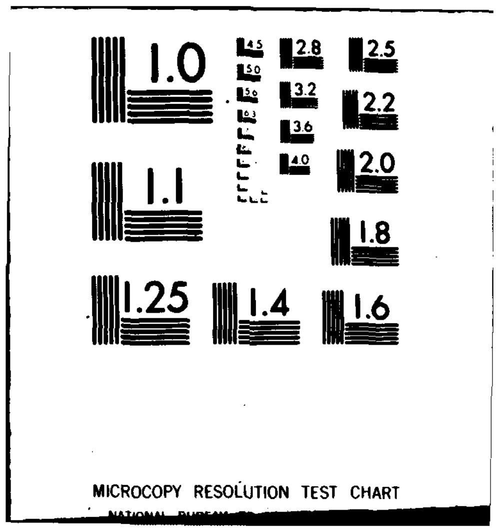
AD-A086 602 ORINCON CORP LA JOLLA CA F/G 17/7
A SCORE FOR CORRECT DATA ASSOCIATION IN MULTI-TARGET TRACKING.(U)
1979 D L ALSPACH, R N LOBBIA N00014-77-C-0296

UNCLASSIFIED NL

1 of 1
AD A
086602

END
DATE FILMED
8-80
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART

ADA086602

1. "A Score for Correct Data Association in Multi-Target Tracking," by D. L. Alspach and R. N. Lobbia, appears in the December 1979 Proceedings of the Decision and Control Conference held in Fort Lauderdale, Florida.

10 D. L. /Alspach R. N. /Lobbia

DOCUMENT 1

6 A SCORE FOR CORRECT DATA ASSOCIATION IN MULTI-TARGET TRACKING

11 1979

15 N00014-77C0296

SEE ALSO ADA 083221

12 26

392446

ORINCON

80 7 14 184

A SCORE FOR CORRECT DATA ASSOCIATION IN MULTI-TARGET TRACKING

D. L. Alspach and R. N. Lobbia
ORINCON CORPORATION
La Jolla, California 92037

SUMMARY

In the real-world multi-target tracking problem, there exists the possibility for many things to go wrong. Typical problems which arise include: too few tracks are formed; too many tracks are formed (false tracks); and inaccurate position, course, and speed estimates are reported. The above difficulties are often the result of incorrect allocation of data to individual tracks. Algorithms, while estimating the motion of a given target, inadvertently mix in clutter and/or measurements from another target. In order for correct allocation of data to a given track to be made, one must have an effective *scoring formula*; that is, some means of determining how likely a given assignment of data is. To be effective, a scoring formula must produce (on the average) a better score for correct assignments than for incorrect assignments. Information useful in the scoring process includes a priori intelligence data (such as initial target locations), models of target motion, models of the transmission channel, and expected moments of clutter for the sensor gain setting being used. Basically, the score is derived from the residuals which come out of the processing of a batch of data with the extended Kalman filter. This is used to evaluate the likelihood of potential tracks. Although the "likelihood" has an intuitive meaning, the term is used here to mean the probability density function $p(\lambda)$ of the track λ. The expected cost of a given assignment is derived with the theory of extremals being used to obtain the expected cost of adding a clutter point in a track. The resulting expected cost is then shown to behave in a quantitative fashion and this can be visualized from a geometric viewpoint.

lambda

1. INTRODUCTION

In the last few years a number of approaches to the problem of tracking multiple targets in a cluttered environment have been published. Some aspects of the problem that have been considered in some subsets of these publications include the problem of false alarms or missing measurements, track initialization, multiple measurement types (MSI) and target classification.

Many of the so-called multi-target trackers described in the open literature really deal only with the problem of one target in clutter. This hypothesis of only one target can greatly restrict the viability of a multi-target tracker to sort out confused situations.

Of the trackers that have been proposed, and in some cases implemented, one can see certain similarities and differences which allow the trackers to be grouped into certain classes. The grouping and certain applications of these groups has led to the following thoughts.

Perhaps the most fundamental aspect of a tracker is how it handles and interacts with the data. The data is after all our handle on the real world and all the information we have about a specific tracking realization is contained in the data. A second fundamental aspect of the multi-target tracking in clutter problem is that of alternate hypotheses. It is possible to derive the probability distribution of tracks and clutter points if one carefully specifies the a priori probability base, i.e., the probabilistic target models, probabilistic measurement models, etc. In very simple cases where one can get the optimal solution, one finds this consists of all possible configurations of the data into the sets. In each configuration each data set represents a possible alternate track or the set of clutter points. A probability measure is assigned to each possible total surveillance region picture. This globally optimal approach is generally not reasonable for implementation and approximate or suboptimal approaches must be considered.

