AD-A045 179 NAVAL POSTGRADUATE SCHOOL MONTEREY CALIF F/G 12/1

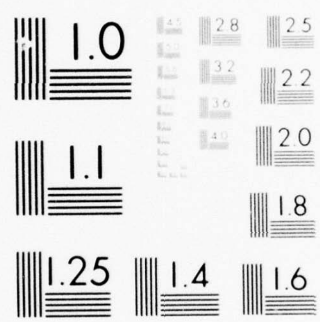
MODELING AND RECURSIVE ESTIMATION OF TWO DIMENSIONAL RANDOM FIE--ETC(U)

JUN 77 M SHACHAR

UNCLASSIFIED NL

1 OF 3

AD
A045179


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

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

MODELING AND RECURSIVE ESTIMATION
OF TWO DIMENSIONAL RANDOM FIELDS
AND APPLICATIONS TO TARGET DETECTION

by

Moshe Shachar

June 1977

Thesis Advisor: H.A. Titus

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 Modeling and Recursive Estimation of Two Dimensional Random Fields and Applications to Target Detection,		5. TYPE OF REPORT & PERIOD COVERED Electrical Engineer; June 1977 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Moshe Shachar		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		11 12. REPORT DATE June 1977 13. NUMBER OF PAGES ~~215~~ 12 214 p.
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Two dimensions, two dimensional recursive estimation, optimal estimation, linear estimation, quadratic form, Markov process, image modeling, target detection, image enhancement, image restoration		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

First an investigation of modeling stochastic processes by difference equations (Markov process) was undertaken. The starting point of the modeling procedure is the knowledge of the spectrum of the process. Two methods are discussed. One is based on optimal estimation theory and leads in most cases to a high-order (perhaps infinite) Markov process. The second method, based on linear system

251 450

(20. ABSTRACT Continued)

theory, leads to a first order Markov process (in matrix representation). Both methods have been extended to two-dimensional processes. Secondly, recursive estimation (filtering) of two-dimensional random fields was addressed. It was shown that a two-dimensional recursive filter cannot be optimal. Therefore, only a sub-optimal solution is available. This solution minimizes the mean square error for a specific structure of a filter. Finally, applications of modeling and recursive filtering are discussed. An image that includes a target, correlated noise and random noise was processed. Some methods of target enhancement (also called "restoration") are discussed.

ACCESSION for		
NTIS	White Section	☒
DDC	Buff Section	☐
UNANNOUNCED		☐
JUSTIFICATION		
BY		
DISTRIBUTION/AVAILABILITY CODES		
Dist.	AVAIL. and/or SPECIAL	
A		

Approved for public release; distribution unlimited.

Modeling and Recursive Estimation of Two Dimensional Random Fields and Applications to Target Detection

Moshe Shachar
Lieutenant, Israeli Navy
Technion, Israel Institute of Technology, 1972

Submitted in partial fulfillment of the requirements for the degree of

ELECTRICAL ENGINEER

from the

NAVAL POSTGRADUATE SCHOOL
June 1977

Author: M. Shachar

Approved by: Harold Titus, Thesis Advisor

John P. Powers, Second Reader

Donald E. Kirk, Chairman, Department of Electrical Engineering

R. R. Fossum, Dean of Science and Engineering

ABSTRACT

First an investigation of modeling stochastic processes by difference equations (Markov process) was undertaken. The starting point of the modeling procedure is the knowledge of the spectrum of the process. Two methods are discussed. One is based on optimal estimation theory and leads in most cases to a high-order (perhaps infinite) Markov process. The second method, based on linear system theory, leads to a first order Markov process (in matrix representation). Both methods have been extended to two-dimensional processes. Secondly, recursive estimation (filtering) of two-dimensional random fields was addressed. It was shown that a two-dimensional recursive filter cannot be optimal. Therefore, only a sub-optimal solution is available. This solution minimizes the mean square error for a specific structure of a filter. Finally, applications of modeling and recursive filtering are discussed. An image that includes a target, correlated noise and random noise was processed. Some methods of target enhancement (also called "restoration" are discussed.

TABLE OF CONTENTS

NOTATION

$x(k,\ell)$	-	correlated field
$u(k,\ell)$	-	modeling error (random wlak of the Markov process)
$v(k,\ell)$	-	white noise field.
$\hat{x}(k,\ell)$	-	estimation of $x(k,\ell)$
$y(k,\ell)$	$\triangleq$	$\left(\hat{x}(k,\ell) - x(k,\ell)\right)$ = estimation error
$\underset{\sim}{\mathcal{E}}^T(k,\ell)$	$\triangleq$	$\left(\mathcal{E}(k+1,\ell) \quad \mathcal{E}(k,\ell) \quad \mathcal{E}(k,\ell+1)\right)$
T	-	target intensity
σ_s, σ	-	variance of correlated process
ρ_1, ρ_2	-	correlation coefficients that determine the bandwidth of the process
θ_1, θ_2	-	correlation coefficients that determine the center frequency of the process
TH	-	value of threshold in the decision process
$\sigma_e(k,\ell)$	-	variance of error
$P(k,\ell)$	$\triangleq$	$\sigma_e^2(k,\ell)$ = mean of square erros
$\underset{\sim}{P}(k,\ell)$	$\triangleq$	$E\{\underset{\sim}{\mathcal{E}}(k,\ell)\cdot\underset{\sim}{\mathcal{E}}^T(k,\ell)\}$
σ_n	-	variance of random noise
R	$\triangleq$	σ_n^2
σ_u	-	variance of modeling error (random walk)
Q	$\triangleq$	σ_u^2
$P_{i,j}(k,\ell)$	$\triangleq$	$E\{\mathcal{E}(k,\ell)\cdot\mathcal{E}(k+i,\ell+j)\}$
$(\cdot)^T$	-	means the transpose of the vector $(\cdot)$
$\overline{x}$	-	mean of correlated field
G	-	steady state gain

I. INTRODUCTION

1. The goal in this thesis was to solve the problem described in the following paragraphs. Consider an image sensing device that has an array of NxN sensing elements. The images that are sensed by the device include some targets that suffer from degradation due to background (for example, clouds). The background noise eliminates the possibility of direct detection of the target. The background is assumed to be a correlated noise source.

Further interference to the output of the imager might be the internal noise of the device. It is assumed that the noise of each sensing element is uncorrelated to the others (white noise). Therefore, the output of the imager, $y(k,\ell)$ includes three types of processes:

1. the target $T(k,\ell)$
2. The correlated background $x(k,\ell)$
3. The white noise $v(k,\ell)$

$$y(k,\ell) = T(k,\ell) + x(k,\ell) + v(k,\ell) .$$

A typical image is seen in Figure 2. The problem in this thesis is to detect the intensity and location of the target by using recursive techniques that are applicable for real-time hardware.

2. The solution to this problem is carried out in three steps:

a) Mathematical modeling of the background.

b) Estimation of the background.

c) Detection of the target.

The main idea was to eliminate the background by subtraction of the estimated background from the original imager's output. Then the residual image includes only the target and the white noise. The detection is, therefore, easier. Still, for detection with small errors (false alarms and misses) the target must have an intensity higher than the white noise. The detection procedure is shown in Fig. 1.

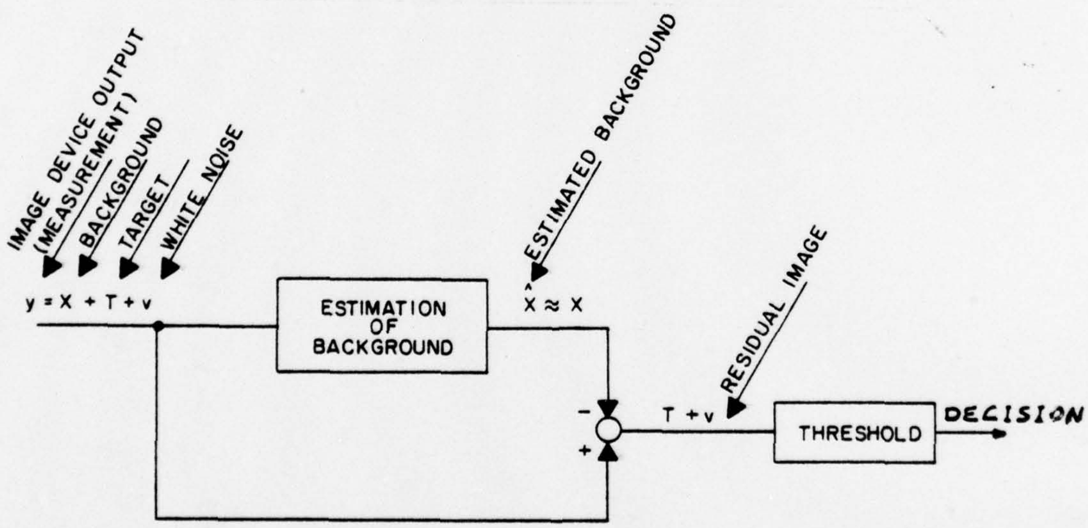


Fig. 1: Target Detection from a Noisy Image

In each of the three steps, some original ideas have been developed.

3. In the "mathematical modeling of the background" the starting point is that there is a correlation between points of the background. The two dimensional autocorrelation function was assumed to be known. Only stationary random fields were treated.

The main idea of the mathematical modeling is to represent the image by difference equations that are forced by white noise. The conventional way of representing a process whose autocorrelation function is given is by using a "Markov Process." The value of a point, $x(k,\ell)$, is represented by its neighbors:

$$\hat{x}(k,\ell) = \Sigma\alpha_{p,q}x(p,q)$$

$\{(p,q)\}$ is a group of "neighbors" near the point (k,ℓ). And:

$$\begin{aligned} x(k,\ell) &= \hat{x}(k,\ell) + u(k,\ell) \\ &= \Sigma\alpha_{p,q}x(p,q) + u(k,\ell). \end{aligned}$$

$u(k,\ell)$ is called "modeling error."

The weighting coefficients $\alpha_{p,q}$ are chosen to make the variance of the modeling error minimum. This is done by using the well known "orthogonality principle" in optimal estimation theory [see ch. II section H]. This technique suffers from some difficulties as follows:

- The weighting coefficients $\alpha_{p,q}$ have no simple expressions.

- The number of neighbors one has to use is theoretically infinite for some types of processes.

- The coefficients are found by solving many algebraic equations, especially in the two-dimensional case. The number of equations is equal to the number of coefficients.

- It is difficult to describe the two-dimensional difference equation in a "state-vector structure".

The optimal-estimation approach to linear modeling is summarized in Ch. II section H.

The method of modeling that was derived in this thesis is to find a recursive, linear, invariant filter $H(z_1,z_2)$ such as:

- when forcing the input with white noise, the output of the filter will be a random process that has the same autocorrelation function as the given field.

So, instead of dealing with the spectrum of the process, one has to deal with white noise and a linear filter. This is well known.

Rosser [Ref. 12], in 1975, suggested a state-space-structure for two dimensional fields and also derived some of the properties for this structure [see Ch. II Section G]. All of the examples that were chosen in this thesis led to a filter, $H(z_1,z_2)$, whose state-space-structure fitted easily into Rosser's structure.

Moreover, this method of modeling doesn't suffer from any of the disadvantages of the previous method. It should be emphasized that this method is <u>limited</u> only to separable, two-dimensional autocorrelation functions and to causal models.

This method is called "Filter Response Method" and is summarized in Ch. II Section F.

4. The estimation of the background is actually an extension of the one dimensional Kalman Filter. The problem is roughly defined: "Given a noisy measurement

$$y(k,\ell) = x(k,\ell) + v(k,\ell)$$

where $x(k,\ell)$ is a correlated field and $v(k,\ell)$ is white noise. The filter should estimate $x(k,\ell)$, denoted $\hat{x}(k,\ell)$. It was shown that a recursive, two dimensional filter cannot be optimal in the sense that the error cannot be made orthogonal to the measurements, as some researchers have tried to do [Ref. 10]. There is a distinct difference between one dimensional processing and two dimensional processing. Therefore the method in this thesis is to define a reasonable structure for the recursive filter, and then to calculate the parameters of the filter to minimize the mean of the square error. Recursive equations for calculating the filter parameters (gain) and variance of error were developed.

The results were checked in two ways:

1) By comparing with the optimal, non recursive estimator.

2) By simulation. A correlated image was added to a field of white noise. Then the correlated part of the combined image was estimated by using the recursive filter. The variance of error was compared with the theoretical variance of error.

The results for all cases that were checked show good coincidence.

5. In the study of target detection, the problem was to detect lines. Most of the algorithms suffer from one disadvantage. When detecting a line, an a-priori assumption must be made about the direction of the line. Therefore, in order to detect lines in several directions, the image has to be scanned several times, or during one scanning to do several calculations for each point. No doubt that for real time applications all of those algorithms have a great disadvantage. The algorithm developed in this thesis detects lines, regardless of their direction. The key point in this algorithm is feedback from the detection (which is a decision process) to the filter. Most algorithms of filtering and detection do it in two separated steps. Fig. 3 shows the target that was detected from the original image in Fig. 2.

6. Outline Of Chapters

Chapter II discusses the problem of two dimensional random fields. In order to make this chapter "stand alone" for reading, some background material was included in Sections B and C. This background material includes information on two-dimensional operations and random processes. Section E describes models for a one-dimensional random process. Section F extends the methods of modeling for two dimensional processes.

Chapter III is a review of estimation theory. It explains the "position" of recursive, linear estimation, with quadratic form, in estimation theory.

Chapter IV solves the problem of estimating a two-dimensional process, when, originally, this process is combined with white noise.

Chapter V shows the application of Ch. II and IV (modeling and estimation) to the specific problem of target detection.

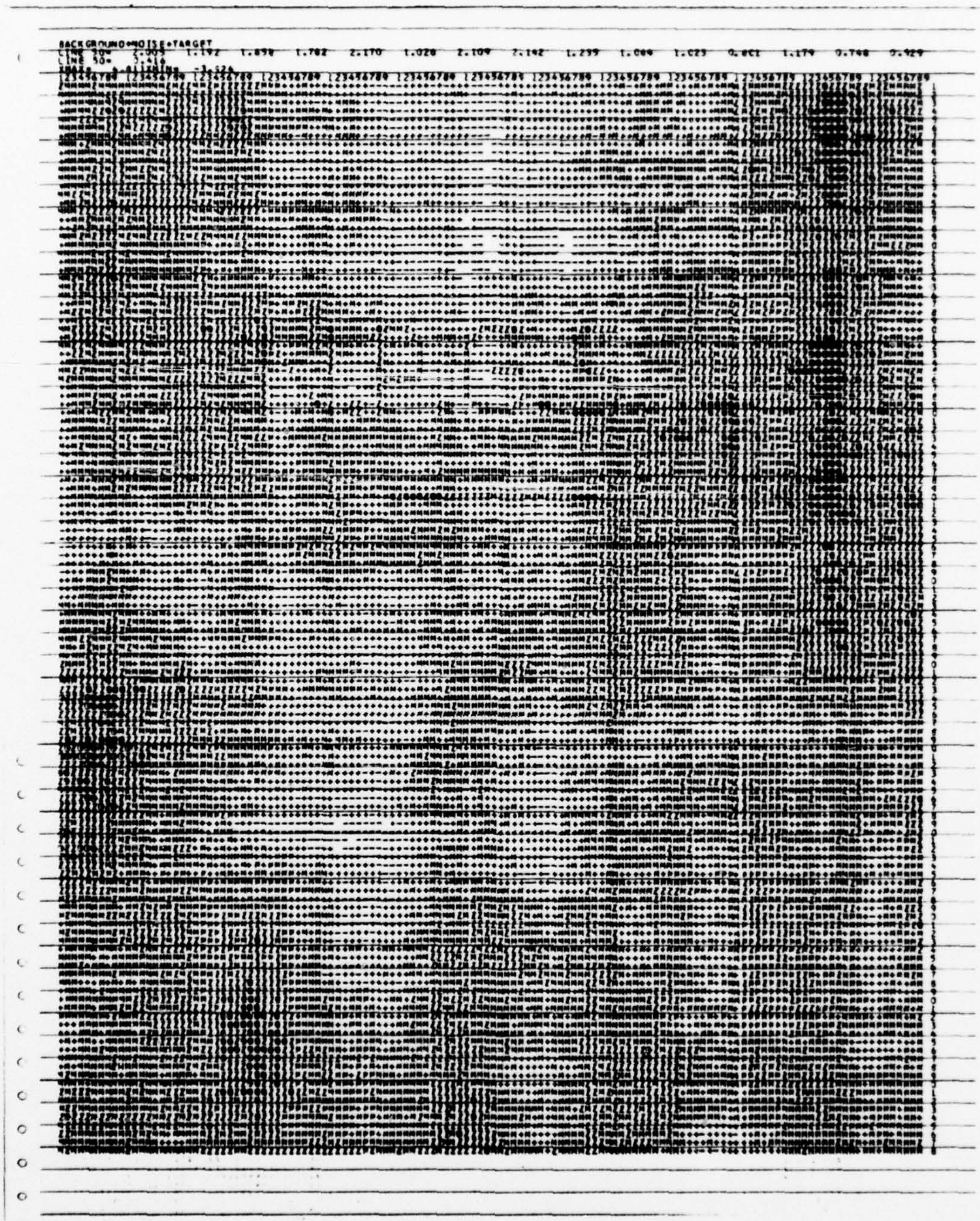


Fig. 2: A Typical Image, Includes a Target, Correlated Noise (Background) and Random Noise

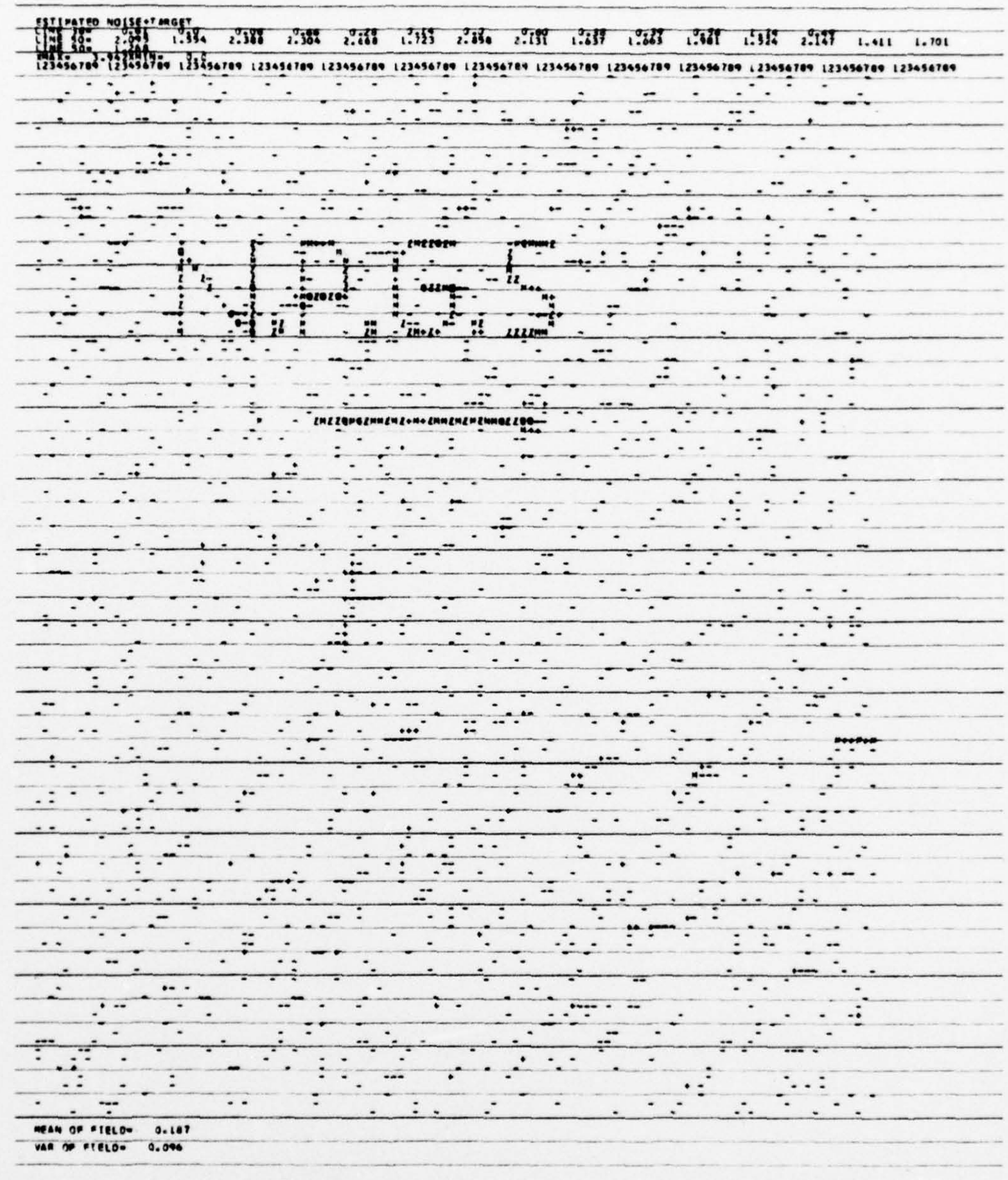


Fig. 3: The Target That Results From The Original Image That Is Shown in Fig 2

II. <u>LINEAR MODELING OF TWO-DIMENSIONAL RANDOM FIELDS</u>

A. INTRODUCTION

1. This chapter contains the basic principles of modeling a random process by linear equations. The <u>basic assumption</u> will be that the <u>statistical relationship</u> between each two measurements is known. This relationship is the so-called <u>autocorrelation-function</u>. The techniques that will be used are mainly extensions of the principles that are used in the one-dimensional case. Terms like "Markov Model", "state variables", "Factorization" will be used.

2. In the one-dimensional case the state-equations have the form:

$$\underset{\sim}{x}(k+1) = \underset{\sim}{A} \cdot \underset{\sim}{x}(k) + \underset{\sim}{B} \cdot \underset{\sim}{u}(k)$$

The words "one-dimensional" refer to the variable k (but the state vector $\underset{\sim}{x}(k)$ can be a multidimensional vector). It will be shown how to use state-equations for a two dimensional random field, in which the variables are k and ℓ .

<u>Note</u>: When a random variable depends only on one dimension, k, it is called a "stochastic process". When the dimensionality is two or more, as k,ℓ, it is called a "random field".

3. What is the philosophy behind describing a random process by state equations? The answer; it is difficult to deal with random fields in which there is correlation between adjacent points. Therefore, the main idea is to show how a correlated field can be described as an output of a linear system which is driven by white noise. So, instead of handling a correlated field, one has to handle white noise and a linear system. The last problem is well known from system theory. Figure 4 illustrates the concept of modeling by the "filter response method".

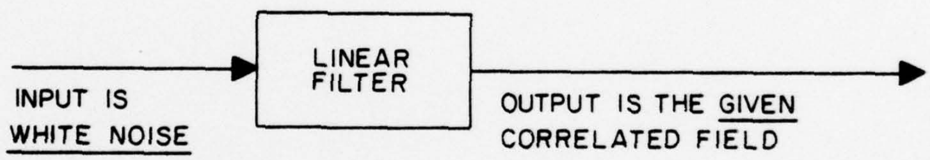


Fig. 4: Concept of "Modeling" a Random Field as an Output of a Linear-Filter

4. Although this thesis mainly delas with discrete fields, continuous models will be described also, in some cases.

5. In order to make this chapter "stand alone" for reading, some background material was included in sections B, C.

6. Section D describes the "Filter Response Method" of modeling a random process.

7. Section E describes the one dimensional processes, and Section F is its extension to two dimensions. The Markov-Process is the key to the modeling procedure. It will be shown that from the same starting point (the given auto-correlation function) two models can be developed.

8. The modeling by the "Filter Response Method" leads to models that have the structure

$$\begin{bmatrix} M(k+1,\ell) \\ \cdots\cdots \\ N(k,\ell+1) \end{bmatrix} = \left[\begin{array}{c:c} A_1 & A_2 \\ \hdashline A_3 & A_4 \end{array}\right] \cdot \begin{bmatrix} M(k,\ell) \\ \cdots\cdots \\ N(k,\ell) \end{bmatrix} + \begin{bmatrix} B_1 \\ \cdots\cdots \\ B_2 \end{bmatrix} \cdot u(k,\ell)$$

Roesser [12] described the properties of that structure and section G is a summary of those properties.

9. Section H summarizes another method of linear modeling, using optimal estimation theory (the orthogonality principle). This method leads to different results than the filter response method with the exception of one case.

10. So, two methods of modeling will be shown (sections F, H). Section I is a summary of this chapter. It also compares the two modeling methods.

B. BACKGROUND MATERIAL:
TWO DIMENSIONAL OPERATIONS AND FILTERS

This section is a summary of two dimensional operations as Fourier transformation, $\mathcal{Z}$ transformation, two dimensional filters, etc. Some of these operations are used in the next sections.

1. Fourier Transformation

Definition: given a function h(x,y), the Fourier transformation $H(\omega_x,\omega_y)$ is

$$H(\omega_x,\omega_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) \cdot e^{-j\omega_x x} \cdot e^{-j\omega_y y} \cdot dxdy$$

$$= \mathcal{F}\left\{h(x,y)\right\} \qquad 2\text{-}1$$

and the inverse transform

$$h(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(\omega_x,\omega_y)\cdot e^{j\omega_x x}\cdot e^{j\omega_y y}\cdot d\omega_x \, d\omega_y \qquad 2\text{-}2$$

$$= \mathcal{F}^{-1}\left\{H(\omega_x,\omega_y)\right\}$$

The two dimensional Fourier Transformation is mostly used in spatial operations, and ω_x, ω_y are called spatial frequencies.

The properties of the Fourier transformation are:

Linearity $\quad \mathcal{F}\left\{a\cdot f_1(x,y) + b\cdot f_2(x,y)\right\}$

$$= a\cdot F_1(\omega_x,\omega_y) + b\cdot F_2(\omega_x,\omega_y) \qquad 2\text{-}3$$

Scaling

$$\mathcal{F}\left\{f(ax, by)\right\} = \frac{1}{a\cdot b}\cdot F\left(\frac{\omega_x}{a}, \frac{\omega_y}{b}\right) \qquad 2\text{-}4$$

Shift Operation

$$\mathcal{F}\left\{f(x-a, y-b)\right\} = e^{-j(\omega_x a+\omega_y b)}\cdot F(\omega_x, \omega_y) \qquad 2\text{-}5$$

Convolution

$$\mathcal{F}\left\{f(x,y)*h(x,y)\right\} = F(\omega_x, \omega_y)\cdot H(\omega_x, \omega_y) \qquad 2\text{-}6$$

Parsaval Theorem

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\cdot g^*(x,y)\; dx\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\omega_x, \omega_y)\cdot G^*(\omega_x, \omega_y)\cdot d\omega_x d\omega_y \qquad 2\text{-}7$$

Autocorrelation

$$\mathcal{F}\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi,\eta)\cdot f(\xi-x, \eta-y)\right\} = \left|F(\omega_x, \omega_y)\right|^2 \qquad 2\text{-}8$$

Gradient

$$\mathcal{F}\left\{\frac{\partial f(x,y)}{\partial x}\right\} = j\omega_x \cdot F(\omega_x,\omega_y) \qquad 2\text{-}9$$

$$\mathcal{F}\left\{\frac{\partial f(x,y)}{\partial y}\right\} = j\omega_y \cdot F(\omega_x,\omega_y) \qquad 2\text{-}10$$

Inversion

$$\begin{aligned}\mathcal{F}\left\{\mathcal{F}^{-1}\left\{f(x,y)\right\}\right\} &= \mathcal{F}^{-1}\left\{\mathcal{F}\left\{f(x,y)\right\}\right\} \\ &= f(x,y)\end{aligned} \qquad 2\text{-}11$$

Rotation

$$\mathcal{F}\left\{\mathcal{F}\left\{f(x,y)\right\}\right\} = f(-x,-y) \qquad 2\text{-}12$$

Equation 2-9 will be useful in the next sections and the proof for this equation is given as follows:

$$\begin{aligned}\mathcal{F}\left\{\frac{\partial f(x,y)}{\partial x}\right\} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial f(x,y)}{\partial x} \cdot e^{-j\omega_x x} \cdot e^{-j\omega_y y}\, dxdy \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} \frac{\partial f(x,y)}{\partial x} \cdot e^{-j\omega_x x}\, dx\right] \cdot e^{-j\omega_y y}\, dy\end{aligned}$$

Using integration by parts:

$$\mathcal{F}\left\{\frac{\partial F(x,y)}{\partial x}\right\} = \int_{-\infty}^{\infty}\left[\left(f(x,y)\cdot e^{-j\omega_x x}\right)\Big|_{-\infty}^{\infty}\right.$$

$$\left. + j\omega_x \cdot \int_{-\infty}^{\infty} f(x,y)\cdot e^{-j\omega_x x}\right]\cdot e^{-j\omega_y y}\cdot dy$$

If $f(x,y) = 0$ at $x = \pm\infty$, $y = \pm\infty$

$$\mathcal{F}\left\{\frac{\partial f(x,y)}{\partial x}\right\} = j\omega_x \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\cdot e^{-j\omega_x x}\cdot e^{-j\omega_y y}\cdot dy$$

$$= \underline{j\omega_x \cdot F(\omega_x, \omega_y)}$$

<u>Note</u>: In the one dimensional case most equations appear with "t" as the variable. In order to make the equations in the two dimensional case similar to the one dimensional case, the notations x, y, ω_x, ω_y will be replaced by t_1, t_2, ω_1, ω_2 .

2. Two Dimensional Z Transformation

Definition: Given a discrete field $X(n_1,n_2)$, the two dimensional z transformation is defined:

$$\mathcal{Z}\{x(n_1,n_2)\} = X(z_1,z_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} x(n_1,n_2)\cdot z_1^{-n_1} z_2^{-n_2} \qquad 2\text{-}13$$

An important property of the transformation which will be used later is

$$\text{if:} \quad \mathcal{Z}\{x(n_1,n_2)\} = X(z_1,z_2) \qquad 2\text{-}14$$

then:

$$\mathcal{Z}\{x(n_1-1,n_2)\} = z_1^{-1}X(z_1,z_2) \qquad 2\text{-}15$$

$$\mathcal{Z}\{x(n_1,n_2-1)\} = z_2^{-1}X(z_1,z_2) \qquad 2\text{-}16$$

$$\text{Proof:} \quad \{x(n_1,n_2)\} = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} x(n_1-1,n_2)\ z_1^{-n_1}\cdot z_2^{-n_2}$$

$$= \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} x(m\ ,n_2)\cdot z_1^{-(m+1)}\cdot z_2^{-n_2}$$

$$= z_1^{-1} \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} x(m\ ,n_2)\cdot z_1^{-m}\cdot z_2^{-n_2}$$

$$= z_1^{-1} \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(n_1,n_2)\cdot z_1^{-n_1} z_2^{-n_2}$$

$$= z_1^{-1} \cdot \mathcal{Z}\{x(n_1,n_2)\}$$

3. Response of a Linear System

If $h(t_1,t_2)$ is the "point spread function" of a two dimensional system, then the response to a given input $x(t_1,t_2)$ will be the two dimensional convolution:

$$y(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t_1-\tau_1,t_2-\tau_2)\cdot x(t_1,t_2)\cdot d\tau_2 d\tau_1 \qquad 2\text{-}17$$

and the transfer function is defined:

$$H(\omega_1,\omega_2) = \frac{Y(\omega_1,\omega_2)}{X(\omega_1,\omega_2)} \qquad 2\text{-}18a$$

In the discrete case, the convolution has the form:

$$y(k,\ell) = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} h(n,m)\cdot x(k-n,\ell-m) \qquad 2\text{-}19$$

where $h(n,m)$ is the discrete "point spread function" and the discrete transfer function:

$$H(z_1,z_2) = \frac{Y(z_1,z_2)}{X(z_1,z_2)} \qquad 2\text{-}18b$$

4. Separable Functions

Definition: $f(t_1,t_2)$ is called "separable" if there are two one-dimensional functions $f_1(t)$, $f_2(t)$ such as:

$$f(t_1,t_2) = f_1(t_1)\cdot f_2(t_2) \quad .$$

Theorem

If: $f(t_1,t_2) = f_1(t_1)\cdot f_2(t_2)$

Then: $F(\omega_1,\omega_2) = F_1(\omega_1)\cdot F_2(\omega_2)$ 2-20

where: $$F_1(\omega_1) = \int_{-\infty}^{\infty} f_1(t_1)\cdot e^{-j\omega_1 t_1}\cdot dt_1$$

$$F_2(\omega_2) = \int_{-\infty}^{\infty} f_2(t_2)\cdot e^{-j\omega_2 t_2}\cdot dt_2$$

In the discrete case:

If: $x(n_1,n_2) = x_1(n_1)\cdot x_2(n_2)$

Then: $\mathcal{Z}\{x(n_1,n_2)\} = \mathcal{Z}\{x(n_1)\}\cdot\mathcal{Z}\{x(n_2)\}$ 2-21

C. BACKGROUND MATERIAL:
TWO DIMENSIONAL LINEAR SYSTEMS WITH RANDOM INPUTS

In this section expressions for the response of two dimensional filters to random inputs are developed. It will be shown that the expressions are very similar to the one dimensional case. First, a brief review of analysis of random signals (in two dimensions) is introduced.

Given: A "brightness function" of a two dimensional field, $x(t_1,t_2)$ [see fig. 5].

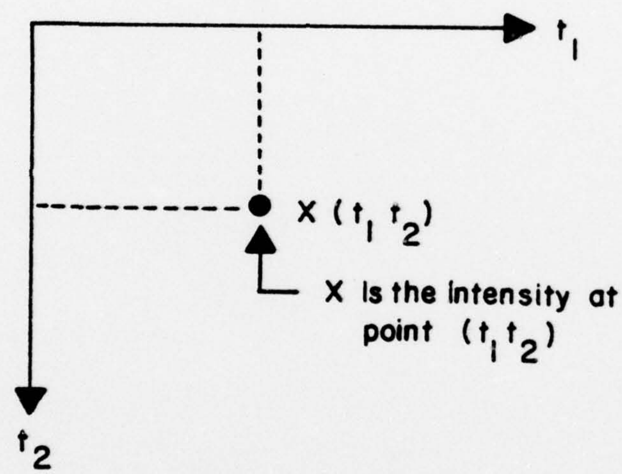


Fig. 5: The Two Dimensional Field

1. The probability that $x(t_1,t_2)$ lies in a finite range $a \leq x(t_1,t_2) \leq b$ at a point in the field, (t_1,t_2), is given by the integral

$$Pr(a \leq x(t_1,t_2) \leq b) = \int_a^b f_1(x,t_1,t_2) \cdot dx \qquad 2\text{-}22$$

$f_1(x,t_1,t_2)$ is the "probability density function."

2. The <u>first moment</u> is the expected value (or mean) and is defined by the integral:

$$\overline{x(t_1,t_2)} = E\{x(t_1,t_2)\} = \int_{-\infty}^{\infty} x \cdot f_1(x,t_1,t_2) \cdot dx \tag{2-23}$$

3. The second order moment (<u>variance</u>) of a point in the field is:

$$E\{x^2(t_1,t_2)\} = \overline{x^2(t_1,t_2)} = \int_{-\infty}^{\infty} x^2 \cdot f_1(x,t_1,t_2) \cdot dx \tag{2-24}$$

4. The <u>autocorrelation</u> function, between points (t_1,t_2) and (t_1',t_2') is defined as:

$$R_{xx}(t_1,t_2,t_1',t_2') = \overline{x(t_1,t_2) \cdot x(t_1',t_2')}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot x', f_2(x_1(t_1,t_2), x_1'(t_1',t_2'))\, dx dx' \tag{2-25}$$

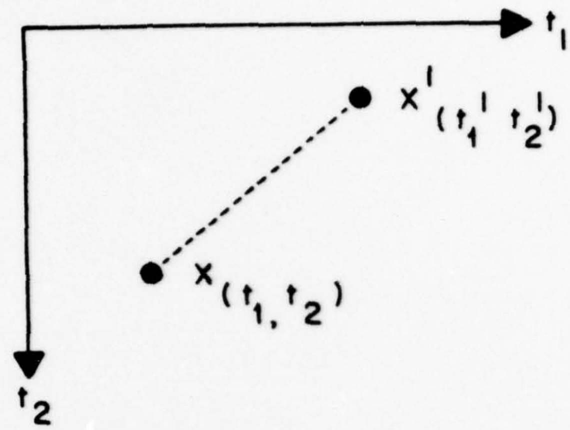


Fig. 6: The Parameters That Take Part in the Autocorrelation Equation

$f_2(\cdot)$ is the "joint probability density function".

Fig. 6 shows the parameters that take part in the autocorrelation equation.

5. Stationary Fields.

The assumption that the field is stationary means that the statistics of a point in the field is not dependent on the location of the point.

In this case the mean and variance have the form

$$E\{x(t_1,t_2)\} = \bar{x} \; , \; E\{x^2(t_1,t_2)\} = \sigma_s^{\,2} \qquad 2\text{-}26$$

and the autocorrelation function:

$$R_{xx}(t_1-t_1',t_2-t_2') = R_{xx}(\tau_1,\tau_2)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \cdot x' \cdot f_2(x\,,x',\tau_1,\tau_2)\,dx\cdot dx' \qquad 2\text{-}27$$

Such a field is called (in the two dimensional case) a "Homogeneous Field."

6. Ergodicity

A further simplifying assumption which is usually done during the analysis of stationary signals is the ergodic property. This hypothesis states that under certain conditions, present in many cases, the statistical averaging of x at a given point is equal to the spatial averaging of all points. That means:

$$E(\overline{x}_{(t_1,t_2)}) = \overline{x} = \langle x \rangle \qquad 2\text{-}28$$

where by definition:

$$\langle x \rangle = \lim_{\substack{T_1\to\infty \\ T_2\to\infty}} \frac{1}{4\cdot T_1\cdot T_2} \int_{-\infty}^{\infty}\int_{\infty}^{\infty} x(t_1,t_2)\cdot dt_1\cdot dt_2 \qquad 2\text{-}29$$

$\langle x \rangle$ is the spatial averaging.

The autocorrelation function will be:

$$R_{xx}(\tau_1,\tau_2) = E\{x(t_1,t_2)\ x(t_1+\tau_1,t_2+\tau_2)\} \qquad 2\text{-}30$$

$$= \lim_{\substack{T_1\to\infty \\ T_2\to\infty}} \frac{1}{4\cdot T_1\cdot T_2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1,t_2)\ x(t_1+\tau_1,t_2+\tau_2)dt_1dt_2$$

7. Power Spectrum Density Function

In the study of random signals the concept of power spectral density function takes place. For the purpose of this review the spectral density of two dimensional fields is defined as the two dimensional Fourier transofrmation of the autocorrelation function. It is called $P_{xx}(\omega_1,\omega_2)$:

$$P_{xx}(\omega_1,\omega_2) \triangleq \mathcal{F}\{R_{xx}(\tau_1,\tau_2)\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{xx}(\tau_1,\tau_2)\cdot e^{-j\omega_1\cdot\tau_1}\, e^{-j\omega_2\cdot\tau_2}\cdot d\tau_1\cdot d\tau_2 \qquad 2\text{-}31$$

and the inverse transform gives the result:

$$R_{xx}(\tau_1,\tau_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P_{xx}(\omega_1,\omega_2)\cdot e^{j\omega_1\cdot\tau_1}\cdot e^{j\omega_2\cdot\tau_2}\cdot d\omega_1\cdot d\omega_2 \qquad 2\text{-}32$$

From the last expression, $P_{xx}(\omega_1,\omega_2)$ gives the density of mean square value of the variable over the spectrum of the real frequencies.

