State University of New York
at Stony Brook

College of Engineering

Technical Report No. 301

by

Sheldon S. L. Chang
Professor of Electrical Engineering

August 26, 1977

"DIGITAL LINEAR PROCESSOR THEORY and OPTIMUM
MULTIDIMENSIONAL IMAGE RECONSTRUCTION

Abstract

This paper introduces frame recursive image processing as a new algorithm for processing of blurred or unblurred pictorial information with additional noise. It gives improved image which approaches optimum in least mean square error sense. The method represents a new direction in two dimensional digital filtering from the current trend of using generating equations and Kalman filter which requires artificial introduction of a causal order of data points.

Applications include two dimensional image processing, three dimensional image reconstruction from two dimensional projections and from two dimensional cross-sections, and real time image processing of a moving object. In all cases the optimum linear processor utilizes all available information on second statistical moments to give least mean square error, and is realized by frame recursive processing in successive approximation with exponentially decaying error.

I. Introduction

Image enhancement and restoration is a subject of broad interest because of its potential in improving the quality of a picture which has been blurred or contaminated by noise or both. The improved picture is then used either as an end product or as an intermediate step for further computerized processing and identification. Important theoretical and experimental progress has been made in recent years as the consequence of a linear assumption: that the light or gray level at each point is representable as a signal, blurring is representable as two-dimensional convolution or sum, and the picture noise is additive. Exact analysis is then further facilitated by an invariance assumption on the blurring and noise processes. [1]- [12].

The above mentioned linear assumption places the image enhancement problem right into the domain of classical optimum linear filter theory with the consequent natural branching into Wiener and Kalman filters. With the Wiener filter or spatial frequncy domain approach, derivation of the optimum filter is a complete parallel to the known temporal case. However realization of the optimum filter is not as straightforward. It can be done in two ways: (i) optical realization [9], [10], and (ii) Fourier transform from image to spatial frequency domain, filtering, and inverse transform back to image domain [4], [5]. Optical realization requires elaborate laboratory setup and very high grade experimental skill. Fourier transform realization requires extensive computation [13] and is also open to the criticism that while deblurring and image enhancement are usually local operations, Fourier transfrom back and forth involves every point in the picture domain, and quantization errors are accumulative. The Kalman filter approach leads directly to a recursive digital filter which is easily implementable [1]-[3], [8]. But it requires artificial introduction of a causal order of data points. The filter is not optimal as information contained in the artificially assigned future data points are already known but not utilized. The sub-optimality can be improved but not removed by averaging output levels obtained by filtering with various causal directional assignments. Roesser in the concluding section of his work [8] stated as number one future research objective generalization to bilateral models, meaning models without causality. The difficulty is that point by point recursive processing implies point by point causality, unless the newly computed value of a data point is not used in computing the values of other data points. However there is not causality constraint on using computed values of all data points for a succeeding round of computation on all data points, which will be referred to as recursive frame processing.

The present paper gives a frame recursive method for realizing Wiener filter. A simple finite impulse response filter (FIR) is obtained readily from the optimum Wiener filter, and will be referred to as frame processor (FP). By repeated operation of FP on the multi-dimensional signal, optimaility can be approximated to any desired degree, and the additional mean square error (in excess that of the optimum Wiener filter) decreases exponentially with the number of repeated operations. An integral expression for the additional mean square error and its upper bound are derived in the paper. While any finite number of operations of the FP is equivalent to another FIR, the number of computing operations with frame recursive processing is an order of magnitude lower than that of the equivalent FIR.

In addition to image enhancement, the paper also gives solutions to two other significant applications: (i) optimum reconstruction of a three dimensional object from two dimensional slices, which is of high interest to microbiology and medical sciences, and (ii) optimum detection or reconstruction of a moving object from a succession of two-dimensional pictures. Solution to (i) is an extension of the classical z-transform theory with multiple sampling periods. Solution to (ii) is a combination Wiener-Kalman filter. Both optimum solutions can be realized by the frame recurisve method in the successive approximation sense as discussed above.