Several approaches focus entirely on the construction of the sets of data (the construction of feasible tracks). It is this latter philosophy that is being addressed in this paper.

Conceptually, using maximum likelihood techniques, various combinations of data are tried and then "scored" using log-likelihood functions. The best fit to the target model gets the lowest score and this is considered to be one of the tracks in the region if no other low score track competes for the same measurements. If two or more targets compete for the same measurements, several situations can occur. These include the possibilities that the two targets are lumped together, one target is rejected, the targets get mixed with track points assigned to clutter and two "bad" tracks reported, or the case that all points are assigned to clutter. It is possible, though perhaps not normal, to find situations where the choice of the best nonoverlapping feasible track does not correspond to a best surveillance picture. This is quite easy to do if the tracks overlap and compete for the same measurements.

Many trackers consider alternate hypotheses as far as assigning measurements to a track. However, once a decision has been made that a measurement belongs with another group of measurements, this decision is not re-examined. Once the decision has been made to assign a piece of data to a "track," that decision is final. This is done because the system usually requires "an answer." Also, there is always new data coming into the tracker allowing new hypothesis tests. In addition, there is a limited amount of computer response. One could say that the hypothesis testing is directed to make a decision on the proper surveillance picture or that the tracker is "decision directed."

In the next section, we will see how effective scoring algorithms are developed—ones which can handle the alternate scenarios posed above. Following this, in Section 3, a refinement to this scoring algorithm is proposed and it is seen that the average cost incurred for assigning measurements to tracks can be visualized from a geometrical viewpoint. In particular, it will be shown quantitatively that there exists a unique number of points (measurements) in a given track that yields an average minimum cost. Incorrectly assigning clutter points to this track and wrongly assigning points to clutter will, on the average, increase the cost in a well-defined manner.

2. SCORING ALGORITHMS

In order for a correct assignment of measurement data to a given track to be made, we must have an effective scoring formula, i.e., some means of determining how likely a given assignment of data is. To be effective, a scoring formula must produce (on the average) a better score for correct assignments than for incorrect assignments. Information useful in the scoring process includes a priori intelligence data (such as initial target locations), models of target motion, and expected amounts of clutter for the sensor gain setting being used. Basically, the score is derived from the residuals which come out of the processing of a batch of data with the extended Kalman filter. This is used to evaluate the likelihood of potential tracks. Although "likelihood" has a useful intuitive meaning, we use the term to mean the probability density function $p(\lambda)$ of the track λ. The concepts we use are well-known, since most of the work in estimation theory pertains to situations where all the observations $Z = \{z_1, z_2, \cdots, z_n\}$ are due to a single target. An obvious example is the stochastic linear system

$$x_{k+1} = A_k x_k + B_k u_k, \qquad k = 0, 1, \cdots, n, \tag{1}$$

$$z_k = C_k x_k + v_k, \qquad k = 1, \cdots, n, \tag{2}$$

with states $\{x_k\} \subset R^x$, observations $\{z_k\} \subset R^z$, process noise $\{u_k\} \subset R^u$, and measurement noise $\{v_k\} \subset R^v$. A_k, B_k, C_k are matrices of appropriate dimension that may vary with time. The initial state x_0 is a Gaussian random vector with covariance P_0, independent of the processes $\{u_k\}$ and $\{v_k\}$, which are themselves zero mean white Gaussian noise with covariances $\{Q_k\}$ and $\{R_k\}$. Under these assumptions, the well known Kalman equations provide minimum variance unbiased estimates $\{\hat{x}_k\}$ of the states based on all past data:

$$\hat{x}_{k+1} = A_k\hat{x}_k + K_k(z_{k+1} - C_{k+1}A_k\hat{x}_k), \tag{3}$$

$$K_k = \bar{P}_kC_{k+1}^t(C_{k+1}\bar{P}_kC_{k+1} + R_{k+1})^{-1}, \tag{4}$$

$$\bar{P}_k = A_kP_kA_k^t + B_kQ_kB_k^t, \tag{5}$$

$$P_{k+1} = (I - K_kC_{k+1})\bar{P}_k, \quad k = 0, 1, \cdots, n. \tag{6}$$

When nonlinear measurements are involved, a simple linearization process (the extended Kalman filter) is used. The equations remain exactly the same, except that the term