8. Response Of A Linear System To Random Inputs

Given: A linear, two-dimensional linear filter with a transfer function $H(j\omega_1,j\omega_2)$. The input to the filter is a stationary process $x(t_1,t_2)$ with an autocorrelation function $R_{xx}(\tau_1,\tau_2)$.

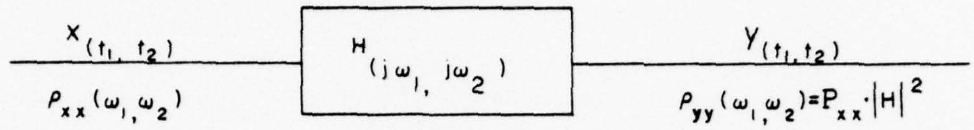


Fig. 7: Response Of A Filter To Random Input

<u>Question</u>: Find the autocorrelation of the output signal. The answer will be presented in next theorems.

<u>*Theorem:*</u> The mean of the output is given by

$$\boxed{\bar{y} = \bar{x} \cdot H(j\omega_1, j\omega_2)\Big|_{\substack{\omega_1=0 \\ \omega_2=0}}} \qquad 2\text{-}33$$

Proof:

$$\overline{y(t_1,t_2)} = E\{\int_{-\infty}^{\infty}\int_{\infty}^{\infty} x(t_1-\alpha_1, t_2-\alpha_2)\cdot h(\alpha_1,\alpha_2)\cdot d\alpha_1 \cdot d\alpha_2\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\{x(t_1-\alpha_1, t_2-\alpha_2)\}\cdot h(\alpha_1,\alpha_2)\cdot d\alpha_1 \cdot d\alpha_2$$

and because the process is stationary:

$$E\{x(t_1-\alpha_1, t_2-\alpha_2)\} = E\{x(t_1,t_2)\} = \bar{x}$$

$$\overline{y} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overline{x} \cdot h(\alpha_1,\alpha_2) \cdot d\alpha_1 \cdot d\alpha_2$$

$$= \overline{x} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\alpha_1,\alpha_2) \cdot d\alpha_1 \cdot d\alpha_2$$

$$= \overline{x} \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\alpha_1,\alpha_2) \cdot e^{-j\omega_1 \cdot t_1} \cdot e^{-j\omega_2 \cdot t_2} \cdot d\alpha_1 \cdot d\alpha_2 \right] \Bigg|_{\substack{\omega_1=0 \\ \omega_2=0}}$$

$$= \overline{x} \cdot H(j\omega_1,j\omega_2) \Bigg|_{\substack{\omega_1=0 \\ \omega_2=0}}$$

Q.E.D.

<u>Theorem</u>: The cross-spectra of x,y is given by

$$\boxed{P_{xy}(\omega_1,\omega_2) = P_{xx}(\omega_1,\omega_2) \cdot H(j\omega_1,j\omega_2)} \qquad 2\text{-}34$$

Proof:

$$R_{xy}(t_1,t_2) = E\{y(t_1+\tau_1,t_2+\tau_2) \cdot x(t_1,t_2)\}$$

$$= E\{y(t_1,t_2) \cdot x(t_1-\tau_1,t_2-\tau_2)\}$$

$$R_{xy}(\tau_1,\tau_2) = E\{x(t_1-\tau_1,t_2-\tau_2)\cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1-\alpha_1,t_2-\alpha_2)\cdot h(\alpha_1,\alpha_2)d\alpha_1 d\alpha_2\}$$

$$= E\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1-\tau_1,t_2-\tau_2)\cdot x(t_1-\alpha_1,t_2-\alpha_2)\cdot h(\alpha_1,\alpha_2)\ d\alpha_1 d\alpha_2\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\{x(t_1-\tau_1,t_2-\tau_2)\cdot x(t_1-\alpha_1,t_2-\alpha_2)\}\cdot h(\alpha_1,\alpha_2)d\alpha_1 d\alpha_2\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{xx}(\tau_1-\alpha_1,\tau_2-\alpha_2)\cdot h(\alpha_1,\alpha_2)\ \cdot\ d\alpha_1\ \cdot\ d\alpha_2$$

$$R_{xy}(\tau_1,\tau_2) = R_{xx}(\tau_1,\tau_2) * h(\tau_1,\tau_2) \qquad \text{2-35}$$

The last expression is a convolution. Therefore, taking the Fourier transform leads to:

$$P_{xy}(\omega_1,\omega_2) = P_{xx}(\omega_1,\omega_2)\cdot H(j\omega_1,j\omega_2)$$

Q.E.D.

<u>Theorem</u>: The spectral density of the output is given by:

$$\boxed{\begin{aligned} P_{yy}(\omega_1,\omega_2) &= P_{xx}(\omega_1,\omega_2)\cdot H(j\omega_1,j\omega_2)\cdot H^*(j\omega_1,j\omega_2) \\ &= P_{xx}(\omega_1,\omega_2)\cdot |H(j\omega_1,j\omega_2)|^2 \end{aligned}} \qquad \text{2-36}$$

Proof:

$$R_{yy}(\tau_1,\tau_2) = E\{y(t_1,t_2)\cdot y(t_1+\tau_1,t_2+\tau_2)\}$$

$$= E\{y(t_1+\tau_1,t_2+\tau_2)\cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1-\alpha_1,t_2-\alpha_2)\cdot h(\alpha_1,\alpha_2)\cdot d\alpha_1 d\alpha_2\}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{xy}(\alpha_1+\tau_1,\alpha_2+\tau_2)\cdot h(\alpha_1,\alpha_2)\cdot d\alpha_1\cdot d\alpha_2$$

define: $\alpha_1 = -\theta_1 \qquad \alpha_2 = -\theta_2$

$d\alpha_1 = -d\theta_1 \qquad d\alpha_2 = -d\theta_2$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{xy}(\tau_1-\theta_1,\tau_2-\theta_2)h(-\theta_1,-\theta_2)\, d\theta_1 d\theta_2$$

$$R_{yy}(\tau_1,\tau_2) = R_{yx}(\tau_1,\tau_2) * h(-\tau_1,-\tau_2) \qquad \text{2-37}$$

Taking the Fourier transform of 2-37:

$$P_{yy}(\omega_1,\omega_2) = P_{yx}(\omega_1,\omega_2)\cdot H(-j\omega_1,-j\omega_2)$$

But:

$$P_{yx}(\omega_1,\omega_2) = P_{xx}\cdot H(j\omega_1,j\omega_2)$$

and:

$$P_{yy}(\omega_1,\omega_2) = P_{xx}(\omega_1,\omega_2) \cdot H(j\omega_1,j\omega_2) \cdot H(-j\omega_1,-j\omega_2)$$

$$= P_{xx}(\omega_1,\omega_2) \cdot H(j\omega_1,j\omega_2) \cdot H^*(j\omega_1,j\omega_2)$$

$$= P_{xx}(\omega_1,\omega_2) \cdot \left| H(j\omega_1,j\omega_2) \right|^2$$

Q.E.D.

The Discrete Case:

In the discrete case, a stationary, two-dimensional field has an autocorrelation function defined as:

$$R_{xx}(n,m) \triangleq E\{x(k,\ell)x(k+n,\ell+m)\} \qquad 2\text{-}38$$

Equivalent to the power density function in the continuous case, we define:

$$P_{xx}(z_1,z_2) \triangleq \mathcal{Z}\{R_{xx}(n,m)\} \qquad 2\text{-}39$$

Since it is known that the z transformation is related to the Fourier transform by:

$$z \longleftrightarrow e^{j\cdot\omega\cdot T}$$

$$z^{-1} \longleftrightarrow e^{-j\cdot\omega\cdot T}$$

It is, therefore, clear that after passing a random process through a two-dimensional discrete filter, the statistics of the output are:

$$\boxed{P_{yx}(z_1,z_2) = P_{xx}(z_1,z_2) \cdot H(z_1,z_2)} \qquad 2\text{-}40$$

$$\boxed{P_{yy}(z_1,z_2) = P_{xx}(z_1,z_2) \cdot H(z_1,z_2) \cdot H(z_1^{-1},z_2^{-1})} \qquad 2\text{-}41$$

The last two expressions can be proven exactly in the same way as in the continuous case, by starting from the discrete convolution.

9. Isotropic Fields

The homogeneous (stationary) field was defined as the case where the statistics of the field do not depend on the location. If the statistics are not only independent of the location, but also independent of the direction, the field is called isotropic.

Example:

If:

$$R_{xx}(\tau_1,\tau_2) = e^{-\alpha_1\cdot\tau_1} \cdot e^{-\alpha_2\cdot\tau_2}$$

the field is not isotropic. The correlation is greater in the directions of the system axis, than in other directions.

But if:

$$R_{xx}(\tau_1,\tau_2) = e^{-\alpha\sqrt{\tau_1^2+\tau_2^2}}$$

the field is isotropic. The correlation between two points depends only on the distance between these two points and does not depend on the direction.

D. BACKGROUND MATERIAL: THE CONCEPT OF MODELING BY "FILTER RESPONSE METHOD" AND THE FACTORIZATION PROBLEM

1. Continuous, One Dimensional Case

Given: A covariance function of a process, $R_{xx}(t,\tau)$.

Find: A linear system (differential equation model) such as:

- when the input is white noise
- the output has an autocorrelation function $R_{xx}(t,\tau)$.

Fig. 8 defines the problem:

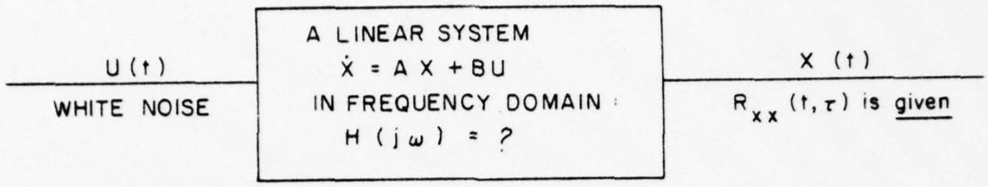


Fig. 8: Definition of the Modeling Problem in the Continuous Case

Solution:

First, the solution exists only if x(t) is stationary, say:

$$R_{xx}(t,\tau) = R_{xx}(\tau)$$

In this case, by using Eq. 2-36 and by assuming that the transfer function of the required filter is $H(j\omega)$, the power spectral density of the output is:

$$P_{xx}(\omega) = P_{uu}(\omega) \cdot H(j\omega) \cdot H^*(j\omega)$$

u(t) is assumed to be white noise, therefore:

$$P_{uu}(\omega) = 1 \qquad 2\text{-}42$$

and:

$$P_{xx}(\omega) = H(j\omega) \cdot H^*(j\omega) = \begin{bmatrix} \text{given power} \\ \text{spectrum} \end{bmatrix} \qquad 2\text{-}43$$

In this problem $P_{xx}(\omega)$ is <u>given</u>. Therefore, the solution for the required filter is <u>to find a function $H(j\omega)$ that satisfies Equation 2-43</u>.

Such a solution exists for symmetric autocorrelation functions (then $P_{xx}(\omega)$ is a real function of ω).

2. <u>Discrete, One Dimensional Case</u>.

In this case $R_{xx}(n)$ is given.

It is required: to find a discrete filter, $H(z)$, so that when the input is white noise, the output will have the given autocorrelation function $R_{xy}(n)$.

Here, the solution is by using the equation:

$$P_{xx}(z) = P_{vv}(z) \cdot H(z) \cdot H(z^{-1}).$$

If the <u>input is white noise then</u>:

$$P_{uu}(z) = 1$$

Therefore the solution for the required filter is to find a function $H(z)$ that satisfies:

$$P_{xx}(\omega) = H(z) \cdot H(z^{-1})$$

where $P_{xx}(\omega)$ is <u>given</u> in our problem.

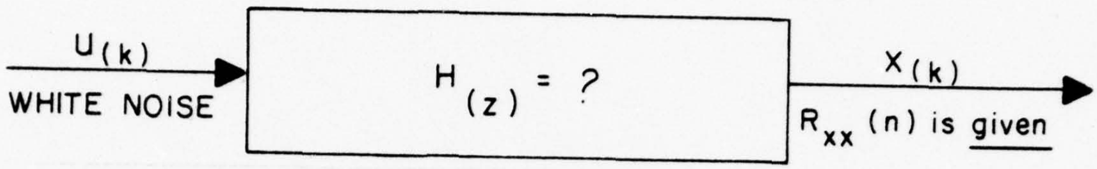


Fig. 9: The Modeling Problem in the Discrete Case

2. The Two-Dimensional Case

In the two dimensional case the extension of this method requires factoring $P_{xx}(\omega_1,\omega_2)$ to $H(\omega_1,\omega_2)\cdot H^*(\omega_1,\omega_2)$ and $P_{xx}(z_1,z_2)$ to $H(z_1,z_2)\cdot H(z_1^{-1},z_2^{-1})$. The problem is that there is no factorization technique in the two-dimensional case. Therefore this modeling method is limited to separable autocorrelation functions. Section H shows another method of modeling (optimal estimation approach) that does not suffer from this limitation, but has other disadvantages.

E. BACKGROUND MATERIAL: MODELING OF ONE DIMENSIONAL RANDOM PROCESSES (BY "FILTER RESPONSE METHOD")

There is a special class of stationary random processes, which is very common. The basic assumption of these random processes is that the correlation between two points decreases exponentially with respect to the distance between the two points.

This section considers the one-dimensional case. The next section will be an extension of these processes to the two dimensional fields.

1. Markov Process (continuous)

This is a stationary random process with an autocorrelation function:

$$R_{xx}(\tau) = \sigma^2 \cdot e^{-\alpha \cdot |\tau|} + m^2 \qquad \text{2-44}$$

m is the mean of the process, and will be taken zero without any loss of generality:

$$\boxed{R_{xx}(\tau) = \sigma^2 \cdot e^{-\alpha \cdot |\tau|}} \qquad \text{2-45}$$

and the power spectrum density function

$$\boxed{P_{xx}(\omega) = \mathcal{F}\{R_{xx}(\tau)\} = \frac{2 \cdot \alpha \cdot \sigma^2}{\omega^2 + \alpha^2}} \qquad \text{2-46}$$

Fig. 10 is a plot of $R_{xx}(\tau)$ and $P_{xx}(\omega)$.

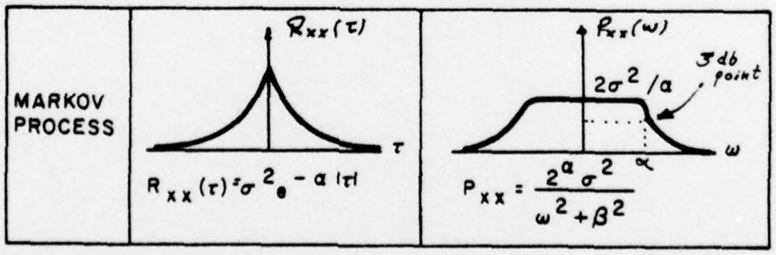


Fig. 10: Markov Process

Theorem: The process in 2-45 can be generated by passing white noise through a simple filter. The spectral density of the white noise is:

$$N_o = 2\alpha\sigma^2$$

and the filter is:

$$H(j\omega) = \frac{1}{\alpha + j\omega}$$

See Fig. 11.

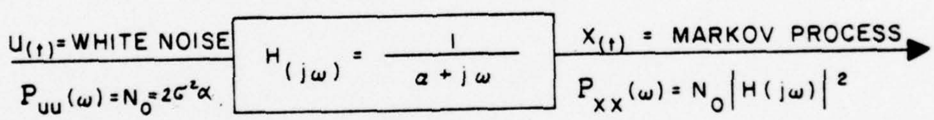


Fig. 11: The Linear System, $H(j\omega)$, for Generating First-Order, One-Dimensional Markov Process

Proof:

Using 2.36 for one dimensional case:

$$\begin{bmatrix} \text{spectral density} \\ \text{at the filter} \\ \text{output} \end{bmatrix} = R_{xx}(\omega) =$$

$$\begin{bmatrix} \text{spectral density} \\ \text{at the filter} \\ \text{output} \end{bmatrix} = R_{VV}(\omega) \; |H(j\omega)|^2$$

$$= 2\sigma^2 \cdot \frac{1}{\alpha + j\omega} \cdot \frac{1}{\alpha - j\omega}$$

$$= \frac{2 \cdot \alpha \cdot \sigma^2}{\alpha^2 + \omega^2}$$

$$= \begin{bmatrix} \text{given spectral density of} \\ \text{the Markov Process} \end{bmatrix}$$

Q.E.D.

The filter in Fig. 11 can be represented by a differential equation, which will be associated with the process:

$$\boxed{\dot{x}(t) = -\alpha \cdot x(t) + u(t)} \qquad \text{2-47}$$

where $u(t)$ is white noise with autocorrelation function

$$\boxed{R_{uu}(\tau) = 2 \cdot \sigma^2 \cdot \alpha\delta(\tau)} \qquad \text{2-48}$$

If a restriction is added, that the probability density function of $u(t)$ is Gaussian, the process is called "Gauss-Markov process", and is completely described by the auto-correlation function. In this case $x(t)$ is also Gaussian.

Another term which is used in connection with the Markov Process is the "correlation time" ($\frac{1}{e}$ point). This time is $1/\alpha$.

2. <u>Markov Sequence (Discrete)</u>

Define:

$$e^{-\alpha T} \triangleq \rho \qquad \text{2-49}$$

$$\tau \triangleq T \cdot n \quad n = 0,1,2,\ \ldots \qquad \text{2-50}$$

The autocorrelation will exist only for discrete points:

$$\boxed{R_{xx}(n) = \sigma^2 \cdot \rho^{|n|}, \quad n = 0,1,2,\ \ldots} \qquad \text{2-51}$$

and taking the z transform of 2-51:

$$\boxed{\mathcal{Z}\{R_{xx}(n)\} = \frac{\sigma^2 \cdot (1 - \rho^2)}{(1 - \rho \cdot z^{-1}) \cdot (1 - \rho \cdot z)}} \qquad \text{2-52}$$

<u>Theorem</u>: The discrete process of Eq. 2-51 can be generated by passing white noise through a discrete filter. The "discrete spectral density" P(z) of the noise is $\sigma^2(1 - \rho^2)$. Fig. 12 shows the filter:

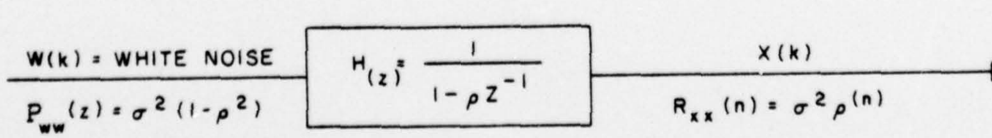


Fig. 12: Discrete Filter — The First Order Markov Sequence

Proof:

Using 2-41 for one dimensional case:

$$\begin{bmatrix} \text{z transform of} \\ \text{the filter} \\ \text{output} \end{bmatrix} = R_{ww}(z)\cdot H(z)\cdot H(z^{-1})$$

$$= \sigma^2\cdot(1-\rho^2)\cdot\frac{1}{1-\rho z^{-1}}\cdot\frac{1}{1-\rho z}$$

$$= \begin{bmatrix} \text{z transform of the } \underline{\text{given}} \\ \text{Markov Process} \end{bmatrix}$$

Q.E.D.

The difference equation that describes the filter is:

$$x(k+1) = \rho\cdot x(k) + w(k+1)$$

It is convenient to define

$$u(k) \triangleq w(k+1)$$

$u(k)$ is also white noise with the same statistics as $w(k)$. From the last definition it follows:

$$R_{uu}(n) = \begin{cases} \sigma^2 \cdot (1-\rho^2) & \text{if} \quad n = 0 \\ 0 & \text{if} \quad n \neq 0 \end{cases} \qquad \text{2-53}$$

$$\boxed{x(k+1) = \rho \cdot x(k) + u(k)} \qquad \text{2-54}$$

The meaning of equation 2-54 is that the value x(k) depends only on the value of the last step in the past, x(k-1).

3. <u>The "Band-Limited", Continuous Case</u>

Another typical autocorrelation function is:

$$\boxed{R_{xx}(\tau) = \sigma^2 \cdot e^{-\alpha|\tau|} \cdot \cos(\omega_o \tau)} \qquad \text{2-55}$$

The spectral density will be:

$$\boxed{R_{xx}(\omega) = \mathcal{F}\{R_{xx}(\tau)\} = \frac{2\alpha\sigma^2}{(\omega - \omega_o)^2 + \alpha^2}} \qquad \text{2-56}$$

This is the case of a "band limited" signal. The process x(t) is limited in a frequency band that is seen in fig. 13.

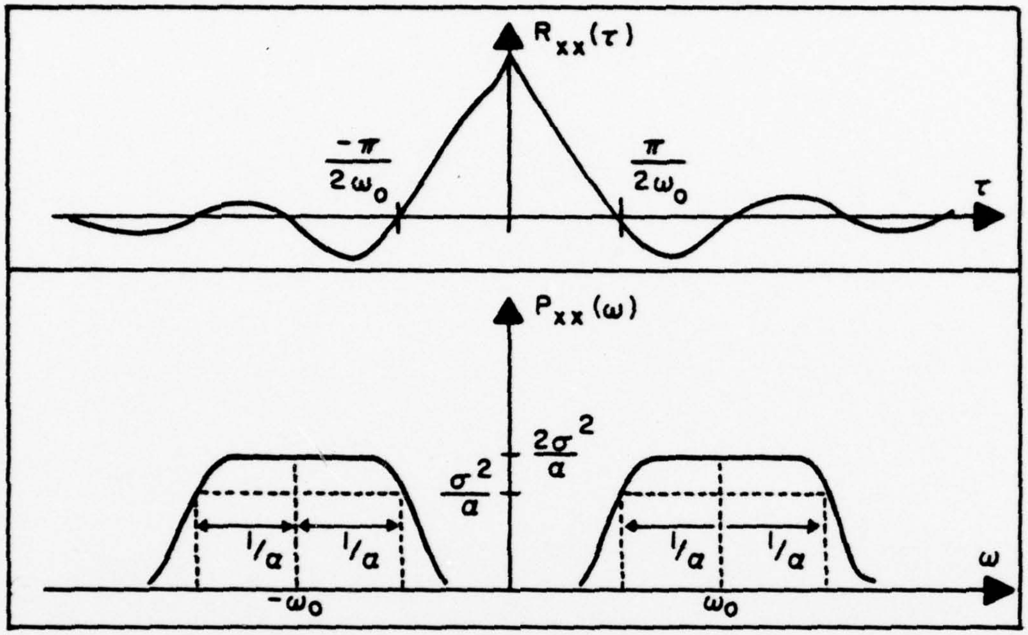


Fig. 13: "Band-Limited" Markov Process

In this thesis, the interest is mainly in discrete cases. The continuous case is given here because one can easily see in the continuous case the frequency band of the random signal and to understand that the autocorrelation function in 2-55 represents an important class of random signals. In order to make a complete summary, the dynamic equations or this case are:

$$\boxed{\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\beta^2 & -2\alpha \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ \beta - 2\alpha \end{bmatrix} x(t)} \qquad 2\text{-}57$$

$$\boxed{\beta^2 \triangleq \alpha^2 + \omega_o^2} \qquad 2\text{-}58$$

$$\boxed{R_{uu}(\tau) = 2\alpha\sigma^2\delta(\tau)} \qquad 2\text{-}59$$

Now, the discrete case will be discussed in detail.

4. <u>The "Band-Limited" Discrete Case</u>

By using the definitions of 2-49, 2-50:

$$e^{-\ T} \triangleq \rho$$

$$\tau \triangleq T \cdot n \qquad n = 0,1,2,...$$

and define:

$$\omega_o \cdot T \triangleq \theta \qquad 2\text{-}60$$

Equations 2-55 has the form:

$$\boxed{R_{xx}(n) = \sigma^2 \cdot \rho^{|n|} \cdot \cos(\theta \cdot n) \quad , \quad n = 0,1,2,...}$$

2-61

and the Z transform of 2-61:

$$P_{xx}(z) = \mathcal{Z}\{R_{xx}(n)\}$$

$$= \frac{(1-\rho^2)\cdot[-Z\cdot\rho\cos\theta + (1+\rho^2) - Z^{-1}\cdot\cos\theta]\sigma^2}{(1-2\rho Z\cos\theta+\rho^2 Z^2)(1-2\rho Z^{-1}\cos\theta+\rho^2 z^{-2})} \qquad 2\text{-}62$$

In order to find a set of different equations that will describe the given autocorrelation function one has to use the procedure that is described in section D.

The first step is to find a factored expression for $P_{xx}(z)$:

$$P_{xx}(z) = P_1(z)\cdot P_1(z^{-1}) \quad .$$

Therefore $P_{xx}(z)$ is assumed to have the form:

$$P_{xx}(z) = \sigma^2(1-\rho^2)\left(\frac{az^{-1}+b}{1-2\rho z^{-1}\cos\theta+\rho^2 z^{-2}}\right)$$

$$\bullet\left(\frac{a\cdot z+b}{1-2\rho z\cos\theta+\rho^2 z^2}\right) \qquad 2\text{-}62a$$

The comparison of 2-62 to 2-62a leads to a pair of algebraic equations:

$$a^2 + b^2 = 1 + \rho^2$$

$$a\cdot b = -\rho\cos\theta \qquad 2\text{-}63$$

and the solution for a,b:

$$\boxed{a = \frac{1}{2} \cdot (\delta^{1/2} + \varepsilon^{1/2})}$$

$$\boxed{b = \frac{1}{2} \cdot (\delta^{1/2} - \varepsilon^{1/2})} \qquad 2\text{-}64$$

where:

$$\boxed{\delta = 1 - \rho \cdot \cos\theta + \rho^2}$$

$$\boxed{\varepsilon = 1 + \rho \cdot \cos\theta + \rho^2} \qquad 2\text{-}65$$

Next, by using 2-41, one can see that the random signal discussed here can be generated by passing white noise through a discrete filter:

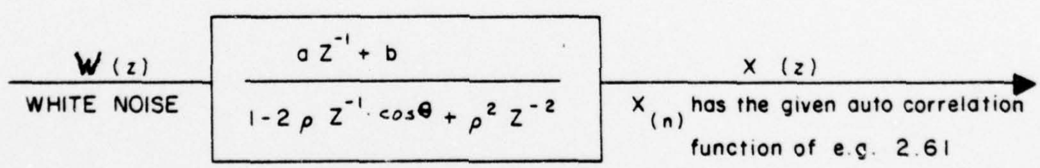


Fig. 14: Filter for One-Dimensional "Band Pass" Process

where

$$R_{ww}(n) = \begin{cases} (1 - \rho^2)\sigma^2 & \text{if } n = 0 \\ 0 & \text{if } n \neq 0 \end{cases}$$

It is convenient to define:

2-66a

$$W(z) = Z^{-1}\, U(z) \tag{2-66a}$$

$u(k)$ is also white noise with the same autocorrelation function.

$$\boxed{R\ \ (n) = \begin{cases} (1-\rho^2)\sigma^2 & \text{if} \quad n = 0 \\ 0 & \text{if} \quad n \neq 0 \end{cases}}$$

and the filter will be:

$$\frac{X(z)}{W(z)} = \frac{a \cdot Z^{-1} + b}{1 - 2\rho Z^{-1}\cos\theta + \rho^2 Z^{-2}} \tag{2-66}$$

define

$$X_1(z) = -\rho^2 Z^{-1} X_2(z)$$

$$X_2(z) = \frac{W(z)}{1 - Z^{-1} 2\rho\cos\theta + \rho^2 Z^{-2}}$$

From the last definition

$$x_1(k) = -\rho^2 x_2(k-1)$$

$$x_2(k) = x_1(k-1) + 2\rho\cos\theta \cdot x_2(k-1) + u(k-1)$$

In matrix form:

$$\begin{vmatrix} x_1(k) \\ \\ x_2(k) \end{vmatrix} = \begin{vmatrix} 0 & -\rho^2 \\ \\ 1 & 2\rho\cos\theta \end{vmatrix} \cdot \begin{vmatrix} x_1(k-1) \\ \\ x_2(k-1) \end{vmatrix} + \begin{vmatrix} 0 \\ \\ 1 \end{vmatrix} \cdot u(k-1) \qquad 2\text{-}67$$

and:

$$X(z) = X_2(z)(a \cdot z^{-1} + b)$$

$$x(k) = a \cdot x_2(k-1) + b \cdot x_2(k)$$

$$= -\rho^{-2} \cdot a \cdot x_1(k) + b \cdot x_2(k)$$

$$\boxed{x(k) = (-\rho^{-2} \cdot a \quad b) \cdot \begin{bmatrix} x_1(k) \\ \\ x_2(k) \end{bmatrix}} \qquad 2\text{-}68$$

An Approximated Model

Using the following procedure one can derive an approximated model, which is simpler than the previous model. Define:

$$(a \cdot Z^{-1} + b) \cdot W(z) \triangleq Z^{-1} \cdot U_a(z)$$

From the last definition:

$$u_a(k-1) = b \cdot W(k) + a \cdot W(k-1) \quad .$$

The autocorrelation of $u_a(k)$ is

$$R_{u_a u_a}(n) = E\{u_a(k) \cdot u_a(k+1)\} = \begin{cases} \sigma^2(1+\rho^2)(1-\rho^2) & \text{if } n = 0 \\ \sigma^2(-\rho\cos\theta)(1-\rho^2) & \text{if } n = 1 \\ 0 & \text{if } n \geq 2 \end{cases} \qquad 2\text{-}69$$

Because $R_{u_a u_a}(1) \neq 0$, u_a is <u>not</u> white noise. Actually, u_a is a <u>linear</u> <u>combination</u> of white noises.

Now, an approximation is done

$$R_{u_a u_a}(1) = 0$$

In this case:

$$\boxed{R_{u_a u_a}(n) = \begin{cases} (1+\rho^2)(1-\rho^2)\sigma^2 & \text{if } n = 0 \\ 0 & \text{if } n \neq 0 \end{cases}} \qquad 2\text{-}70$$

and now $u_a(n)$ is white noise.

Figure 15 shows the correct and the approximated noise.

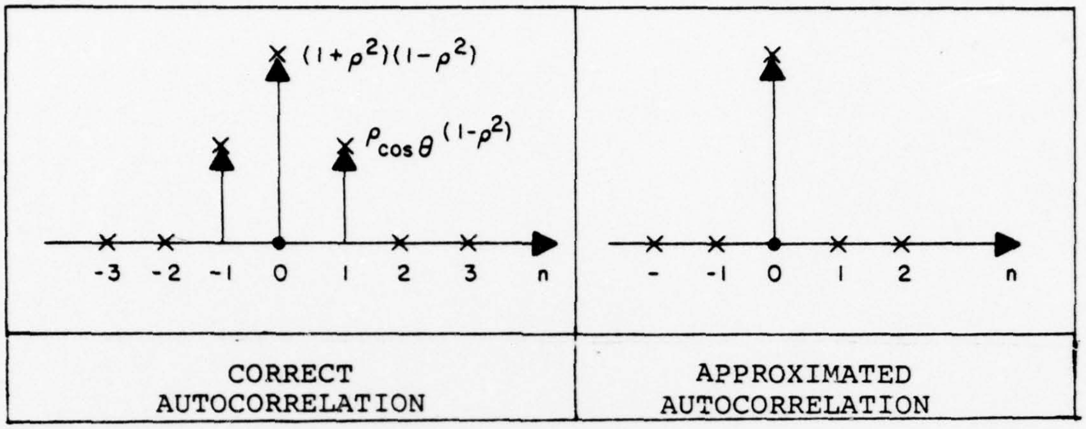


Fig. 15: The Autocorrelation of u_a, $R_{u_a u_a}(n)$

The filter that generates the "band limited" random signal is (compare with 2-66)

$$X(z) = \frac{Z^{-1} \cdot u_a(z)}{1 - 2\rho Z^{-1} \cos\theta + \rho^2 Z^{-2}}$$

Define:

$$x_1(k) \triangleq -\rho^2 \cdot x_2(k-1)$$

$$x_2(k) \triangleq x(k)$$

$$\boxed{\begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} = \begin{bmatrix} 0 & -\rho^2 \\ 1 & 2\rho\cos\theta \end{bmatrix} \begin{bmatrix} x_1(k-1) \\ x_2(k-1) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot u_a(k-1)} \quad 2\text{-}71$$

$$x(k) = \begin{bmatrix} 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} \qquad \text{2-72}$$

This model is simpler than the one previously described.

5. Second Order Markov Process

In the continuous case, the process is defined by

$$R_{xx}(\tau) = \sigma^2 \cdot e^{-\alpha|\tau|}(1 + \alpha|\tau|)$$

define:

$$e^{-\alpha T} \triangleq \rho$$

$$\tau \triangleq n \cdot T$$

$$\alpha T = \ell n \rho = \theta$$

It leads to the discrete autocorrelation function:

$$\boxed{R_{xx}(n) = \sigma^2 \rho^{|n|}(1 + \theta|n|)}$$

The z transform of this process is

$$P_{xx}(z) = \mathcal{Z}\{R_{xx}(n)\} = P_1(z) + P_2(z)$$

where:

$$P_1(z) = \frac{-\sigma^2\theta\rho[4\rho - (Z + Z^{-1})(1 + \rho^2)]}{(1 - \rho Z^{-1})^2(1 - \rho Z)^2}$$

$$P_2(z) = \frac{\sigma^2(1 - \rho^2)}{(1 - \rho Z^{-1})(1 - \rho Z)}$$

This process can be generated by a combination of two filters, that are forced by two uncorrelated white noises:

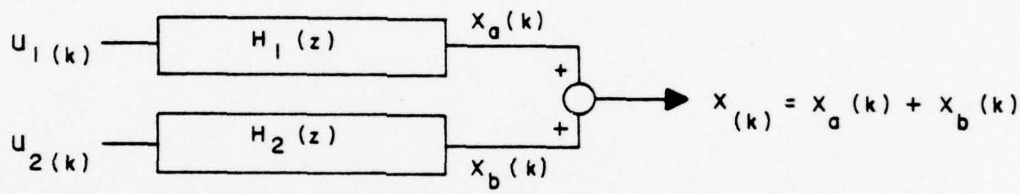


Fig. 16: Filter to Generate Second Order Markov Process

$x_a(k)$ and $x_b(k)$ are uncorrelated (because $u_1(k)$ $u_2(k)$ are uncorrelated). Therefore the spectrum of $x(k)$ is the sum of the spectrums of $x_a(k)$ and $x_b(k)$.

Now: $P_1(z)$ can be written:

$$P_1(z) = \frac{-\sigma^2\theta\rho(aZ^{-1} - b)(aZ - b)}{(1 - \rho Z^{-1})^2(1 - \rho Z)^2}$$

where:

$$\left.\begin{aligned} a^2 + b^2 &= 4\rho \\ ab &= 1 + \rho^2 \end{aligned}\right\}$$

$a = \frac{1}{2}(\varepsilon^{\frac{1}{2}} + \delta^{\frac{1}{2}})$	$\varepsilon = 4\rho + 1 + \rho^2$
$b = \frac{1}{2}(\varepsilon^{\frac{1}{2}} - \delta^{\frac{1}{2}})$	$\delta = 4\rho - 1 - \rho^2$

the filter $H_1(z)$ is:

$$H_1(z) = \frac{(aZ^{-1} - b)}{(1 - \rho Z^{-1})^2} = \frac{X_a(z)}{U_1(z)}$$

$$\boxed{R_{u_1u_1}(m) = \begin{cases} -\theta\rho\sigma^2 & \text{if} \quad m = 0 \\ 0 & \text{if} \quad m \neq 0 \end{cases}}$$

the state variable expression that follows $H_1(z)$:

$$\begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} = \begin{bmatrix} 0 & -\rho^2 \\ 1 & 2\rho \end{bmatrix} \cdot \begin{bmatrix} x_1(k+1) \\ x_2(k-1) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k-1)$$

$$x_a(k) = (-\rho^2 a \quad -b) \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$$

and for $H_2(z)$:

$$H_2(z) = \frac{1}{1 - \rho Z^{-1}} = \frac{X_b(z)}{W_2(z)}$$

define:

$$u_2(k) = w_2(k+1)$$

then:

$$\boxed{R_{u_2u_2}(m) = \begin{cases} \sigma^2(1-\rho^2) & \text{if} \quad m = 0 \\ 0 & \text{if} \quad m \neq 0 \end{cases}}$$

and

$$x_3(k+1) = \rho x_3(k) + u_2(k)$$

$$x_b(k) = x_3(k)$$

The combination of x_a, x_b:

$$\boxed{\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} = \begin{bmatrix} 0 & -\rho^2 & 0 \\ 1 & 2\rho & 0 \\ 0 & 0 & \rho \end{bmatrix} \cdot \begin{bmatrix} x_1(k-1) \\ x_2(k-1) \\ x_3(k-1) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}}$$

$$\boxed{x(k) = \begin{bmatrix} -\rho^2 a & -b & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1(k-1) \\ x_2(k-1) \\ x_3(k-1) \end{bmatrix}}$$

6. Conclusions

1) It was shown in this section that some stochastic processes can be generated by passing white noise (or a

combination of white noises) through a linear filter of order P.

2) Also it is seen that the order P is a finite number [in the first two examples P was 1, in the last two examples P was 2].

3) P_1, the order of the process, tells the number of "neighbors" near the point k, upon which the value X(k) depends.

F. STATE VARIABLE MODELING OF TWO DIMENSIONAL FIELDS BY "FILTER RESPONSE METHOD"

1. Introduction

The starting point of the modeling procedure will be an extension of the *Markov Process* to a two dimensional case.

It is emphasized that the technique used in this section is good only for separable autocorrelation functions. In Section H another modeling method which is valid for any homogeneous random field will be discussed. [A comparison between the two methods is given in Section I].

2. Model For First Order, Continuous, Markov Process

Given a two dimensional field, with an autocorrelation function:

$$R_{xx}(\tau_1,\tau_2) = \sigma^2 e^{-\alpha_1|\tau_1|} e^{-\alpha_2|\tau_2|} \qquad 2\text{-}73$$

α_1 is the correlation factor in the t_1 direction

α_2 is the correlation factor in the t_2 direction.

