n(\gamma_1(u) - s_j) \ ,$$

and

$$I_2(i,j,u) = (b-r_i)\ell n(b-r_i) + (\gamma_2(u) - \beta_2(u))\ell n(\gamma_2(u)-\beta_2(u)) - (b-\beta_2(u))\ell n(b-\beta_2(u)) - (\gamma_2(u)-r_i)\ell n(\gamma_2(u)-r_i) \ .$$

1.7 Case 3. $u = r(w) + s(w)$.

Proceeding as in the earlier sections, we obtain

$$h(u) = \int_W k(w) \left[\int_{T'(u,w)} f_1(r|w)\, f_2(u-r(w))\, dr\right] dw$$

1.8 Case 4. $u = wr(w,S_1) + (1-w)s(w,S_1)$ and $u = r(w,S_1) + s(w,S_1)$.

The computation of $h(u)$ for these expressions seems to be very difficult. One requires, initially, a characterization of P_1, that part of the subpicture containing target elements only. For complex (say non-convex) target shapes and for arbitrary target orientations within the subpicture this task appears to be extremely complicated.

1.9 Extension to a Set of Measurements.

Let us suppose that for each positioning of the receptor, one obtains several measurements, $x_1, x_2, \ldots x_d$. Let $X = (x_1, \ldots, x_d)$. We are interested in determining $p(X|R_1)$, $p(X|R_2)$ and $p(X|R_3)$. Let us assume that for each of the classes R_1, R_2, and R_3 the components of X are statistically independent. This allows us to write

$$p(X|R_i) = p(x_1|R_i)\, p(x_2|R_i) \ldots p(x_d|R_i) \; ; \; i=1,2,3$$

and the theory discussed in sections 1.1 to 1.8 is applicable.

2. CLASSIFICATIONS OF PATTERNS

2.1 Bayes Analysis.

Here we use decision surfaces of our pattern classifier which are defined by a set of functions $g_1(x)$, $g_2(x)$, and $g_3(x)$. These functions, called discriminant functions are chosen such that for all x in R_i, $g_i(x) > g_j(x)$ for $i,j=1,2,3$, $i \neq j$.

The patterns in each of the three categories R_1, R_2, and R_3 are random variables governed by the probability functions $p(x|R_i) = f_1(x)$, $p(x|R_2) = f_2(x)$, and $p(x|R_3) = h(x)$. An additional set of values which is needed in order to construct the discriminant functions is the set of a priori probabilities, $\{p(R_i)\}$, i=1,2,3. The discriminant functions will be expressed in terms of the $p(x|R_i)$ and $p(R_i)$. Our pattern classifier will be optimum in a Bayesian sense [5], [6] if we let

$$g_i(x) = - \sum_{j=1}^{3} \lambda(i|j)\ p(x|R_j)\ p(R_j),$$

where $\lambda(i|j)$, the "loss function" represents the loss incurred when the classifier places a pattern actually belonging to the class R_j into category R_i. Therefore the pattern classifier makes its classifications by the following steps:

1. the measurement x is presented to the classifier.

2. the classifier computes Max $[g_1(x),\ g_2(x),\ g_3(x)]$ and decides in favor of the category associated with the function yielding the largest value.

A loss function is said to be symmetric if it is of the form

$$\lambda(i|j) = 1 - \delta_{ij}$$

where

$$\delta_{ij} = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}$$

For such loss functions, the discriminant functions turn out to be

$$g_i(x) = p(x|R_i)\ p(R_i)\ ;\ i=1,2,3.$$

If all of the classes are equally likely a priori, i.e.

$$p(R_1) = p(R_2) = p(R_3) = \frac{1}{3}$$

we need only compute $p(x|R_i)$ for i=1,2,3 and select the maximum. This decision is called the maximum likelihood decision.

Returning to the numerical example in Section 1.4, we shall determine the maximum likelihood decision for each x in the interval [1,4]. The results are

$$\text{decide in favor of}\begin{cases}\text{class } R_1 & \text{if} \quad x \in [1,2] \\ \text{class } R_3 & \text{if} \quad x \in [2,3] \\ \text{class } R_2 & \text{if} \quad x \in [3,4]\end{cases} .$$

2.2 Sequential Analysis.

The result of a measurement for the subpicture is the value x which has the probability density function $p(x|R_1)$ or $p(x|R_2)$ or $p(x|R_3)$ depending on whether the subpicture contains target elements only, non-target elements only, or both target and non-target elements. Following each measurement, the observer makes one of the following four decisions:

D_1: decide target is present.