$$z_{k+1} - C_{k+1}A_k\hat{x}_k \tag{7}$$

in the first equation above is replaced by

$$z_{k+1} - h_{k+1}(A_k\hat{x}_k), \tag{8}$$

where $h_{k+1}(\cdot)$ denotes the nonlinear relationship between the measurements and state vector, and the C_k-matrix becomes

$$C_k = \left.\frac{\partial h_k(x)}{\partial x}\right|_{A_k\hat{x}_{k-1}} \tag{9}$$

It is natural to compute the likelihood function $p(\lambda)$ *for the track* $\lambda \subset Z$ based on the Kalman filter state estimates. The innovations sequence is an integral part of this computation, which is a sequence of the measurement residuals:

$$\delta_{k+1} = z_{k+1} - C_{k+1}A_k\hat{x}_k, \quad k = 0, 1, \cdots, n. \tag{10}$$

The (negative) log likelihood function is given in terms of $\{\delta_k\}$ by

$$c(\lambda) = n\dim(z)\ln 2\pi + \frac{1}{2}\sum_{k=1}^{n}\ln|V_k| + \frac{1}{2}\sum_{k=1}^{n}\delta_k^t V_k^{-1}\delta_k. \tag{11}$$

The covariance matrix for the measurement residual, V_k, *can be computed directly:*

$$V_k = C_{k+1}\bar{P}_kC_{k+1}^t + R_{k+1}, \quad k = 1, 2, \cdots, n. \tag{12}$$

Each feasible track is the result of a hypothesis test that uses the track likelihood function $p(\lambda)$ (or equivalently, the negative log likelihood $c(\lambda)$) determined from the Kalman filter. Since the density function of the alternative hypothesis (that λ is not a track) is unknown, the decision rule is simply

$$-c(\lambda) = \ln p(\lambda \mid \{\hat{x}_k\}_{k=1}^{n}) > \alpha_n \rightarrow \lambda \in F, \tag{13}$$

$$-c(\lambda) = \ln p(\lambda \mid \{\hat{x}_k\}_{k=1}^{n}) < \alpha_n \rightarrow \lambda \notin F. \tag{14}$$

Based on the log likelihood decision function, the feasible track set is

$$F = \left\{\lambda \mid \lambda\ Zm \ln p(\lambda \mid \{\hat{x}_k\}_{k=1}^{n}) > \alpha_n\right\}. \tag{15}$$

Primarily, the only random component of $c(\lambda)$ is

$$\Sigma\delta^t V^{-1}\delta, \tag{16}$$

which for real tracks is a chi-squared random variable with $n \cdot$ dimension (z) degrees of freedom.* Therefore, error probabilities can easily be computed for the hypothesis test to predict the accuracy of feasible track construction. This has a critical impact on the ultimate accuracy of the tracking algorithm, since a real track mistakenly exclused from F cannot be used in the subsequent Bayesian decision process.

*Due to the random nature of the measurement arrival time and of the sensor which senses "the next measurement," the development of V_k is also random in nature but it is hard to compare its meaning from one realization to the next.

Before any new point is added to a partial track, it should pass a coarse test,

$$\|\delta_k\|_\infty < \beta_k \tag{17}$$

a fine test,

$$\delta_k^t V_k^{-1}\delta_k < \gamma_k \tag{18}$$

and, finally, the likelihood test. The coarse test checks the magnitude of the maximum component of the vector $\delta_k \in R^2$ against β_k, and is included because it is computationally cheaper to perform than the fine test. The constants β_k and γ_k can be chosen so that

$$\left\{z_k \mid \delta_k^t V_k^{-1}\delta_k < \gamma_k\right\}\left\{z_k \mid \|\delta_k\|_\infty < \beta_k\right\}. \tag{19}$$

One difficulty with this approach is that because of the "deterministic" terms in the likelihood function

$$n\,dm(z)\ln 2\pi + \frac{1}{2}\sum_{k=1}^{n}\ln|V_k| \tag{20}$$

it is difficult to compare some of the tracks of different length. It is also difficult to assess the absolute goodness of a score. Therefore, an alternate score with a more absolute meaning can be defined. This is described in the next section.