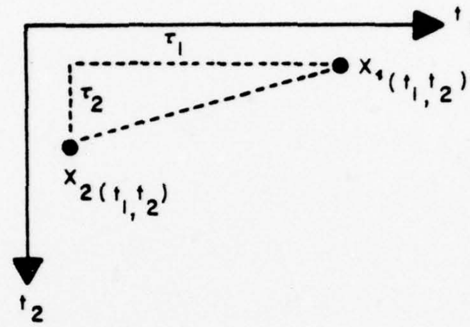


Fig. 17: System Coordinates For The Field in 2-73

The power spectral density function (using 2-46, 2-20):

$$P_{xx}(\omega_1, \omega_2) = \frac{(2\alpha_1)(2\alpha_2)\sigma^2}{(\omega_1^2 + \alpha_1^2)(\omega_2^2 + \alpha_2^2)} \qquad 2\text{-}74$$

This autocorrelation function might be generated by passing white noise through a two dimensional filter:

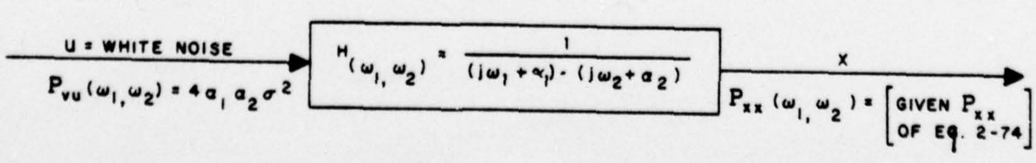


Fig. 18: Generation of First Order, Two Dimensional, Markov Process

The proof is by using 2-36. One can see that the spectral density at the filter-output is the same as in 2-74. Therefore:

$$\frac{X(\omega_1,\omega_2)}{U(\omega_1,\omega_2)} = \frac{1}{(j\omega_1+\alpha_1)(j\omega_2+\alpha_2)}$$

Define:

$$M(\omega_1,\omega_2) = \frac{U(\omega_1,\omega_2)}{j\omega_2+\alpha_2}$$

$$N(\omega_1,\omega_2) = X(\omega_1,\omega_2)$$

Then:

$$\frac{dM(t_1,t_2)}{dt_2} = -\alpha_2 M(t_1,t_2) + u(t_1,t_2)$$

$$N(\omega_1,\omega_2) = \frac{M(\omega_1,\omega_2)}{\alpha_1+j\omega_1}$$

$$\frac{dN(t_1,t_2)}{dt_1} = -\alpha_1 N(t_1,t_2) + M(t_1,t_2)$$

and in matrix form:

$$\left[\begin{array}{c} \dfrac{dN(t_1,t_2}{dt_1} \\ \hline \dfrac{dM(t_1,t_2)}{dt_2} \end{array}\right] = \left[\begin{array}{c|c} -\alpha_1 & +1 \\ \hline 0 & -\alpha_2 \end{array}\right] \cdot \left[\begin{array}{c} N(t_1,t_2) \\ \\ M(t_1,t_2) \end{array}\right] + \left[\begin{array}{c} 0 \\ \\ 1 \end{array}\right] \cdot u(t_1,t_2)$$

2-75

$$x(t_1,t_2) = (1 \quad 0) \cdot \left[\begin{array}{c} N(t_1,t_2) \\ \\ M(t_1,t_2) \end{array}\right]$$

2-76

$u(t_1,t_2)$ is white noise:

$$\boxed{R_{uu}(\tau) = 4\alpha_1\alpha_2\sigma^2\delta(\tau_1)\delta(\tau_2)}$$ 2-77

3. Model For First Order, Discrete, Markov Sequence

Define:

$$\rho_1 = e^{-\alpha T}$$

$$\rho_2 = e^{-\alpha_2 T}$$

$$t_1 = \ell T \qquad \ell = 0,1,2, \ldots$$

$$t_2 = kT \qquad k = 0,1,2, \ldots$$

$$\tau_1 = nT \qquad n = 0,1,2, \ldots$$

$$\tau_2 = mT \qquad m = 0,1,2, \ldots$$

Eq. 2-73 has now the form:

$$\boxed{R_{xx}(n,m) = \sigma^2 \rho_1^{|n|} \rho_2^{|m|} \qquad \begin{array}{l} n = 0,1,2, \ldots \\ m = 0,1,2, \ldots \end{array}}$$

2-78

and the two dimensional z transformation of 2-78:

$$P_{xx}(z_1,z_2) = \frac{\sigma^2(1-\rho_1^2)(1-\rho_2^2)}{(1-\rho_1 z_1^{-1})(1-\rho_1 z_1)(1-\rho_2 z_2^{-1})(1-\rho_2 z_2)}$$

2-79

This discrete autocorrelation function can be generated by passing white noise through a two-dimensional discrete filter:

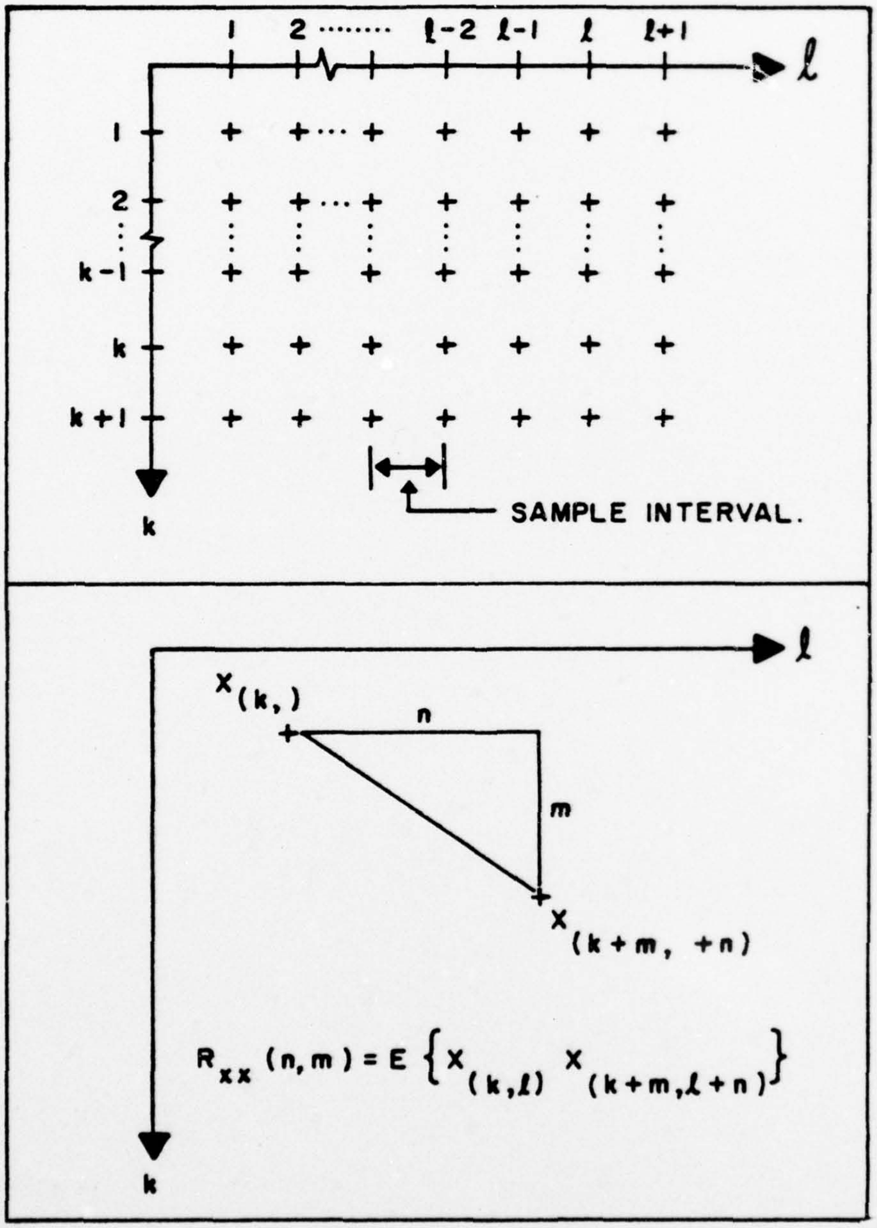


Fig. 19: The Discrete Area

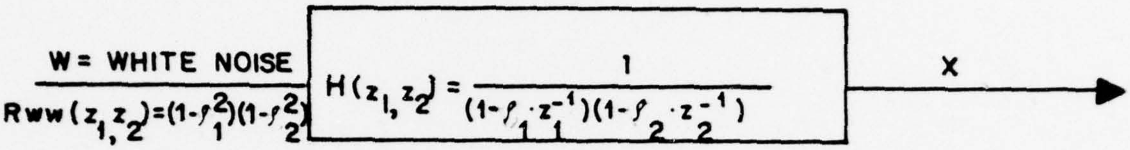


Fig. 20: The Filter to Create First Order, Two Dimensional, Markov Field

The proof is, again, by using 2-41. One can see that $P_{xx}(z_1,z_2)$ at the output of the filter will be as in 2-79. The filter expression is, therefore:

$$\frac{X(z_1,z_2)}{W(z_1,z_2)} = \frac{1}{(1-\rho_1 z_1{}^{-1})(1-\rho_2 z_2{}^{-1})}$$

It is convenient to define:

$$W(z_1,z_2) \triangleq z_1{}^{-1}\, z_2{}^{-1}\, U(z_1,z_2)$$

and the filter has not the form

$$\frac{X(z_1,z_2)}{U(z_1,z_2)} = \frac{z_1{}^{-1}\, z_2{}^{-1}}{(1-z_1{}^{-1}\rho_1)(1-z_2{}^{-1}\rho_2)} \qquad \text{2-80}$$

One can continue in two directions:

Direction 1: (Model 1)

From 2-80:

$$x(z_1,z_2)\cdot[1 - z_1^{-1}\rho_1 - z_2^{-1}\rho_2 + z_1^{-1}z_2^{-1}\rho_1\rho_2] = z_1^{-1}z_2^{-1}\cdot u(z_1,z_2)$$

and a difference equation can be written:

$$\boxed{x(k+1,\ell+1) = \rho_1\cdot x(k+1,\ell) - \rho_1\rho_2\cdot x(k,\ell) + \rho_2\cdot x(k,\ell+1) + x(k,\ell)} \quad \text{2-81}$$

$$\boxed{R_{VV}(n,m) = \begin{cases} \sigma^2(1-\rho_1^2)(1-\rho_2^2) & \text{if } n=0 \text{ and } m=0 \\ 0 & \text{if } n\neq 0 \text{ or } m\neq 0 \end{cases}}$$

2-82

The point $k+1,\ell+1$ is connected to its three nearest neighbors left and above:

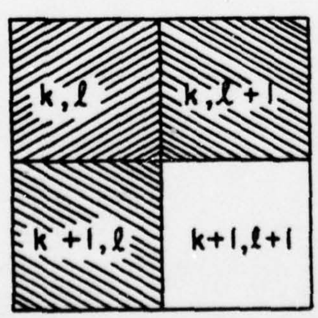


Fig. 21: The Neighbors that are Connected to Point $k+1,\ell+1$

Direction 2: (Model 2)

Define

$$N(z_1,z_2) = \frac{z_1^{-1}\,U(z_1,z_2)}{1 - z_1^{-1}\rho_1} \qquad 2\text{-}83$$

$$M(z_1,z_2) = X(z_1,z_2) \qquad 2\text{-}84$$

From 2-83

$$N(z_1,z_2) = z_1^{-1}\rho_1 N(z_1,z_2) + z_1^{-1}\,U(z_1,z_2)$$

$$N(k,\ell) = \rho_1 N_1(k,\ell-1) + u(k,\ell-1) \qquad 2\text{-}85$$

Note: from Eq. 2-78 it is obvious that z_2 corresponds to the k direction and z_1 corresponds to the ℓ direction.
Substituting 2-83, 2-84 into 2-80:

$$M(z_1,z_2) = \frac{z_2^{-1}N(z_1,z_2)}{1 - z_2^{-1}\rho_2}$$

$$M(z_1,z_2) = z_2^{-1}\rho_2 M(z_1,z_2) + z_2^{-1}N(z_1,z_2)$$

$$M(k,\ell) = \rho_2 M(k-1,\ell) + N(k-1,\ell) \qquad 2\text{-}86$$

By changing indices of 2-85, 2-86 one can write these two equations in a state vector form

$$\boxed{\begin{bmatrix} M(k+1,\ell) \\ N(k,\ell+1) \end{bmatrix} = \begin{bmatrix} \rho_2 & 1 \\ 0 & \rho_1 \end{bmatrix} \cdot \begin{bmatrix} M(k,\ell) \\ N(k,\ell) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot u(k,\ell)}$$

2-87

$$\boxed{x(k,\ell) = (\ 1 \quad 0\ ) \cdot \begin{bmatrix} M(k,\ell) \\ N(k,\ell) \end{bmatrix}} \qquad \text{2-88}$$

By comparison 2-87, 2-88, to 2-81:

$$\boxed{M(k,L) = x(k,\ell)} \qquad \text{2-89}$$

$$\boxed{N(k,L) = x(k+1,\ell) - \rho_2 \cdot x(k,\ell)} \qquad \text{2-90}$$

It is obvious that model 1 and model 2 describe the same field.

An interesting property of this model:

$$\boxed{E\{M(k,\ell)N(i,j)\} = 0} \qquad \text{2-90a}$$

for any k,ℓ,i,j.

Proof:

$$E\{M(k,\ell)N(i,j)\} = E\{x(k,\ell)[x(i+1,j) - \rho_2 x(i,j)]\}$$

$$= E\{x(k,\ell)[\rho_2 x(i,j) + u(i,j) - \rho_2 x(i,j)]\}$$

u(i,j) is white noise and therefore it is uncorrelated to $x(k,\ell)$. Therefore:

$$E\{M(k,\ell)N(i,j)\} = 0.$$

Q.E.D.

4. Model For "Band Limited", Discrete Case

Equivalently to the previous discussion, the two-dimensional, "band limited." discrete Markov Process is defined by the autocorrelation function:

$$\boxed{R_{xx}(n,m) = \sigma^2 [\rho_1{}^{|n|}\cos(\theta_1 n)][\rho_2{}^{|m|}\cos(\theta_2 m)] \quad \begin{array}{l} n = 0,1,2,\ \ldots \\ m = 0,1,2,\ \ldots \end{array}}$$

2-91

The z transform of this autocorrelation function:

$$P_{xx}(z_1,z_2) = \sigma^2(1-\rho_1{}^2)(1-\rho_2{}^2)\frac{A(z_1,z_2)}{B(z_1,z_2)} \qquad 2\text{-}92$$

where

$$A(z_1,z_2) = [-z_1\cos\theta_1+(1+\rho_1^{\,2})-z_1^{\,-1}\cos\theta_1][-z_2\cos\theta_2+(1+\rho_2^{\,2})-z_2^{\,-1}\cos\theta_2] \qquad \text{2-93}$$

$$B(z_1,z_2) = [(1-2\rho_1z_1\cos\theta_1+\rho_1^{\,2}z_1^{\,2})(1-2\rho_1z_1^{\,-1}\cos\theta_1+\rho^2z_1^{\,-2})]$$
$$\cdot\ [(1-2\rho_2z_2\cos\theta_2+\rho_2^{\,2}z_2^{\,2})(1-2\rho_2z_2^{\,-1}\cos\theta_2+\rho_2^{\,2}z_2^{\,-2})] \qquad \text{2-94}$$

$A(z_1,z_2)$ can be written in a form:

$$A(z_1,z_2) = [(a_1z_1^{\,-1}+b_1)(a_2z_2+b_1)][(a_2z_2^{\,-1}+b_2)(a_2z_2+b_2)] \qquad \text{2-95}$$

where:

$$a_1^{\,2} + b_1^{\,2} = 1 + \rho_1^{\,2} \qquad \text{2-96a}$$

$$a_2^{\,2} + b_2^{\,2} = 1 + \rho_2^{\,2} \qquad \text{2-96b}$$

$$a_1b_1 = -\rho_1\cos\theta_1 \qquad \text{2-96c}$$

$$a_2b_2 = -\rho_2\cos\theta_2 \qquad \text{2-96d}$$

The solution for the a's and b's:

$$\boxed{\begin{aligned} a_1 &= \tfrac{1}{2}(\delta_1^{1/2} + \varepsilon_1^{1/2}) \\ a_2 &= \tfrac{1}{2}(\delta_2^{1/2} + \varepsilon_2^{1/2}) \\ b_1 &= \tfrac{1}{2}(\delta_1^{1/2} - \varepsilon_1^{1/2}) \\ b_2 &= \tfrac{1}{2}(\delta_2^{1/2} - \varepsilon_2^{1/2}) \end{aligned}} \qquad \text{2-97}$$

where:

$$\boxed{\begin{aligned} \delta_1 &= 1 - \rho_1\cos\theta_1 + \rho_1^2 \\ \delta_2 &= 1 - \rho_2\cos\theta_2 + \rho_2^2 \\ \varepsilon_1 &= 1 + \rho_1\cos\theta_1 + \rho_1^2 \\ \varepsilon_2 &= 1 + \rho_2\cos\theta_2 + \rho_2^2 \end{aligned}} \qquad \text{2-98}$$

Now: This random process can be generated by passing white noise through a discrete filter of the form:

$$\underset{\text{WHITE NOISE}}{\omega(Z_1, Z_2)} \longrightarrow \boxed{\frac{(a Z_1^{-1} + b_1)(a_2 Z_2^{-1} + b_2)}{(1 - 2\rho_1 Z_1^{-1} \cos\theta_1 + \rho^2 Z_1^{-2})(1 - 2\rho_2 Z_2^{-1} \cos\theta_2 + \rho^2 Z_2^{-2})}} \longrightarrow x(Z_1, Z_2)$$

Fig. 22: Filter to generate "Band Pass" random field.

where:

$$R_{ww}(n,m) = \begin{cases} \sigma^2(1 - \rho_1^2)(1 - \rho_2^2) & , \text{ if } n = 0 \text{ and } m = 0 \\ 0 & , \text{ if } n \neq 0 \text{ or } m \neq 0 \end{cases} \qquad 2\text{-}99$$

It is convenient to define:

$$W(z_1, z_2) = z_1^{-1} \cdot z_2^{-1} \cdot U(z)$$

$u(k,\ell)$ is also white noise with the same statistics as $w(k,\ell)$.

$$\boxed{R_{uu}(n,m) = \begin{cases} \sigma^2(1 - \rho_1^2)(1 - \rho_2^2) & , \text{ if } n = 0 \text{ and } m = 0 \\ 0 & , \text{ if } n \neq 0 \text{ or } m \neq 0 \end{cases}} \qquad 2\text{-}100$$

the filter has the form:

$$X(z_1,z_2) = \frac{(a_1 z_1^{-1} + b_1)(a_2 z_2^{-1} + b_2)\, z_1^{-1} z_2^{-1} \cdot U(z_1,z_2)}{(1 - 2\rho_1 z_1^{-1}\cos\theta_1 + \rho_1^2 z_1^{-2})(1 - 2\rho_2 z_2^{-1}\cos\theta_2 + \rho^2 z_2^{-2})} \qquad \text{2-101}$$

Now the following definitions are done:

$$N_1(z_1,z_2) \triangleq -\rho_1^2 \cdot z_1^{-1} \cdot N_2(z_1,z_2)$$

$$N_2(z_1,z_2) \triangleq \frac{z_1^{-1} \cdot U(z_1,z_2)}{1 - 2\rho_1 z_1^{-1}\cos\theta_1 + \rho_1^2 z_1^{-2}}$$

$$X_1(z_1,z_2) \triangleq (a_1 z_1^{-1} + b_1) W_2(z_1,z_2)$$

From the last three definitions one can write a set of difference equations:

$$\begin{bmatrix} N_1(k,\ell) \\ N_2(k,\ell) \end{bmatrix} = \begin{bmatrix} 0 & -\rho_1^2 \\ 1 & 2\rho_1\cos\theta_1 \end{bmatrix} \cdot \begin{bmatrix} N_1(k,\ell-1) \\ N_2(k,\ell-1) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot u(k,\ell-1) \qquad \text{2-102}$$

$$x_1(k,\ell) = -\rho_1^{-2} \cdot a_1 N_1(k,\ell) + b_1 N_2(k,\ell) \qquad \text{2-103}$$

Now, other definitions are done:

$$M_1(z_1,z_2) = -\rho_2^{2} \cdot z_2^{-1} \cdot M_2(z_1,z_2)$$

$$M_2(z_1,z_2) = \frac{z_2^{-1} \cdot X_1(z_1,z_2)}{1 - 2\rho_2 \cdot z_2^{-1} \cdot \cos\theta_2 + \rho_2^{2} \cdot z_2^{-2}}$$

From these definitions it follows:

$$X(z_1,z_2) = (a_2 \cdot z_2^{-1} + b_2) \cdot M_2(z_1,z_2) \qquad \text{2-104}$$

$$\begin{bmatrix} M_1(k,\ell) \\ \\ M_2(k,\ell) \end{bmatrix} = \begin{bmatrix} 0 & -\rho_2^{2} \\ \\ 1 & 2\rho_2 \cos\ \theta_2 \end{bmatrix} \begin{bmatrix} M_1(k-1,\ell) \\ \\ M_2(k-1,\ell) \end{bmatrix} + \begin{bmatrix} 0 \\ \\ 1 \end{bmatrix} \cdot x_1(k-1,\ell)$$

2-105

$$x(k,\ell) = -\rho_2^{-2} \cdot a_2 \cdot M_1(k,\ell) + b_2 \cdot M_2(k,\ell) \qquad \text{2-104a}$$

Equations 2-102, 2-103, 2-105 can be written together:

$$\begin{bmatrix} M_1(k+1,\ell) \\ M_2(k+1,\ell) \\ N_1(k,\ell+1) \\ N_2(k,\ell+1) \end{bmatrix} = \begin{bmatrix} 0 & -\rho_2^{\,2} & 0 & 0 \\ 1 & 2\rho_2 \cos\theta_2 & 0 & 0 \\ 0 & 0 & 0 & -\rho_1^{\,2} \\ 0 & 0 & 1 & 2\rho_1 \cos\theta_1 \end{bmatrix} \begin{bmatrix} M_1(k,\ell) \\ M_2(k,\ell) \\ N_1(k,\ell) \\ N_2(k,\ell) \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot x_1(k,\ell) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \cdot u(k,\ell)$$

Using 2-103

$$\begin{bmatrix} M_1(k+1,\ell) \\ M_2(k+1,\ell) \\ \hdashline N_1(k,\ell+1) \\ N_2(k,\ell+1) \end{bmatrix} = \left[\begin{array}{cc:cc} 0 & -\rho_2^{\,2} & & 0 \\ 1 & 2\rho_2\cos\theta_2 & -\rho_1^{\,2}a_1 & b_1 \\ \hdashline 0 & 0 & 0 & -\rho_1^{\,2} \\ 0 & 0 & 1 & 2\rho_1\cos\theta_1 \end{array}\right] \cdot \begin{bmatrix} M_1(k,\ell) \\ M_2(k,\ell) \\ \hdashline N_1(k,\ell) \\ N_2(k,\ell) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \hdashline 0 \\ 1 \end{bmatrix} u(k,\ell)$$

2-106

MODEL

3

and using 2-104a:

$$x(k,\ell) = \left[\begin{array}{cc:cc} -\rho_2^{-2}\cdot a_2 & b_2 & 0 & 0 \end{array}\right] \cdot \begin{bmatrix} M_1(k,\ell) \\ M_2(k,\ell) \\ \hdashline N_1(k,\ell) \\ N_2(k,\ell) \end{bmatrix} \qquad \text{2-107}$$

Approximated Model

One can derive for the "band limited" case an approximated model, which has a simpler structure. The starting point is, again, 2-101.

Define:

$$u_a(z_1,z_2) = (a_1 z_1^{-1} + b_1)(a_2 z_2^{-1} + b_2)\, u(z_1,z_2) \qquad \text{2-108}$$

From eq. (2-87) one can see that $u_a(k,\ell)$ is the noise-input to the filter that creates the "band limited" random field. Now, what are the statistics of $u_a(k,\ell)$? Equation
2-108
(2-108) tells us that $u_a(k,\ell)$ is a linear combination of white noises:

$$u_a(k,\ell) = a_1 a_2 u(k-1,\ell-1) + b_1 a_2 u(k-1,\ell) + a_1 b_2 u(k,\ell-1) + b_1 b_2 u(k,\ell) \qquad \text{2-109}$$

Using eq. (2-108), together with the definitions of the a's and b's in eq. (2-96), the autocorrelation of $u_a(k,\ell)$ can be calculated.

$$R_{u_a u_a}(n,m) = \begin{cases} (1+\rho_1^2)(1+\rho_2^2)(1-\rho_1^2)(1-\rho_2^2) & \text{if } m = 0 \text{ and } n = 0 \\ \text{(a number that is } \underline{\text{not}} \text{ zero)} & \text{if } m \le 1 \text{ and } n \le 1 \\ 0 & \text{if } m > 1 \text{ or } n > 1 \end{cases} \qquad \text{2-110}$$

For example, the calculation for $R_{u_a u_a}(0,0)$ is shown:

$$R_{u_a u_a}(0,0) = E\{u_a^2(k,\ell)\}$$

$$= [(a_1a_2)^2 + (b_1a_2)^2 + (a_1b_2)^2 + (b_1b_2)^2] \cdot [(1-\rho_1^2)(1-\rho_2^2)]$$

$$= (a_1^2+b_1^2)(a_2^2+b_2^2)(1-\rho_1^2)(1-\rho_2^2)$$

$$= (1-\rho_1^2)(1+\rho_2^2)(1-\rho_1^2)(1-\rho_2^2) \qquad \text{2-111}$$

Here is the place to make an approximation: One can see that $R_{u_a u_a}$ is not white noise, because $R_{u_a u_a}(n,m) \neq 0$ when n or m are equal to one. Now, arbitrarily, it is decided to let this value be zero (for n or m equal one). In that case u_a is white noise:

$$\boxed{R_{u_a u_a}(n,m) = \begin{cases} (1+\rho_1^2)(1+\rho_2^2)(1-\rho_1^2)(1-\rho_2^2), & \text{if } n = 0 \text{ and } m = 0 \\ 0, & \text{if } n \neq 0 \text{ or } m \neq 0 \end{cases}} \qquad \text{2-112}$$

The filter to create the random field will be:

$$X(z_1,z_2) = \frac{z_1^{-1}z_2^{-1} \cdot U_a(z_1,z_2)}{(1-2\rho_1 z_1^{-1}\cos\theta_1 + \rho_1^2 z_1^{-2}) \cdot (1-2\rho_2 z_2^{-1}\cos\theta_2 + \rho_2^2 z_2^{-2})} \qquad \text{2-113}$$

Comparing 2-113 with 2-101, one can write for 2-113 a state variable model that will be similar to eq. 2-106:

Define:

$$N_1(z_1,z_2) = -\rho_1{}^2 \cdot z_1{}^{-1} \cdot N_2(z_1,z_2)$$

$$N_2(z_1,z_2) = \frac{z_1{}^{-1} \cdot U_a(z_1,z_2)}{1 - 2\rho_1 z_1{}^{-1} \cdot \cos\theta_1 + \rho_2{}^2 \cdot z_1{}^{-2}}$$

$$M_1(z_1,z_2) = -\rho_2{}^2 z_2{}^{-1} \cdot M_2(z_1,z_2)$$

$$M_2(z_1,z_2) = \frac{z_2{}^{-1} \cdot N_2(z_1,z_2)}{1 - 2\rho_2 z_2{}^{-1} \cdot \cos\theta_2 + \rho_2{}^2 \cdot z_2{}^{-2}}$$

From these definitions it follows:

2-114

$$\boxed{\begin{vmatrix} M_1(k+1,\ell) \\ M_2(k+1,\ell) \\ \hline N_1(k,\ell+1) \\ N_2(k,\ell+1) \end{vmatrix} = \left|\begin{array}{cc|cc} 0 & -\rho_2{}^2 & 0 & 1 \\ 1 & 2\rho_2\cos\theta_2 & 0 & 0 \\ \hline 0 & 0 & 0 & -\rho_1{}^2 \\ 0 & 0 & 1 & 2\rho_1\cos\theta_1 \end{array}\right| \begin{vmatrix} M_1(k,\ell) \\ M_2(k,\ell) \\ N_1(k,\ell) \\ N_2(k,\ell) \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ 1 \\ 0 \end{vmatrix} \cdot u_a(k,\ell)}$$ MODEL 4

$$\boxed{x(k,\ell) = \left|\begin{array}{cc|cc} 0 & 1 & 0 & 0 \end{array}\right| \cdot \begin{vmatrix} M_1(k,\ell) \\ M_2(k,\ell) \\ \hline N_1(k,\ell) \\ N_2(k,\ell) \end{vmatrix}}$$ 2-115

By comparing the difference equation that follows from 2-113 to the structure of 2-114, 2-115, it is found:

$$M_2(k,\ell) = x(k,\ell)$$

$$N_2(k,\ell) = x(k,\ell+1) - 2\rho_2\cdot\cos\theta_2\cdot x(k,\ell) + \rho_2\cdot x(k,\ell-1)$$

$$M_1(k,\ell) = -\rho_2{}^2\cdot x(k-1,\ell)$$

$$N_1(k,\ell) = -\rho_1{}^2\cdot x(k,\ell) - 2\rho_2\cdot\cos\theta_2\cdot x(k,\ell-1) + \rho_2\cdot x(k,\ell-2) \ .$$

5. Model for Second Order Markov Process

The extension of the one dimensional case to the two dimensional case is:

$$R_{xx}(n,m) = \rho_1{}^{|n|}(1+\theta_1|n|)\ \rho_2{}^{|m|}(1+\theta_2|m|)$$

and following very similar procedures to all previous cases, the state space model in this case has the form:

$$
\begin{vmatrix} M_1(k+1,\ell) \\ M_2(k+1,\ell) \\ M_3(k+1,\ell) \\ \hdashline N_1(k,\ell+1) \\ N_2(k,\ell+1) \\ N_3(k,\ell+1) \end{vmatrix}
=
\left|\begin{array}{ccc:ccc} 0 & -\rho_2^{\,2} & 0 & 0 & 0 & 0 \\ 1 & 2\rho_2 & 0 & -\rho_1^{\,2}a_1 & b_1 & 0 \\ 0 & 0 & \rho_2 & 0 & 0 & 1 \\ \hdashline 0 & 0 & 0 & 0 & -\rho_1^{\,2} & 0 \\ 0 & 0 & 0 & 1 & 2\rho_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_1 \end{array}\right|
\begin{vmatrix} M_1(k,\ell) \\ M_2(k,\ell) \\ M_3(k,\ell) \\ \hdashline N_1(k,\ell) \\ N_2(k,\ell) \\ N_3(k,\ell) \end{vmatrix}
$$

$$
+ \left|\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \hdashline 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{array}\right|
\begin{vmatrix} u_1(k,\ell) \\ u_2(k,\ell) \end{vmatrix}
$$

$$
x(k,\ell) = \left|\begin{array}{ccc:ccc} -\rho_2^{\,2}\cdot a_2 & -b_2 & 1 & 0 & 0 & 0 \end{array}\right|
\begin{vmatrix} M_1(k,\ell) \\ M_2(k,\ell) \\ M_3(k,\ell) \\ N_1(k,\ell) \\ N_2(k,\ell) \\ N_3(k,\ell) \end{vmatrix}
$$

$$\varepsilon_1 = 4\rho_1 + 1 + \rho_1^2 \qquad a_1 = \frac{1}{2}\left(\varepsilon_1^{\frac{1}{2}} + \delta_1^{\frac{1}{2}}\right)$$

$$\varepsilon_2 = 4\rho_2 + 1 + \rho_2^2 \qquad a_2 = \frac{1}{2}\left(\varepsilon_2^{\frac{1}{2}} + \delta_2^{\frac{1}{2}}\right)$$

$$\delta_1 = 4\rho_1 - 1 - \rho_1^2 \qquad b_1 = \frac{1}{2}\left(\varepsilon_1^{\frac{1}{2}} - \delta_1^{\frac{1}{2}}\right)$$

$$\delta_2 = 4\rho_2 - 1 - \rho_2^2 \qquad b_2 = \frac{1}{2}\left(\varepsilon_2^{\frac{1}{2}} - \delta_2^{\frac{1}{2}}\right)$$

$$R_{u_1u_1}(m,n) = \begin{cases} \theta_1\theta_2\rho_1\rho_2\sigma^2 & \text{if } m = n = 0 \\ 0 & \text{if } m \neq 0 \text{ or } n \neq 0 \end{cases}$$

$$R_{u_2u_2}(m,n) = \begin{cases} (1-\rho_1^2)(1-\rho_2^2)\sigma^2 & \text{if } m = n = 0 \\ 0 & \text{if } n \neq 0 \text{ or } m \neq 0 \end{cases}$$

G. THE STATE SPACE STRUCTURE

1. Introduction

Section F showed that one can arrive at a Discrete-Space-Model for linear Image Processing, that has the form:

$$\begin{vmatrix} \underset{\sim}{M}(k+1,\ell) \\ \hdashline \underset{\sim}{N}(k,\ell+1) \end{vmatrix} = \left|\begin{array}{c:c} A_1 & A_2 \\ \hdashline A_3 & A_4 \end{array}\right| \cdot \begin{vmatrix} M(k,\ell) \\ \hdashline N(k,\ell) \end{vmatrix} + \begin{vmatrix} B_1 \\ \hdashline B_2 \end{vmatrix} \cdot u(k,\ell)$$

$$x(k,\ell) = \left|\begin{array}{c:c} C_1 & C_2 \end{array}\right| \begin{vmatrix} M(k,\ell) \\ N(k,\ell) \end{vmatrix}$$

where:

k: An integer valued vertical coordinates

ℓ: An integer valued horizontal coordinates

M: A vector which conveys information vertically

N: A vector which conveys information horizontally

u: A vector that acts as an input

x: A vector that acts as an output.

A_1, A_2, A_3, B_1, B_2, C_1, C_2 are matrices of appropriate dimensions. Boundary conditions $N(k,0)$, $M(0,\ell)$ are also inputs. $u(k,\ell)$ is specified externally.

This section will summarize the properties of the State-Space model. The discussion is based on Ref. [12].

2. Realization of a Discrete Filter By The State Space Structure

The following realization result has been demonstrated: Given an arbitrary two dimensional digital filter whose 2-D z transfer function is:

$$H(z_1,z_2) = \frac{n(z_1,z_2)}{z_1^{m}\cdot z_2^{n} + a_{10}\cdot z_1^{m-1}\cdot z_2^{n} + \ldots a_{mn}} \qquad 2\text{-}117$$

then matrices A_1, A_2, A_3, A_4 and vectors B_1, B_2 exist with dimensions A_1 - mxn, A_2 — mxn, A_3 — nxm, A_4 — nxn, B_1 — mxl, B_2 - nxl such that Eq. 2-117 may be put into the form of Eq. 2-116.

In addition, the following canonical form for Eq. 2-116 has been developed: If the denominator $d(z_1,z_2)$ of Eq. 2-117 factors as follows (such a factorization is denoted "doubly factorable"):

$$d(z_1,z_2) = d_1^{1}(z_1)\cdot d_2^{1}(z_2) + d_1^{2}(z_1)\cdot d_2^{2}(z_2)$$

then the canonical form for the A_i (i = 1,...4) of Eq. 2-116 is:

$$A_1 = \begin{vmatrix} 0 & 0 & 0 & \cdots & 0 & -a_1 \\ 1 & 0 & 0 & \cdots & 0 & -a_2 \\ 0 & 1 & 0 & \cdots & 0 & -a_3 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 1 & -a_m \end{vmatrix}$$

$$A_2 = \begin{vmatrix} 0 & 0 & 0 & \cdots\cdots & -b_1 \\ 0 & 0 & 0 & \cdots\cdots & -b_2 \\ 0 & 0 & 0 & \cdots\cdots & -b_3 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots\cdots & -b_n \end{vmatrix}$$

$$A_3 = \begin{vmatrix} 0 & 0 & 0 & \cdots\cdots & -c_1 \\ 0 & 0 & 0 & \cdots\cdots & -c_2 \\ 0 & 0 & 0 & \cdots\cdots & -c_3 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots\cdots & -c_n \end{vmatrix}$$

$$A_4 = \begin{vmatrix} 0 & 0 & 0 & \cdots & 0 & -d_1 \\ 1 & 0 & 0 & \cdots & 0 & -d_2 \\ 0 & 1 & 0 & \cdots & 0 & -d_3 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 1 & -d_n \end{vmatrix}$$

where the coefficients a_i, b_i, c_i and d_i are determined from $d(z_1,z_2)$ by

$$d(z_1,z_2) = (z_1^m + a_m z_1^{m-1} + a_{m-1} z_1^{m-2} + \ldots a_1)(z_2^n + d_n z_1^{n-1} + \ldots d_1)$$

$$- (b_m z_1^{m-1} + b_{m-1} z_1^{m-2} + \ldots b_1)(c_n z_1^{n-1} + \ldots c_1)$$

2. The State Transition Matrix

Definitions:

$$A \triangleq \begin{vmatrix} A_1 & A_2 \\ A_3 & A_4 \end{vmatrix} \qquad \text{2-118a}$$

$$B \triangleq \begin{vmatrix} B_1 \\ B_2 \end{vmatrix} \qquad \text{2-118b}$$

$$C \triangleq \begin{vmatrix} C_1 & C_2 \end{vmatrix} \qquad \text{2-118c}$$

$$T(k,\ell) \triangleq \begin{vmatrix} M(k,\ell) \\ N(k,\ell) \end{vmatrix} \qquad \text{2-118d}$$

$$T'(k,\ell) = \begin{vmatrix} M(k+1,\ell) \\ \\ N(k,\ell+1) \end{vmatrix} \qquad 2\text{-}118e$$

Then:

$$T'(k,\ell) = AT(k,\ell) + Bu(k,\ell) \qquad 2\text{-}119$$

$$y(k,\ell) = CT(k,\ell) \qquad 2\text{-}120$$

Now, for $k > 0$, $\ell > 0$, the next definitions are done:

$$A^{1,0} = \begin{vmatrix} A_1 & A_2 \\ 0 & 0 \end{vmatrix} \qquad 2\text{-}121a$$

$$A^{0,1} = \begin{vmatrix} 0 & 0 \\ \\ A_1 & A_2 \end{vmatrix} \qquad 2\text{-}121b$$

$$A^{k,\ell} = A^{1,0}\,A^{k-1,\ell} + A^{0,1}\,A^{k,\ell-1} \qquad 2\text{-}121c$$

$$A^{0,0} = I$$

$$A^{-k,\ell} = A^{k,-\ell} = 0$$

$A(k,\ell)$ is called "State Transition Matrix". Examination of these definitions leads to formulas of system-response that are similar to the one-dimensional discrete case.

Table 1 compares the formulas for the one dimensional case to the two dimensional case.