D_2: decide target is not present.

D_3: decide both target and non-target elements are present.

W: wait for another measurement.

The decisions D_1, D_2, and D_3 are terminal decisions, completing the process. The losses of the searcher making decision R_i given R_j is true are $\lambda(i|j)$. We shall assume that $\lambda(1|1) = \lambda(2|2) = \lambda(3|3) = 0$. The decision W allows the process to continue at least one more time step (i.e. permits the observer to make at least one more measurement). The loss incurred by this delay will be assumed to be dependent on which class R_i we are measuring. Let W_i equal the delay loss incurred if R_i is truly the class being measured. The objective of the searcher is to minimize the expected cost of a search. The decision policy that achieves the minimum expected cost is called the "optimal" policy. Following Pollock [2] we can write a functional equation which will yield the optimal policy. The equation is an application of Bellman's Principle of Optimality [7]. Let $f(p_1,p_2)$ be the minimum cost of search obtained using an optimal policy where $p_i = p(R_i)$ is the a priori probabilities associated with the occurrence of the class R_i.

We are interested in determining the function $f(p_1,p_2)$ which is a function of the current information about probabilities of the occurrence of R_1 and R_2. We are equally interested in determining $D(p_1,p_2)$ the decision that one should make as a function of p_1 and p_2. A more realistic problem arises when one considers that there exist at most only n available observations remaining before a terminal decision MUST be made. If the decision W is made, at the next decision there will be only n-1 possible observations left. If n=0, then one of the terminal decisions must be made. Letting $f_n(p_1,p_2)$ be the cost of search using an optimal policy given there are n available observations remaining before a terminal decision must be made, we get

$$f_n(p_1,p_2) = \min\begin{cases} p_2\lambda(1|2) + p_3\lambda(1|3) & ; \text{ decision } D_1 \\ p_1\lambda(2|1) + p_3\lambda(2|3) & ; \text{ decision } D_2 \\ p_1\lambda(3|1) + p_2\lambda(3|2) & ; \text{ decision } D_3 \\ p_1W_1 + p_2W_2 + p_3W_3 + \int_0^\infty \sum_{i=1}^{3} p_i p(x|R_i)\, f_{n-1}(p_1',p_2')\, dx & ; \\ \qquad\qquad \text{decision } W. \end{cases}$$

For n=0,

$$f_o(p_1,p_2) = \min\begin{cases} p_2\lambda(1|2) + p_3\lambda(1|3) & ; \ D_1 \\ p_1\lambda(2|1) + p_3\lambda(2|3) & ; \ D_2 \\ p_1\lambda(3|1) + p_2\lambda(3|2) & ; \ D_3 \end{cases}$$

If the decision is $\begin{Bmatrix} D_1 \\ D_2 \\ D_3 \end{Bmatrix}$ respectively, the expected losses incurred due to terminal incorrect decisions are $\left\{\begin{array}{l} p_2\lambda(1|2) + p_3\lambda(1|3) \\ p_1\lambda(2|1) + p_3\lambda(2|3) \\ p_1\lambda(3|1) + p_2\lambda(3|2) \end{array}\right\}$ respectively.

If the decision is W, the expected cost is $p_1W_1 + p_2W_2 + p_3W_3$ plus the cost of continuing from that point on, having observed some value of x. The probability of observing a value between x and x + dx is $[p_1p(x|R_1) + p_2p(x|R_2) + p_3p(x|R_3)]\,dx$. Having observed the value x, the probabilities

$$p_1' = p(R_1|x) = \frac{p(x|R_1)\,p_1}{p_1p(x|R_1) + p_2p(x|R_2) + p_3p(x|R_3)},$$

and

$$p_2' = p(R_2|x) = \frac{p(x|R_2)\,p_2}{p_1p(x|R_1) + p_2p(x|R_2) + p_3p(x|R_3)}$$

are obtained from Bayes rule.

Although the basic formalism of dynamic programming carries over without change from the two class detection problem, we have introduced an additional "state" variable into the functional equations. This introduction causes some further numerical difficulties which may be resolved by any one of several techniques recorded in [7].

ACKNOWLEDGMENTS

I wish to thank Professor A. Rosenfeld, Dr. A. Celmins, and Dr. A. Kiwan who read the manuscript carefully and proposed several changes.

WILLIAM J. SACCO

REFERENCES

1. Middleton, D. An Introduction to Statistical Communication Theory, McGraw Hill, New York (1960).

2. Pollock, S. Sequential Search and Detection, Tech Report No. 5 Operations Res Center MIT Cambridge, Massachusetts AD 601 641.