The mathematical basis of the frame recursive method is an algebraic expansion theorem which states that if a polynomial $D(\vec{z})$ is real and non-zero on the unit circle product space $C\{\vec{z}: |z_i| = 1\}$, then a unique and convergent series expansion of $1/D(z)$ exists, which can be successively approximated term by term by expanding $^1/D(z)$ in a prescribed manner. The first part of the theorem on the existence of a convergent expansion has been proved by Justice and Shanks [14] using a Tauberian theorem due to Wiener in a more general context. However, nowhere in Justice and Shanks' proof is an indication of how the expansion is to be made nor how the speed of convergence is to be calculated. Proof of the present version of the theorem is given in the paper as an Appendix.

Section II gives the main results of linear processor theory in two-dimensions. Section III gives its application in two-dimensional image processing, and demonstrates frame recursive processing scheme for a numerical example. Section IV generalizes linear processor to N spatial dimensions with a signal vector on each spatial point. It gives a solution to the problem of optimum reconstruction (least mean square error) of a three dimensional object from its two-dimensional projections or two dimensional cross-sections which can be blurred as well as contaminated with additive noise. Section V gives optimum real-time processing of successive images of a moving object. Section VI gives concluding remarks, and mathematical development of the linear processor theory is given in Section VII.

In the paper, ξ, σ, ρ, λ, r, and s represent spatial coordinates and z, and w represent transform variables. Signal and observed signal are denoted by x and y while noise and other random processes are represented by n, u, and v. The impulse response function of linear processors are represented by f, g, and p. The corresponding transformed variables and functions are represented by capitalized letters.

An arrow over head $\vec{\xi}$, $\vec{z}$, etc, represent spatial vectors while bold face symbols $\underset{\sim}{x}$, $\underset{\sim}{X}$, $\underset{\sim}{f}$ and $\underset{\sim}{F}$ etc represent signal vectors and matrices. An overhead bar represents $z_i \rightleftarrows z_i^{-1}$, i = 1, 2, ...N, in a function, and upperscript T represents transpose of a matrix or vector, while † represents both operations. It is the Hermitian adjoint symbol for matrix functions with $\vec{z}$ on C.

Symbols are introduced as they are used except in Section VII, where the symbols introduced in Section IV are assumed.

II. Two Dimensional Z-transform

Let $x(r,s)$ be defined for integer values of r and s. The Z-transform of $x(r,s)$ is defined as

$$X(z,w) = \sum_{r=-\infty}^{r=\infty} \sum_{s=-\infty}^{s=\infty} x(r,s)\, z^{-r} w^{-s} \tag{1}$$

Capital letters are used to denote corresponding transformed variables. A linear processor (L.P.) is a noncausal filter. The processed output of $x(r,s)$ with invariant L.P. $f(\rho,\sigma)$ is given by

$$y(r,s) = \sum_{\rho=-\infty}^{\rho=\infty} \sum_{\sigma=-\infty}^{\sigma=\infty} f(\rho,\sigma)\, x(r-\rho,\ s-\sigma) \tag{2}$$

In terms of transformed variables, (2) can be expressed as

$$Y(z,w) = F(z,w)\, X(z,w) \tag{3}$$

A L.P. $f(\rho,\sigma)$ is said to be stable if there are constants A and γ such that $A>0$, $0 < \gamma < 1$, and

$$|f(\rho,\sigma)| < A\gamma^{|\rho| + |\sigma|}$$

In the product space of the complex planes of z and w, the product of the two unit circles: $\{(z,w);\ |z| = |w| = 1\}$ is denoted as C. The open set $\{(z,w);\ 1-\varepsilon < |z| < 1+\varepsilon,\ 1-\varepsilon < |w| < 1+\varepsilon\}$, $\varepsilon > 0$ is referred to as ε - neighborhood of C, $N_c(\varepsilon)$, or simply N_c. The existence of some $\varepsilon > 0$ is then implied. For a stable L.P., the inverse of (1) is

$$x(r,s) = \frac{1}{(2\pi j)^2} \oint\!\!\oint_{(z,w)\text{ on } C} X(z,w)\, z^{r-1} w^{s-1}\, dz\,dw \tag{4}$$

The following theorems are proved in Section VII.

<u>Theorem 2</u> Every stable L.P. is convergent unconditionally, absolutely in N_c.