3. Properties Of $A^{k,\ell}$

1. $$A = \begin{vmatrix} A_1 & A_2 \\ A_3 & A_4 \end{vmatrix} = \begin{vmatrix} A_1 & A_2 \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 \\ A_3 & A_4 \end{vmatrix}$$

$$\boxed{A = A^{1,0} + A^{0,1}} \qquad 2\text{-}122$$

2. $$A^{k,0} = A^{1,0} A^{k-1,0} + A^{0,1} A^{k,-1}$$

$$= A^{1,0} A^{k-1,0}$$

Thus, $$\boxed{A^{k,0} = \left(A^{1,0}\right)^k} \qquad 2\text{-}123$$

and $$\boxed{A^{0,\ell} = \left(A^{0,1}\right)^\ell} \qquad 2\text{-}124$$

3. $$I = \begin{vmatrix} I & 0 \\ 0 & I \end{vmatrix}$$

	ONE DIMENSION	TWO DIMENSIONS
MODEL	$X_{(k+1)} = AX_{(k)} + BU_{(k)}$	$\left[\begin{array}{c} M_{(k+1,\ell)} \\ \hline N_{(k,\ell+1)} \end{array}\right] = \left[\begin{array}{c\|c} A_1 & A_2 \\ \hline A_3 & A_4 \end{array}\right] \left[\begin{array}{c} M_{(k,\ell)} \\ \hline N_{(k,\ell)} \end{array}\right] + \left[\begin{array}{c} B_1 \\ \hline B_2 \end{array}\right] U_{(k,\ell)}$ $X_{(k,\ell)} = \left[\begin{array}{c\|c} C_1 & C_2 \end{array}\right] \left[\begin{array}{c} M_{(k,\ell)} \\ \hline N_{(k,\ell)} \end{array}\right]$
STATE TRANSITION MATRIX	$\phi_{(k)} = A^k = A^1 \; A^{k-1}$ $\phi_{(o)} = A^0 = I$	$A^{k,\ell} \triangleq A^{1,0} \; A^{k-1,\ell} + A^{0,1} \; A^{k,\ell-1}$ $A^{0,0} = I$
RESPONSE FOR ZERO INPUT: $U \equiv 0$	$\boxed{X_{(k)} = \phi_{(k)} \; X_{(o)} = A^k X_{(o)}}$	DEFINE: $T_{(k,\ell)} \triangleq \left[\begin{array}{c} M_{(k,\ell)} \\ \hline N_{(k,\ell)} \end{array}\right]$ $T'_{(k,\ell)} \triangleq \left[\begin{array}{c} M_{(k+1,\ell)} \\ \hline N_{(k,\ell+1)} \end{array}\right]$ THEN: $T'_{(k,\ell)} = A \; T_{(k,\ell)} + BU_{(k,\ell)}$ $X_{(k,\ell)} = C \; T_{(k,\ell)}$ $\boxed{T_{(k,\ell)} = A^{k\ell} \; T_{(o,o)}}$

Table 1: Comparison Between the Response of Two Dimensional Systems to One Dimensional Systems

where I is the identity matrix with appropriate dimensions. Thus,

$$I^{1,0} = \begin{vmatrix} I & 0 \\ 0 & 0 \end{vmatrix}$$

$$I^{0,1} = \begin{vmatrix} 0 & 0 \\ 0 & I \end{vmatrix}$$

4. $$I^{1,0} A = \begin{vmatrix} I & 0 \\ 0 & 0 \end{vmatrix} \begin{vmatrix} A_1 & A_2 \\ A_3 & A_4 \end{vmatrix} = \begin{vmatrix} A_1 & A_2 \\ 0 & 0 \end{vmatrix}$$

$$= A^{1,0}$$

or:

$$\boxed{I^{1,0} A = I^{1,0} A^{1,0} = A^{1,0}} \qquad \text{2-125a}$$

$$\boxed{I^{0,1} A = I^{0,1} A^{0,1} = A^{0,1}} \qquad \text{2-125b}$$

5. $$I^{0,1} A^{1,0} = \begin{vmatrix} 0 & 0 \\ 0 & I \end{vmatrix} \begin{vmatrix} A_1 & A_2 \\ 0 & 0 \end{vmatrix} = 0$$

Briefly,

$$\boxed{I^{0,1} A^{1,0} = I^{1,0} A^{0,1} = 0} \qquad 2\text{-}126$$

In the next discussion we shall use some figures for better understanding. The figures include arrows that are going from one pixel to its neighbor. For example: The arrows in Fig. 23 means that the values $M(k,\ell)$, $N(k,\ell)$ contribute the values $N(k,\ell+1)$ and $M(k+1,\ell)$.

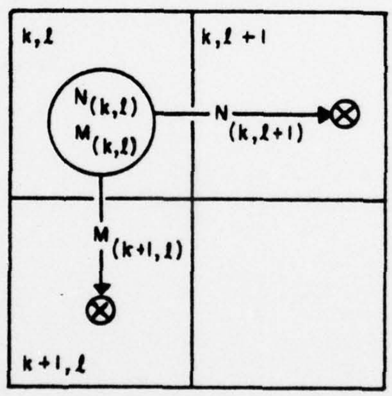


Fig. 23: The Propagation of the States M and N. [The values $M(k,\ell)$, $N(k,\ell)$ contribute the values $N(k,\ell+1)$ and $M(k+1,\ell)$].

4. <u>General Response Formula</u>

<u>Lemma</u>: Let $u(k,\ell)$ be zero for all (k,ℓ).

$$M(0,\ell) = N(k,0) = 0 \quad \text{for} \quad (i,j) \neq 0 .$$

(only $M(0,0)$, $N(0,0)$ are not zero).

Then: $$\boxed{T(k,\ell) = A^{k,\ell}\, T(0,0)} \qquad 2\text{-}127$$

Proof: (By induction)

$$T(0,0) = I \cdot T(0,0) = A^{0,0}\, T(0,0)$$

Now assume it is correct for any $(0,0) \leq (i,j) \leq (k,\ell)$, then:

$$T(k,\ell) = \begin{vmatrix} M(k,\ell) \\ \\ N(k,\ell) \end{vmatrix}$$

$$= \begin{vmatrix} A_1\, M(k-1,\ell) + A_2\, N(k-1,\ell) + B_1 \cdot 0 \\ \\ A_3\, M(k,\ell-1) + A_3\, N(k,\ell-1) + B_2 \cdot 0 \end{vmatrix}$$

$$= \begin{vmatrix} A_1 & A_2 \\ \\ 0 & 0 \end{vmatrix} T(k-1,\ell) + \begin{vmatrix} 0 & 0 \\ \\ A_3 & A_4 \end{vmatrix} T(k,\ell-1) \; .$$

$$= A^{1,0}\, A^{k-1,\ell}\, T(0,0) + A^{0,1}\, A^{k,\ell-1}\, T(0,0$$

$$= A^{k,\ell}\, T(0,0)$$

Q.E.D.

Figure 24 shows the propagation of the field due to $M(0,0)$ and $N(0,0)$.

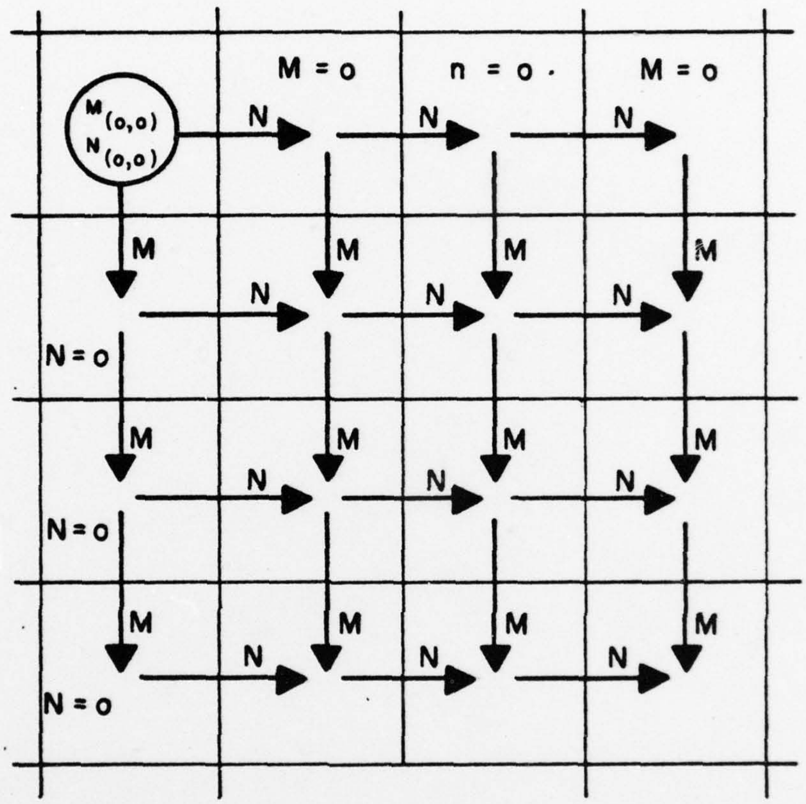


Fig. 24: The propagation of states due to M(0,0) N(0,0).

Equation 2-127 will be used for the next two cases:

Effect of M(0,j):

Assume: M(0,j) is the only non-zero boundary condition and that all inputs are zero.

Using Eq. (2-118d):

$$T(0,j) = \begin{vmatrix} M(0,j) \\ 0 \end{vmatrix}$$

Therefore T(0,j) can be used as initial condition for

2-128: 2-128:

$$\boxed{T(k,\ell) = A^{k,\ell-j} \cdot T(0,j) = A^{k,\ell-j} \cdot \begin{vmatrix} M(0,j) \\ 0 \end{vmatrix}} \quad 2\text{-}128$$

Effect of N(i,0):

Similarly to M(0,j) the effect of N(i,0) is:

$$\boxed{T(k,\ell) = A^{k-i,\ell} \begin{vmatrix} 0 \\ N(i,0) \end{vmatrix}} \quad 2\text{-}129$$

Effect of $u(k,\ell)$:

Assume: $u(i,j)$ for some $(i,j) < (k,\ell)$ is the only non-zero input. All boundary conditions are zero.

Then:

$$T(i+1,j) = \begin{vmatrix} M(i+1,j) \\ N(i+1,j) \end{vmatrix} =$$

$$= \begin{vmatrix} A_1 M(i,j) + A_2 N(i,j) + B_i u(i,j) \\ A_3 M(i+1,j-1) + A_3 N(i+1,j-1) + B_2 u(i+1,j-1) \end{vmatrix}$$

AD-A045 179 NAVAL POSTGRADUATE SCHOOL MONTEREY CALIF F/G 12/1

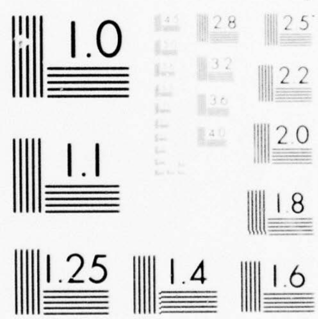
MODELING AND RECURSIVE ESTIMATION OF TWO DIMENSIONAL RANDOM FIE--ETC(U)

JUN 77 M SHACHAR

UNCLASSIFIED NL

2 OF 3

AD A045179



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

$$T(i+1,j) = \begin{vmatrix} A_1 \cdot 0 + A_2 \cdot 0 + B_1 \cdot u(i,j) \\ \\ 0 \end{vmatrix}$$

$$= \begin{vmatrix} B_1 \\ \\ 0 \end{vmatrix} u(i,j)$$

and:

$$T(i,j+1) = \begin{vmatrix} 0 \\ \\ B_2 \end{vmatrix} u(i,j)$$

$T(i+1,j)$, $T(i,j+1)$ might be substituted in 2-127 as boundary conditions. Therefore, by using superposition:

$$\boxed{T(k,\ell) = \left(A^{i-i-1,\ell-j} \cdot \begin{vmatrix} B_1 \\ \\ 0 \end{vmatrix} + A^{k-i,\ell-j-1} \cdot \begin{vmatrix} 0 \\ \\ B_2 \end{vmatrix} \right) u(i,j)} \tag{2-130}$$

Fig. 25 shows the effect of $u(i,j)$.

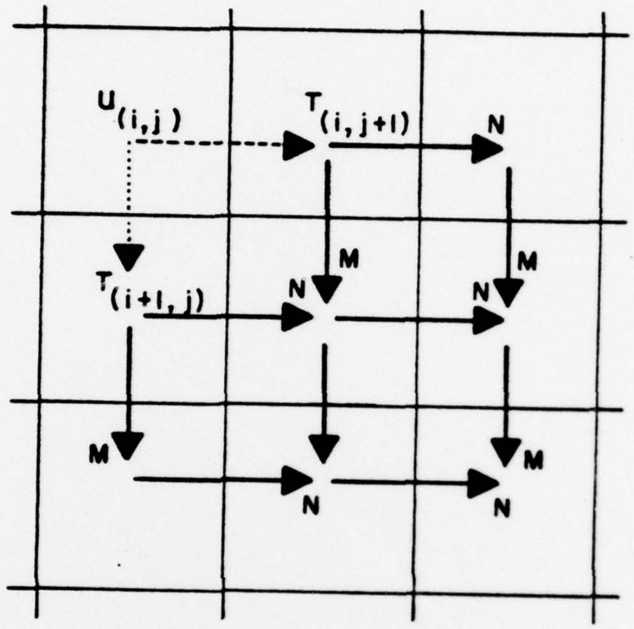


Fig. 25. Effect of u(i,j)

<u>Superposition:</u>

Theorem: for all $(k,\ell) \geq 0$:

$$T(k,\ell) = \sum_{j=0}^{\ell} A^{k,\ell-j} \cdot \begin{vmatrix} M(o,j) \\ \\ 0 \end{vmatrix} + \sum_{i=0}^{k} A^{k-i,\ell} \cdot \begin{vmatrix} 0 \\ \\ N(i,0) \end{vmatrix}$$

$$+ \sum_{(0,0) \leq (i,j) \leq (k,\ell)} \sum \left(A^{k-i-1,\ell-j} \cdot \begin{vmatrix} B_1 \\ \\ 0 \end{vmatrix} + A^{k-k,\ell-j-1} \cdot \begin{vmatrix} 0 \\ \\ B_2 \end{vmatrix} \right)$$

Proof: By superposition of the effects of all inputs and boundary conditions.

Q.E.D.

5. Characteristic Function

In the one dimensional case the eigen values of the system

$$x(k+1) = Ax(k)$$

are defined as those values which satisfy the algebraic equation:

$$Ax = \lambda x$$

or the characteristic equation

$$|\lambda I - A| = 0$$

Now, for the "State Space Structure" let's define in the same way: given:

$$\begin{vmatrix} M(k+1,\ell) \\ N(k,\ell+1) \end{vmatrix} = \begin{vmatrix} A_1 & A_2 \\ A_3 & A_4 \end{vmatrix} \cdot \begin{vmatrix} M(k,\ell) \\ N(k,\ell) \end{vmatrix}$$

define operators E, F such that:

$$\begin{vmatrix} A_1 & A_2 \\ A_3 & A_4 \end{vmatrix} \cdot \begin{vmatrix} M(k,\ell) \\ N(k,\ell) \end{vmatrix} = \begin{vmatrix} EM(k,\ell) \\ FN(k,\ell) \end{vmatrix}$$

or:

$$\begin{vmatrix} E\cdot I-A_1 & A_2 \\ A_3 & F\cdot I-A_4 \end{vmatrix} \cdot \begin{vmatrix} M(k,\ell) \\ N(k,\ell) \end{vmatrix} = 0 \qquad \text{2-131}$$

To have a nontrivial solution of this equation we require that the matrix in 2-131 is singular. Therefore the determinant should be zero.

$$\boxed{\begin{vmatrix} E\ I-A_1 & A_2 \\ A_3 & F\ I-A_4 \end{vmatrix} = |E\ I^{1,0} + f\ I^{0,1} - A| = 0}$$

2-131a

6. <u>Stability</u>

The stability criteria of Huang [24] can be generalized in a straightforward manner to systems represented in state variable form. This generalization allows the use of standard one-dimensional routines in the determination of two dimensional system stability.

The stability criteria will be:

1) The eigenvalues of the matrices A_1, A_4 are less than one in magnitude.

2) The eigenvalues of the complex matrix:

$$A_4 + A_3(z_1 I - A_1)^{-1} A_2 \qquad 2\text{-}132$$

are all less than 1 in magnitude as z_1 varies about the unit circle ($|z_1| = 1$). If any eigenvalues of Eq. 2-132 has magnitude greater than one for $|z_1| = 1$, then the system is unstable.

H. MODELING BY USING OPTIMAL ESTIMATION THEORY

So far we have seen a technique for modeling random fields that uses z transformation, properties of linear filters and special types of correlation functions (separable). That method was called "Filter Response Method". In this section optimal estimation theory is used to solve the modeling problem. The advantage of this method is that it is not limited. For comparison between the two methods (see Section I).

1. The Basic Principle

Suppose a discrete random field $x(k,\ell)$ is given and assumed to be homogeneous (stationary). Given the values $x(k,\ell)$ at all points $\Omega(k,\ell)$, the problem is to estimate the value $x(k,\ell)$. In other words, if $\hat{x}(k,\ell)$ is the optimum estimate for $x(k,\ell)$, then:

$$\hat{x}(k,\ell) = \sum_{\substack{\text{for all } (i,j) \\ \text{so that} \\ (k-i,n-j)\in\Omega(k,\ell)}} \alpha_{ij}\, x(k-i,\ell-j) \qquad \text{2-133}$$

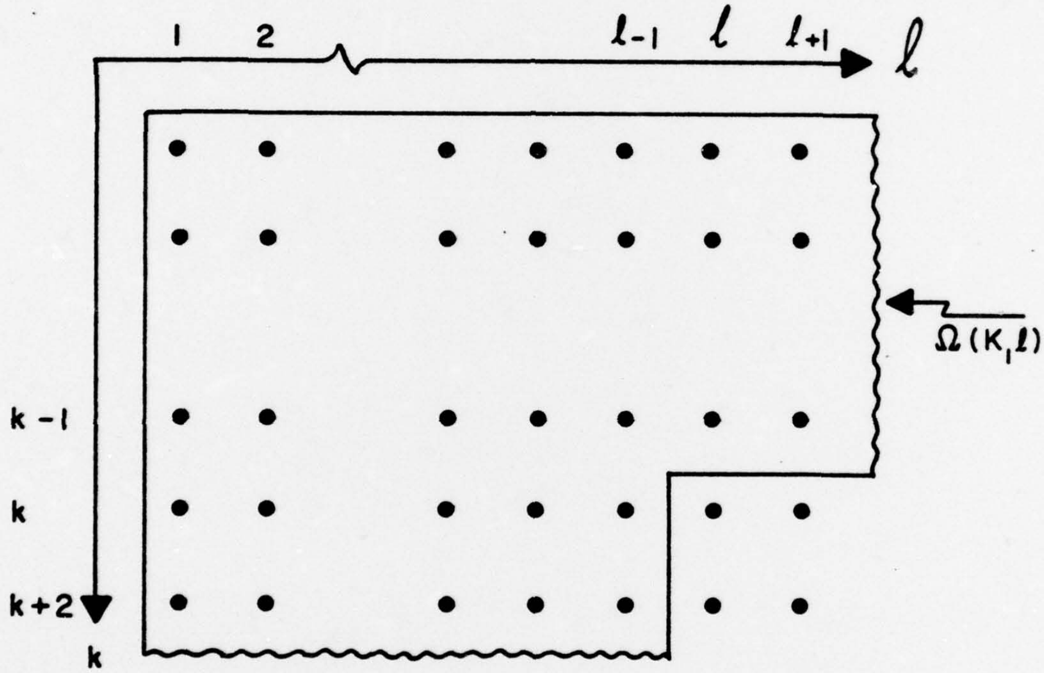


Fig. 26: Definition of $\Omega(k,\ell)$

The coefficients of $\alpha_{i,j}$ must be determined so that the "mean square error"

$$\varepsilon(k,\ell) = E\{[x(k,\ell) - \hat{x}(k,\ell)]^2\} \qquad \text{2-134}$$

is minimized. $\hat{x}(k,\ell)$ will be the "linear least square estimate" of $x(k,\ell)$.

Substituting 2-133 in 2-134, differentiating with respect to each $\alpha_{i,j}$, and setting each derivative equal to zero, we obtain the following set of simultaneous equations for the unknowns $\alpha_{i,j}$:

$$\boxed{E\{[x(k,\ell) - \hat{x}(k,\ell)]\ x(i,j)\} = 0 \qquad \text{for all } (i,j)\varepsilon\Omega} \tag{2-135}$$

which says that the coefficients $\alpha_{i,j}$ must be such that the estimation error $x(k,\ell) - \hat{x}(k,\ell)$ is statistically orthogonal to each $x(i,j)$ that is used to form the linear estimate. This is known as the <u>orthogonality principle</u>, in linear least square estimation.

Let D represent the following collection of pairs (i,j):

$$D = \{(0,1),\ (1,1),\ (1,0)\} \tag{2-136}$$

Now comes the definition of a first order process.

<u>Definition 1:</u>

A random field will be called <u>first order Markov</u> if the coefficients $\alpha_{i,j}$ in 2-133 are such that $\hat{x}$ is of the form:

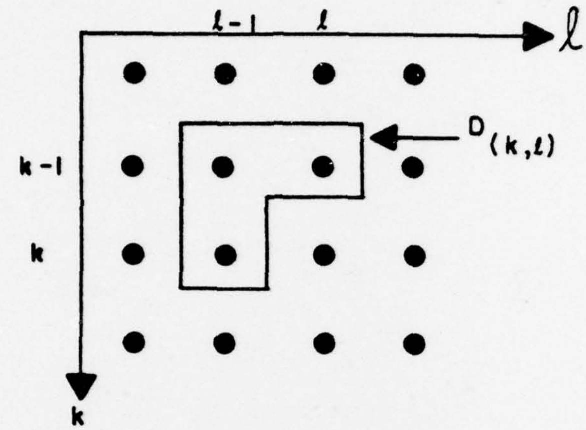


Fig. 27: Definition of Reg D

$$\hat{x}(k,\ell) = \sum_{(i,j)\in D(k,\ell)} \alpha_{ij}\, x(k-i,\ell-j) \qquad \text{2-137}$$

(all other α's are zero).

That is, the least square estimate of $x(k,\ell)$ in terms of $\Omega(k,\ell)$ is the same as that in terms of only the three immediate neighbors left and above the point (k,ℓ).

Substituting 2-137 in 2-135, the following conditions for the Markov field must be satisfied:

$$E\{[x(k,\ell) - \sum_{(i,j)\in D(k,\ell)} \alpha_{i,j}\, x(k-i,\ell-j)]\, x(p,q)\} = 0 \qquad \text{2-138}$$

for all $(p,q) \in \Omega$.

For a Markov field the coefficients $\alpha_{i,j}$ must be such that 2-138 is satisfied for all $(p,q) \in \Omega(k,\ell)$. In particular, 2-138 must be satisfied for the following values of (p,q):

$$(k-1,\ell),\ (k-1,\ell-1),\ (k,\ell-1).$$

Substituting the values of (p,q) in 2-138, the following equations are obtained, for $\alpha_{i,j}$:

$$\alpha_{1,0}\, R_{xx}(0,0) + \alpha_{1,1}\, R_{xx}(-1,0) + \alpha_{0,1}\, R_{xx}(-1,1) = R_{xx}(-1,0)$$

$$\alpha_{1,0}\, R_{xx}(1,0) + \alpha_{1,1}\, R_{xx}(0,0) + \alpha_{0,1}\, R_{xx}(0,1) = R_{xx}(-1,1)$$

$$\alpha_{1,0}\, R_{xx}(1,-1) + \alpha_{1,1}\, R_{xx}(0,-1) + \alpha_{0,1}\, R_{xx}(0,0) = R_{xx}(0,-1)$$

2-139

where $R_{xx}(\alpha,\beta) = E\{x(k,\ell)\ X(k+\alpha,\ell+\beta)\}$. For system coordinates in this problem see Fig. 19.

<u>Example 1</u>

Given

$$R_{xx}(n,m) = \rho_1^{|n|}\ \rho_2^{|m|} \qquad \text{2-140}$$

by substituting 2-140 into 2-139, the solution for the $\alpha_{i,j}$ is:

$$\alpha_{0,1} = \rho_1$$

$$\alpha_{1,1} = -\rho_1 \rho_2$$

$$\alpha_{1,0} = \rho_2$$

and by substituting into 2-137:

$$\hat{x}(k,\ell) = \rho_1 \, x(k-1,\ell) - \rho_1\rho_2 \, x(k-1,\ell-1) + \rho_2 \, x(k,\ell-1). \tag{2-141}$$

Note: The discussion in this section will be limited mainly to "one-side process" (causal difference equation) in order to compare it to the method of modeling that is described in Section F. In the end of this section there will be some examples of non-causal models.

2. The Modeling Error

Definition 2: The modeling error is the difference between the true value, $x(k,\ell)$ and the estimate, $\hat{x}(k,\ell)$.

$$u(k,\ell) \triangleq x(k,\ell) - \hat{x}(k,\ell)$$

$$= x(k,\ell) - \sum_{(i,j)\varepsilon\Omega} x(k-i,\ell-j)\,\alpha_{i,j} \tag{2-142}$$

It is obvious that:

$$x(k,\ell) = \hat{x}(k,\ell) + u(k,\ell)$$

$$\boxed{x(k,\ell) = \sum_{(i,j)\epsilon\Omega} \alpha_{i,j}\, x(k-i,\ell-j) + u(k,\ell)} \qquad 2\text{-}143$$

The error $u(k,\ell)$ creates also a random field. The question is what kind of random field? We are concerned in the variance and the autocorrelation-function of this error.

a) <u>The Variance</u> (of the modeling error)

Using 2-137:

$$Q \triangleq E\{u^2(k,\ell)\}$$

$$= E\{[x(k,\ell) - \hat{x}(k,\ell)\}$$

$$= E\{x^2(k,\ell) - 2\hat{x}(k,\ell)\cdot x(k,\ell) + \hat{x}^2(k,\ell)\}$$

$$= E\{x^2(k,\ell) - \hat{x}(k,\ell)\cdot x(k,\ell) + \cancel{\hat{x}(k,\ell)\,[\hat{x}(k,\ell) - x(k,\ell)]}^{\,0}\}$$

The zero in the last term is set by using the orthogonal principle (Eq. 2-138).

Therefore:

$$Q = E\{x^2(k,\ell) - [\sum_{(i,j)\epsilon\Omega} \alpha_{(i,j)}\, x(k-i,\ell-j)]\; x(k,\ell)\}$$

and recall that:

$$R(i,j) = E\{x(k-i,\ell-j)\; x(k,\ell)\} = \text{autocorrelation function}$$

$$\boxed{Q = E\{u^2(k,\ell)\} = R(0,0) - \sum_{(i,j)\in\Omega} \alpha_{i,j} \, R(i,j)}$$ 2-144

b) <u>The Autocorrelation</u> (of the modeling error)

<u>Theorem</u>: The modeling error creates a random field that is white noise, e.g.:

$$R\ (n,m) = \begin{cases} Q & \text{if} \quad m = 0 \quad \text{and} \quad n = 0 \\ 0 & \text{if} \quad m \neq 0 \quad \text{or} \quad n \neq 0 \end{cases}$$

<u>Proof</u>: By using 2-143, Eq. 2-137 can be written in a non-recursive way as follows:

$$\hat{x}(k,\ell) = \sum_{m=1}^{k} \sum_{n=1}^{\ell} \beta_{m,n} u(m,n)$$

$$+ \underbrace{\sum_{j=1}^{\ell} \gamma_{0,j}\ x(0,\ell-j) + \sum_{i=1}^{\ell} \gamma_{i,0}\ x(i-i,0)}_{\text{Initial conditions}}$$ 2-145

But using the orthogonality principle:

$$E\{u(k,\ell)\ x(i,j)\} = 0 \qquad (i,j) \ \varepsilon\ \Omega$$ 2-146

and especially:

$$E\{u(k,\ell)\ \hat{x}(k,\ell)\} = 0 \qquad \text{2-147}$$

substituting 2-145 into 2-147, the result is obvious:

$$E\{u(k,\ell)\ u(m,n)\} = 0 \qquad (m,n) \in \Omega$$

Q.E.D.

Note: As previously determined, this discussion concerns causal models. In this case we proved that the random process is forced by white noise. That is not the case in a non-causal model.

Example 2

Problem: Find the variance of the modeling error for the model of Example 1.

Solution:

$$Q = 1 - \alpha_{0,1} R(0,1) - \alpha_{1,1} R(1,1) - \alpha_{1,0} R(1,0)$$

$$= 1 - \rho_1^2 + (\rho_1\rho_2)^2 - \rho_2^2$$

$$Q = (1 - \rho_1^2)(1 - \rho_2^2) \qquad \text{2-148}$$

Conclusion: For the autocorrelation function of 2-140

$$R_{xx}(n,m) = \rho_1^{|n|}\rho_2^{|m|}$$

the model is:

$$x(k+1,\ell+1) = \rho_1 x(k+1,\ell) + \rho_1\rho_2 x(k,\ell) + \rho_2 x(k,\ell+1) + u(k,\ell)$$

$$E\{u^2(k,\ell)\} = Q = (1-\rho_1^{\,2})(1-\rho_2^{\,2})$$

2-149

Now, comparing 2-149 to 2-81, 2-82 it is seen that:

For the First Order Markov Field the orthogonality principle (Minimum Mean Square Error) leads to the same model as the "filter response method".

3. Advantages Of The "Orthogonality Principle" Method

There are three distinct advantages of this method:

a) This method can be applied to non-separable autocorrelation functions. This is impossible to do with the Linear Filter Response method.

b. This method can be extended to non-causal models. The technique is very similar to the causal-modeling technique. The only differences are that the modeling error is not white any more and that the solution doesn't exist for all cases.

c. It is optimal in the sense that the modeling error, $u(k,\ell)$, is minimum. It was not proved that the method of Section F, Linear Filter Response, leads to an optimal model. [We see it only for one special case, the first order Markov field.] For further discussion of this point see Ref. 23.

The next examples will highlight these three advantages.

Example 3 (for advantage a):

Given: $R_{xx} = \rho^{n^2+m^2}$ 2-150

Find: A linear model for the field that uses region D (defined in Eq. 2-136 and described in Fig. 27)

Solution: As in Example 1, the set of three algebraic equations of 2-139 has to be used:

$$\begin{vmatrix} R_{xx}(0,0) + R_{xx}(-1,0) + R_{xx}(-1,1) \\ R_{xx}(1,0) + R_{xx}(0,0) + R_{xx}(0,1) \\ R_{xx}(1,-1) + R_{xx}(0,-1) + R_{xx}(0,0) \end{vmatrix} \cdot \begin{vmatrix} \alpha_{1,0} \\ \alpha_{1,1} \\ \alpha_{0,1} \end{vmatrix} = \begin{vmatrix} R_{xx}(-1,0) \\ R_{xx}(-1,1) \\ R_{xx}(0,-1) \end{vmatrix}$$

and by substituting 2-150:

$$\begin{vmatrix} 1 & \rho & \rho^{\sqrt{2}} \\ \rho & 1 & \rho \\ \rho^{\sqrt{2}} & \rho & 1 \end{vmatrix} \cdot \begin{vmatrix} \alpha_{1,0} \\ \alpha_{1,1} \\ \alpha_{0,1} \end{vmatrix} = \begin{vmatrix} \rho \\ \rho^{\sqrt{2}} \\ \rho \end{vmatrix}$$

The α's are solutions of these three equations.

Note: The process in Eq. 2-150 is not first-order, if we use definition 1. There is not a finite set of elements on which the point $x(k,\ell)$ depends. Therefore the choice to model the process by region D was arbitrary in this example.

Example 4: (for advantage b):

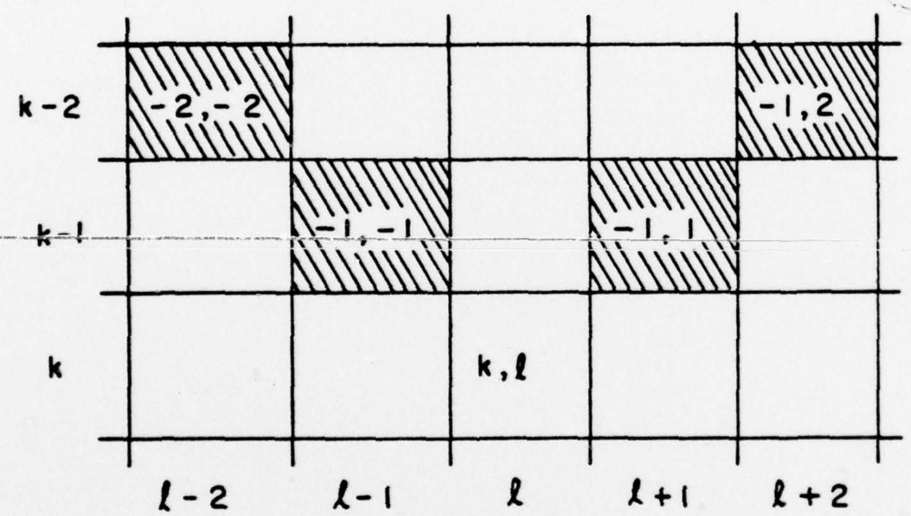


Fig. 28: A Non-causal Region For Modeling (in example 4)

Problem: Find the non-causal representation of $x(k,\ell)$ by using the dotted elements of Fig. 28, for the autocorrelation function

$$R_{xx}(n,m) = \rho_1^{|n|}\rho_2^{|m|}$$

Solution: By using Eq. 2-138 for

$$(i,j) = \{(-2,-2),\ (-1,-1),\ (-1,1),\ (-1,2)\}$$

we have the next set of equations:

$$\begin{vmatrix} R_{xx}(0,0) & R_{xx}(-1,-1) & R_{xx}(-3,-1) & R_{xx}(-4,0) \\ R_{xx}(1,1) & R_{xx}(0,0) & R_{xx}(-2,0) & R_{xx}(-3,1) \\ R_{xx}(3,1) & R_{xx}(2,0) & R_{xx}(0,0) & R_{xx}(-1,1) \\ R_{xx}(4,0) & R_{xx}(3,-1) & R_{xx}(1,-1) & R_{xx}(0,0) \end{vmatrix} \cdot \begin{vmatrix} \alpha_{-2,-2} \\ \alpha_{-1,-1} \\ \alpha_{-1,1} \\ \alpha_{-1,2} \end{vmatrix} = \begin{vmatrix} R_{xx}(-2,2) \\ R_{xx}(-1,1) \\ R_{xx}(-1,1) \\ R_{xx}(-1,2) \end{vmatrix}$$

$$\begin{vmatrix} 1 & \rho_1\rho_2 & \rho_1^{3}\rho_2 & \rho_1^{4} \\ \rho_1\rho_2 & 1 & \rho_1^{2} & \rho_1^{3}\rho_2 \\ \rho_1^{3}\rho_2 & \rho_1^{2} & 1 & \rho_1\rho_2 \\ \rho_1^{4} & \rho_1^{3}\rho_2 & \rho_1\rho_2 & 1 \end{vmatrix} \begin{vmatrix} \alpha_{-2,-2} \\ \alpha_{-1,-1} \\ \alpha_{-1,1} \\ \alpha_{-1,2} \end{vmatrix} = \begin{vmatrix} \rho_1^{2}\rho_2^{2} \\ \rho_1\rho_2 \\ \rho_1\rho_2 \\ \rho_1^{2}\rho_2^{2} \end{vmatrix}$$

The solution:

$$\alpha_{-2,-2} = \alpha_{-1,2} = 0$$

$$\alpha_{-1,-1} = \alpha_{-1,1} = \frac{\rho_1 \rho_2}{1 + \rho_1^2}$$

It is not surprising that $\alpha_{-2,-2}$ and $\alpha_{-1,2}$ are zero. The given autocorrelation function in this example is a <u>first-order</u> Markov process.

<u>Example 5</u>:

This example compares between the two methods of modeling (by optimal-estimation and by "filter-response-method"). The results turn out to be different. In this example we consider a second-order process. The optimal estimation approach for the <u>two-dimensional</u> case requires solving eight algebraic equations, which is complicated. Therefore this example deals with a one-dimensional autocorrelation function. The extension to a similar two-dimensional problem is straightforward.

<u>Given</u>: The one dimensional "Band Pass" autocorrelation function (see Chpt. II, Section E, parts 3,4).

$$R_{xx}(n) = \rho^{|n|} \cos(\theta n)$$

Find: A model that describes this process by:

a) optimal estimation approach.

b) by filter response method.

Solution:

a) In the optimal-estimation-approach one has to determine the order of the model (order of the difference equation). So let's choose a second order model:

$$x(k) = \alpha_1 x(k-1) + \alpha_2 x(k-2) + u(k)$$

where:

$$\begin{vmatrix} 1 & \rho \cos \theta \\ \rho \cos \theta & 1 \end{vmatrix} \begin{vmatrix} \alpha_1 \\ \alpha_2 \end{vmatrix} = \begin{vmatrix} \rho \cos \theta \\ \rho^2 \cos \theta \end{vmatrix}$$

Therefore:

$$\alpha_1 = \frac{\rho \cos\theta (1 - \rho^2 \cos 2\theta)}{1 - \rho^2 \cos^2\theta}$$

$$\alpha_2 = \frac{-\rho^2 \sin^2\theta}{1 - \rho^2 \cos^2\theta}$$

and by using 2-144: The variance of $u(k,\ell)$ is

$$E\{u^2(k,\ell)\} = \underline{Q = 1-\alpha_1 \rho \cos\theta - \alpha_2\rho^2 \cos^2\theta}$$

for:

$$\rho = 0.96$$

$$\theta = 8°$$

$$\alpha_1 = 1.127$$

$$\alpha_2 = -0.1854$$

$$Q = 0.0911 \Longrightarrow \underline{\underline{\left[\begin{array}{l}\text{variance of}\\ \text{modeling error}\end{array}\right] = \sqrt{Q} = 0.302}}$$

b) Eq. 2-71, 2-72, lead to a model:

$$x(k) = \rho^2 x(k-1) - \rho\cos\theta\ x(k-2) + u_a(k)$$

$$= \alpha_1' x(k-1) + \alpha_2' x(k-2) + u_a(k)$$

for:

$$\rho = 0.96$$

$$\theta = 8°$$

$$\alpha_1' = 0.9216$$

$$\alpha_2' = -0.95$$

$$Q' = 0.150 \Longrightarrow \left[\begin{matrix}\text{Variance of} \\ \text{modeling error}\end{matrix}\right] = \sqrt{Q'} = 0.387$$

Q' is found from: $Q' = (1 - \rho^2)(1 + \rho^2)$. Table 2 summarizes the results.

Table 2:

$R_{xx}(n) = \rho^{\lvert n \rvert} \cos(\theta n)$ $\rho = 0.96$ $\theta = 8°$	
Optimal estimation approach	Filter Response method
One can choose to represent x(k), theoretically, by a set of infinite previous points (the order of the difference equation can be infinite).	Inherently from the modeling process it follows that the difference equation that expresses the process is of second-order.
$x(k)=\alpha_1 x(k-1)+\alpha_2 x(k-2)+u(k)$	$x(k)=\alpha_1' x(k-1)+\alpha_2' x(k-2)+u_a(k)$
$\alpha_1 = \frac{\rho \cos\alpha(1-\rho^2\cos 2\alpha)}{1-\rho^2\cos^2\alpha} = 1.127$	$\alpha_1' = \rho = 0.9216$
$\alpha_2 = \frac{-\rho^2 \sin^2 \theta}{1-\rho^2\cos^2\alpha} = -0.1854$	$\alpha_2' = -\rho \cos \theta = -0.95$
$\text{var}\{u\}$ = modeling error $= \sqrt{1-\alpha_1\rho\cos\theta - \alpha_2\rho^2\cos 2\theta}$ $= 0.302$	$\text{var}\{u\}$ = modeling error $= \sqrt{(1-\rho^2)(1+\rho^2)}$ $= 0.387$

One can see that the modeling error in the optimal solution is not much smaller than in the Filter Response Method (0.302 compared with 0.387). Therefore, in the sense of optimality, the model in Eq. 2-67, 2-68 is "almost good" as the optimal solution. It seems that the "nature" of the process $R_{xx}(n) = \rho^{|n|}\cos(\theta n)$ is to be modeled by Eq. 2-67, 2-68. These equations are simpler, and easy to handle.