3. Helstrom, C. W. Statistical Theory of Signal Detection, Pergamon Press, 1960.

4. Gagliardi, R. and Reed, I. On Sequential Detection of Targets IEEE Transactions on Information Theory, IT-11 (1965).

5. Nilsson, N. Learning Machines, McGraw-Hill, 1965.

6. Chu, J. Optimal Decision Functions for Computer Character Recognition, Journal of the Association for Computing Machinery. Vol. 12, No. 2. April, 1965.

7. Bellman, R. and Dreyfus, S. Applied Dynamic Programming, Princeton University Press, 196, Chapters I and II.

DISTRIBUTION LIST

No. of Copies	Organization
20	Commander Defense Documentation Center ATTN: TIPCR Cameron Station Alexandria, Virginia 22314
2	Director of Defense Research and Engineering (OSD) ATTN: Asst Dir/Tac Msl & Ord Asst Dir/Comm & Elct Washington, D.C. 20301
1	Director Advanced Research Projects Agency Department of Defense Washington, D.C. 20301
1	Chairman Defense Science Board Washington, D.C. 20301
1	Commanding General U.S. Army Materiel Command ATTN: AMCRD-TE Washington, D.C. 20315
1	Commanding General U.S. Army Materiel Command ATTN: AMCPM-AI, Mr. E. Gardner Washington, D.C. 20315
1	Commanding General U.S. Army Materiel Command ATTN: AMCRD-M Washington, D.C. 20315
4	Commanding General U.S. Army Electronics Command ATTN: AMSEL-BL-D AMSEL-HL-CT Mr. V. Friedrick Mr. R. Pierce Mr. A. Zarin Fort Monmouth, New Jersey 07703
3	Commanding General U.S. Army Missile Command ATTN: AMSMI-REO, Mr. N. Mangus AMSMI-R, Mr. F. Durall AMSMI-RDS, Mr. D. Kilbourn Redstone Arsenal, Alabama 35809
1	Commanding General U.S. Army Missile Command ATTN: AMSMI-RBL Redstone Arsenal, Alabama 35809
2	Commanding Officer U.S. Army Engineer Research and Development Laboratories ATTN: STINFO Div (1 cy) Fort Belvoir, Virginia 22060
3	Commanding Officer U.S. Army Frankford Arsenal ATTN: SMUFA-N4100 Mr. R. Pfeilsticker Mr. C. Kerensky Mr. K. Raisner Philadelphia, Pennsylvania 19137
1	Commanding General U.S. Army White Sands Missile Range White Sands Missile Range New Mexico 88002
1	Commanding Officer U.S. Army Watertown Arsenal Watertown, Massachusetts 02172
1	Commanding General U.S. Army Combat Developments Command Fort Belvoir, Virginia 22060

DISTRIBUTION LIST

No. of Copies	Organization
1	Under Secretary of the Army ATTN: Ofc of Op Rsch Mr. Golub Washington, D.C. 20310
2	Commanding Officer U.S. Army Engineer, Geodesy Intelligence and Mapping Research and Development Agency ATTN: Mr. B. Scheps Mr. A. Anson Fort Belvoir, Virginia 22060
1	Chief of Research and Development Department of the Army Washington, D.C. 20310
1	Director U.S. Army Research Office ATTN: CRDPES 3045 Columbia Pike Arlington, Virginia 22204
1	Commanding Officer U.S. Army Research Office (Durham) Box CM, Duke Station Durham, North Carolina 27706
1	Chief of Naval Operations ATTN: Op-722C Department of the Navy Washington, D.C. 20350
5	Commander U.S. Naval Air Systems Command Headquarters ATTN: AIR-604 (3 cys) RM-3 RMGA-612 Washington, D.C. 20360
1	Commander U.S. Naval Missile Center ATTN: Code 112 Point Mugu, California 93041
1	Commander U.S. Naval Ordnance Laboratory Silver Spring, Maryland 20910
3	Commanding Officer U.S. Naval Ordnance Laboratory ATTN: Code 454, Mr. H. Johnson Code 724 Mr. W. G. Hueber Corona, California 91720
2	Commander U.S. Naval Ordnance Test Station ATTN: Code 753 Code 3512, Mr. H. Gregory China Lake, California 93557
1	Director U.S. Naval Research Laboratory Washington, D.C. 20390
1	Commanding Officer U.S. Naval Avionics Facility ATTN: Mr. P. Brink 21 St. & Arlington Avenue Indianapolis, Indiana 46218
1	HQ USAF (AFRDPC) Washington, D.C. 20330
1	RTD (RTTW) Bolling AFB Washington, D.C. 20332
1	APGC (PGBPS-12) Eglin AFB Florida 32542