<u>Theorem 3</u> Let $D(z,w)$ be a polynomial of z and w, and $D(z,w)$ is real

and positive (non zero) on C. Then the following:

(i) There exist D_u, and $D_\ell > 0$:

$$D_u = \underset{\vec{z}\varepsilon C}{\text{Max}}\ D(z, w)$$

$$D_\ell = \underset{\vec{z}\varepsilon C}{\text{Min}}\ D(z,w)$$

(ii) The expansion

$$\frac{1}{D(z,w)} = \frac{1}{A-[A-D(z,w)]}$$

$$= \frac{1}{A} \sum_{n=0}^{n=\infty} \left(1 - \frac{D(z,w)}{A}\right)^n \qquad (5)$$

where $A > \frac{D_u}{2}$, converges for every z on C.

(iii) Let f(r,s) and $f_L(r,s)$ be defined by

$$f(r,s) = \frac{1}{(2\pi j)^2} \underset{(z,w)\varepsilon C}{\oint\oint} \frac{z^{r-1}\ w^{s-1}}{D(z,w)}\ dzdw$$

$$f_L(r,s) = \text{Coefficient of } z^{-r}\ w^{-s} \text{ in } \frac{1}{A} \sum_{n=0}^{n=L} \left(1- \frac{D(z,w)}{A}\right)^n$$

Then

$$\lim_{L\to\infty} f_L(r,s) \to f(r,s)$$

(iv) Let F(z,w) be defined in terms of f(r,s) following the convention expressed in (1). F(z,w) is the unique, stable, expansion of 1/D(z,w) in N_c.

Theorem 4 A L.P. of finite terms is stable. Let $F(z,w)$ and $G(z,w)$ be stable L.P.'s. Then the L.P. $F(z,w)\ G(z,w)$ is stable.

Theorem 5 Let N and D denote polynomials of z and w such that

$$D(z,w) \neq 0$$

for (z,w) on C. Then the rational function

$$P(z,w) \triangleq \frac{N(z,w)}{D(z,w)}$$

has a unique, stable expansion.

Let $x(r,s)$, $y(r,s)$ be stationary random variables with zero mean, a correlation function $\phi_{xy}(\rho,\sigma)$ can be defined:

$$\phi_{xy}(\rho,\sigma) = E\{x(r,s)\ y(r-\rho,\ s-\sigma)\} \tag{6}$$

The spectral density function $\Phi_{xy}(z,w)$ is the transformed ϕ_{xy}

$$\Phi_{xy}(z,w) = \sum_{\rho}\sum_{\sigma}\ \phi_{xy}(\rho,\sigma)\ z^{-\rho}\ w^{-\sigma} \tag{7}$$

Theorem 6 Let x, y denote processed output of stationary signals u, v with invariant stable L.P.'s f and g:

$$X(z,w) = F(z,w)\ U(z,w)$$

$$Y(z,w) = G(z,w)\ V(z,w)$$

Then

$$\Phi_{xy}(z,w) = F(z,w)\ \bar{G}(z,w)\Phi_{uv} \text{ where } \bar{G}(z,w) \triangleq (z^{-1},w^{-1}). \tag{8}$$

Let $[H]_0$ denote the constant term in the series expansion of $H(z,w)$. Then

$$E[x^2] = \phi_{xx}(0,0) = [\Phi_{xx}]_0 > 0 \tag{9}$$

III Optimum Image Processing

Let y be observed x with mask M and linear additive noise n:

$$Y = MX + N \tag{10}$$

where n is uncorrelated to x.

A linear processor P is to be constructed to estimate x from y with least mean square error:

$$\hat{X} = PY \tag{11}$$

$$E(x-\hat{x})^2 = E\,\tilde{x}^2 = \text{minimum} \tag{12}$$

From (10) and (11)

$$X - \hat{X} = (1-PM)\,X - PN \tag{13}$$

From (8)

$$\begin{aligned} \Phi_{\tilde{x}\tilde{x}} &= (1-PM)\,(1-\bar{P}\bar{M})\,\Phi_{xx} + P\bar{P}\,\Phi_{nn} \\ &= \Phi_{xx} - PM\,\Phi_{xx} - \bar{P}\bar{M}\,\Phi_{xx} + P\bar{P}\,\Phi_{yy} \end{aligned} \tag{14}$$

Let P_0 denote the processor $\bar{M}\,\Phi_{xx}/\Phi_{yy}$. Equation (14) can be rewritten as

$$\Phi_{\tilde{x}\tilde{x}} = \Phi_{yy}\,(P-P_0)\,(\bar{P}-\bar{P}_0) + \frac{\Phi_{xx}\Phi_{nn}}{\Phi_{yy}} \tag{15}$$