3. REFINED SCORING

For track i at stage k define the stagewise chi-squared score

$$S_k^i = (z_k^i - \hat{z}_k^i)^T V_k^{-1}(z_k^i - \hat{z}_k^i). \tag{21}$$

This has the features—if all measurements have been assigned to the correct targets and all filter parameters chosen correctly—that

$$E\{S_k^i\} = E\{S_k\} = 2 \tag{22}$$

$$E\{(S_k - 2)^2\} = 4 \tag{23}$$

$$\sigma_{S_k} = 2 \tag{24}$$

$$f(S) = e^{-S/2}U(S); \quad \begin{array}{l} U(S) = 0,\ S < 0 \\ U(S) = 1,\ S > 0 \end{array} \tag{25}$$

where $E\{\cdot\}$ indicates the expected value operator, σ_{S_k} is the standard deviation of S_k, and $f(S)$ is the appropriate density function for a two-dimensional random variable S. Define the cumulative chi-squared score as

$$S^i = \sum_{k=1}^{n_i} S_k^i. \tag{26}$$

For easy evaluation on a single track an evaluation cost that would be quite meaningful would be

$$C_i = \frac{1}{N_i}S^i. \tag{27}$$

For this cost function the statistical parameters are

$$E(C_i) = 2 \tag{28}$$

$$E\left((C_i - 2)^2\right) = 4/N_i \qquad \sigma_{C_i} = 2/N_i^{1/2} \tag{29}$$

$$f_C(y) = \frac{1}{2^{N_i}\Gamma(N_i)}\,y^{(N_i-1)}e^{-y/2}U(y). \tag{30}$$

For display purposes the use of C_i as a value measure of a single track makes a great deal of intuitive sense.

Values of $C_i \ll 2$ for reasonable length tracks tend to imply that the filter parameters are set too loose. Thus, by reducing Q_k and/or R_k, one could obtain tighter tracks. Tracks for which $C_i \gg 2$ clearly represent bad data assignment which should not be kept. More precisely, if

$$C_i > 2 + \alpha 2/N_i^{1/2} \tag{31}$$

when $\alpha = 3$ (for a 3-sigma case) indicates either that incorrect data has been assigned to the track or that the filter parameters (Q,R) are too small. While

$$C_i \leq 2 - \alpha 2/N_i^{1/2}, \quad N_i > 3 \text{ for } \alpha = 3 \tag{32}$$

indicates that the filter parameters (Q,R) are set too large.

For a total surveillance region picture of M measurement points and L tracks, one has:

Track No.	Score	Number of Points in Track
1	C_1	N_1
2	C_2	N_2
.	.	.
.	.	.
.	.	.
L	C_L	N_L

The score for the total area should be made up of these scores and the cost for assigning a point to clutter. The number of points in clutter is N_C, the number of points assigned to targets is N_P, and the total number of measurements are M. These are related by

$$M = N_C + N_P \qquad N_P = \sum_{i=1}^{L} N_i . \tag{33}$$

A meaningful score could be defined as:

$$S = \sum_{i=1}^{L} N_i C_i + N_C S_C \tag{34}$$

where S_C is a score defined for a clutter point. If all the measurements are correctly assigned to the track:

$$E(S) = 2 \sum_{i=1}^{L} N_i + N_C S_C \tag{35}$$

$$= 2(M - N_C) + N_C S_C . \tag{36}$$

If we define the clutter score, S_C, and a total surveillance score as

$$C = \frac{1}{M} S \tag{37}$$

where, if all the measurement points belong to the track,

$$E(C) = \frac{1}{M} E(S) = \frac{1}{M} [2(M - N_C) + N_C S_C] \tag{38}$$

$$E(C) = 2 \frac{M - N_C}{M} + \frac{N_C}{M} S_C \tag{39}$$

so that a plot of the score for a global surveillance region for a real case of N_P measurements assigned to N^* targets with all points assigned correctly and N_c clutter points can be geometrically described by Figure 1.

The curve from N=0, that is, all points assigned to clutter, to $N = N_P$, the correct number of points assigned to clutter, is just a straight line described by:

$$E(S) = \frac{2(M - N_C)}{M} + \frac{N_C}{M} S_C . \tag{40}$$

where

N_C varies from M to $M - N_P$.

Beyond the point ($N = N_P$), a clutter point must be assigned to the track (assuming just one track in clutter). The cost to assign a single clutter point to the track can be calculated in the following manner: A given track will project to a given point in measurement space. For example, consider the two-dimensional measurement vector (τ, θ) described in Figure 2.