Another point is that the method of Section F [Linear Filter Response] leads to a representation of $x(k,\ell)$ by a finite number of points. Using the optimal estimation approach, one can find a model by using three, four and more points, previous to $x(k,\ell)$. The coefficients of the far-away points are not zero. In other words it seems like the "nature" of this case is not to be presented by orthogonal projections.

Finally, these models are used in recursive estimation. In this application, anyway, the whole set of previous data is used to estimate a point (in a recursive way). Therefore, both models lead to an optimal-estimation-solution. We can summarize, as a consequence of this example. In the sense of optimality, the Linear Filter Response method is worse than the orthogonality principle method. But it has other obvious advantages that make it a good candidate for signal processing and especially image processing.

I. SUMMARY

Note: It is recommended that Example 5 in Section H be studied until the end of that section, before reading Section I.

1. This chapter has developed a state-variable representation of two dimensional random fields.
2. The concept of the modeling is primarily based on passing white noise through a linear filter. The basic equations that were used is Eq. 2-40 and 2-41. This concept was explained in Section D, and then was used in Sections E, F. It was called "filter response method". The method was applied to three types of Markov fields:

   a) The first-order type, where:

$$R_{xx}(n,m) = \sigma^2 \rho_1^{|n|} \rho_2^{|m|}$$

   b) The "band limited" type, where the spectrum of the signal is limited in a certain band:

$$R_{xx}(n,m) = \sigma^2 \rho_1^{|n|} \cos(\theta_1 n) \rho_2^{|m|} \cos(\theta_2 m)$$

   c) The second order type, where:

$$R_{xx}(n,m) = \sigma^2 \rho_1^{|n|} (1+\theta_1 |n|) \rho_2^{|m|} (1+\theta_2 |m|)$$

$$\theta_1 = \ln \rho_1 \qquad \theta_2 = \ln \rho_2$$

3. In all cases it was shown that the random fields can be expressed in a form that was recently suggested by Roesser [Ref. 12].

$$\begin{vmatrix} \underset{\sim}{M}(k+1,\ell) \\ \\ \underset{\sim}{N}(k,\ell+1) \end{vmatrix} = \begin{vmatrix} A_1 & A_2 \\ \\ A_3 & A_4 \end{vmatrix} \begin{vmatrix} \underset{\sim}{M}(k,\ell) \\ \\ \underset{\sim}{N}(k,\ell) \end{vmatrix} + \begin{vmatrix} B_1 \\ \\ B_2 \end{vmatrix} u(k,\ell)$$

The nature of models with such a structure was discussed in Section G.

4. The "filter response method" of modeling as it was discussed in Section F has two disadvantages:

   a) The difference equation which is used to represent the field (i.e.: the linear filter) turns out, always, to be a <u>causal equation</u>. But in image processing, there might be reasons to use non-causal filters. It is obvious that representing *a random field by a non-causal filter* will be better than a causal filter

   b) This method is limited to separable autocorrelation functions (smaller modeling error).

5. The modeling method by the orthogonal principle has its disadvantages:

   a) The weighting factors have no simple expressions.

   b) The number of neighbors one has to use, is theoretically infinite for some types of autocorrelation functions.

c) The calculation of the coefficients required solving many algebraic equations, especially in the two-dimensional case. The number of equations is equal to the number of coefficients.

in any → many

d) Perhaps, the greatest disadvantage in using this method for modeling two dimensional fields is the fact that it is difficult to represent the two dimensional difference equation that describes the field in the "state-space-structure" [Ref. 12].

The "Filter Response" method does not suffer from these disadvantages.

6. Although there are two disadvantages mentioned above, it seems that the "Filter Response Method" is a good candidate for modeling random fields.

7. Table 3 summarizes the properties of the two methods:

Table 3: Summary of Two Methods of Modeling

	optimal estimation approach	"filter response" approach	remarks
Optimality (minimum modeling error)	The model is optimal. The modeling error is white noise (for a causal model).	The model is not necessarily optimal, and the modeling error is not always white noise	see Example 5 - Section H
Deriving a non-causal model	Possible	Impossible	see Example 4 - Section H
Handling of non-separable autocorrelation functions	Possible	Impossible	see Example 3 - Section H
Complication of: - calculation - model form	Calculation - complicated (many algebraic equations) Model form - in some cases complicated.	Calculations are simple and the model has a simple structure	see Example 5 - Section H
Deriving a form of the model by using a state vector	Difficult in most cases [even for separable autocorrelation functions]	Simple	see Example 5 - Section H
Is the recursive equation that describes the process (model) finite?	Theoretically, for most cases, the recursive equation is infinite	Yes	see Example 5 - Section H
How many points to include in the model	Theoretically - infinite. Practically - enough points to make the modeling error small	The order of the recursive equation that describes the process is an inherent result of the modeling procedure	see Example 5 - Section H

III. REVIEW OF ESTIMATION THEORY CONCEPTS

A. INTRODUCTION

1. Definition: Stochastic estimation is the operation of assigning a value to an unknown system state or parameters, based upon noise-corrupted observation involving some function of the state or parameters.

2. For example, consider a random field whose model is the first-order-Markov:

$$x(k+1,\ell+1) = \rho_1 x(k+1,\ell) - \rho_1\rho_2 x(k,\ell) + \rho_2 x(k,\ell+1) + u(k,\ell) \qquad 3\text{-}1$$

where $u(k,\ell)$ is white noise:

$$E\{u(k,\ell)v(i,j)\} = \begin{cases} 0 & \text{if} \quad k \neq i \quad \text{or} \quad \ell \neq j \\ Q & \text{if} \quad k = i \quad \text{or} \quad \ell = j \end{cases} \qquad 3\text{-}2$$

The observation of the field is corrupted by noise: $y(k,\ell)$ is the observation:

$$y(k,\ell) = x(k,\ell) + v(k,\ell). \qquad 3\text{-}3$$

where:

$$E\{v(k,\ell)v(i,j)\} = \begin{cases} 0 & \text{if} \quad m \neq i \quad \text{or} \quad n \neq j \\ R & \text{if} \quad n = i \quad \text{and} \quad n = j \end{cases} \qquad 3\text{-}4$$

The problem is to find the "best" estimate for $x(k,\ell)$ and the so-called $\hat{x}(k,\ell)$, from a set of noisy measurements $y(i,j)$, $i = 1, \ldots k$, $j = 1, \ldots \ell$.

3. In Section B, the criteria for an estimator to be the "best" (optimal) are defined.

- Section C defines the "Linear Estimator".

- Section D defines the most common estimator: the linear estimator with the quadratic-cost criteria. The recursive and non-recursive estimators are discussed. The orthogonality principle is also explained.

This chapter explains the concepts of optimal estimation. Some of these concepts will be used in the next chapter to solve a few particular problems.

B. ESTIMATION CONCEPTS AND CRITERIA

In this section "optimal criteria" will be defined.

1. Baysian Estimation

The problem is to estimate a state X, from a set of measurements Y.

P(x) - is called "a priori" probability density function.

P(x/y) - is called "a posteriori" probability density function. Knowing the density function P(x,y), various types of estimators can be determined. Fig. 29 will be used to define them:

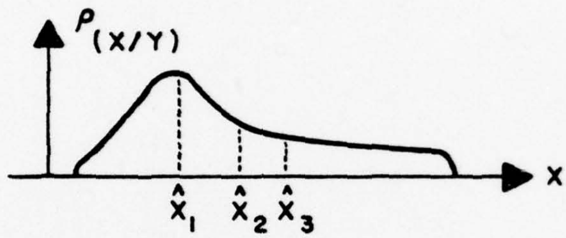


Fig. 29: The Various Types of Estimators

a) The most probable estimate is $\hat{x}_1$ (the x that is most likely to happen).

b) The estimator is $\hat{x}_2$, the solution of minimizing a conditional mean. The problem is to find

$$\min \int_{-\infty}^{\infty} \phi(x-\hat{x})\; P(x/y)\; dx \qquad 3\text{-}5$$

$\phi(x-\hat{x})$ is called the cost function. The following are typical examples.

1) $\phi(x-\hat{x}) = |x-\hat{x}| =$ absolute value.

2) $\phi(x-\hat{x}) = (x-\hat{x})^2 =$ square error.

Both types are the so-called equal risk estimators. The probability of x being larger or smaller than $\hat{x}_2$ is the same.

The quadratic form is the most commonly used cost function.

c) The minimax estimate is $\hat{x}_3$ - the estimate that minimizes the maximum probability of error, $|x - \hat{x}_3|$. This is simply the median.

The Bayes Rule:

Bayes rule is:

$$P(x/y) = \frac{P(x,y)}{P(y)} = \frac{P(y/x)P(x)}{P(y)}$$

$P(x)$ = a priori density function of x

$P(y)$ = density function of the measurements

$P(x,y)$ = joint density function.

$P(y/x)$ = conditional density function of the measurements y.

2. Maximum Likelihood Estimation

The fundamental idea is to define a "likelihood function" and to maximize this function.

The estimation problem assumes the form of the probability density function of y, P(y), is known:

$$P(y) = P(y,x)$$

where:

x is unknown parameter

y is a set of measurements.

Therefore, one can ask what is the best estimate of x based on these measurements? To find the answer it has to be found what is the set of parameters x, that will cause the set of measurements $y = (y_1 \ldots y_n)$ most likely to occur. The following procedure of computation is done:

a) Set up a likelihood function

$$L(y,x) = P(y/x)$$

b) Find the parameter x which maximizes this likelihood function. Hence the solution must satisfy the condition:

$$\frac{\partial L}{\partial x} = 0 , \qquad \frac{\partial^2 L}{\partial x^2} < 0$$

C. LINEAR ESTIMATION

Optimal estimation theory is the subject of finding an estimator that minimizes a given "cost function", or that is "most likely" to occur. It is very common to add an additional requirement to the estimator: The requirement of <u>linearity</u>. The linearity requirement will be introduced here in two ways: Non-Recursive and Recursive.

1. Non-Recursive Estimation

If Y is a given set of measurements, then the estimator will be a linear combination of this measurement:

$$\hat{x} = \sum_{i \in Y} \alpha_i y_i \qquad 3\text{-}6$$

As a particular example, assume an image that is scanned top to bottom, left to right, and that a causal estimator is required. In that case:

$$\hat{x}(k,\ell) = \sum_{(p,q) \in \Omega_1(k,\ell)} \alpha_{p_I q}\, y(p,q) \qquad 3\text{-}7$$

The region $\Omega_1(k,\ell)$ is seen in Fig. 30.

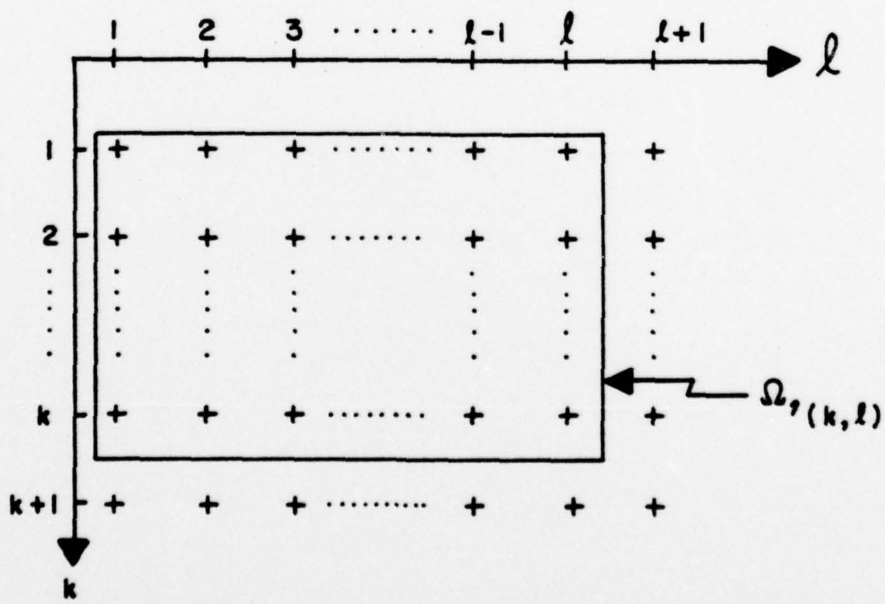


Fig. 30: Definition of Region $\Omega_1(k,\ell)$

The problem is to find the coefficient $\alpha(p,q)$ that will make the estimator optimal in the sense that was described in the previous section.

2. Recursive Estimation

For easy implementation it would be desirable to express $\hat{x}(k,\ell)$ in terms of previous estimators. For example:

$$\hat{x}(k,\ell) = \sum_{(p,q)\varepsilon D(k,\ell)} d_{k,\ell}(p,q)\hat{x}(p,q) + G_{k,\ell}y(k,\ell) \tag{3-8}$$

The *region* $D(k,\ell)$ is seen in Fig. 22. Now the problem is to find the d's and $G(k,\ell)$ for each point in the image, satisfying an optimal solution (with a given cost function). It is easy to see that 3-7 could be expressed in the form of 3-8 (the α's in 3-7 could be expressed in terms of the d's and the G's).

There is no doubt that the non-recursive estimator is the best that could be done, simply because it includes the whole set of given data in the "best" way. The recursive estimator also uses the whole set of data to estimate a point $x(k,\ell)$. It's only done in a recursive way. <u>Therefore the best recursive estimator that can be found is the one in which the "cost" is equal to the non-recursive counterpart</u>.

Of course one can ask if such a recursive estimator (whose cost is equal to the non-recursive counterpart) exist ? The answer to this question emphasizes the difference between one-dimensional estimation and two dimensional estimation.

In the one dimensional case, if the model of the process is linear and the cost has a quadratic form, both ways of estimation have the same cost (same variance of error). In that case the recursive estimator is the well known Kalman filter.

In the two dimensional case, under the same conditions, there is no way to find a recursive estimator that is as good as the non-recursive estimator. It will be proven in the next chapter. On the other hand it will be shown that one can find a recursive estimator whose variance of error is almost as low as the variance of the non-recursive counterpart. Therefore the recursive estimator will be a good candidate for estimation of two dimensional random fields.

The next section will define the most commonly used estimator.

D. LINEAR ESTIMATION COMBINED WITH QUADRATIC COST FUNCTION

1. Definition Of The Estimator

Let the cost function $\phi(x - \hat{x})$ in 3-5 be a quadratic function (or "quadratic form") of $(x - \hat{x})$. The optimal estimate is obtained by minimizing the cost:

$$E\{[x(k,\ell) - \hat{x}(k,\ell)]^2\} \qquad 3\text{-}9$$

Now, this cost is combined with another restriction: The estimator should be linear.

2. The Non Recursive Case

The linear estimator is given in 3-6

$$\hat{x}(k,\ell) = \sum_{(p,q)\in\Omega_1(k,\ell)} \alpha_{p,q}\, y(p,q)$$

The error:

$$\mathcal{E}(k,\ell) = x(k,\ell) - \hat{x}(k,\ell) = x(k,\ell) - \sum_{(p,q)\in\Omega_1(k,\ell)} \alpha_{p,q} y(p,q)$$

and the variance of the error:

$$E\{\mathcal{E}^2(k,\ell)\} = E\{[x(k,\ell) - \sum_{(p,q)\in\Omega_1(k,\ell)} \alpha_{p,q} y(p,q)]^2\}$$

Now, differentiating with respect to $\alpha_{p,q}$, $p = 1, \ldots k$, $q = 1, \ldots \ell$, and letting the derivative equal to zero. the following set of orthogonal conditions must be satisfied for optimal solution:

$$E\{[x(k,\ell) - \hat{x}(k,\ell)]\, y(p,q)\} = 0$$

or

$$\boxed{E\{[x(k,\ell) - \sum_{(p_1q)\varepsilon\Omega_1(k,\ell)} \alpha_{p,\dot{q}}\, y(p,q)]\; y(k,j)\} = 0}$$

$$(i,j)\varepsilon\Omega_1(k,\ell) \qquad 3\text{-}10$$

This is the "orthogonal principle" in optimal estimation. The meaning of 3-10 is that the estimation error $x(k,\ell) - \hat{x}(k,\ell)$ is uncorrelated to the set of data that is used to estimate the value $x(k,\ell)$

<u>To summarize</u>: An optimal estimator will be <u>only</u> the one that satisfies the orthogonality principle.

3. <u>The Recursive Case</u>

It is clear that the <u>optimal</u> recursive estimator will be the one that satisfies the orthogonality principle. The problem is <u>that</u> <u>it</u> <u>is</u> <u>impossible</u> to find a two-dimensional recursive estimator that does it. In the next chapter it will be shown that such an estimator does not exist.

Therefore, the approach in this thesis is to <u>define</u> a structure for the filter, say as in 3-8 [it will be shown that the <u>definition</u> of the structure itself leads to a conclusion that the estimator is not optimal], and then to find the parameters of the filter to minimize the variance of the error.

IV. RECURSIVE ESTIMATION OF TWO-DIMENSIONAL FIELDS

A. INTRODUCTION

1. Problem Definition

In this chapter the general description of the problem will be:

a) $x(k,\ell)$ is a stationary random field. The autocorrelation function $R_{xx}(n,m)$ is known.

b) From the autocorrelation-function one can find a dynamic model (state variable representation) of the field.

c) The measurement of the field includes noise:

$$y(k,\ell) = x(k,\ell) + v(k,\ell)$$

$x(k,\ell)$ is the correlated image. $v(k,\ell)$ is white noise. $y(k,\ell)$ is the measurement.

d) The problem is to estimate $x(k,\ell)$ from the measurement, using a <u>linear</u> <u>recursive</u> <u>filter</u> that minimizes the <u>quadratic</u> <u>form</u>

$$E\{(\hat{x} - x)^2\}$$

$\hat{x}$ is the estimate of x.

2. General Remarks About The Solution

1) As a consequence of Chpt. III one thing has to be kept in mind: That the optimal solution is defined only in a non-recursive form, by Eq. 3-7, to minimize the mean of the square error (3-9). This definition leads to the orthogonal principle. Therefore the optimal recursive filter will be only the one that is proved to satisfy the orthogonality principle [the estimation error is uncorrelated to the data that is used to find the estimator]. In the one-dimensional case, a recursive filter that satisfies the orthogonal principle exists [this is the well known Kalman Filter]. But in the two-dimensional case, if one defines a structure for a filter that is equivalent to the one dimensional counterpart, the orthogonal principle cannot be satisfied.

2) The conclusion is that it is impossible to find an optimal recursive filter, and we have to look for an sub-optimal solution [in the two-dimensional problem].

3) In a sub-optimal solution, the way to work is to define a reasonable structure for the filter, and then to find the gain in the filter structure to minimize the mean of the square error.

3. Previous Work in Two-Dimensional Recursive Estimation

There are two earlier works in this subject [Ref. 10 and 13]. The approach in these algorithms was to assume that the orthogonality condition is satisfied for all points (i,j) where $i \leq k$, $j \leq \ell$, and then to find the gain for point k, ℓ , by induction. These algorithms are incorrect because the assumption of the induction is incorrect: the orthogonal principle cannot be satisfied for point (k,ℓ) in a recursive filter.

4. The Primary Difficulty of the Solution That is Suggested in this Thesis

The algorithm that is suggested in this thesis is not "clear" from difficulties. As in the Kalman Filter, the suboptimal gain is found by using a recursive set of equations [these equations calculate the variance of error and the gain]. But the problem was that it was impossible to find a complete "closed" set of recursive equations, and one approximation had to be done. In general, it is dangerous to make an approximation in a recursive algorithm; a small error in one step can go through an "integration process", and the solution might "blow up" to an incorrect solution. The approximation that was done in this thesis was tested carefully and was found to lead to excellent results.

B. A RECURSIVE FILTER FOR THE PROCESS $R_{xx}(n,m) = \rho_1^{|n|}\rho_2^{|m|}$

1. Introduction

First, let's summarize the details from previous chapters about this process [using the definition of system coordinates in Fig. 26].

1) The autocorrelation function is

$$R_{xx}(n,m) = \sigma_s^2\rho_1^{|n|}\rho_2^{|m|} \qquad \text{(Eq. 2-78)}$$

2) The "model" of this process:

$$x(k+1,\ell+1) = \rho_1 x(k+1,\ell) - \rho_1\rho_2 x(k,\ell) + \rho_2 x(k,\ell+1) + u(k,\ell)$$

$$\begin{bmatrix}\text{variance of} \\ \text{modeling error}\end{bmatrix} = \mathrm{var}\{u(k,\ell)\} = \sqrt{Q} = \sigma_s^2 \; (1-\rho_1^2)(1-\rho_2^2)$$

(eq's. 2-81, 2-82). $x(k+1,\ell+1)$ depends on his neighbors in region $D(k+1,\ell+1)$. [Fig. 27.]

3) Each line and column of the field has a one-dimensional model:

$$x(k+1) = \rho_2 x(k) + u_2(k) \qquad \text{for any } \ell$$

$$x(\ell+1) = \rho_1 x(\ell) + u_1(\ell) \qquad \text{for any } k$$

$$\mathrm{var}\{u_2\} = \sqrt{1 - \rho_2^2} = \sqrt{Q_2}$$

$$\mathrm{var}\{u_1\} = \sqrt{1 - \rho_1^2} = \sqrt{Q_1}$$

(eqs. 2-51, 2-53, 2-54)

2. <u>Statement of the Problem</u>

Given: a discrete, random, two dimensional process:

$$\boxed{x(k+1,\ell+1) = \rho_1 x(k+1,\ell) - \rho_1\rho_2 x(k,\ell) + \rho_2 x(k,\ell+1) + u(k,\ell)} \qquad 4\text{-}1$$

and the noisy observation, starting at $(k,\ell) = (1,1)$

$$\boxed{y(k+1,\ell+1) = cx(k+1,\ell+1) + v(k+1,\ell+1)} \qquad 4\text{-}2$$

where:

$$E\{u(k,\ell)\} = E\{v(k,\ell)\} = 0 \quad \text{for all } (k,\ell) \qquad 4\text{-}3a$$

$$E\{u(i,j)v(k,\ell)\} = 0 \quad \text{for all } i,j,k,\ell \qquad 4\text{-}3b$$

$$E\{u(k,\ell)u(i,j)\} = \begin{cases} Q & \text{if } k = i \text{ and } \ell = j \\ 0 & \text{if } k \neq i \text{ or } \ell \neq j \end{cases} \qquad 4\text{-}4$$

$$E\{x(k,\ell)v(i,j)\} = 0 \quad \text{for all } k,\ell,i,j \qquad 4\text{-}5$$

$$E\{x(k,0) = E\{x(0,\ell)\} = \bar{x} \quad \text{for all } k,\ell \qquad 4\text{-}6a$$

$$E\{[x(k,0)-\bar{x}]^2\} = E\{[x(0,\ell)-\bar{x}]^2\} = \sigma_s^2 \text{ for all } k,\ell \qquad 4\text{-}6b$$

$$E\{v(k,\ell)v(i,j)\} = \begin{cases} R & \text{if } k = i \text{ and } \ell = j \\ 0 & \text{if } k \neq i \text{ or } \ell \neq j \end{cases} \qquad 4\text{-}7$$

It is convenient to define:

$$\underset{\sim}{\rho} \triangleq (\rho_1 \quad -\rho_1\rho_2 \quad \rho_2) \qquad 4\text{-}8$$

$$\underset{\sim}{x}(k,\ell) \triangleq \begin{vmatrix} x(k+1,\ell) \\ x(k,\ell) \\ x(k,\ell+1) \end{vmatrix} \qquad 4\text{-}9$$

using 4-1, 4-9, 4-10:

$$\boxed{x(k+1,\ell+1) = \underset{\sim}{\rho}\ \underset{\sim}{x}(k,\ell) + u(k,\ell)} \qquad 4\text{-}10$$

<u>Problem</u>: Find an estimate of $x(k,\ell)$, denoted $x(k,\ell)$ which is a linear function of all observations $y(i,j)$, minimizing the quadratic-form

$$E\{x(k,\ell) - x(k,\ell)\} \qquad 4\text{-}11$$

The estimate should be done in a <u>recursive</u> filter.

3. <u>The One-Step Predictor</u>

In the one-dimensional filter, the one-step predictor is defined:

$$x(k|k-1) = x(k-1|k-1)$$

$x(k|k-1)$ means the estimate of $x(k)$ given k-1 points

$x(k-1|k-1)$ means the estimate of $x(k-1)$ given k-1 points.

Therefore, it will be reasonable to define the one step predictor in the two-dimensional case:

$$\boxed{\hat{x}^{(1)}(k+1,\ell+1) = \rho_1\hat{x}(k+1,\ell) - \rho_1\rho_2\hat{x}(k,\ell) + \rho_2\hat{x}(k,\ell-1)} \qquad 4\text{-}12$$

where:

$\hat{x}^{(1)}(k+1,\ell+1)$	is the estimate of $x(k+1,\ell+1)$ by using the data of $\Omega_1(k+1,\ell+1)$
$\hat{x}(k+1,\ell)$	is the estimate of $x(k+1,\ell)$ by using data of $\Omega_2(k+1,\ell)$
$\hat{x}(k,\ell)$	is the estimate of $x(k,\ell)$ by using data of $\Omega_2(k,\ell)$
$\hat{x}(k,\ell+1)$	is the estimate of $x(k,\ell+1)$ by using data of $\Omega_2(k,\ell+1)$

Defining:

$$\underset{\sim}{\hat{x}}(k,\ell) = \begin{vmatrix} \hat{x}(k+1,\ell) \\ \hat{x}(k,\ell) \\ \hat{x}(k,\ell+1) \end{vmatrix} \qquad 4\text{-}13$$

and by using 4-9, 4-12, 4-13:

$$\boxed{\hat{x}^{(1)}(k+1,\ell+1) = \underset{\sim}{\rho}\ \underset{\sim}{\hat{x}}(k,\ell)} \qquad 4\text{-}14$$

But there is a fundamental difference between the one dimensional and the two dimensional one-step predictor. This difference is emphasized in the next theorem.

Theorem: 1) The one-dimensional one-step-predictor is optimal if $\hat{x}(k)$ is optimal.

2) The two dimensional "one-step predictor" is not optimal, even if $\hat{x}(k+1,\ell)$, $\hat{x}(k,\ell)$, $\hat{x}(k,\ell+1)$ in equation 4-12 are optimal.

Proof: See Appendix A.

Conclusion: The recursive filter cannot be optimal.

4. Estimator

When one has the measurements of the point $(k+1,\ell+1)$ then the estimator $\hat{x}(k+1,\ell+1)$ will be:

$$\boxed{\hat{x}(k+1,\ell+1) = \hat{x}^{(1)}(k+1,\ell+1) + G(k+1,\ell+1)[y(k+1,\ell+1) - Cx^{(1)}(k+1,\ell+1)]} \quad 4\text{-}15$$

The effect of the GAIN, $G(k,\ell)$, is to weight the correction term $[y(k,\ell) - Cx(k+1,\ell+1)]$.

Remark: Because the Estimator and the one step predictor are recursive in their nature, it is easy to see that the value $x(k+1,\ell+1)$ is actually formed by the measurements in $\Omega_2(k+1,\ell+1)$ and $\hat{x}^{(1)}(k+1,\ell+1)$ is formed by the measurement in $\Omega_1(k+1,\ell+1)$.

$$\hat{x}^{(1)}(k+1,\ell+1) = \sum_{(i,j)\in\Omega_1(k+1,\ell+1)} \alpha_{i,j} y(i,j)$$

$$\hat{x}(k+1,\ell+1) = \sum_{(i,j)\in\Omega_2(k+1,\ell+1)} \beta_{i,j} y(i,j)$$

Let us define:

$$\boxed{F(k+1,\ell+1) = 1 - C \cdot G(k+1,\ell+1)} \qquad 4\text{-}16$$

Using 4-16, 4-15:

$$\boxed{\hat{x}(k+1,\ell+1) = F(k+1,\ell+1)\hat{x}^{(1)}(k+1,\ell+1) + G(k+1,\ell+1)y(k+1,\ell+1)}$$

4-17

Now, because the one step predictor cannot be optimal, the estimator is also not optimal (there does not exist a value $G(k+1,\ell+1)$ to make $\hat{x}(k+1,\ell+1)$ optimal). Of course one can minimize the variance of error of the filter that is defined in 4-12, 4-15 and get a suboptimal estimator.

5. Estimation Error

Define:

$$\left[\begin{matrix}\text{Estimation}\\ \text{Error}\end{matrix}\right] \triangleq \epsilon(k+1,\ell+1)$$

$$= \hat{x}(k+1,\ell+1) - x(k+1,\ell+1) \qquad 4\text{-}18$$

and:

$$\underset{\sim}{\mathcal{E}}(k,\ell) \triangleq \begin{vmatrix} \mathcal{E}(k+1,\ell) \\ \mathcal{E}(k,\ell) \\ \mathcal{E}(k,\ell+1) \end{vmatrix} \qquad 4\text{-}19$$

substituting 4-1, 4-2, 4-9, 4-15, 4-19 into 4-18:

$$\mathcal{E}(k+1,\ell+1) = F(k+1,\ell+1)\underset{\sim}{\rho}\underset{\sim}{\mathcal{E}}(k,\ell) - F(k+1,\ell+1)u(k,\ell) + G(k+1,\ell+1)v(k+1,\ell+1)) \qquad 4\text{-}20$$

In the situation at hand, the estimate is a random process and so is the estimation error. Next, we shall discuss the statistics of the estimation error.

6. Unbiased Estimator

Using 4-20, let us compute the mean of $\underset{\sim}{\mathcal{E}}(k+1,\ell+1)$:

$$E\{\mathcal{E}(k+1,\ell+1)\} = F(k+1,\ell+1)E\{\underset{\sim}{\rho}\underset{\sim}{\mathcal{E}}(k,\ell)\} - F(k+1,\ell+1)E\{u(k,\ell)\} + G(k+1,\ell+1)E\{v(k+1,\ell+1)\}$$

The last two terms in the above equation are zero (using 4-3a). Therefore:

$$E\{\mathcal{E}(k+1,\ell+1)\} = F(k+1,\ell+1)E\{\underset{\sim\sim}{\rho\mathcal{E}}(k,\ell)\}$$

$$= F(k+1,\ell+1)[\rho_1\overline{\mathcal{E}(k+1,\ell)} - \rho_1\rho_2\overline{\mathcal{E}(k,\ell)} + \rho_2\overline{\mathcal{E}(k,\ell-1)}]$$

Now: we noticed that the estimates of points $x(0,\ell)$, $x(k,0)$ are not based upon measurements (recall that measurements are given for $k,\ell \geq 1$). These are arbitrary choices.

Therefore let us choose:

$$\hat{x}(o,\ell) = \hat{x}(k,0) = \bar{x} = \left[\begin{matrix}\text{mean of the}\\ \text{random field}\end{matrix}\right]$$

In this case:

$$\mathcal{E}(k,0) = \bar{x} - x(k,0)$$

$$\mathcal{E}(0,\ell) = \bar{x} - x(0,\ell)$$

and using 4-6a

$$\overline{\mathcal{E}(k,0)} = \overline{\mathcal{E}(0,k)} = 0$$

Because 4-20 is a recursive equation, it is easy to see:

$$\underline{\underline{E\{\mathcal{E}(k,\ell)\} = 0}} \qquad \text{for all } k,\ell.$$

This is the case of an unbiased estimator.

7. Variance of Estimation Error

Definitions:

$P(k+1,\ell+1) \triangleq E\{\mathcal{E}^2(k+1,\ell+1)\}$	4-23
$P^{(1)}(k+1,\ell+1) \triangleq E\{[x(k+1,\ell+1) - \hat{x}^{(1)}(k+1,\ell+1)]^2\}$	4-24
$\underset{\sim}{P}(k,\ell) \triangleq E\{\underset{\sim}{\mathcal{E}}(k,\ell)\,\underset{\sim}{\mathcal{E}}^T(k,\ell)\}$	4-25
$P_{i,j}(k,\ell) \triangleq E\{\mathcal{E}(k,\ell)\,\mathcal{E}(k+i,\ell+j)\}$	4-26

One has to be careful: 4-23 is a definition of a <u>scalar</u> $P(k+1,\ell+1)$. 4-25 is a definition of a <u>matrix</u> $\underset{\sim}{P}(k+1,\ell+1)$. The sign $\underset{\sim}{(\cdot)}^T(k,\ell)$ is the transpose of $\underset{\sim}{(\cdot)}(k,\ell)$. $P(k+1,\ell+1)$ is the variance of the estimation error at point $(k+1,\ell+1)$. $P^{(1)}(k+1,\ell+1)$ is the variance of the "one step predictor" error at point $(k+1,\ell+1)$. $\underset{\sim}{P}(k,\ell)$ is a 3x3 symmetric matrix. Three of its values are the variances of the estimation errors of region $D(k+1,\ell+1)$. Three other values of the matrix are correlations between the estimation erros of points in region $D(k+1,\ell+1)$. For better understanding see also Eq. 4-27. $P_{i,j}(k,\ell)$ is the correlation between the estimation errors of points (k,ℓ) and $(k+i,\ell+j)$.

Remark: The definition in Eq. 4-23 is a special case of Eq. 4-26;

$$P(k,\ell) = P_{0,0}(k,\ell).$$

8. Variance Calculations

Using the definitions in 4-23, 4-24, 4-25, 4-26, one can derive recursive equations as follows:

$\underset{\sim}{P}(k,\ell)$ calculation

$$\underset{\sim}{P}(k,\ell) = E\{\underset{\sim}{\mathcal{E}}(k,\ell)\,\underset{\sim}{\mathcal{E}}^T(k,\ell)\}$$

$$= E\begin{Bmatrix} \mathcal{E}(k+1,\ell)\mathcal{E}(k+1,\ell) & \mathcal{E}(k+1,\ell)\mathcal{E}(k,\ell) & \mathcal{E}(k+1,\ell)\mathcal{E}(k,\ell+1) \\ \mathcal{E}(k,\ell)\mathcal{E}(k+1,\ell) & \mathcal{E}(k,\ell)\mathcal{E}(k,\ell) & \mathcal{E}(k,\ell)\mathcal{E}(k,\ell+1) \\ \mathcal{E}(k,\ell+1)\mathcal{E}(k+1,\ell) & \mathcal{E}(k,\ell+1)\mathcal{E}(k,\ell) & \mathcal{E}(k,\ell+1)\mathcal{E}(k,\ell+1) \end{Bmatrix}$$

using 4-26, 4-23:

$$\boxed{\underset{\sim}{P}(k,\ell) = \begin{vmatrix} P(k+1,\ell) & P_{1,0}(k,\ell) & P_{1,-1}(k,\ell+1) \\ P_{1,0}(k,\ell) & P(k,\ell) & P_{0,1}(k,\ell) \\ P_{1,-1}(k,\ell+1) & P_{0,1}(k,\ell) & P(k,\ell+1) \end{vmatrix}} \qquad \text{4-27}$$

$P^{(1)}(k+1,\ell+1)$ calculation

$$P^{(1)}(k+1,\ell+1) = E\{[\hat{x}^{(1)}(k+1,\ell+1) - x(k+1,\ell+1)]^2\}$$

$$= E\{[\underset{\sim}{\rho}\underset{\sim}{\hat{x}}(k,\ell) - \underset{\sim}{\rho}\underset{\sim}{x}(k,\ell) - u(k,\ell)]^2\}$$

Because the term in the rectangular brackets is a scalar we can rewrite:

$$P^{(1)}(k+1,\ell+1) = E\{[\underset{\sim}{\rho}\underset{\sim}{\mathcal{E}}(k,\ell) - u(k,\ell)][\underset{\sim}{\rho}\underset{\sim}{\mathcal{E}}(k,\ell) - u(k,\ell)]^T\}$$

$$= E\{\underset{\sim}{\rho}\underset{\sim}{\mathcal{E}}(k,\ell)\underset{\sim}{\mathcal{E}}^T(k,\ell)\underset{\sim}{\rho}^T - u(k,\ell)\underset{\sim}{\mathcal{E}}^T\underset{\sim}{\rho}^T - \underset{\sim}{\rho}\underset{\sim}{\mathcal{E}}(k,\ell)u(k,\ell) + u^2(k,\ell)\}$$

The expectation of the two middle terms in the last expression is zero [see Appendix B]. Therefore,

$$P^{(1)}(k+1,\ell+1) = \underset{\sim}{\rho}E\{\underset{\sim}{\mathcal{E}}(k,\ell)\underset{\sim}{\mathcal{E}}^T(k,\ell)\}\underset{\sim}{\rho}^T + E\{u^2(k,\ell)\}$$

By using 4-4, 4-25:

$$\boxed{P^{(1)}(k+1,\ell+1) = \underset{\sim}{\rho}\underset{\sim}{P}(k,\ell)\underset{\sim}{\rho}^T + Q} \qquad 4\text{-}28$$

$P(k+1,\ell+1)$ calculation

By substituting 4-20 in 4-23 and noting that $\mathcal{E}^T(k+1,\ell+1) = \mathcal{E}(k+1,\ell+1)$, since it is a scalar:

$$P(k+1,\ell+1) = E\{\varepsilon^2(k+1,\ell+1)\}$$

$$= E\{[F(k+1,\ell+1)\underset{\sim\sim}{\rho}\mathcal{E}(k,\ell) - F(k+1,\ell+1)u(k,\ell) + G(k+1,\ell+1)v(k+1,\ell+1)]$$

$$\cdot [F(k+1,\ell+1)\underset{\sim\sim}{\rho}\mathcal{E}(k,\ell) - F(k+1,\ell+1)u(k,\ell) + G(k+1,\ell+1)v(k+1,\ell+1)]^T\}$$

The last equation includes the three next terms that are zero:

$$E\{\mathcal{E}(k,\ell)v(k+1,\ell+1)\} = 0 \qquad \text{(see Appendix B)}$$

$$E\{\mathcal{E}(k,\ell)u(k,\ell)\} = 0 \qquad \text{(see Appendix B)}$$

$$E\{u(k,\ell)v(k+1,\ell+1)\} = 0 \qquad \text{(by using 3-3b)}$$

Therefore:

$$P(k+1,\ell+1) = F^2(k+1,\ell+1)\underset{\sim}{\rho}\,\overline{\underset{\sim}{\mathcal{E}}(k,\ell)\underset{\sim}{\mathcal{E}}^T(k,\ell)}\,\underset{\sim}{\rho} + F^2(k+1,\ell+1)\overline{u^2(k,\ell)}$$

$$+ G^2(k+1,\ell+1)\overline{v^2(k+1.\ell+1)}$$

and by using 4-4, 4-25, 4-7, 4-28:

$$\boxed{P(k+1,\ell+1) = F^2(k+1,\ell+1)\; P^{(1)}(k+1,\ell+1) + G^2(k+1,\ell+1)\cdot R} \tag{4-29}$$

$\underline{P_{0,1}(k,\ell),\ P_{1,0}(k,\ell)\ \text{calculation}}$

From 4-26:

$$P_{i+1,j+1}(k,\ell) \triangleq E\{\mathcal{E}(k,\ell)\mathcal{E}(k+i+1,\ell+j+1)$$

substituting $\mathcal{E}(k+i,\ell+j)$ from 4-20:

$$P_{i+1,j+1}(k,\ell) = E\{\mathcal{E}(k,\ell)F(k+i+1,\ell+j+1)\underset{\sim}{\rho}\mathcal{E}(k+i,\ell+j)$$

$$+ E\{\mathcal{E}(k,\ell)G(k+i+1,\ell+j+1)v(k+i+1,\ell+j+1)$$

$$- E\{\mathcal{E}(k,\ell)F(k+i+1,\ell+j+1)u(k+i,\ell+j)\}$$

The last two terms are zero [see Appendix B]. Now, by using 4-26, 4-20, 4-19 and 4-9, the last term becomes:

$$P_{i+1,j+1}(k,\ell) = F(k+i+1,\ell+j+1)[\rho_1 P_{i+i,j}(k,\ell) - \rho_1\rho_2 P_{i,j}(k,\ell)$$

$$+ \rho_2 P_{i,j+1}(k,\ell)]$$

Substituting in the last equation the obvious identities:

$$P_{-i,-j}(k,\ell) = P_{i,j}(k-i,\ell-j)$$

$$P_{-i,j}(k,\ell) = P_{i,-j}(k-i,\ell+j)$$

we arrive at:

$$\boxed{P_{0,1}(k,\ell) = F(k,\ell+1)[\rho_1 P_{0,0}(k,\ell) - \rho_1\rho_2 P_{1,0}(k-1,\ell) + \rho_2 P_{1,-1}(k-1,\ell+1)]} \tag{4-30}$$

$$\boxed{P_{1,0}(k,\ell) = F(k+1,\ell)[\rho_1 P_{1,-1}(k,\ell) - \rho_1\rho_2 P_{0,1}(k,\ell+1) + \rho_2 P_{0,0}(k,\ell)]} \tag{4-31}$$

calculation of $P_{-1,1}(k,\ell)$

It is not possible to find a recursive equation of $P_{-1,1}(k,\ell)$ that includes just previous values of $P(i,j)$, $P_{0,1}(i,j)$, $P_{1,0}(i,j)$. Trying to evaluate $P_{-1,1}(k,\ell)$ leads to expressing $P_{-1,1}(k,\ell)$ in terms of correlation of points farther removed from the region $D(k,\ell)$. This problem is solved by an approximation. But first let's see some special cases:

a. $P_{1,-1}(0,1)$

$$P_{1,-1}(0,1) = \overline{\mathcal{E}(0,1)\cdot\mathcal{E}(1,0)}$$

$$= \overline{[\hat{x}(0,1) - x(0,1)][\hat{x}(1,0) - x(0,1)]}$$

and using 2-78

$$\underline{P_{1,-1}(0,1) = P_{-1,1}(1,0) = \sigma_s^2 \rho_1\rho_2}$$

b. $\underline{P_{1,-1}(0,\ell)}$

$$P_{1,-1}(0,\ell) = E\{\mathcal{E}(1, -1)\mathcal{E}(0,\ell)$$

$$= E\{\mathcal{E}(1,\ell-1)[x(0,\ell) - \hat{x}(0,\ell)]\} \qquad 4\text{-}33$$

Using 2-53, 2-54, and 4-21:

$$P_{1,-1}(0,\ell) = E\{\mathcal{E}(1,\ell-1)[\rho_1 x(0,\ell-1) + u_1(0,\ell-1) - \overline{x}]$$

Now, using the same procedure as in Appendix B it can be shown:

$$E\{\mathcal{E}(1,\ell-1)u_1(0,\ell-1)\} = 0$$

and using Eq. 4-22:

$$E\{\mathcal{E}(1,\ell-1) - \overline{x}\} = 0$$

Therefore:

$$P_{1,-1}(0,\ell) = E\{\mathcal{E}(1,\ell-1)\ \rho_1 x(0,\ell-1)\}$$

$$= E\{\mathcal{E}(1,\ell-1)\mathcal{E}(0,\ell-1)\rho_1\} \qquad 4\text{-}34$$

using 4-26:

$$P_{1,-1}(0,\ell) = \rho_1 P_{1,0}(0,\ell-1) = \frac{P_{1,0}(0,\ell-1)P_{0,1}(0,\ell-1)}{P(0,\ell-1)} \qquad 4\text{-}35$$

c. $P_{1,-1}(1,k)$

In the same procedure:

$$P_{1,-1}(,1) = \rho_2 P_{0,1}(k,0) = \frac{P_{1,0}(k,0)\ P_{0,1}(k,0)}{P(k,0)} \qquad 4\text{-}36$$

d. In The General Case

Comparing Eq. 4-1 and 4-20 one can see that the error is a random field that has the same nature as the original field $x(k,\ell)$. Therefore one can assume that the autocorrelation function of the error has a similar form to 4-1.

$$R_{ee}(n,m) = \sigma_e^{2} . \rho_1^{|n|} . \rho_2^{|m|} \qquad 4\text{-}37$$

This autocorrelation function is correct only in the steady-state, when $F(k,\ell)$ is constant. Therefore Eq. 4-37 is an approximation that is correct in steady state.