From (9)

$$E\,\tilde{x}^2 = \left[\frac{\Phi_{xx}\Phi_{nn}}{\Phi_{yy}}\right]_0 + \left[(P-P_0)(\bar{P}-\bar{P}_0)\,\Phi_{yy}\right]_0 \tag{16}$$

The first term on the RHS of (16) is positive and represents the minimum error to be obtained with $P = P_0$. The second term is also positive if $P \neq P_0$, and represents the additional mean square error when a processor other than P_0 is used. <u>P_0 is the optimum processor.</u>

Approximate Realization of P_O by Frame Recursive Processing

In general P_o is of the form

$$P_o = \frac{N(z,w)}{D(z,w)} \tag{17}$$

where D(z) is real and positive on C. From (5), P_o can be expanded as

$$P_o = \frac{N(z,w)}{A} \sum_{n=0}^{n=\infty} Q(z,w)^n \tag{18}$$

where

$$Q(z,w) = 1 - \frac{D(z,w)}{A} \tag{19}$$

and, for (z,w) on C:

$$|Q(z,w)| \leq Q_m < 1 \tag{20}$$

Equation (20) defines Q_m as the modulus of Q(z,w). While (20) is satisfied by choosing any A greater than $\frac{D_u}{2}$, the lowest Q_m is obtained with

$$A = \frac{1}{2}(D_u + D_\ell) \tag{21}$$

and the resultant Q_m is:

$$Q_m = \frac{D_u - D_\ell}{D_u + D_\ell} \tag{22}$$

Let P be defined by

$$P_\ell(z,w) \triangleq \frac{N(z,w)}{A} \sum_{n=0}^{n=\ell} Q(z,w)^n \tag{23}$$

Then

$$P_o - P_\ell = Q(z,w)^{\ell+1} P_o \tag{24}$$

Equation (24) shows that the optimum processor P_o can be approximated to any desired degree with P_ℓ by choosing ℓ sufficiently large. From (16) and (24), the additional mean square error Δ_ℓ with P_ℓ is given by

$$\Delta_\ell = [(Q\bar{Q})^{\ell+1} P_o \bar{P}_o \Phi_{yy}]_0 \tag{25}$$

Since $P_o \bar{P}_o \Phi_{yy}$ is real and non-negative on C, it is readily shown from (20) and (25)

$$\Delta_\ell \le Q_m^{2\ell+2} [P_o \bar{P}_o \Phi_{yy}]_0 \tag{26}$$

Obviously P is a finite impulse response filter (FIR), and can be realized as such. An alternative method which requires considerably less computation is the frame recursive processing (FRP) method which consists of the following steps:

$$\begin{aligned}
&(1)\ X_0 \leftarrow \frac{N(z_1 w)}{A} \circledast Y \\
&(2)\ Y_1 \leftarrow Q(z, w)\, X_0 \\
&(3)\ X_0 \leftarrow X_0 + Y_1 \\
&(4)\ Y_1 \leftarrow Q(z,w) Y_1
\end{aligned} \tag{27}$$

In the first step, $N(z,w)Y/A$ is entered into the memory space for X_0. In the second step $Q(z,w)X_0$ is entered into the memory space for Y_1. In the third and fourth steps X_0 is replaced by X_0+Y_1, and then Y_1 by $Q(z,w)Y_1$. The last steps are repeated ℓ -1 times or until no further improvement is noticeable.

The savings in computation can be quantified as follows: Let multiplying two numbers and adding the product to a third number be defined as one computation. The number of computations per data point with FRP is

$$C_{FR} = T_N + \ell\, T_Q \tag{28}$$

where T_N is the number of terms in $N(z,w)$, and T_Q is the number of terms in $Q(x,w)$. The number of computations per data point with FIR is equal to the number of terms in P_ℓ (T_p) which is usually much larger.