If a clutter point is added to the track, the increase in score on the track will be given by Equations (21) and (37). The difference is that now the value of the score can be written as

$$C = \frac{S}{M} = \frac{1}{M} \left\{ \underbrace{\sum_{i=1}^{L} N_i C_i + (N_C - 1) S_C}_{\text{assume all track points are properly assigned}} + C_1^c \right\} . \tag{41}$$

Here C_1^c is the cost for assigning one clutter point to the track. C_1^c is a random variable and its distribution depends on the distribution of the clutter points. From Equation (21),

$$C_1^c = (z_k^c - \hat{z}_k^i) V_k^{-1} (z_k^c - \hat{z}_k^i) \tag{42}$$

where z_k^c are the τ, θ points for the clutter point. The distribution of this random variable depends on the random nature of the clutter.

If the distance of the clutter point from the predicted point is assumed to be Gaussian in τ and θ with zero mean, the distance from the predicted point will be Rayleigh distributed and the cost or score C_1^c, the weighted square of the distance (42), will have an exponential density

$$f_{c_1}(c) = \frac{1}{2\sigma^2} e^{-c/2\sigma^2} u(c); \quad \begin{array}{l} u(c) = 0, c < 0 \\ u(c) = 1, c > 0 \end{array} \tag{43}$$

where σ^2 is a measure of the dispersion of the clutter points with respect to the mid- or tracker-predicted point. The uncertainty, σ, takes into account the unequal variance in τ and θ and is given by:

$$\sigma^2 = \frac{(\text{range of } \theta \text{ in surface})\,(\text{range of } \tau \text{ in surface})}{\sigma_\theta \sigma_\tau} \tag{44}$$

The score for the closest clutter point to a point predicted by the tracker (closest in the sense of having the smallest score defined by (42)) will be distributed as follows (x = score of the closest point):

$F_x(x)$ = probability that the closest clutter point out of N_C having a score less than or equal to x.

$$F_x(x) = 1 - e^{-x/2\sigma^2} . \tag{45}$$

From the theory of extremals, it is obvious that this distribution is identical to the probability distribution of the event defined below,

$F_x(x)$ = {probability that all of the N_C clutter points have scores less than or equal to x}

or

$F_x(x)$ = 1 - {probability that all N_C clutter points have scores greater than x}

From (43), it follows that the probability that the score of any one of the N_C clutter points being greater than x is given by:

$$e^{-x/2\sigma^2} . \tag{46}$$

Assuming the clutter points are independent of one another, the probability that all of the N_C clutter points have scores greater than x is given by

$$e^{-N_C x/2\sigma^2}$$

and the probability that all of these clutter points have scores less than or equal to x is:

$$1 - e^{-N_C x/2\sigma^2} \tag{47}$$

and density function:

$$f_x(x) = \frac{N_C}{2\sigma^2}\, e^{-N_C x/2\sigma^2}\,. \tag{48}$$

The expected value of the cost of assigning the closest clutter point to a track is then given by:

$$E(x) = \int_{-\infty}^{\infty} x f(x)\, dx = \frac{2\sigma^2}{N_C}\,. \tag{49}$$

The expected cost for assigning the next nearest point to a track is

$$\frac{2\sigma^2}{(N_C - 1)} \tag{50}$$

and, finally, the expected cost for adding the k^{th} closest clutter point out of N_C to a track is given by:

$$\frac{2\sigma^2}{N_C + 1 - k}\,. \tag{51}$$

Note that in computing the expected value of the total cost function defined in (41), one must also account for the linear decrease in cost caused by the decreasing weighting coefficient $(N_C\text{-}k)$ on S_C when k clutter points are incorrectly assigned to tracks. This amounts to a decrease of S_C/M for every clutter point assigned to a track. Therefore, the net increase in the expected cost by assigning the k closest clutter points out of N_C to tracks is given by:

$$\frac{1}{M}\left\{\sum_{j=1}^{k} \frac{2\sigma^2}{N_C + 1 - j} - kS_C\right\}. \tag{52}$$

In effect, this is the sum of an increasing hyperbolic function and a decreasing linear function.

Since, on the average, $\sigma^2 \gg S_C$, the shape of the right-hand side of the curve in Figure 1 is concave and monotonically increasing.