From 4-37 it is seen that in steady state:

$$P_{1,0}(k,\ell) = P(k,\ell) \cdot \rho_2 \qquad 4\text{-}38a$$

$$P_{0,1}(k,\ell) = P(k,\ell)\cdot\rho_1 \qquad \text{4-38b}$$

$$P_{1,-1}(k,\ell-1) = P(k,\ell)\cdot\rho_1\rho_2 \qquad \text{4-38c}$$

From 4028a,b,c:

$$\boxed{P_{1,-1}(k,\ell+1) = \frac{P_{1,0}(k,\ell)\cdot P_{0,1}(k,\ell)}{P(k,\ell)}} \qquad \text{4-39}$$

Note that this equation is accurate for first line and colume (e.g. 4-35, 4-36).

9. Minimum Variance Condition

We have found the condition for an unbiased estimator. This condition doesn't require any constraint upon the gain $G(k,\ell)$. Thus we wish to minimize the value of $P(k+1,\ell+1)$, with respect to $G(k+1,\ell+1)$. Differentiating the expression for $P(k+1,\ell+1)$, set the result to zero and solve for $G(k+1,\ell+1)$. From 4-29:

$$0 = \frac{\partial}{\partial G(k,\ell)} P(k,\ell)$$

$$= \frac{\partial}{\partial G(k,\ell)} \{[1 - G(k,\ell)C]^2\, P^{(1)}(k,\ell) + G^2(k,\ell)\, R\}$$

$$\boxed{G(k,\ell) = P^{(1)}(k,\ell)C[C^2P^{(1)}(k,\ell) + R]^{-1}} \qquad \text{4-40}$$

And substituting 4-40 into 4-29:

$$\boxed{P(k,\ell) = (1 - G(k,\ell)\ P^{(1)}(k,\ell)} \quad 4\text{-}41$$

10. <u>Initial Condition</u>

To start the recursive process of Gain calculation, we need the next initial conditions:

1. $\hat{x}(0,\ell)$ $\quad 0 \leq \ell \leq N$

2. $\hat{x}(k,\ell)$ $\quad 0 \leq k \leq N$

and then we are able to calculate:

3. $P(0,\ell)$ $\quad 0 \leq \ell \leq N$

4. $P(k,0)$ $\quad 0 \leq k \leq N$

5. $P_{0,1}(0,\ell)$ $\quad 0 \leq \ell \leq N-1$

6. $P_{1,0}(k,0)$ $\quad 0 \leq k \leq N-1$

7. $P_{1,-1}(0,1)$

Fig. 30 shows what is meant by "initial condition."

For an unbiased estimator:

$$\hat{x}(0,\ell) = \hat{x}(k,0) = \overline{x}$$

By using 4-16, 4-21, 4-23:

$$P(0,\ell) = E\{\varepsilon^2(0,\ell)\} = E\{(\hat{x}(0,\ell) - x(0,\ell))^2\}$$

$$= E\{[\bar{x} - x(0,\ell)]^2$$

by using 4-6b:

$$\underline{\underline{P(0,\ell) = \sigma_s^2}}$$

and similarly for P(k, 0):

$$\underline{\underline{P(k,0) = \sigma_s^2}} \qquad 4\text{-}43$$

For $P_{0,1}(0,\ell)$:

$$P_{0,1}(0,\ell) = E\{\varepsilon(0,\ell)\varepsilon(0,\ell+1)$$

$$= E\{(\hat{x}(0,\ell) - x(0,\ell))(\hat{x}(0,\ell+1) - x(0,\ell+1))$$

$$= E\{(\bar{x} - x(0,\ell))(\bar{x} - x(0,\ell))$$

$$\underline{\underline{P_{0,1}(0,\ell) = \rho_2\,\sigma_s^2}} \qquad 4\text{-}44$$

and in the same form:

$$\underline{\underline{P_{1,0}(k,0) = \rho_1 \sigma_s^2}} \qquad 4\text{-}45$$

$$\underline{\underline{P_{1,-1}(0,1) = \sigma_s^2 \rho_1\rho_2}} \qquad 4\text{-}46$$

11. Summary

The proposed filter for the process that is defined in 4-1, 4-2 is given in 4-12 and 4-15. The recursive equation for GAIN calculations are

$$G(k,\ell) = P^{(1)}(k,\ell)C[C^2P^{(1)}(k,\ell) + R]^{-1}$$

$$P^{(1)}(k,\ell) = \underset{\sim}{\rho}\ \underset{\sim}{P}(k-1,\ell-1)\ \underset{\sim}{\rho}^T + Q$$

$$P(k,\ell) = (1 - G(k,\ell))\ P^{(1)}(k,\ell)$$

where:

$$\underset{\sim}{P}(k,\ell) = \begin{pmatrix} P(k+1,\ell) & P_{1,0}(k,\ell) & P_{1,-1}(k,\ell+1) \\ P_{1,0}(k,\ell) & P(k,\ell) & P_{0,1}(k,\ell) \\ P_{1,-1}(k,\ell+1) & P_{0,1}(k,\ell) & P(k,\ell+1) \end{pmatrix}$$

$$\underset{\sim}{\rho} = (\rho_1 \quad -\rho_1\rho_2 \quad \rho_2)$$

and:

$$P_{1,0}(k-1,\ell) = F(k,\ell)[\rho_1 P_{1,-1}(k-1,\ell) - \rho_1\rho_2 P_{0,1}(k-1,\ell-1) + \rho_2 P(k-1,\ell)]$$

$$P_{0,1}(k,\ell-1) = F(k,\ell)[\rho_1 P(k,\ell-1) - \rho_1\rho_2 P_{1,0}(k-1,\ell-1) + \rho_2 P_{1,-1}(k-1),\ell)]$$

$$P_{1,-1}(k,\ell+1) = \frac{P_{1,0}(k,\ell)\ P_{0,1}(k,\ell)}{P(k,\ell)}$$

Actually it is impossible to find a complete set of recursive equations. The equation for $P_{1,-1}(k,\ell)$ is only an approximation for the case where $\ell \geq 2$ and $k \geq 1$. Therefore we have an algorithm that is correct for the first line and column and converges to a correct suboptimal solution in steady state.

Fig. 35 helps to understand the procedure of the recursive calculation. Assume, we have finished the calculations for point $(k+1,\ell)$ and we wish to calculate the gain $G(k+1,\ell+1)$. In order to do this we need to know $P^{(1)}(k+1,\ell+1)$ [variance of one step predictor]. Now:

- The value $P^{(1)}(k+1,\ell+1)$ depends on six values that are seen on the boundaries of triangle I in Fig. 29:

$$P(k,\ell+1),\ P(k,\ell),\ P(k+1,\ell),\ P_{0,1}(k,\ell),\ P_{1,0}(k,\ell),\ P_{1,-1}(k,\ell+1)$$

- The value $P_{0,1}(k,\ell)$ depends on three values that create the boundaries of triangle II in Fig. 29:

$$P_{0,1}(k,\ell-1),\ P(k,\ell),\ P_{1,-1}(k,\ell)$$

- The value $P_{1,0}(k,\ell)$ depends on the three values that create the boundaries of triangle III in Fig. 29:

$$P_{1,-1}(k,\ell),\ P_{0,1}(k,\ell-1),\ P_{0,0}(k,\ell).$$

The value of $P_{1,-1}(k,\ell+1)$ is given in Eq. 4-39. As one moves along a line, this structure is also moving.

11. Results

Figures 36, 37, 38 show typical results. Steady state is reached after 4-5 points. The steady state for the first line and column is higher than for points in the middle of the field. The reason for this fact is that points in the first line are estimated only by their neighbors to the left. It was found that the results for the first line and column are exactly the same as one finds from the one-dimensional Kalman Filter, using the model in Eq. 2-53, 2-54.

12. Checking Of The Results

Two things must be checked: First, it should be checked if the algorithm developed above is correct. Recall that it was impossible to develop a complete accurate algorithm and an approximation had to be used. In any recursive calculation an approximation can "blow up" the results. Therefore it is very important to check if this is

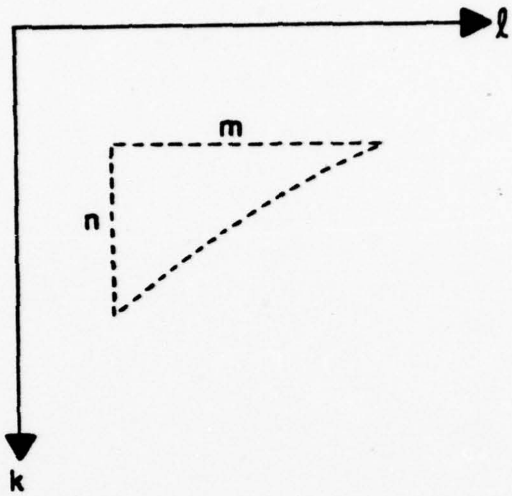


Fig. 31: Definition of System Coordinates

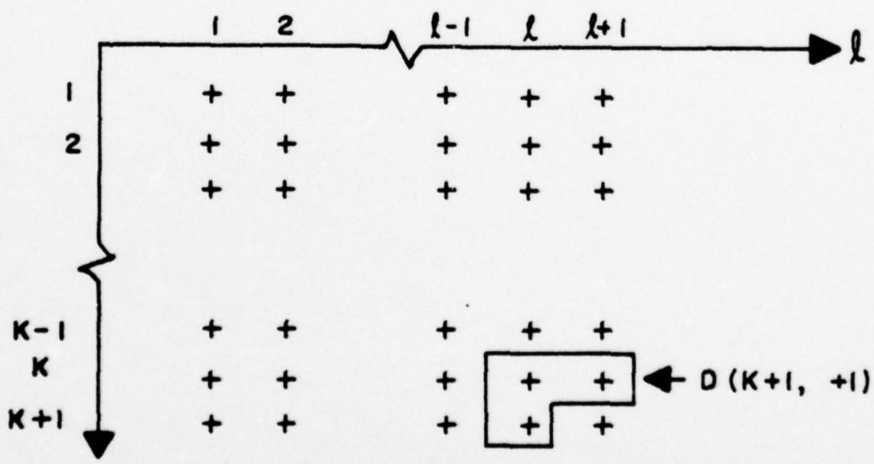


Fig. 32: Definition of Region D(k+1,ℓ+1)

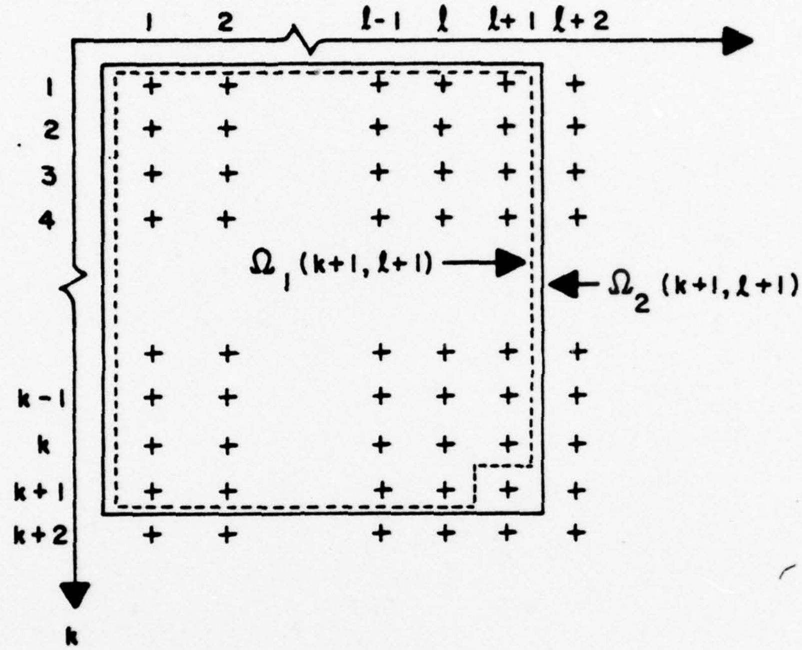


Fig. 33: Definition of Regions $\Omega_1(k+1,\ell+1),\Omega_2(k+1,\ell+1)$

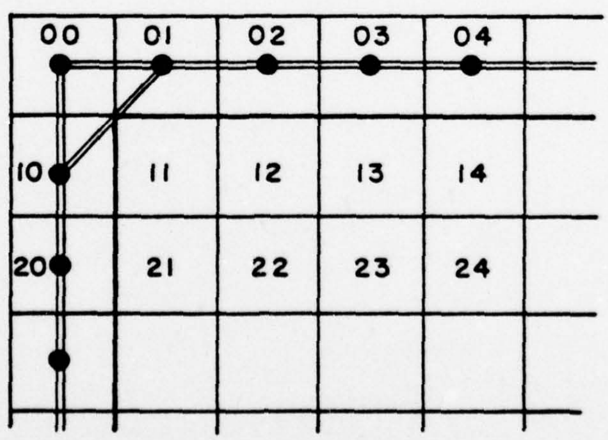
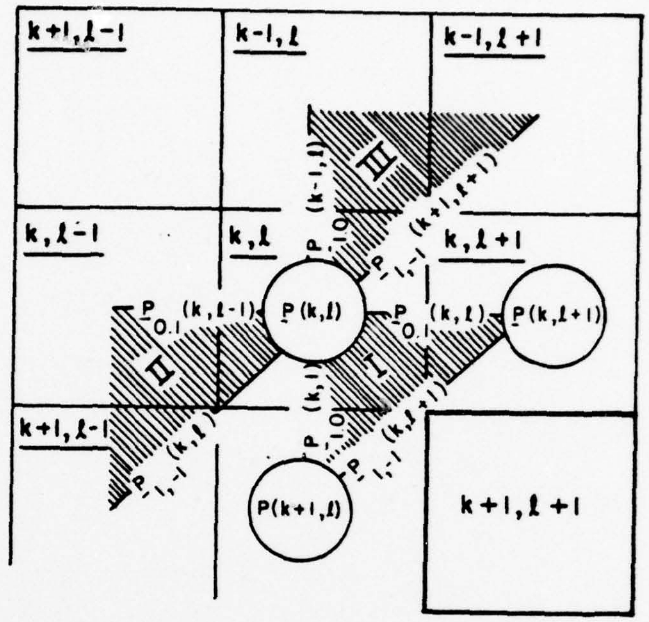
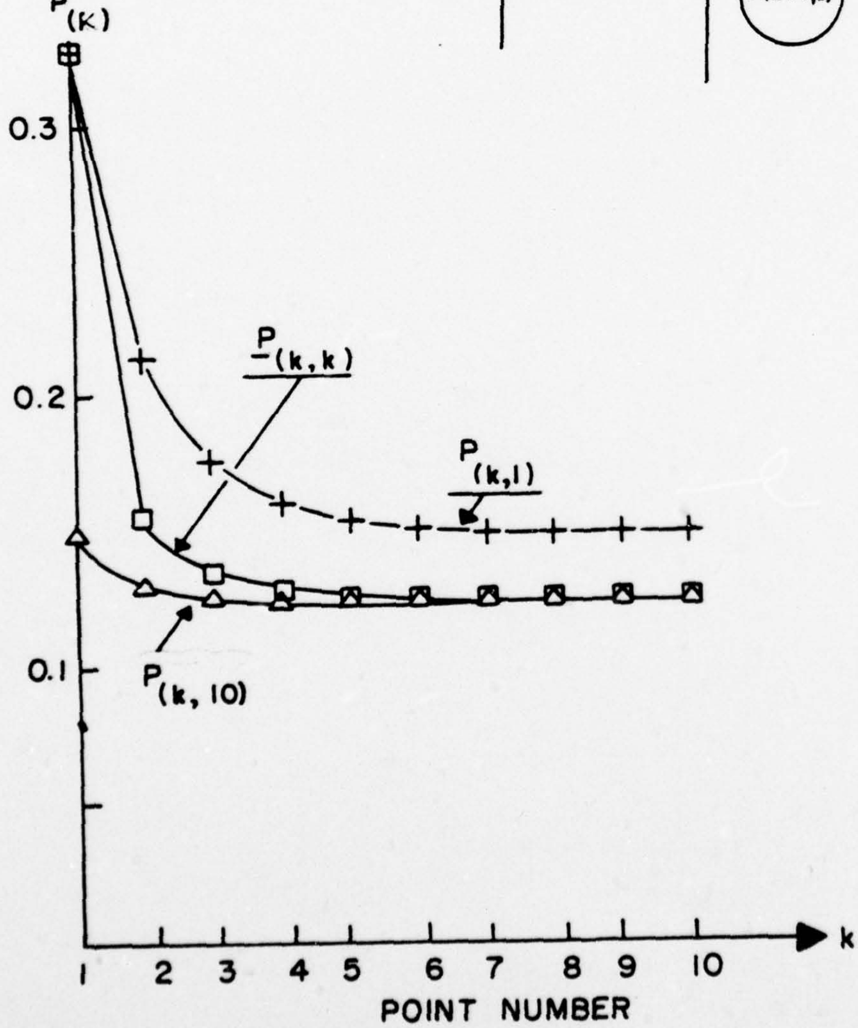


Fig. 34: Initial Conditions. A black point in a cell (i,j) represents the initial value $\hat{x}(i,j)$ and the value $P(i,j)$. A line between two black points represents the correlation between the estimation error of the two pixels: $P_{i,j}(k,\ell)$.

(Right) Fig. 35: The values that are needed to calculate G(k+1,ℓ+1)





(Left) Fig. 36: Estimation Error for

$\rho_1 = \rho_2 = 0.96$

$R = 0.5$

$\sigma^2 = 1$

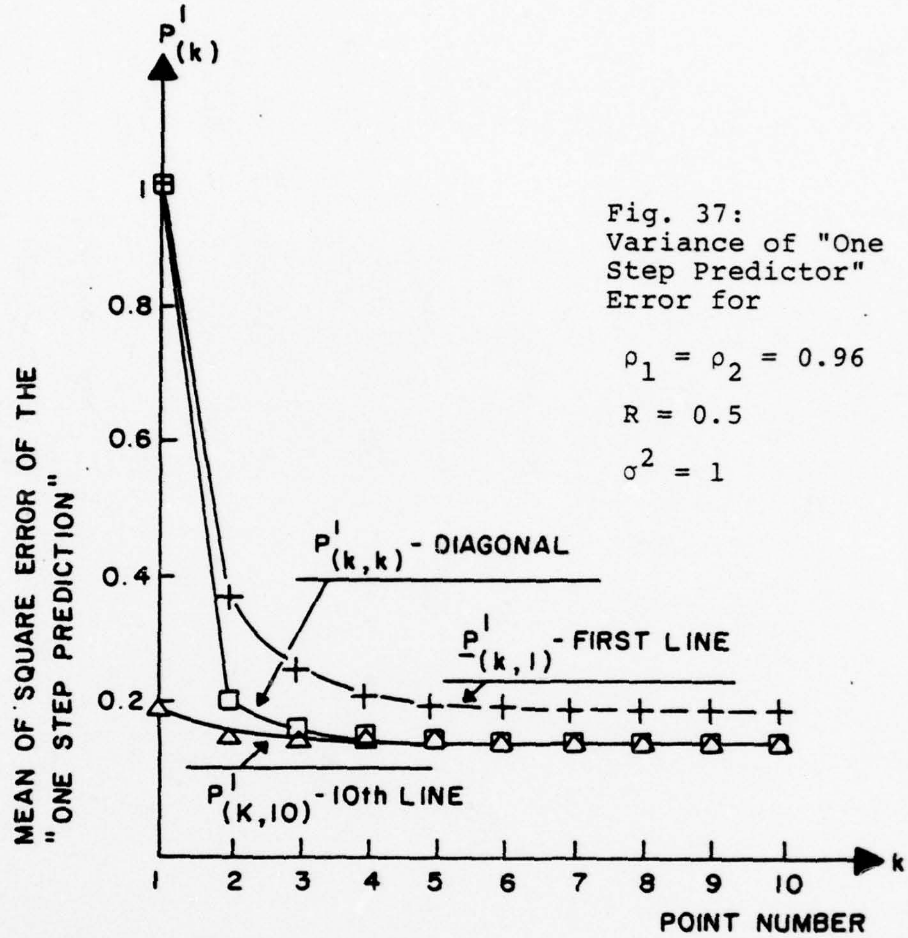


Fig. 37:
Variance of "One Step Predictor" Error for

$\rho_1 = \rho_2 = 0.96$

$R = 0.5$

$\sigma^2 = 1$

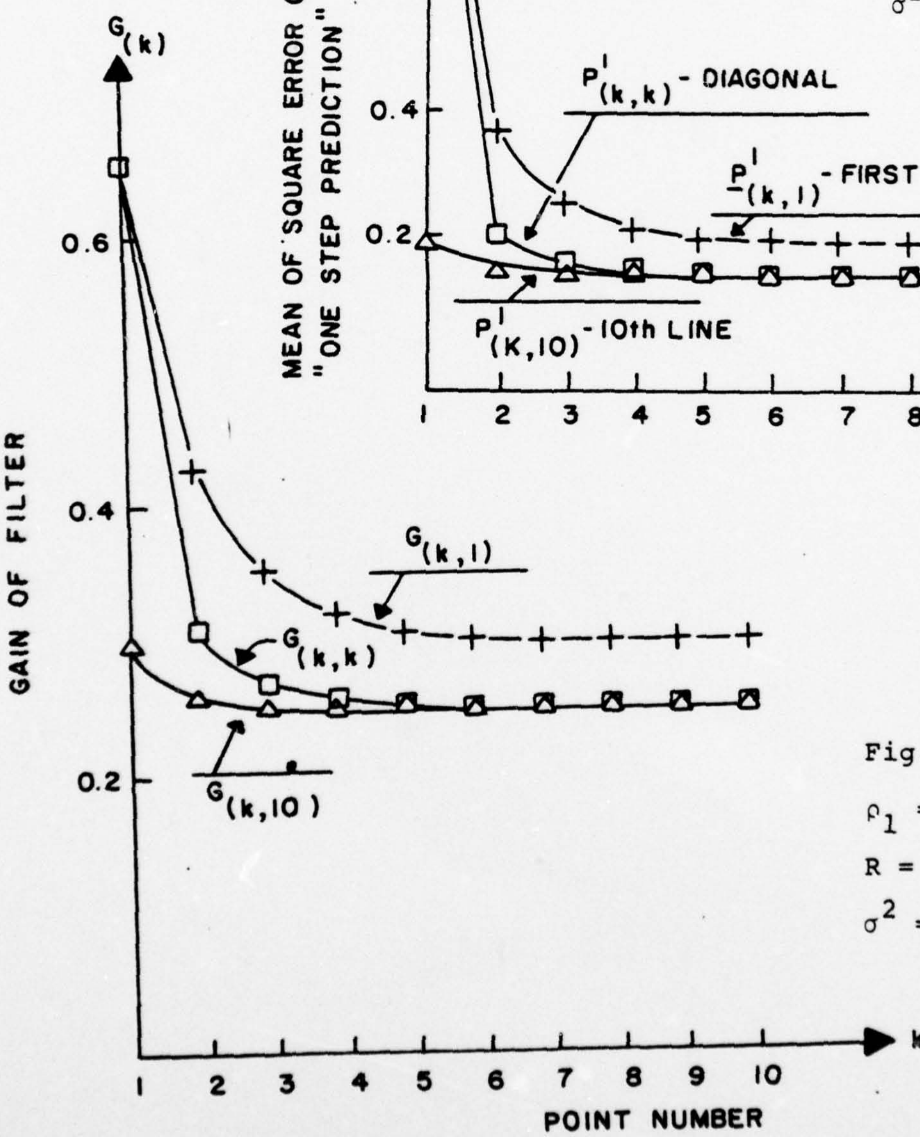


Fig. 38: Gain for

$\rho_1 = \rho_2 = 0.96$

$R = 0.5$

$\sigma^2 = 1$

not the case here. It was found that the results for the first line and column are the same as the one dimensional Kalman filter. Now, we want to check the steady state, after MxM points. In order to solve this problem, the next procedure was done:

1) An image, of size 128 x 128 with autocorrelation function

$$R_{xx}(n,m) = \rho_1^{|n|}\rho_2^{|m|}$$

was created.

2) White noise with variance $\sqrt{R}$ was added to the correlated image.

3) The filter developed above was used in order to estimate the correlated process. The gain of the filter was found by substituting ρ_1, ρ_2, R, into the recursive algorithm of gain calculation.

4) The variance of error of the simulation was compared with the theoretical variance of error.

This procedure is described in Fig. 39.

The basic assumption here is that if the simulated variance of error is similar to the theoretical variance of error, then the algorithm is correct. A mathematical justification for this assumption was not proved. Still one can ask if the mean square error is minimum. But the theoretical variance of error is an inherent result of gain calculation, and part of the algorithm. Therefore if it is equal to the simulated result, it is reasonable to assume that the algorithm is correct.

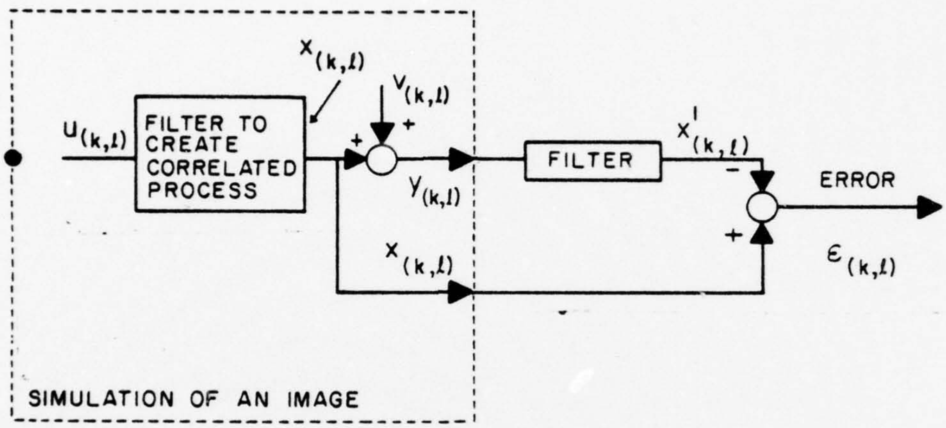


Fig. 39: Experiment to Check the Algorithm for Estimation by Using a Simulated Image

The second question arises from the knowledge that the two-dimensional filter is not optimal. Therefore it is important to compare the optimal non-recursive estimation error to the sub-optimal recursive filter.

Results:

Fig. 40 shows three types of results:

- optimal non-recursive estimation error (in steady state).

- theoretical steady state variance of error of the recursive filter.

- experimental steady state variance of error of the recursive filter.

This experiment was carried out for $\rho_1 = \rho_2 = 0.96$, $\sigma^2 = 1$. The variance of random noise varies from 0.1 to 0.8.

It was found excellent coincidence (less than 5%) between the theoretical estimation error and the simulated error. The recursive filter shows variance of error not bigger than 2% than the optimal non-recursive.

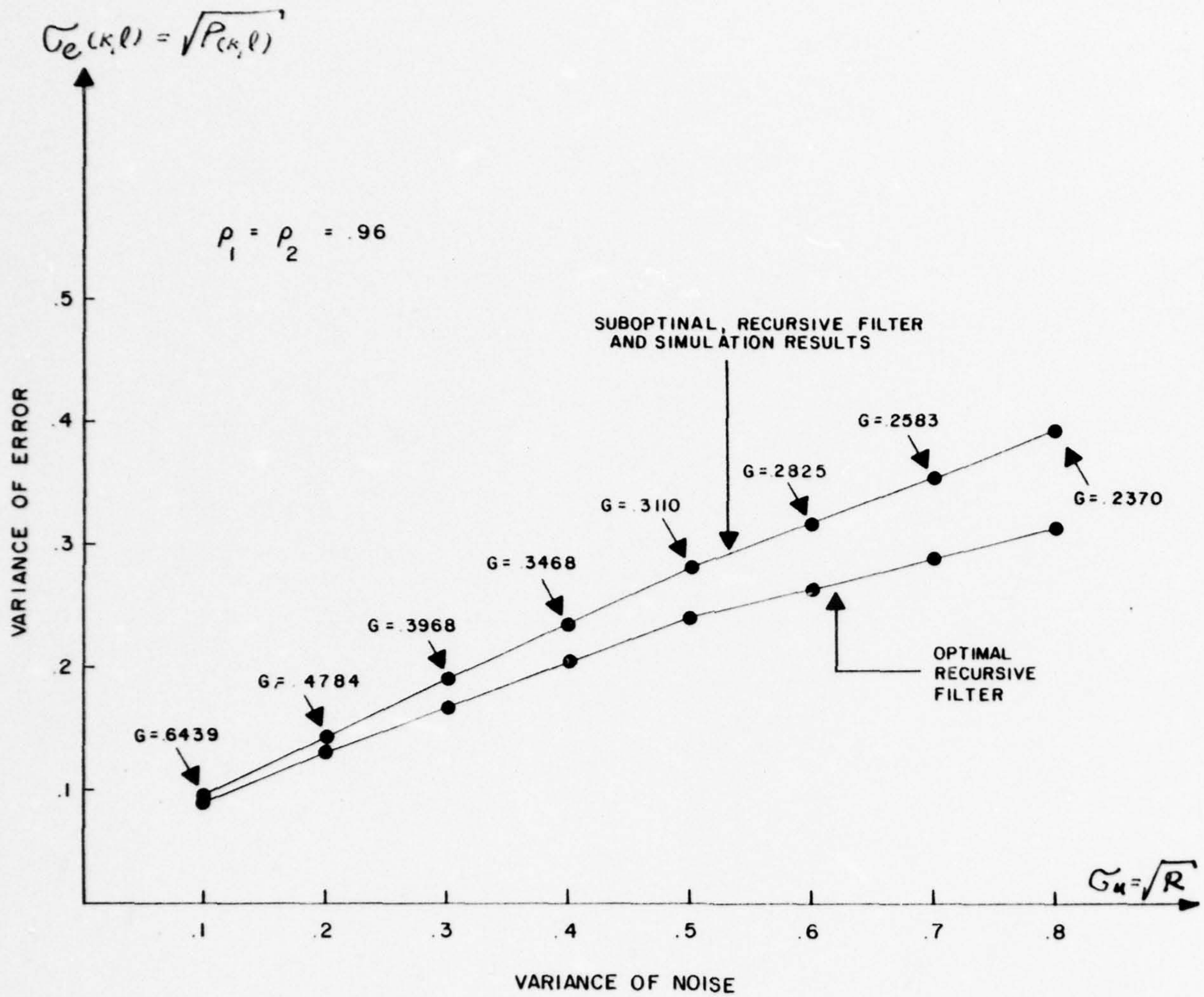


Fig. 40: Steady State Variance of Error for:
- theoretical recursive filter
- experimental results using the recursive filter
- optimal, non-recursive estimation (theoretical)

V. APPLICATIONS: DETECTION OF TARGETS IN PRESENCE OF CORRELATED NOISE

A. INTRODUCTION

Assume a problem of detecting an infrared target. In this case one has interest in the location of the target and the intensity of the target. In infrared detection the intensity has a special importance, because the target intensity is proportional to the temperature, and by knowing the intensity one can conclude whether the target is a missile, or an aircraft, etc. The target detection can be interfered by two types of noise:

1. Correlated Noise. An example for correlated noise is clouds.

2. White Noise. An example for white noise is measurement noise.

Now, the problem is to find the location and intensity of that target.

B. STATEMENT OF THE PROBLEM

In this chapter the next problem is solved:

Given: A set of measurements of a two dimensional field, $y(k,\ell)$. These measurements are composed of three types of signals:

1) A target that has an intensity function $T(k,\ell)$.

2) The target is corrupted in a correlated background $x(k,\ell)$. The autocorrelation function is, for example:

$$R_{xx}(n,m) = \sigma^2 \rho_1^{|n|} \rho_2^{|m|} \qquad 5\text{-}1$$

3) White noise, $v(k,\ell)$, that is uncorrelated to the background or target, so that:

$$y(k,\ell) = x(k,\ell) + T(k,\ell) + v(k,\ell) \qquad 5\text{-}2$$

Find:

1) The location of the target.

2) The intensity of the target.

C. BASIC CONCEPT

Assuming the autocorrelation of the background is known, one can build an estimator for the background [see Chpt. IV]. The filter is applied with steady state gain. We use the "one step predictor" and compare it with the actual measurement.

After subtracting the estimated background from the measured image, the residual image includes mainly the noise and the target. If the target's intensity is higher enough from the noise, the detection can be done in the residual image.

A further improvement can be achieved if one knows the shape of the target. In that case classical methods like matched filters will increase the processing gain. Figure 41 shows the basic concept.

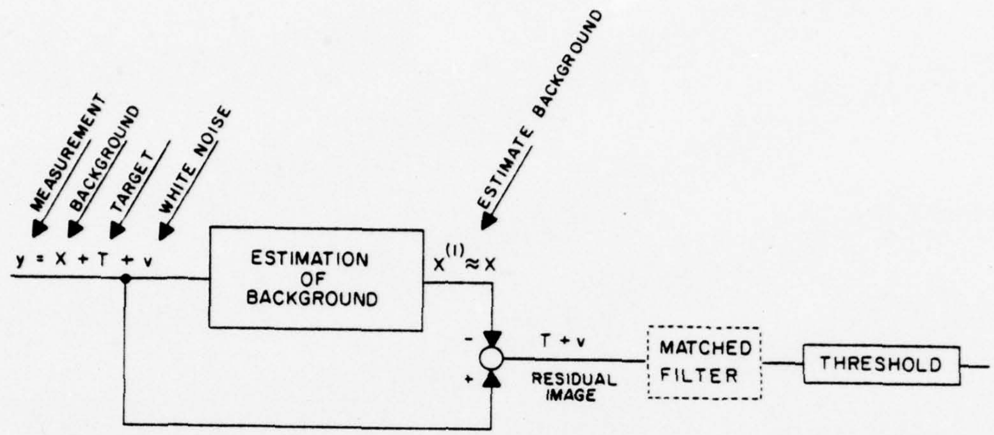


Fig. 41: Basic Concept of Target Detection

D. DETECTION OF A POINT TARGET

1. The Two Hypotheses

From Fig. 41 it is seen that the "residual image" is composed of white noise and the target. As a consequence one can see that a point target is detectable with a low probability-of-error only if its intensity is significantly higher from the white noise, say three times more than the variance of the noise.