As an example, let $Q(z,w)$ be a polynomial with powers of z between $-n_1$ and n_1, and powers of w between $-n_2$ and n_2. Similarly, let $N(z,w)$ be a polynomial with powers n_1' and n_2'. Then

$$T_Q = (2n_1 + 1)\ (2n_2 + 1)$$

$$T_N = (2n_1' + 1)\ (2n_2' + 1)$$

and from (28)

$$C_{FRP} = (2n_1' + 1)\ (2n_2' + 1) + \ell(2n_1 + 1)\ (2n_2 + 1) \tag{29}$$

The number of computations for each data point with FIR realization is equal to T_p

$$C_{FIR} = [2(n_1' + \ell n_1)+1]\ [2(n_2' + \ell n_2)+1] \tag{30}$$

As a typical example, let $n_1' = n_2' = 1$, $n_1 = n_2 = 2$, and $\ell = 40$. Then

$$C_{FRP} = 9 + 40 \times 25 = 1{,}009$$

$$C_{FIR} = 163^2 = 26{,}569$$

In addition to the numerical difference, frame recursive processing has the following advantages in computation:

(i) Its repetitiousness in instruction steps and operand locations can be utilized to save computation time

(ii) M does not have to be selected beforehand. Steps (3) and (4) can be repeated until no further improvement is noticeable. It is a natural adaptive system.

Example 1 [3], The signal $x(r,s)$ is generated by the following process:

$$x(r+1,\ s+1) = \rho_1\ x(r+1,\ s) + \rho_2 x(r,\ s+1) - \rho_1\rho_2\ x(r,s)$$
$$+ \sqrt{(-\rho_1^2)(1-\rho_2^2)}\ u\ (r,s) \tag{31}$$

The observed signal y is

$$y(r,s) = x(r,s) + v(r,s)$$

where u, v are white processes with zero mean and mean square values S and N respectively. Determine the optimum processor P_o.

Solution The spectral functions for u and v are $\Phi_{uu} = S$ and $\Phi_{vv} = N$. Equation (31) can be written in terms of transformed variables:

$$F(z,w) = \frac{\sqrt{(1-\rho_1^2)(1-\rho_2^2)}}{(1-\rho_1\ z^{-1})(1-\rho_2 w^{-1})}$$

$$X(z,w) = F(z,w)\ U\ (z,w)$$

from (8)

$$\Phi_{xx} = \frac{(1-\rho_1^2)\ (1-\rho_2^2)\ S}{(1-\rho_1\ z^{-1})(1-\rho_2 w^{-1})(1-\rho_1 z)(1-\rho_2 w)}$$

$$= \frac{B}{[1-\alpha_1(z^{-1}+z)]\ [1-\alpha_2(w^{-1}+w)]}$$

where

$$B = \frac{(1-\rho_1^2)(1-\rho_2^2)S}{(1+\rho_1^2)(1+\rho_2^2)}$$

$$\alpha_{1,2} = \frac{\rho_{1,2}}{1+\rho_{1,2}^2}$$

and M = 1. The optimum processor is

$$P_o = \frac{\Phi_{xx}}{\Phi_{yy}} = \frac{B}{B+[1-\alpha_1(z^{-1}+z)][1-\alpha_2(w^{-1}+w)]N} \tag{32}$$

Let $N(z,w) = 1$ in (32), then,

$$D(z,w) = 1 + \frac{N}{B}[1-\alpha_1(z^{-1}+z)][1-\alpha_2(w^{-1}+w)]$$

From (19) and (21)

$$A = 1 + \frac{N}{B}(1+ 4\alpha_1\alpha_2)$$

$$Q(z,w) = \frac{N[4\alpha_1\alpha_2+\alpha_1(z^{-1}+z) + \alpha_2(w^{-1}+w) -\alpha_1\alpha_2(z^{-1}+z)(w^{-1}+w)]}{B + N(1+4\alpha_1\alpha_2)}$$

$$Q_m = \frac{2N(\alpha_1+\alpha_2)}{B+N(1+4\alpha_1\alpha_2)}$$

The processor P_o can be approximated to any desired degree by steps (i) to (iv) of (27)

For a typical numerical illustration,

let

$\rho_1 = \rho_2 = 0.726$

$S = 6.61$, $N = 9$

Then $B = 0.634$, $\alpha_1 = \alpha_2 = 0.475$

$Q_m = 0.963$, and $Q(z,w)$ is given by the array of Table 1:

Table 1: Array Q(z,w)

r / S	-1	0	1
-1	-.1144	.2407	-.1144
0	.2407	.4574	.2407
1	-.1144	.2407	-.1144

The mean square error from the optimum processor is calculated from (4) and (16) with r = s = o and $P - P_o = 0$. Its value is $E_o = 1.964$. The mean square error from FRP is $E_o + \Delta_\ell$ with Δ_ℓ calculated from (25). The result is then plotted as the solid curve in Figure 1 versus the number of recursive operations $\ell+1$. The upper bound of the mean square error is obtained from (26) and is plotted as the broken curve in Fig. 1. It is noted that (i) the mean square error from FRP converges to within 1 db of the optimum with $\ell=24$, and (ii) the broken curve gives a close upper bound to the solid curve for all value of ℓ.