DISTRIBUTION LIST

No. of Copies	Organization
1	AFATL (ATP) Eglin AFB Florida 32542
2	RADC Griffiss AFB New York 13442
1	AFWL (WLL) Kirtland AFB New Mexico 87117
2	AFCRL (CRBD, Mr. H. Blum) L. G. Hanscom Fld Bedford, Massachusetts 01731
3	AFAL (AVNT, LT Romhorst/ Mr. G. Kettering/ Mr. A. Welburn) Wright-Patterson AFB Ohio 45433
1	AFAL (AVWW, Mr. R. Leasure) Wright-Patterson AFB Ohio 45433
1	SEG (SEAFC, Mr. J. Muhleman) Wright-Patterson AFB Ohio 45433
3	Director National Bureau of Standards ATTN: Mr. C. McCarny (2 cys) Miss M. Stevens U.S. Department of Commerce Washington, D.C. 20234
1	Director National Aeronautics and Space Administration Goddard Space Flight Center Greenbelt, Maryland 20771

No. of Copies	Organization
1	Director National Science Foundation ATTN: Tech Lib Washington, D.C. 20550
2	Honeywell, Inc. Military Products Group ATTN: Mr. C. Johnson Mr. G. Swanland 2345 Walnut Street St. Paul, Minnesota 55408
2	Stanford Research Institute ATTN: Dr. N. Nilsson Dr. C. Rosen 333 Ravenswood Avenue Menlo Park, California 94025
3	University of Michigan Institute of Science and Technology IRIA-Infrared Laboratory Ann Arbor, Michigan 48104
1	Dr. R. Asendorf Hughes Research Laboratory 3011 South Malibu Canyon Road Malibu, California 90265
1	Dr. Donald Birx Franklin Institute Benjamin Franklin Pkwy at 20th Street Philadelphia, Pennsylvania 19103
1	Dr. R. Cosgriff Ohio State University Columbus, Ohio 43210
1	Dr. H. Freeman Electrical Engineering Department New York University New York, New York 10053

DISTRIBUTION LIST

No. of Copies	Organization
1	Mr. M. Halberstam Newtek, Incorporated 39-25 62nd Street Woodside, New York 10077
1	Dr. R. D. Joseph Astropower Laboratory Douglas Aircraft Company, Inc. 2121 Paularino Avenue Newport Beach, California 92660
1	Dr. L. Kanal Applied Research Laboratory Philco Corporation Union Metting and Jolly Roads Blue Bell, Pennsylvania 19422
1	Dr. G. Lendaris Defense Research Laboratories General Motors Corporation 6767 Hollister Avenue Goleta, California 93017
1	Dr. C. McKinney Defense Research Laboratory University of Texas P.O. Drawer 8029 Austin, Texas 78712
2	Dr. M. Minsky Massachusetts Institute of Technology 77 Massachusetts Avenue Cambridge, Massachusetts 02139
1	Dr. J. Muerle Cornell Aeronautical Laboratory Inc. P.O. Box 235 Buffalo, New York 14221
2	Dr. F. Rosenblatt Cornell University Ithaca, New York 14850

No. of Copies	Organization
1	Dr. A. Rosenfeld Computer Science Center University of Maryland College Park, Maryland 20740
1	Dr. M. Spooner Cornell Aeronautical Laboratory Inc. P.O. Box 235 Buffalo, New York 14221
2	Dr. B. Widrow Stanford University Stanford, California 94305
1	Mr. J. Hawkins Aeronutronics Division Philco Corporation Ford Road Newport Beach, California 92663
1	Mr. S. Krulikowski Bendix Research Laboratory Southfield, Michigan 48075
1	Mr. D. Mello Lowell Technological Institute 450 Aiken Street Lowell, Massachusetts 01850
1	Mr. M. Uffelman Scope, Inc. 121 Fallfax Drive Falls Church, Virginia 22043
1	Mr. S. Viglione Astropower Laboratory Douglas Aircraft Company, Inc. 2121 Paularino Avenue Newport Beach, California 92660

Aberdeen Proving Ground

Ch, Tech Lib
Air Force Ln Ofc
Navy Ln Ofc
CDC Ln Ofc