The point target detection is done as follows: When arriving at a new point, say $(k+1, \ell+1)$, during the scanning process, the situation at hand is that there are two estimators of the same point:

One - is the "one step predictor", $\hat{x}^{(1)}(k,\ell)$, that does not include the measurement at the point.

The Second is the measurement itself, $y(k+1,\ell+1)$. Now, if $y(k+1,\ell+1)$ is significantly higher than the "one step predictor", it is most likely due to a target. The target detection is, therefore, a result of comparing the measurement at a point to the prediction of the value at that point, by using all of the neighbors near the point. It is clear that this difference can be due to the measurement noise $v(k,\ell)$. Therefore a reasonably good detection will be if the target is three times higher than the variance of the noise.

Now, the procedure will be explained mathematically: Assume a specific problem:

1) The autocorrelation function of the background:

$$R_{xx}(n,m) = \sigma_s^2 \rho_1^{|n|} \rho_2^{|m|}$$

2) The measurement noise has a Gaussian, white noise statistic:

$$R_{vv}(n,m) = \begin{cases} R & \text{if } n = m = 0 \\ 0 & \text{if } n \neq 0 \text{ or } m \neq 0 \end{cases}$$

3) The target has a Gaussian probability density function with mean T_o and variance $\sqrt{R_T}$.

Note: It is more reasonable to assume a Rayleigh distribution. The Gaussian assumption is done because it simplifies the discussion.

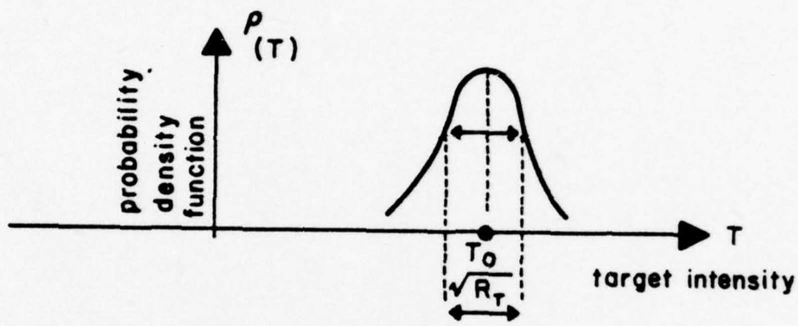


Fig. 42: Probability Density Function Of the Target

The estimation of the background is done by using the filter that is described in Chpt. Iv. Also assume that the procedure of estimation/detection is finished for point $(k+1,\ell)$ and now it is repeated for point $(k+1,\ell+1)$.

Recall:

$\hat{x}^{(1)}(k+1,\ell+1)$ = one step predictor.

$\hat{x}(k+1,\ell+1)$ = estimator of point $(k+1,\ell+1)$.

using Eq. 4-15, assuming C = 1:

$$\hat{x}(k+1,\ell+1) = \hat{x}^{(1)}(k+1,\ell+1) + G(k+1,\ell+1)[y(k+1,\ell+1) - x^{(1)}(k+1,\ell+1)]$$

$$= (1 - G(k+1,\ell+1))x^{(1)}(k+1,\ell+1) + G(k+1,\ell+1)y(k+1,\ell+1)$$

$$= F(k+1,\ell+1)x^{(1)}(k+1,\ell+1) + G(k+1,\ell+1)y(k+1,\ell+1)$$

The residual image:

$$r^{(1)}(k+1,\ell+1) = \hat{x}(k+1,\ell+1) - y(k+1,\ell+1)$$

$$= F(k+1,\ell+1)\ (x^{(1)}(k+1,\ell+1) - y(k+1,\ell+1)) \qquad 5\text{-}1$$

<u>Define</u>:

1. $$r(k+1,\ell+1) \triangleq \frac{r^{(1)}(k+1,\ell+1)}{F(k+1,\ell+1)} \qquad 5\text{-}2$$

$$= x^{(1)}(k+1,\ell+1) - y(k+1,\ell+1)$$

2. $$x^{(1)}(k+1,\ell+1) = x(k+1,\ell+1) + \epsilon^{(1)}(k+1,\ell+1) \qquad 5\text{-}3$$

$\epsilon^{(1)}(k+1,\ell+1)$ = (error of one step predictor).

3. The variance of the one step predictor

$$E\{(\mathcal{E}^{(1)})^2\} = P^{(1)} \qquad 5\text{-}4$$

4. H_1 - represents hypothesis no. 1, the <u>presence</u> of the target at point $(k+1,\ell+1)$.

H_2 - represents hypothesis no. 2, the <u>absence</u> of the target at point $(k+1,\ell+1)$.

Due to these hypotheses:

$$y(k+1,\ell+1) = x(k+1,\ell+1) + v(k+1,\ell+1) + \begin{cases} T & \text{when } H_1 \\ 0 & \text{when } H_2 \end{cases} \qquad 5\text{-}5$$

and:

$$r(k+1,\ell+1) = -y(k+1,\ell+1) + x^{(1)}(k+1,\ell+1)$$

$$= \mathcal{E}^{(1)}(k+1,\ell+1) + v(k+1,\ell+1) + \begin{cases} T & \text{when } H_1 \\ 0 & \text{when } H_2 \end{cases} \qquad 5\text{-}6$$

$r(k+1,\ell+1)$ is a random variable with the following statistics:

when H_1	when H_2
mean $[r(k+1,\ell+1)] = T_o$	mean $[r(k+1,\ell+1)] = 0$
var $[r(k+1,\ell+1)] = P^{(1)} + R + R_T$	var $[r(k+1,\ell+1)] = P^{(1)} + R$

It is shown in Fig. 43.

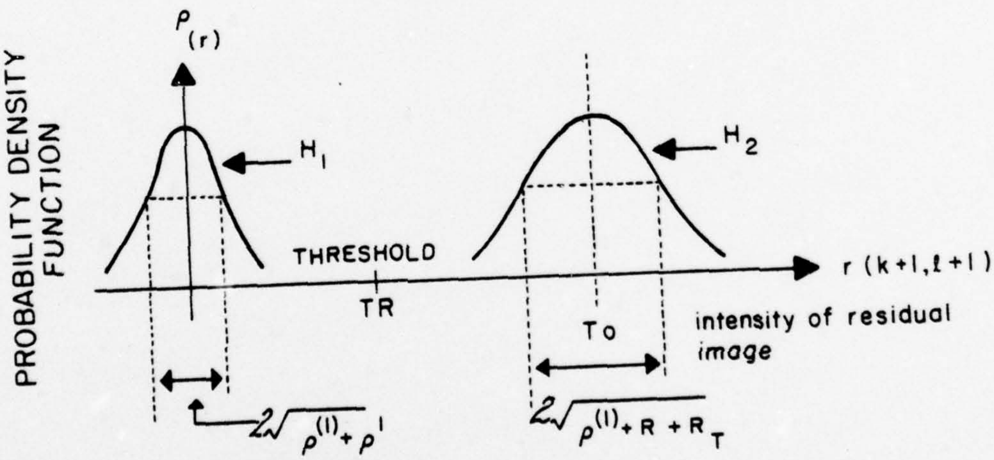


Fig. 43: Statistics of the Residual Due to the Two Hypotheses

The Processing Gain

Definition:

$$\left(\begin{matrix}\text{processing}\\ \text{gain}\end{matrix}\right) = \frac{(S/N)_{out}}{(S/N)_{in}} \qquad 5\text{-}7$$

In our case: The intensity of the target, T, as the residual image is equal to the intensity at the input (at the measurement device).

$$S_{out} = S_{in}$$

and therefore:

$$\left(\begin{matrix}\text{processing}\\ \text{gain}\end{matrix}\right) = \frac{N_{out}}{N_{in}}$$

The variance of the input noise:

$$N_{in} = \sqrt{\sigma^2 + R}$$

σ = variance of the correlated field.

$\sqrt{R}$ = variance of the white noise.

The variance of the output noise:

$$N_{out} = P^{(1)} + R$$

$P^{(1)}$ = variance of "one step predictor" error.

$$\boxed{\binom{\text{processing}}{\text{gain}} = \frac{\sigma^2 + R}{P^{(1)} + R}} \qquad 5\text{-}8$$

<u>The Threshold Device</u>

Finally, one has to define two regions on the $r(k+1,\ell+1)$ axis:

- a region of values where the <u>decision</u> will be for presence of a target.

- a region of values where the <u>decision</u> will be for <u>absence</u> of a target.

The input to the threshold device is the "one step predictor"

<u>Case 1</u>: The mean T_o is not known.

In this case the threshold has to refer only to the variance of r when H_2 occurs. A reasonable threshold is, for example

$$\text{threshold} = 3\sqrt{P^{(1)} + R}$$

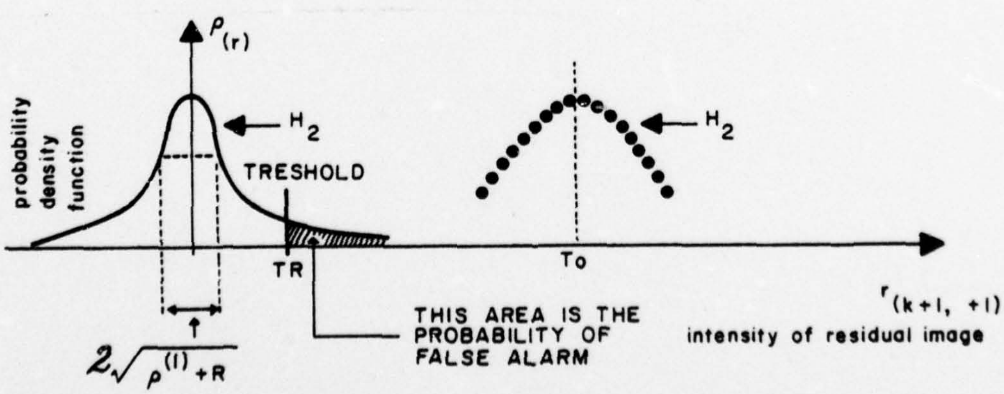


Fig. 44: Threshold for case 1 (T_o is unknown).
if r > TR → decide for "target"
if r < TR → decide for "no target"

Case 2: The mean T_o is known.

In this case the threshold has to refer to T_o. In this case one can use a "window threshold". See Fig. 45.

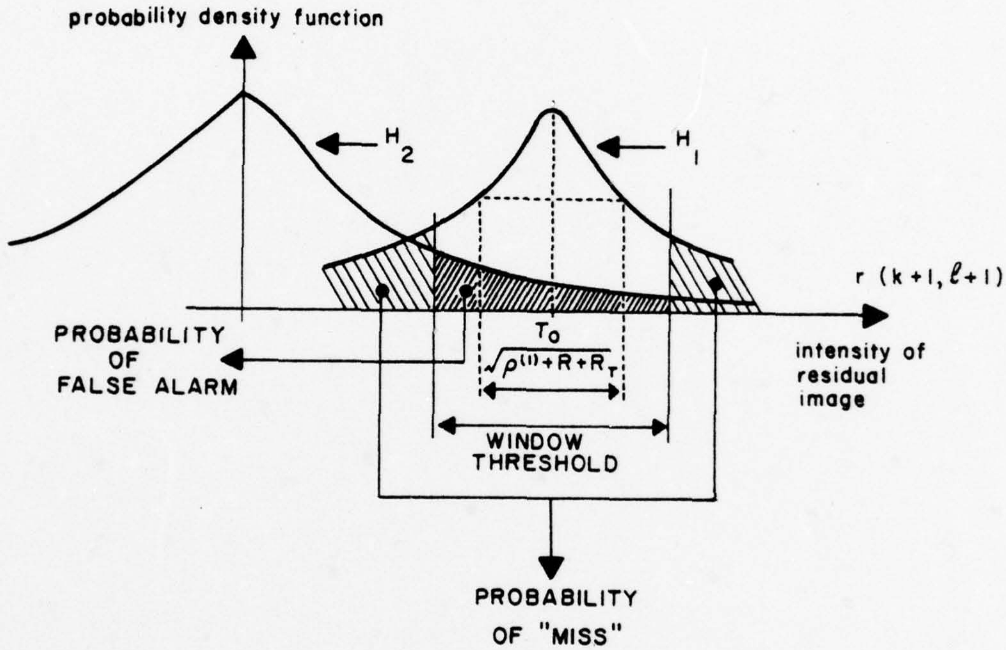


Fig. 45: Threshold for Case 2 (T_o is Known).

E. DETECTION OF LINES

1. The Problem

There are many algorithms to detect lines. Most of them suffer from one disadvantage: that to detect a line, an a-priori assumption must be about the direction of the line. Therefore, in order to detect lines in several directions, the image has to be scanned several times or, during one scan many calculations must be done for each point in the scan. There is no doubt that, for "real time" applications, all of those algorithms have a great disadvantage. The goal here is to develop an algorithm that will detect lines, regardless of their direction.

2. The Algorithm

The first, intuitive, feeling is that the algorithm of Section D, for point-detection, is correct for line detection. But, if this algorithm is used, only diagonal lines and the edge points of vertical and horizontal lines can be detected. Therefore, it was necessary to improve the algorithm for line-detection. The problem lies in the estimator-equation:

$$\hat{x}(k+1,\ell+1) = \hat{x}^1(k+1,\ell+1) + \underbrace{G(k+1,\ell+1)[y(k+1,\ell+1) - \hat{x}^1(k+1,\ell+1)]}_{\text{correction term}}$$

where the correction term depends on the measurement, $y(k+1,\ell+1)$. The "correction term" causes an error in a point where a target is present (because in this case $y = x + v + T$).

From another point of view, looking on Fig. 41, one can see that the basic assumption is that the output of the filter includes the background alone. In that case, the subtraction of the measured image, $y(k+1,\ell+1)$, from $\hat{x}$ (the output of the filter) will result with a residual image that is $T + v$. But it was shown that part of the target intensity appears in the output of the filter. Hence, after subtraction, the residual image doesn't include the target.

This discussion indicates the improvement required. Recall that the one step predictor is used in the threshold device to decide whether or not a target is present. This decision is used to improve the estimation procedure. The new algorithm will be a combination of the detection and filtering process.

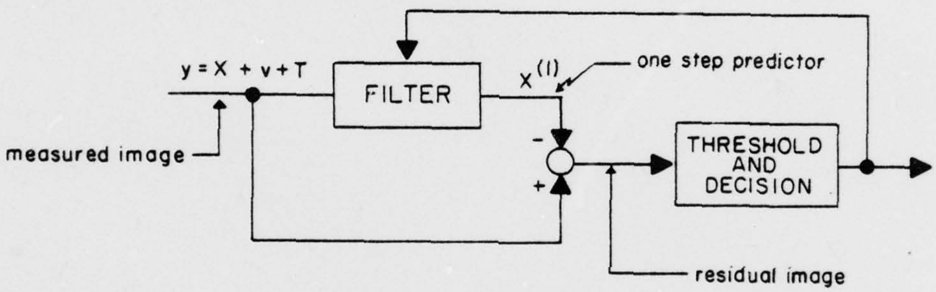


Fig. 46: The Combination Of Detection And Estimation For Improved Line Detection

The improved estimator equation will be

$$\hat{x}(k+1,\ell+1) = \begin{cases} x^{(1)}(k+1,\ell+1) \; \ldots\ldots\ldots\ldots & \text{if } r(k+1,\ell+1) > \text{threshold} \\ x^{(1)}(k+1,\ell+1) + G(k+1,\ell+1) \cdot [y(k+1,\ell+1) - x^{(1)}(k+1,\ell+1)] & \text{if } r(k+1, +1) < \text{threshold} \end{cases} \qquad 5\text{-}9$$

where, again

$$r(k+1,\ell+1) = y(k+1,\ell+1) - x^{(1)}(k+1,\ell+1).$$

The idea behind this improvement is that if the residual is greater than a certain threshold we know that this is due to a target, and therefore it is wrong to include the measurement $y(k+1,\ell+1)$ in the estimation of that particular point. In this case the best one can do is to use the one-step-predictor as the estimator of the point.

In the case where the mean of the target, T_o, is known, the equation for $x(k+1,\ell+1)$ will be:

$$\hat{x}(k+1,\ell+1) = \begin{cases} x^{(1)}(k+1,\ell+1) + G(k+1,\ell+1) & \text{if } r(k+1,\ell+1) > \text{threshold} \\ \cdot [(y(k+1,\ell+1) - T_o) - x^{(1)}(k+1,\ell+1)] & \\ & \\ x^{(1)}(k+1,\ell+1) + G(k+1,\ell+1) & \text{if } r(k+1,\ell+1) < \text{threshold} \\ \cdot [y(k+1,\ell+1) - x^{(1)}(k+1,\ell+1)] & \end{cases}$$

5-10

F. SUMMARY OF EXPERIMENTS

1. The Simulated Image

In order to check the algorithm of line detection, a simulated image was created (the size: 128 x 128).

a) According to the autocorrelation function

$$R_{xx}(n,m) = \sigma^2 \rho_1^{|n|} \rho_2^{|m|}$$

the correlated background was created by passing white noise through the filter,

$$x(k+1,\ell+1) = \rho_1 x(k+1,\ell) - \rho_1\rho_2 x(k,\ell) + \rho_2 x(k,\ell+1) + u(k,\ell)$$

where

$u(k,\ell)$ is Gaussian white noise, zero mean, and variance of $\sqrt{Q}$, where

$$Q = (1-\rho_1^2)(1-\rho_2^2).$$

b. A Gaussian white noise with zero mean and variance σ_n ($\sigma_n = \sqrt{R}$) was added to the correlated field.

c. A target that has the shape:

N. P. G. S.

was added to the image of b).

d. The simulated image will be

$$y(k+1,\ell+1) = \underbrace{x(k+1,\ell+1)}_{\text{Background}} + \underbrace{v(k+1,\ell+1)}_{\text{white noise}} + \underbrace{T(k+1,\ell+1)}_{\text{target}}$$

2. <u>Detection Of The Target</u>

a) The background was estimated by the filter:

$$\hat{x}^{(1)}(k+1,\ell+1) = \rho_1\hat{x}(k+1,\ell) - \rho_1\rho_2\hat{x}(k,\ell) + \rho_2\hat{x}(k,\ell+1)$$

with

$$x(k+1,\ell+1) = \begin{cases} x^{(1)}(k+1,\ell+1) \text{ .... } & \text{if } r(k+1,\ell+1) > TR \\ x^{(1)}(k+1,\ell+1) + & \text{if } r(k+1,\ell+1) < TR \\ G[y(k+1,\ell+1) - x^{(1)}(k+1,\ell+1)] & \end{cases}$$

where:

$$r(k+1,\ell+1) = y(k+1,\ell+1) - x^{(1)}(k+1,\ell+1).$$

b) The residual image $r(k+1,\ell+1)$ was passed through a half-wave rectifier, and then displayed. The reason for the rectification was the fact that it is known that negative values in the residual image are only noise. The display used a line-printer with letters representing eight gray levels.

c) The residual image was fed into the threshold device to decide if there is a target at point $(k+1,\ell+1)$.

if $r(k+1,\ell+1) > TR \longrightarrow$ presence of target

if $r(k+1,\ell+1) < TR \longrightarrow$ absence of target

This decision was fed back to the filter that estimates the background. Fig. 47 summarizes the procedure of simulating an image and target detection.

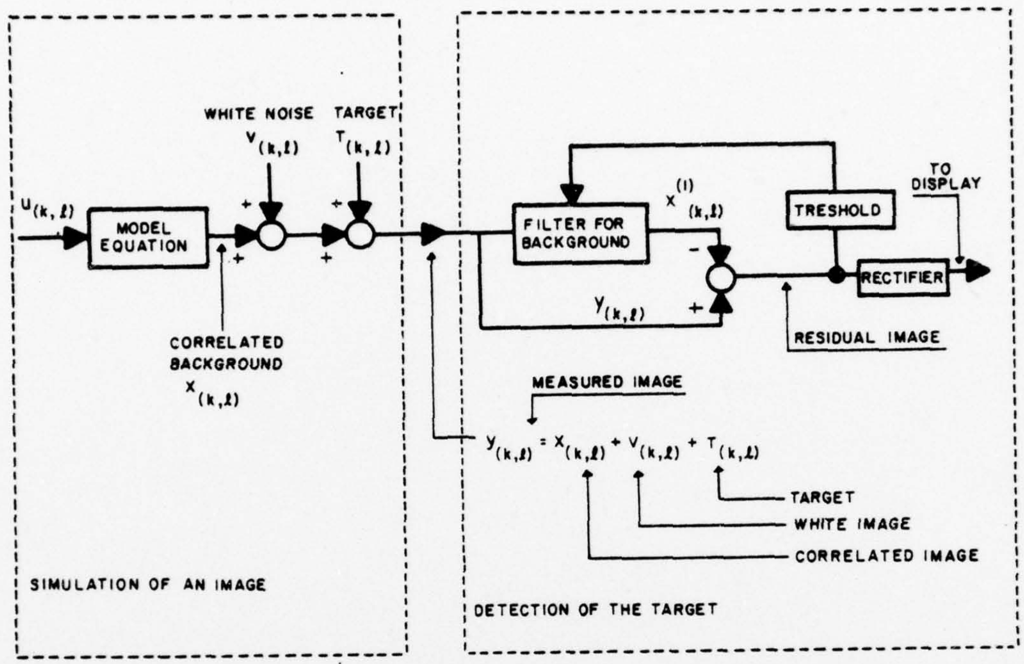


Fig. 47: Simulation of An Image and Detection of Targets

Results are given in Table 4, and Figs. 48-53.

Table 4

	Case 1	Case 2	Case 3
σ	1	1	1
$\rho_1 = \rho_2$	0.96	0.96	0.96
$\sqrt{R} = \sigma_n$	0.0	0.2	0.33
T	1.0	1.0	1.8
TR	0.6	0.6	0.9
G (gain)	1	0.46	0.32
Results Are Shown In Figures	43,44	45,46	47,48

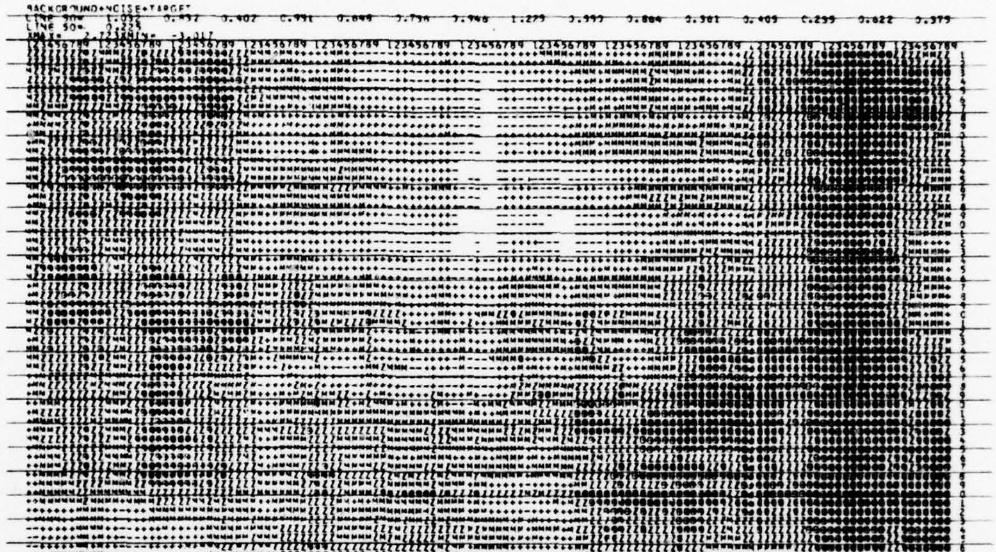


Fig. 48: Target Detection Case 1: The Original Image

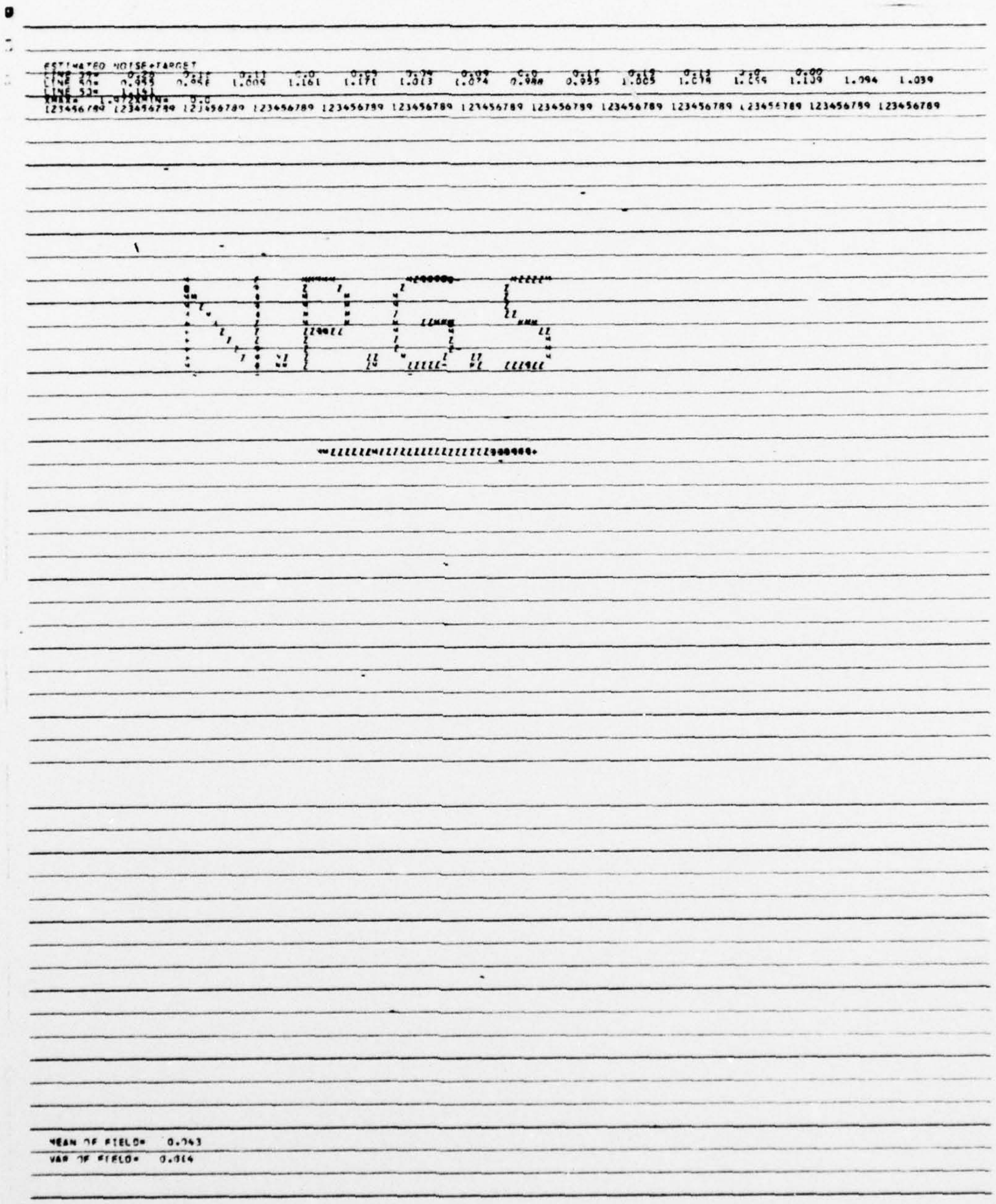


Fig 49 : Target Detection case 1 : The Residual Image

Fig. 49: Target Detection Case 1: The Residual Image

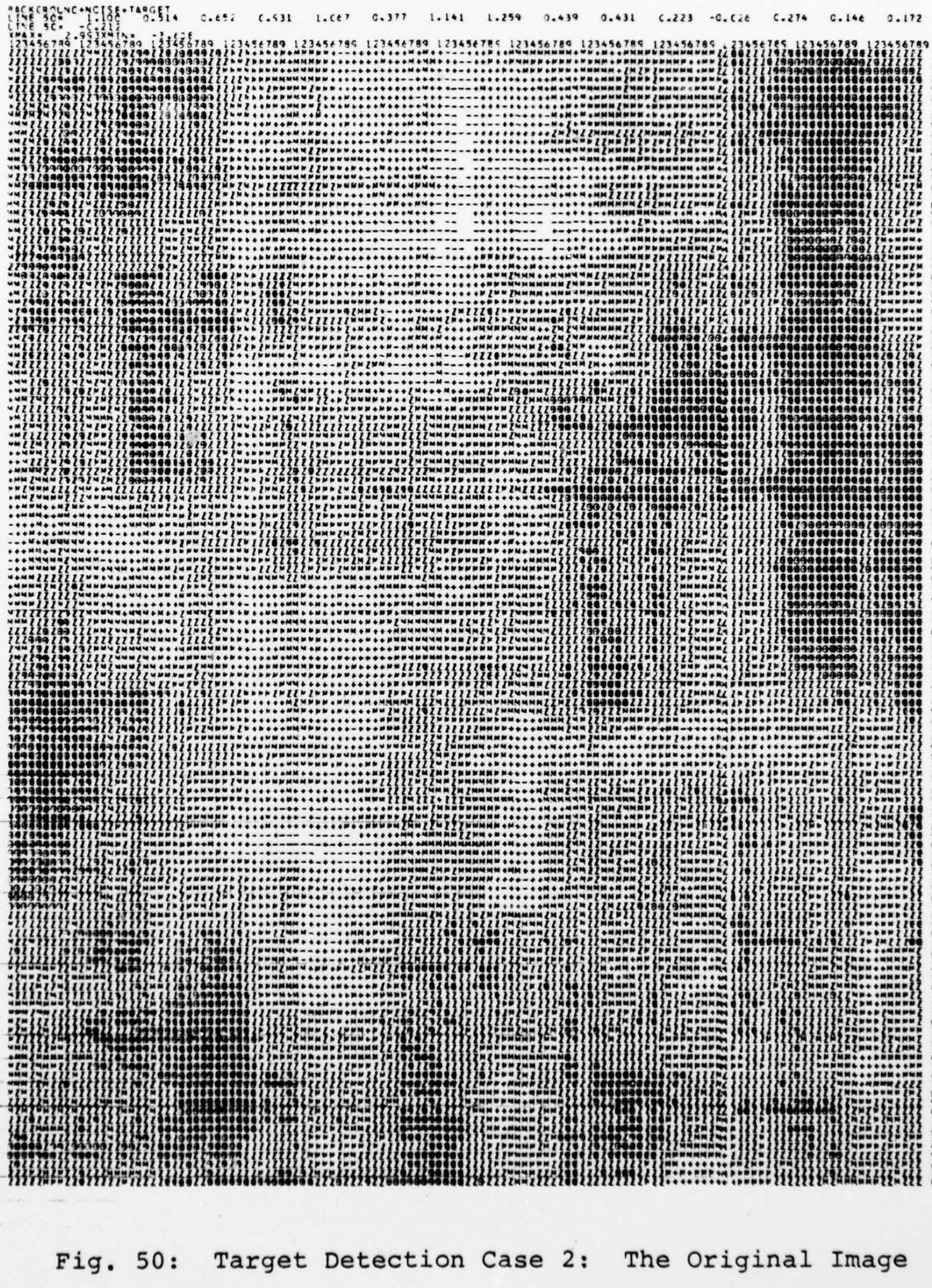


Fig. 50: Target Detection Case 2: The Original Image

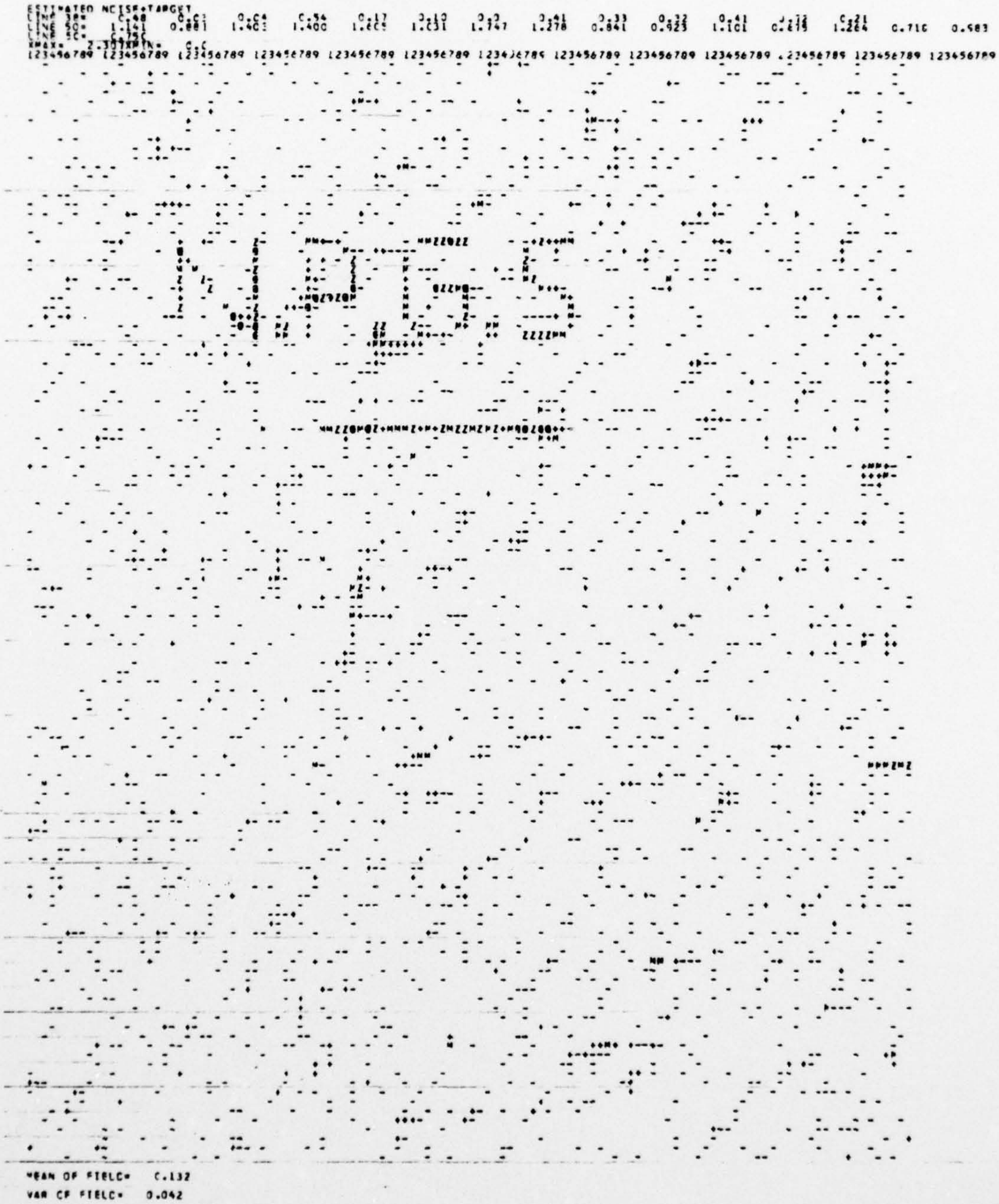


Fig. 51: Target Detection Case 2: The Residual Image

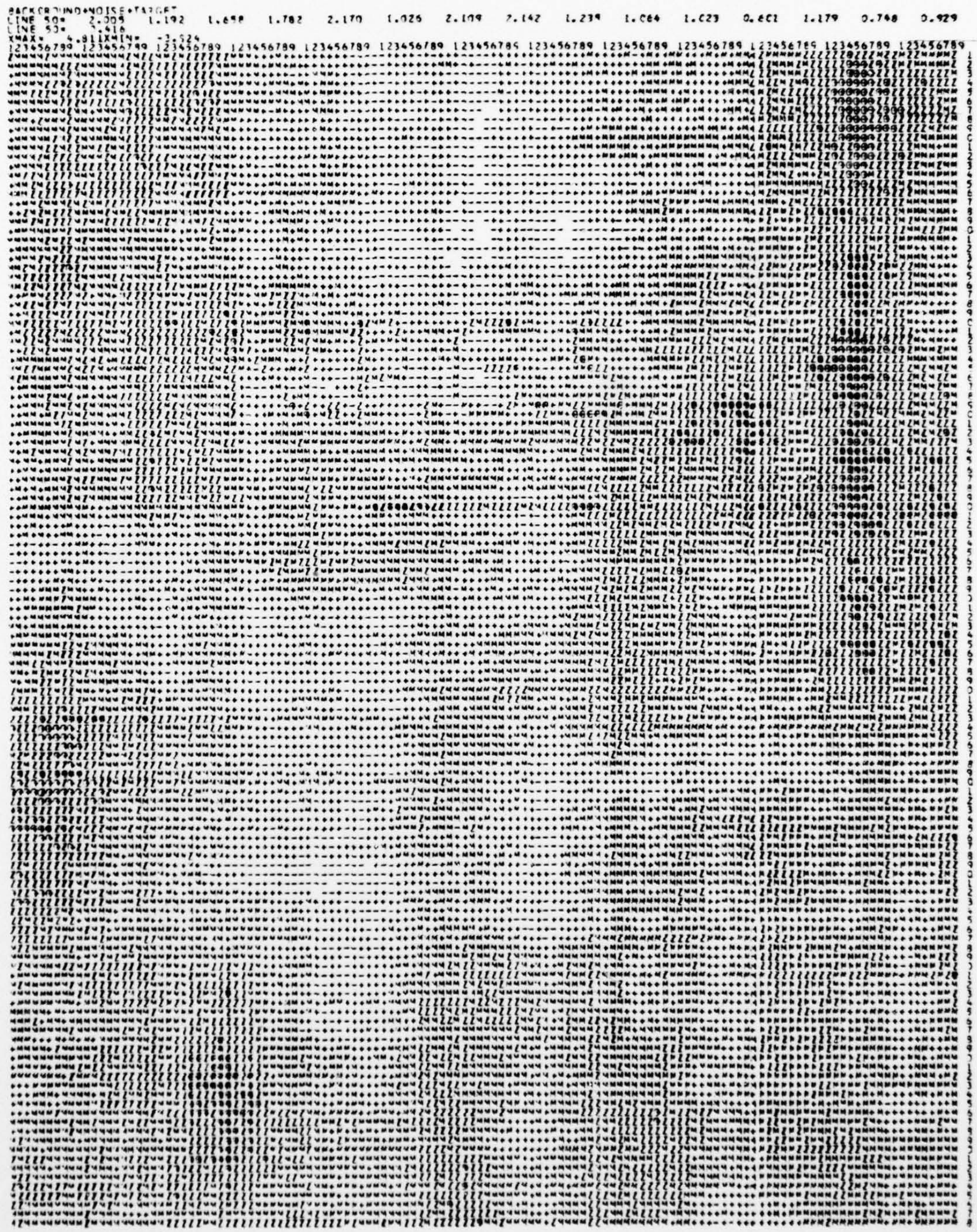


Fig. 52: Target Detection Case 3: The Original Image

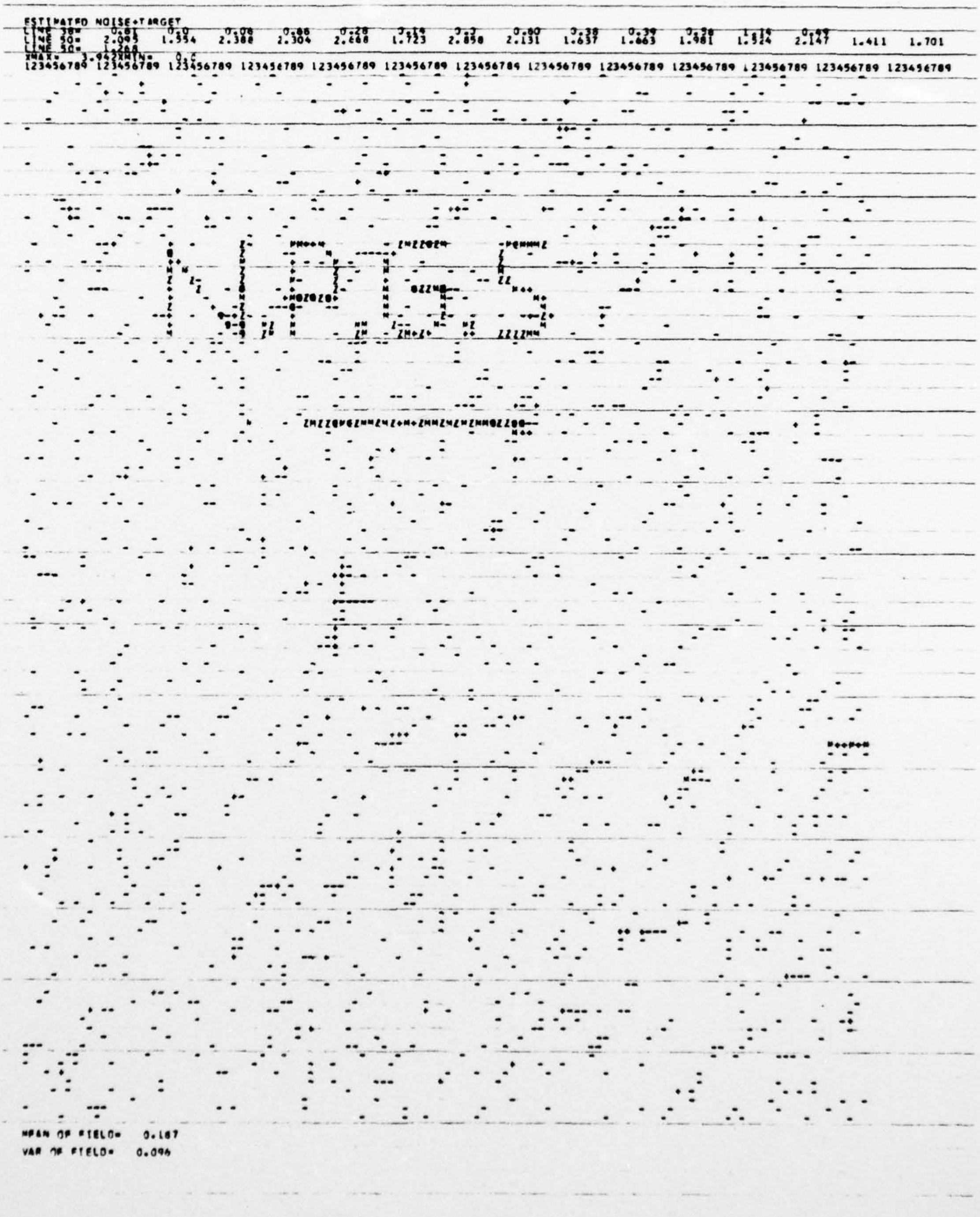


Fig. 53: Target Detection Case 3: The Residual Image

VI. FURTHER IDEAS IN THE SUBJECT OF RECURSIVE IMAGE PROCESSING

The purpose of this chapter is to suggest some additional topics in this research.