Example 2 A picture is blurred with

$$y(r,s) = 0.5X(r,s) + 0.5X(r+1,s) + u(r,s) \tag{33}$$

where $\Phi_{xx} = S$, and $\Phi_{uu} = N$. Determine the optimum processor for the following cases:

(i) $\frac{N}{S} = 0$, and (ii) $\frac{N}{S} = 0.1$.

<u>Solution</u> Equation (33) is transformed into

$$Y = MX + U$$

where $M = 0.5\ (1+z)$. The optimum processor is

$$P_o = \frac{0.5(1+z^{-1})S}{0.25(1+z)(1+z^{-1}) + N}$$

Case (i) $N = 0$

$$P_o = \frac{2}{1+z}$$

There is no convergent expansion in the neighborhood of the unit circle. However, we can use the following approximation:

$$P_{oa} = \frac{1}{\lambda z+1} + \frac{1}{z+\lambda}$$

where $\lambda = 1-\varepsilon$, and ε is very small but positive.

$$P_{oa} = \frac{z^{-1}}{1+\lambda z^{-1}} + \frac{1}{1+\lambda z}$$

$$= (1+z^{-1}) - \lambda\ (z+z^{-2}) + \lambda^2\ (z^2+z^{-3})\ --- \tag{34}$$

Case (ii) $\frac{N}{S} = 0.1$

$$P_o = \frac{0.5(1+z^{-1})}{0.25(1+z)(1+z^{-1}) + 0.1}$$

$$= \frac{0.6985}{z+0.5367} + \frac{0.6985}{1+0.5367z}$$

$$= 0.6985(1+z^{-1}) - 0.3749(z+z^{-2})$$

$$+ 0.2012(z^2+z^{-3}) - 0.1080(z^3+z^{-4}) + \ldots \tag{35}$$

Equation (34) and (35) provide an interesting contrast. While the optimum deblurring filter without noise is not convergent, it converges rapidly in the presence of noise.

IV Generalization to N-Dimensional

The generalization will be in two ways: (i) instead of two-dimensional coordinates r and s, the coordinates will be N-dimensional $\xi_1, \xi_2 \cdots \xi_N$, and (ii) the signal is an n-component vector on each space point:

$$\underset{\sim}{x}(\vec{\xi}) = \begin{pmatrix} x_1(\vec{\xi}) \\ x_2(\vec{\xi}) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n(\vec{\xi}) \end{pmatrix} \tag{36}$$

In (36), an overhead arrow is used to denote a space vector and bold face letter $\underset{\sim}{x}$ is used to denote signal vector. The z-transform of $\underset{\sim}{x}(\vec{\xi})$ is

$$\underset{\sim}{X}(\vec{z}) = \sum_{\xi_1=-\infty}^{\xi_1=\infty} \sum_{\xi_2=-\infty}^{\xi_2=\infty} \cdots \sum_{\xi_N=-\infty}^{\xi_N=\infty} \underset{\sim}{x}(\vec{\xi})\, z_1^{-\xi_1} z_2^{-\xi_2} \cdots z_n^{-\xi_n} \tag{37}$$

Let U_M denote the cube $-M \leq \xi_i \leq M$, $i = 1, 2 \cdots N$. Let $\vec{z}^{\,\vec{\xi}}$ and $d\vec{z}$ denote the products

$$\prod_{i=1}^{i=N} z_i^{\xi_i} \text{ and } dz_1\, dz_2 \cdots dz_n$$

respectively. The definitions of C and N_c are now generalized to the N-fold complex space. Equation (37) can be written simply as

$$\underset{\sim}{X}(\vec{z}) = \sum_{\vec{\xi}\in U_\infty} \underset{\sim}{x}(\vec{\xi})\, \vec{z}^{\,-\vec{\xi}}$$

The linear processor $\underset{\sim}{F}$ is a mxn matrix, with each element F_{ij} being an N-dimensional linear processor:

$$F_{ij}(\vec{z}) = \sum_{\vec{\xi}\in U_\infty} f_{ij}(\vec{\xi})\vec{z}^{\,-\vec{\xi}}$$