Note the two regions of this figure. In one part fewer points are assigned to tracks than actually are available, i.e., too many detections were missed. In the other half of the figure, false alarms or clutter points were assigned to tracks. In both parts of the curve the assumption is made that the data association is done in an optimal manner on the average. Thus, if a clutter point is added to a track, it is the clutter point that lies "closest" to a track of those unassigned. The curve also assumes that all measurements correctly assigned as track points are assigned to the correct track. Wrong assignments of data would, of course, make for even worse scores "on the average."

4. CONCLUSIONS

In the last section, it was shown that the refined scoring algorithm possesses appealing properties from a geometric viewpoint. There exists a unique number of points in a track that results in lower average cost than any other number. Also, the sensitivity of the score to variations in assigned number of track points can be controlled by the clutter score, S_C. This is readily apparent from Figure 1.

This scoring algorithm is therefore a very useful approach in extracting clutter points out of a given target track.

The algorithm is currently being applied to an ocean surveillance problem and the results of this are very encouraging. For a given data set having a false alarm rate of 10^{-4}, i.e., one clutter point in every 10^4 measurements, using the refined scoring algorithm defined in Section 3, we have found that we can effectively decrease this false alarm rate to 10^{-7}. This represents a three-order-of-magnitude decrease and, hence, the detection capabilities of the tracking algorithm have been significantly enhanced.

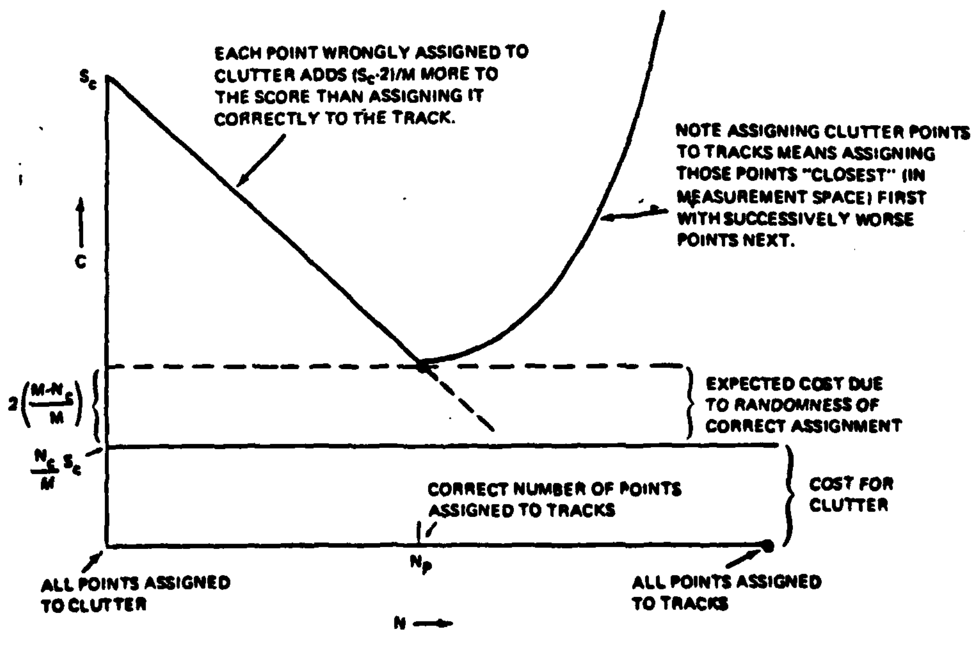


Figure 1. Geometrical description of refined scoring algorithm.

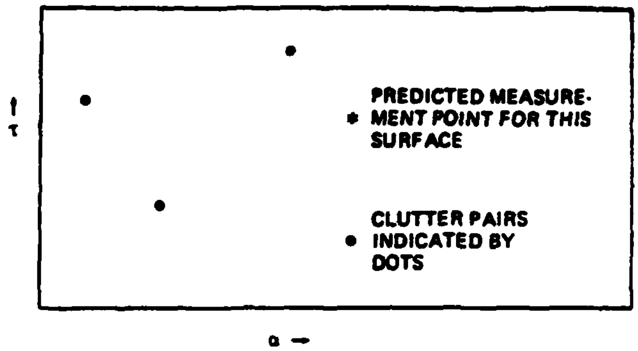


Figure 2. Two-dimensional measurement space.

This work was partially supported by ONR Research Contract No. N00014-77-C-0296.