A. PROBLEMS IN RECURSIVE FILTERING

It is our belief that the filter that was developed in Chpt. IV is the best suboptimal filter in the structure of (4-12, 4-15). It is so because our approximation is assumed to be correct in steady state. But it is still an open question to prove that this filter is really the best.

In recursive estimation of more complicated structures the "state space structure", as in (2-116), should be developed. A proposed filter might be:

$$\begin{vmatrix} \hat{M}^{(1)}(k+1,\ell) \\ \hline \hat{N}^{(1)}(k,\ell+1) \end{vmatrix} = \left|\begin{array}{c:c} A_1 & A_2 \\ \hdashline A_3 & A_4 \end{array}\right| \begin{vmatrix} \hat{M}(k,\ell) \\ \hline \hat{N}(k,\ell) \end{vmatrix} \qquad \text{6-1}$$

and

$$\begin{vmatrix} \hat{M}(k+1,\ell) \\ \hat{N}(k+1,\ell) \end{vmatrix} = \begin{vmatrix} \hat{M}^{(1)}(k+1,\ell) \\ \hat{N}^{(1)}(k,\ell+1) \end{vmatrix} + \begin{vmatrix} G_m(k,\ell) \\ G_n(k,\ell) \end{vmatrix} [y(k,\ell) - (C_1 \quad C_2) \begin{vmatrix} \hat{M}^{(1)}(k,\ell) \\ \hat{N}^{(1)}(k,\ell) \end{vmatrix}]$$

6-2

or:

$$\begin{vmatrix} \hat{M}^{(1)}(k+1,\ell+1) \\ \hline \hat{N}^{(1)}(k+1,\ell+1) \end{vmatrix} = \begin{vmatrix} A_1 & A_2 \\ A_3 & A_4 \end{vmatrix} \cdot \begin{vmatrix} \hat{M}(k,\ell+1) \\ \hline \hat{N}(k+1,\ell) \end{vmatrix} \qquad \text{6-3}$$

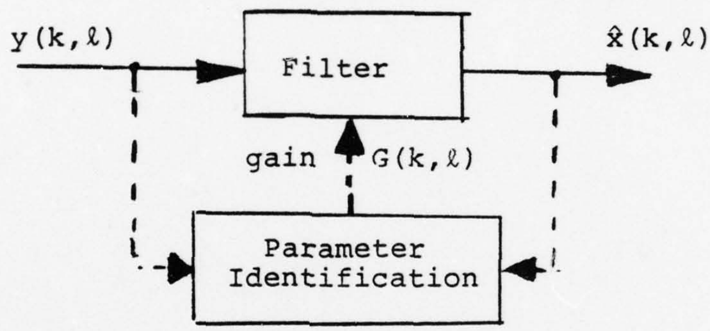
and

$$\begin{vmatrix} \hat{M}(k+1,\ell+1) \\ \hline \hat{N}(k+1,\ell+1) \end{vmatrix} = \begin{vmatrix} \hat{M}^{(1)}(k+1,\ell+1) \\ \hline \hat{N}^{(1)}(k+1,\ell+1) \end{vmatrix} + \begin{vmatrix} G_m(k+1,\ell+1) \\ \hline G_n(k+1,\ell+1) \end{vmatrix} [y(k+1,\ell+1) - (C_1 \quad C_2)$$

$$\cdot \begin{vmatrix} \hat{M}^{(1)}(k+1,\ell+1) \\ \hline \hat{N}^{(1)}(k+1,\ell+1) \end{vmatrix}] \qquad \text{6-4}$$

Following the procedure of Chpt. IV the gains $G_m(k,\ell)$, $G_n(k,\ell)$ can be calculated using recursive equations. Recall that the problem with the filter of Chpt. IV was the non-existance of a "closed set" of recursive formulas for gain calculation (an approximation had to be done). The proposed filter that is suggested in this chapter suffers from the same difficulty. Therefore it is our belief that formulas for the general case cannot be found. Two dimensional filters with the proposed structure above should be derived and checked for specific cases.

The next topic for additional research is sensitivity analysis of the algorithm to changes of the field parameters. In order to improve the estimation in the case where the parameters of the field are not known, one can use an adaptive system as follows.



One can apply the target detection algorithm on images with an autocorrelation function from (2-9) of the form

$$R_{xx}(n,m) = \rho_1^{|n|} \cos(\theta_1 n) \rho_2^{|m|} \cos(\theta_2 m)$$

This can be done only after developing a filter for this case (see proposed structure of the filter in Section A of this chapter).

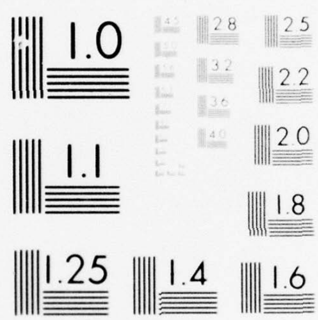
The approach to target detection in this thesis was to use a feedback line from the output of the threshold device back to the filter (see Fig. 41). It was found that for low values of threshold, this approach is unstable. The

AD-A045 179 NAVAL POSTGRADUATE SCHOOL MONTEREY CALIF F/G 12/1
MODELING AND RECURSIVE ESTIMATION OF TWO DIMENSIONAL RANDOM FIE--ETC(U)
JUN 77 M SHACHAR

UNCLASSIFIED NL

3 OF 3
AD
A045179

END
DATE
FILMED
11 -77
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

question of stability criteria for this two-dimensional, non-linear operation remains unsolved. Using a feedback from the threshold device to the filter has other disadvantages. A solution to this problem might be to estimate $x(k,\ell)$ without a feedback from the threshold device. The residual image will then be

$$r(k+1,\ell+1) = y(k+1,\ell+1) - \alpha\hat{x}(k+1,\ell) - \beta x(k,\ell+1)$$

where

α,β have been found by some optimization criteria.

Detection may be improved by processing more than one frame or alternatively tracking in time domain from frame to frame by a regular Kalman filter.

An improvement may be achieved using the algorithm of (5-10), and assuming the intensity of the target is known. If the intensity of the target is not known, one can use this algorithm by adding a "tracking loop" on the target intensity in order to estimate this intensity level on-line.

At last all these ideas should be checked on "real life" images.

APPENDIX A

THE ONE STEP PREDICTOR

Here we compare between the recursive filter and the optimal non-recursive estimation (by direct use of the orthogonality principle). It will be seen that the one-step-predictor is not optimal. In order to prove it a specific case will be shown and then we will extend the result.

Part 1: A Specific Case.

The recursive procedure of gain calculation by the algorithm of Chpt. IV, leads to optimal estimators for the first line and column. That fact is obvious because the first line and column could be treated as a one dimensional case, using Kalman Filtering. Now assume the measurement starts at (0,0). Then, the first place where a difficulty appears is in the calculation of the "one-step-predictor" of point (1,1). We shall calculate $P^{(1)}(1,1)$ using two methods: (1) optimal non-recursive form, and (2) by using the recursive form, in order to compare the results.

Given:

$$R_{xx} = \sigma^2 \rho_1^{|n|} \rho_2^{|m|}$$

$$\sigma = 1$$

$$\rho_1 = \rho_2 = 0.96$$

$$y(k,\ell) = x(k,\ell) + v(k,\ell)$$

$$R_{VV}(n,m) = \begin{cases} 0 & , n \neq 0 \text{ or } m \neq 0 \\ R = 0.64 & , n = 0 \text{ and } m = 0 \end{cases}$$

Find: The estimator of point (1,1) by using its three neighbors (0,0), (0,1), (1,0).

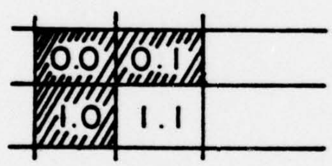


Fig. 43: The one step predictor. The Fig. shows the pixels that take part in the one-step-predictor of pixel 1,1

Solution By Using the Orthogonal Principle

By using Eq. C-1 (Appendix C) the next set of equations is obtained.

$$\begin{pmatrix} R_{xx}(0,0)+R & R_{xx}(1,0) & R_{xx}(0,1) \\ R_{xx}(-1,0) & R_{xx}(0,0)+R & R_{xx}(-1,1) \\ R_{xx}(0,1) & R_{xx}(1,-1) & R_{xx}(0,0)+R \end{pmatrix} \begin{pmatrix} \alpha_{0,0} \\ \alpha_{0,1} \\ \alpha_{1,0} \end{pmatrix} = \begin{pmatrix} R_{xx}(-1,-1) \\ R_{xx}(0,-1) \\ R_{xx}(-1,0) \end{pmatrix}$$

and, for $\rho_1 = \rho_2 = \rho$:

$$\begin{pmatrix} 1+R & \rho & \rho \\ \rho & 1+R & \rho^2 \\ \rho & \rho^2 & 1+R \end{pmatrix} \begin{pmatrix} \alpha_{0,0} \\ \alpha_{0,1} \\ \alpha_{1,0} \end{pmatrix} = \begin{pmatrix} \rho^2 \\ \rho \\ \rho \end{pmatrix}$$

and:

$$\alpha_{0,0} = \frac{(1+R)^2\rho^2 + 2\rho^4 - (1+R)(2\rho^2+\rho^6)}{(1+R)^3 + 2\rho^4 - (1+R)(2\rho^2 + \rho^4)} = 0.2250$$

$$\alpha_{0,1} = \alpha_{1,0} = \frac{(1+R)^2 + \rho^5 - (1+R)2\rho^3}{(1+R)^3 + 2\rho^4 - (1+R)(2\rho^2 + \rho^4)} = 0.2925$$

Now, using Eq. C-2, the error is calculated:

$$P(1,1)_{optimal} = E\{\varepsilon^2(1,1)\}$$

$$= E\{x^2(1,1) - (\Sigma\alpha_{p,q}y(p,q))x(1,1)\} ,$$

where:

$$(p,q) = \{(0,0), (0,1), (1,0)\}$$

$$= 1 - \alpha_1\rho^2 - \alpha_2\rho - \alpha_3\rho$$

$P(1,1)_{optimal} = 0.2360$.

Solution By Recursive Filtering

To simplify the calculations, the G's and P's for points (0,0), (0,1), (1,0) will be calculated from the Kalman filter equation. (Recall: the results for the first line and column by using the algorithm of Chpt. IV are the same as in the Kalman Filtering results.)

$$P(0,0) = 1 \qquad \text{(Initial Condition)}$$

$$G(0,0) = \frac{P^{(1)}(0,0)}{P^{(1)}(0,0) + R} = 0.6098$$

$$P(1,1) = \frac{P^{1}(0,0)R}{P^{1}(0,0) + R} = 0.3902$$

$$P^{1}(1,2) = P^{1}(2,1) = \rho^{2}\, P(1,1) + Q = 0.4380$$

$$G(1,2) = G(2,1) = \frac{P^{1}(1,2)}{P^{1}(1,2) + R} = 0.4063$$

$$P(1,2) = P(2,1) = \frac{P^{1}(1,2)R}{P^{1}(1,2) + R} = 0.2601$$

Also the calculations for the $P_{0,1}(1,1)$ and $P_{1,0}(1,1)$ can be done in a simplified way.

Equivalently to Eq. 4-20, in the one-dimensional filter, one can write:

$$\mathcal{E}(k+1) = F(k+1)\rho_2 \mathcal{E}(k) - F(k+1)u(k) + G(k+1)v(k+1) \quad - \underline{\text{for } \ell = 0}$$

$$\mathcal{E}(\ell+1) = F(\ell+1)\rho_1\mathcal{E}(\ell) - F(\ell+1)u(\ell) + G(\ell+1)v(\ell+1) \quad - \quad \underline{\text{for } k = 0}$$

therefore:

$$P_{0,1}(0,0) = E\{\mathcal{E}(0,0)\ \mathcal{E}(0,1)\}$$

$$= P(0,0)\big(1 - G(0,1)\big)\rho_1 = 0.2224$$

$$P_{1,0}(0,0) = E\{\mathcal{E}(0,0)\cdot\mathcal{E}(1,0)\}$$

$$= P(0,0)(1 - G(1,0))\rho_2 = 0.2224$$

$$P_{1,-1}(0,1) = E\{\mathcal{E}(0,1)\ \mathcal{E}(1,0)\}$$

$$= (1 - G(1,0))(1 - G(0,1))P(0,0)\rho_1\rho_2$$

$$= 0.1268$$

Now, the equation for the one step predictor is:

$$P^1(1,1) = \underset{\sim}{\rho}\ \underset{\sim}{P}(0,0)\underset{\sim}{\rho}^T + Q$$

$$= (\rho_1 \quad -\rho_1\rho_2 \quad \rho_2)\begin{pmatrix} P(1,0) & P_{1,0}(0,0) & P_{1,-1}(0,1) \\ P_{1,0}(0,0) & P(0,0) & P_{0,1}(0,0) \\ P_{1,-1}(0,1) & P_{0,1}(0,0) & P(0,1) \end{pmatrix}\begin{pmatrix} \rho_1 \\ -\rho_1\rho_2 \\ \rho_2 \end{pmatrix} + (1-\rho_1^2)(1-\rho_2^2)$$

$\underline{P^1(1,1) \;=\; 0.2634}$

$\underline{P^1(1,1)}$ is not far away $\underline{\text{from } P^{(1)}(1,1)_{optimal}}$, but anyhow different.

Part 2: Proof of the General Case

First, prove that the optimal solution requires the estimation error of the predictor to be uncorrelated (orthogonal) to the estimators that take part in the equation of the predictor:

$$\hat{x}^{(1)}(k+1,\ell+1) = \alpha_1\hat{x}(k+1,\ell) + \alpha_2\hat{x}(k,\ell) + \alpha_3\hat{x}(k,\ell+1)$$

and the error, $\mathcal{E}^{(1)}(k+1,\ell+1)$ is given by

$$\mathcal{E}^{(1)}(k+1,\ell+1) \triangleq \hat{x}^{(1)}(k+1,\ell+1) - x(k+1,\ell+1)$$

$$= \alpha_1\hat{x}(k+1,\ell) + \alpha_2\hat{x}(k,\ell) + \alpha_3\hat{x}(k,\ell+1) - x(k+1,\ell+1)$$

we want to minimize

$$E\{[\mathcal{E}^{(1)}(k+1,\ell+1)]^2\} = E\{[\alpha_1\hat{x}(k+1,\ell)+\alpha_2\hat{x}(k,\ell)+\alpha_3\hat{x}(k,\ell+1)-x(k+1,\ell+1)]^2$$

which requires

$$\frac{\partial}{\partial\alpha_i} E\{[\mathcal{E}^{(1)}(k+1,\ell+1)]^2\} = 0. \quad i = 1,2,3.$$

We obtain

$$\boxed{E\{\mathcal{E}^{(1)}(k+1,\ell+1)\ \hat{x}(m,n)\}} = 0$$

$$(m,n) = \{(k+1,\ell),\ (k,\ell),\ (k,\ell+1)\}$$

- in the one dimensional case:

$$\boxed{E\{\mathcal{E}^{(1)}(k+1)\cdot\hat{x}(k)\} = 0}$$

Now check if this orthogonality condition can be satisfied.

1) In the one dimensional case:

$$E\{\mathcal{E}^{(1)}(k+1)\ \hat{x}(k)\} = E\{[\hat{x}^{(1)}(k+1) - x(k+1)]\ \hat{x}(k)\}$$

substituting: $x(k+1) = \rho x(k) + u(k)$

$$E\{\mathcal{E}^{(1)}(k+1)\ \hat{x}(k)\} = E\{[\rho\hat{x}(k) - \rho x(k) - u(k)]\ \hat{x}(k)\}$$

$$= E\{[\rho\mathcal{E}(k) - u(k)]\ \hat{x}(k)\}$$

assuming $\hat{x}(k)$ is optimal, we know that:

$$E\{\mathcal{E}(k)\ \hat{x}(k)\} = 0$$

also:

$$E\{u(k)\ \hat{x}(k)\} = 0$$

Therefore:

$$\underline{E\{\varepsilon^{(1)}(k+1)\ \hat{x}(k)\} = 0}$$

<u>Conclusion</u>: The condition is satisfied in the case of the one-dimensional one-step-predictor.

2) In the two dimensional case:

Check if $E\{\varepsilon^{(1)}(k+1,\ell+1)\ \hat{x}(k,\ell+1)\}$ can be made zero.

$$E\{\varepsilon^{(1)}(k+1,\ell+1)\ \hat{x}(k,\ell+1)\} =$$

$$= E\{[(\rho_1\hat{x}(k+1,\ell)-\rho_1\rho_2\hat{x}(k,\ell)+\rho_2\hat{x}(k,\ell+1))-x(k+1,\ell+1)]\hat{x}(k,\ell+1)\}$$

$$= E\{[(\rho_1\hat{x}(k+1,\ell)-\rho_1\rho_2\hat{x}(k,\ell)+\rho_2\hat{x}(k,\ell+1))-(\rho_1\hat{x}(k+1,\ell)$$

$$-\ \rho_1\rho_2 x(k,\ell)+\rho_2 x(k,\ell+1)+u(k,\ell)]\hat{x}(k,\ell+1)\}$$

$$= E\{[\rho_1\varepsilon(k+1,\ell)-\rho_1\rho_2\ \ (k,\ell)+\rho_2\varepsilon(k,\ell+1)]\cdot\hat{x}(k,\ell+1)\}$$

Now, because $\hat{x}(k+1,\ell)$, $\hat{x}(k,\ell)$, $\hat{x}(k,\ell+1)$ are based on different sets of data, and so are also the corresponding errors, the only way to make the last expression zero for each k, is to make each part of it zero.

$$E\{\mathcal{E}(k,\ell+1)\ \hat{x}(k,\ell+1)\} = 0$$

$$E\{\mathcal{E}(k,\ell)\ \hat{x}(k,\ell+1)\} = 0$$

$$E\{\mathcal{E}(k+1,\ell)\ \hat{x}(k,\ell+1)\} = 0$$

But the last expression, $E\{\mathcal{E}(k+1,\ell)\ \hat{x}(k,\ell+1)\}$ cannot be made zero. Note that $\mathcal{E}(k+1,\ell)$ is a result that is obtained from a different set of data than $\hat{x}(k,\ell+1)$. For example one cannot require $\mathcal{E}(1,0)$ and $\hat{x}(0,1)$ to be orthogonal, knowing that:

$$\hat{x}(0,1) = \text{function of } y(0,0),\ y(0,1)$$

$$\mathcal{E}(1,0) = \text{function of } y(0,0),\ y(1,0).$$

Q.E.D.

APPENDIX B

Let us show how one can write nonrecursive equations for $x(k+1,\ell+1)$, $y(k+1,\ell+1)$, $\hat{x}(k+1,\ell+1)$ and $(k+1,\ell+1)$ and then show some properties of the correlation between these values.

From the recursive equation:

$$x(k+1,\ell+1) = \rho_1 x(k+1,\ell) - \rho_1\rho_2 x(k,\ell) + \rho_2 x(k,\ell+1) + u(k,\ell)$$

one can write:

$$x(k+1,\ell+1) = \overbrace{\Sigma a(0,\ell)x(0,\ell) + \Sigma a(k,0)x(k,0)}^{\text{initial conditions}}$$

$$+ \sum_{(i,j)\in\Omega_2(k,\ell)} \alpha(i,j)u(i,j)$$

$$= M + \sum_{(i,j)\in\Omega_2(k,\ell)} \alpha(i,j)u(i,j) \qquad \text{B-1}$$

Now, using Eq. 2:

$$\hat{y}(k+1,\ell+1) = ([M + \sum_{(i,j)\in\Omega_2(k,\ell)} \alpha(i,j)u(i,j)] +$$

$$v(k+1,\ell+1) \qquad \text{B-2}$$

Using 2, 12, 15, A-1, A-2:

$$\hat{x}(k+1,\ell+1) = N + \sum_{(i,j)\varepsilon\Omega_2(k,\Omega)} \beta(i,j)u(i,j)$$

$$+ \sum_{(i,j)\varepsilon\Omega_2(k+1,\ell+1)} \gamma(i,j)v(i,j) \qquad \text{B-3}$$

<u>Theorem</u>: $E\{u(k,\ell)\underset{\sim}{\mathcal{E}}^T(k,\ell)\} = E\{\underset{\sim}{\mathcal{E}}(k,\ell)u(k,\ell)\} = 0$

<u>Proof</u>: Using 4-18, 4-19, B-1, B-3 we see that $\underset{\sim}{\mathcal{E}}(k,\ell)$ does not include the value $u(k,\ell)$. Therefore by using 4-4 we have completed the proof.

<u>Theorem</u>: $E\{v(k+1,\ell+1)\mathcal{E}^T(k,\ell)\} = E\{\underset{\sim}{\mathcal{E}}(k,\ell)v(k+1,\ell+1)\} = 0$

<u>Proof</u>: The proof is in the same procedure as above. By substituting B-1, B-3 into the expression for $\underset{\sim}{\mathcal{E}}(k,\ell)$ we see that $\mathcal{E}(k,\ell)$ does not include the value $v(k+1,\ell+1)$, and therefore, by using Eq. 4-7, we have completed the proof.

In the same way one can show that:

$$E\{\mathcal{E}(k,\ell)\ u(k+i,\ell+j)\} = 0$$

$$E\{\mathcal{E}(k,\ell)\ v(k+i+1,\ell+j+1)$$

if $i \geq 0$

$j \geq 0$

APPENDIX C

NON-RECURSIVE ESTIMATION

1. Motivation For The Non-Recursive Estimation

The goal of this thesis is recursive estimation in image processing. Chpt. IV derived a recursive "least square error" estimator that was not optimal (the error was not orthogonal to the data). So, the motivation to this section is to compare the recursive filter to the optimal solution.

2. The Basic Procedure

The non recursive estimation is based on the orthogonality principle, which is given in Eq. 3-10:

$$E\{[x(k,\ell) - \sum_{(p,q)\varepsilon\Omega_1(k,\ell)} \alpha_{p,q}\, y(p,q)]\, y(i,j)\} = 0 \qquad \text{C-1}$$

and writing it in the form:

$$E\{[x(k,\ell) - \sum_{(p,q)\varepsilon\Omega_1(k,\ell)} \alpha_{p,q}(x(p,q) + v(p,q))] \cdot [x(p,q) + v(p,q)]\} = 0$$

v is the noise that is added to the state x.

Let us clarify the procedure with an example;

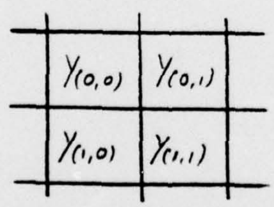
Given: 1) $R_{xx}(n,m) = \rho_1^{|n|} \rho_2^{|m|}$

2) $$R_{vv}(n,m) = \begin{cases} 0 & \text{if } n \neq 0 \quad \text{or} \quad m \neq 0 \\ R & \text{if } n = 0 \quad \text{and} \quad m = 0 \end{cases}$$

$v(k,\ell)$ is uncorrelated to any of the random-field values, $x(i,j)$.

3) $y(k,\ell) = x(k,\ell) + v(k,\ell)$

Find: the four optimal coefficients to estimate the value $x(0,0)$ by the set of measurements in a 2×2 area: $y(0,0)$, $y(0,1)$, $y(1,0)$, $y(1,1)$



$$\hat{x}(1,1) = \alpha_1 y(0,0) + \alpha_2 y(0,1) + \alpha_3 y(1,0) + \alpha_4 y(1,1)$$

Fig. 55: Non Recursive Estimation

Solution: If the noise is uncorrelated to the signal $x(k,\ell)$, then the set of equations that follows from the orthogonality condition is:

$$
\begin{pmatrix} R(-1,-1) \\ R(0,-1) \\ R(-1,0) \\ R(0,0) \end{pmatrix} = \begin{pmatrix} R_{xx}(0,0)+R & R_{xx}(1,0) & R_{xx}(0,1) & R_{xx}(1,1) \\ R_{xx}(-1,0) & R_{xx}(0,0)+R & R_{xx}(-1,1) & R_{xx}(0,-1) \\ R_{xx}(1,1) & R_{xx}(1,-1) & R_{xx}(0,0)+R & R_{xx}(1,0) \\ R_{xx}(-1,-1) & R_{xx}(0,-1) & R_{xx}(-1,0) & R_{xx}(0,0)+R \end{pmatrix} \begin{pmatrix} \alpha_{0,0} \\ \alpha_{0,1} \\ \alpha_{1,0} \\ \alpha_{1,1} \end{pmatrix}
$$

or:

$$
\begin{pmatrix} \rho_1\rho_2 \\ \rho_2 \\ \rho_1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+R & \rho_2 & \rho_1 & \rho_1\rho_2 \\ \rho_1 & 1+R & \rho_1\rho_2 & \rho_2 \\ \rho_2 & \rho_1\rho_2 & 1+R & \rho_1 \\ \rho_1\rho_2 & \rho_2 & \rho_1 & 1+R \end{pmatrix} \begin{pmatrix} \alpha_{0,0} \\ \alpha_{0,1} \\ \alpha_{1,0} \\ \alpha_{1,1} \end{pmatrix}
$$

From this four equation system, one can find the α's.

3. <u>The Estimation Error (variance)</u>

$$
\varepsilon(k,\ell) \triangleq x(k,\ell) - \hat{x} = x(k,\ell) - \Sigma\, \alpha_{p,q}\, y(p,q)
$$

$$
\varepsilon^2(k,\ell) = [x(k,\ell) - \Sigma\alpha_{p,q}\, y(p,q)]^2
$$

$$
E\{\varepsilon^2\} = E\{x^2(k,\ell) - 2x(k,\ell)\,\Sigma\alpha_{p,q}y(p,q)
$$

$$
+ (\Sigma\alpha_{p,q}y(p,q))^2\}
$$

$$E\{\ ^2\} = E\{x^2(k,\ell) - x(k,\ell)\Sigma\alpha_{p,q}y(p,q)$$

$$+ \Sigma\ \alpha_{p,q}y(p,q)(\Sigma\alpha_{p,q}y(p,q) - x)\}$$

The zero was set due to the orthogonality principle.

$$P(k,\ell) = E\{\ ^2(k,\ell)\} = E\{x^2(k,\ell) - (\Sigma\alpha_{p,q}y(p,q))x(k,\ell)\} \qquad \text{C-2}$$

<u>Example</u>: If we continue the previous example (the estimation of a point by a 2 x 2 area):

$$E\{\ ^2\} = 1 - \alpha_{0,0}\rho_1\rho_2 - \alpha_{0,1}\rho_2 - \alpha_{1,0}\rho_1 - \alpha_{1,1}$$

<u>Theorem</u>: If the measurement $y(k,\ell)$ is <u>included</u> in the estimation of the point $x(k,\ell)$, then:

$$P(k,\ell) = E\{\ ^2(k,\ell)\} = \alpha_{k,\ell}\ R \qquad \text{C-3}$$

<u>Proof</u>: Assume the estimation of the point $x(k,\ell)$ is done <u>without</u> the point (k,ℓ), by using a set of measurements Y. Let's call this estimator $\hat{x}_p$ (p = Prediction).

$$\hat{x}_p(k,\ell) = \sum_{(i,j)\epsilon Y} \alpha_{i,j}y(i,j) \qquad (i,j) \neq (k,\ell)$$

$\hat{x}_p(k,\ell)$ is the "predictor" of $x(k,\ell)$.

Also assume that the variance of the error of the predictor is P_o. Now assume another estimator that uses the set Y and the measurement $y(k,\ell)$. In that case we can say:

$$\hat{x}(k,\ell) = \alpha_1 \hat{x}_p(k,\ell) + \alpha_{k,\ell} y(k,\ell).$$

$\hat{x}(k,\ell)$ is supposed to be an optimal combination of $x(k,\ell)$ and $y(k,\ell)$. Because there is no correlation between $v(k,\ell)$ and other values, this is a case of a combination between two estimators that have uncorrelated errors [see Ref. 4, Chapter on optimal smoothing].

estimator A: $\hat{x}_p(k,\ell)$ with variance P_p.

estimator B: $y(k,\ell)$ with variance R.

The optimal combination is [due to Ref. 4]:

$$\hat{x}(k,\ell) = \frac{R}{R + P_p} \hat{x}_p(k,\ell) + \frac{P_p}{R + P_p} y(k,\ell)$$

with variance:

$$P(k,\ell) = \frac{P_p R}{R + P_p} = \left(\frac{P_p}{R + P_p}\right) R = \alpha_{k,\ell} R$$

Q.E.D.

4. Unbiased Estimator

Given: A random field $x(k,\ell)$ with mean $\overline{x}$.

A noise measurement $y(k,\ell) = x(k,\ell) + v(k,\ell)$.

We wish to find an unbaised estimator for this case.

Define: $x(k,\ell) \triangleq x'(k,\ell) + \overline{x}$

$x'(k,\ell)$ is a random field with mean zero. Suppose we want to estimate $x'(k,\ell)$. To do that define

$$y'(k,\ell) \triangleq x'(k,\ell) + v(k,\ell)$$

$$= y(k,\ell) - \overline{x}$$

the estimation of $x'(k,\ell)$ is:

$$\hat{x}(k,\ell) = \Sigma \alpha_{p,q} y'(p,q)$$

$\hat{x}(k,\ell)$ is unbiased because $x'(k,\ell)$ is unbiased.

And now, the estimation of $x(k,\ell)$ is:

$$\hat{x}(k,\ell) = \hat{x}'(k,\ell) + \overline{x} = {}_{p,q} y'(p,q) + \overline{x}$$

$$= [\Sigma\alpha_{p,q} (y(p,q) - \overline{x})] + \overline{x}$$

$$\hat{x}(k,\ell) = \Sigma\alpha_{p,q} y_{p,q} + \overline{x}[1 - \Sigma\alpha_{p,q}]$$

Conclusion: The estimator can be done unbaised regardless what the coefficients $\alpha_{p,q}$ are. Therefore, the calculation of the α's is the same as before. One has only to add one term to the estimator equation. This term depends on the mean.

BIBLIOGRAPHY

1. R. E. Kalman, "A New Approach to Linear Filtering on Prediction Problems", ASME Trans., J. Basic Eng., Vol. 83, March 1960, pp. 34-35.

2. Nahi, N. E., Estimation Theory with Applications, John Wiley and Sons, Inc., New York 1969.

3. Lee, C. K. Robert, "Optimal Estimation, Identification and Control", Book Research Monograph No. 28, The M.I.T. Press, Cambridge, Mass., 1964.

4. By: Technical Staff - The Analytic Science Corporation. Applied Optimal Estimation, the M.I.T. Press, 1971.

5. Franks, L. E., "A Model for Random Video Process", Bell System Tech J., Vol. 45, April 1966, pp. 609-630.

6. N. E. Nahi, "Role of Recursive Estimation in Image Enhancement", Proc. IEEE, Vol. 60, No. 7, July 1972, pp. 872-878 (Special Issue on Digital Picture Processing).

7. N. E. Nahi and T. Assefi, "Bayesian Recursive Image Estimation", IEEE Trans. On Comp., VC-21 No. 7, July 1972, pp. 734-738.

8. S. R. Powell and L. M. Silvermann, "Modeling of Two Dimensional Covariance Functions with Application to Image Restoration", IEEE Trans. Aut. Contr., Vol. AC-19, No. 1, Febr. 1974, pp. 8-12.

9. N. E. Nahi and C. A. Franco, "Bayesian Recursive Image Estimation", IEEE Trans. on Comp., VC-21 no. 7, July 1972, pp. 734-738.

10. Rosenfeld, A. and Kak, C. A., Digital Picture Processing, Academic Press, 1976, pp. 238-249.

11. A. K. Jain and E. Angel, "Image Restoration, Modeling, and Reduction of Dimensionality", IEEE Trans. on Comp., VC-23, no. 5, May 1974.

12. Roesser, R. P., "A Discrete Spare Model for Linear Image Processing" IEEE Trans. on Auto Control, vol. AC-20, no. 1, pp. 1-10, February, 1975.

13. A. Habibi, "Two Dimensional Bayesian Estimation of Images", Proc. IEEE, vol. 60, no. 7, July 1972, pp. 878-884.

14. G. M. Strintzis, "Comments On Two Dimensional Bayesian Estimate of Images", Proceedings of the IEEE, Vol. 64, no. 8, 1976, pp. 1255-1257.

15. T. Assefi, "Bayesian Recursive Image Estimation", Ph.D. dissertation, University of Southern California, September 1973.

16. S. Power, "Modeling of Nonstationary Stochastic Processes with Application to Image Restoration", Ph.D. dissertation, University of Southern California, 1973.

17. M. Naraghi, "An Algorithmic Image Estimation Method Applicable to Non-Linear Observations", University of Southern California (Image Processing Institute), June 1975.

18. B. D. O. Anderson, "The Inverse Problem of Stationary Covariance Generation", Journal of Stat. Physics, vol. 1, no. 1, 1969.

19. B. D. O. Anderson and J. B. Moore, "Spectral Factorisation of Time-Varying Functions", IEEE Trans. on Inf. Th., Vol. IT-15, no. 5, Sept. 1969.

20. N. Weiner, The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications, John Wiley and Sons, Inc., New York, 1949.

21. Duda and Hart, Pattern Classification and Scene Analysis, John Wiley and Sons, 1973, pp. 318-322.

22. P. E. Barry, R. Gran, C. R. Waters, "Techniques for Optimal Image Processing-State Variable Extensions of One Dimensional Observers", Research Department, Gruman Aerospace Corporation, Bethpage, New York, 11714.

23. John W. Woods, "Two Dimensional Discrete Markovian Fields", IEEE, Transaction on Information Theory, March 1972.

24. Huang, T. S., "Stability of Two Dimensional Recursive Filters", IEEE Trans. Audio Electroacous., Vol. AV-20, pp. 158-163, June 1972.

25. M. P. Ekstrom and J. W. Woods: "Two Dimensional Spectral Factorization with Applications in Recursive Digital Filtering", IEEE Trans. ASSP, Vol. ASSP-24, pp. 115-128, 1976.

26. A. Papoulis, Probability, Random Variables and Stochastic Processes, New York: McGraw-Hill, 1965.

27. Donald E. Kirk, "Optimal Estimation: An Introduction to The Theory and Application". Notes for the use of the students at the Naval Postgraduate School.

28. J. S. Meditch, Stochastic Optimal Linear Estimation and Control, McGraw-Hill, 1969.

29. H. W. Sorenson, "Kalman Filtering Techniques", Advances in Control Systems, Vol. 3, C. T. Leondes, Editor, Academic Press, pp. 219-292, 1966.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 1422 Naval Postgraduate School Monterey, California 93940	2
3.	Electrical Engineering Department Chairman Code 62 Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
4.	Professor H. Titus, Code 62Ts Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	10
5.	Professor Sydney R. Parker, Code 62Px Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
6.	Professor Ku, Code 62Ku Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
7.	Professor T. F. Tao, Code 62Tv Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
8.	LT. Shachar Moshe 9 Montefiory, Haifa Israel	2
9.	LT. Zach Shlomo SMC 2059 Naval Postgraduate School Monterey, California 93940	1
10.	LT. Bar Yehoshua David Israeli Navy, c/o Embassy of Israel	1

11. LT Even Or 1
SMC 2190
Naval Postgraduate School
Monterey, California 93940

12. Dr. M.G. Srinzits 1
Electrical Engineering Department
University of Pittsburgh
Pittsburgh, Pa. 15261

13. Professor C. O. Wilde, Code 53Wm 1
Department of Mathematics
Naval Postgraduate School
Monterey, California 93940

14. Professor D.E. Kirk, Code 62 1
Department of Electrical Engineering
Naval Postgraduate School
Monterey, California 93940

15. Dr. John Roese 1
Naval Undersea Center
San Diego, California 92132

16. LT. Martinos Thon Fragiskos 1
SMC 2378
Naval Postgraduate School
Monterey, California 93940

17. LT. Epitas Kayhan 1
SMC 1848
Naval Postgraduate School
Monterey, California 93940