Output of $\underset{\sim}{F}$ is an m-dimensional vector $\underset{\sim}{y}(\vec{\xi})$ with its transform given by the matrix equation:

$$\underset{\sim}{Y}(\vec{z}) = \underset{\sim}{F}(\vec{z})\, \underset{\sim}{X}(\vec{z})$$

Equation (5) is replaced by the matrix equation

$$\underset{\sim}{\Phi}_{xy}(\vec{z}) = \sum_{\vec{\xi}\in U_\infty} \underset{\sim}{\phi}_{xy}(\vec{\xi})\vec{z}^{\,-\vec{\xi}} \tag{38}$$

where the i,jth elements of $\underset{\sim}{\phi}_{xy}$ and $\underset{\sim}{\Phi}_{xy}$ are the correlation function $\phi_{x_iy_j}$, and spectral density function $\Phi_{x_iy_j}$ respectively. Equation (6) becomes

$$\underset{\sim}{\Phi}_{xy}(\vec{z}) = \underset{\sim}{F}(\vec{z})\, \underset{\sim}{\Phi}_{uv}(\vec{z})\underset{\sim}{G}(\vec{z})^{\dagger} \tag{39}$$

where

$\underset{\sim}{G}(\vec{z})^{\dagger} = \underset{\sim}{G}(z_1^{-1}, z_2^{-1}, \ldots z_N^{-1})^T$. It is the Hermitian adjoint of $\underset{\sim}{G}(\vec{z})$ on C. Theorems 1 to 6 remain valid. In fact, they are proved in the Appendix for the N-dimensional case.

For N-dimensional optimum processing (10), (11), and (14) are replaced by

$$\underset{m}{Y} = \underset{m}{M}\,\underset{m}{X} + \underset{m}{N} \tag{40}$$

$$\underset{m}{\hat{X}} = \underset{m}{P}\,\underset{m}{Y} \tag{41}$$

$$\underset{m}{\Phi}_{\tilde{x}\tilde{x}} = (\underset{m}{1} - \underset{m}{P}\,\underset{m}{M})\underset{m}{\Phi}_{xx}\,(\underset{m}{1}-\underset{m}{P}\,\underset{m}{M})^{\dagger} + \underset{m}{P}\,\underset{m}{\Phi}_{nn}\,\underset{m}{P}^{\dagger}$$

$$= \underset{m}{\Phi}_{xx} - \underset{m}{P}\,\underset{m}{M}\,\underset{m}{\Phi}_{xx} - \underset{m}{\Phi}_{xx}\,\underset{m}{M}^{\dagger}\underset{m}{P}^{\dagger} + \underset{m}{P}\,\underset{m}{\Phi}_{yy}\,\underset{m}{P}^{\dagger} \tag{42}$$

where

$$\underset{m}{\Phi}_{yy} = \underset{m}{M}\,\underset{m}{\Phi}_{xx}\,\underset{m}{M}^{\dagger} + \underset{m}{\Phi}_{nn} \tag{43}$$

The optimum processor is given as

$$\underset{m}{P}_{o} = \underset{m}{\Phi}_{xx}\,\underset{m}{M}^{\dagger}\,\underset{m}{\Phi}_{yy}^{-1} \tag{44}$$

Equation (16) becomes

$$E\left[\underset{m}{\tilde{x}}\,\underset{m}{\tilde{x}}^{\dagger}\right] = \left[\underset{m}{\Phi}_{xx} - \underset{m}{\Phi}_{xx}\,\underset{m}{M}^{\dagger}\underset{m}{\Phi}_{yy}^{-1}\,\underset{m}{M}\,\underset{m}{\Phi}_{xx}\right] + \left[(\underset{m}{P} - \underset{m}{P}_{o})\,\underset{m}{\Phi}_{yy}\,(\underset{m}{P} - \underset{m}{P}_{o})^{\dagger}\right] \tag{45}$$

The two terms on the RHS of (45) have the same signifances as the corresponding terms on the RHS of (16). The matrix $(\underset{m}{P} - \underset{m}{P}_{o})^{\dagger}$ is the Hermitian conjugate of $\underset{m}{P} - \underset{m}{P}_{o}$ on C. Any processor other than $\underset{m}{P}_{o}$ gives a positive second term and adds to the error matrix

The Projection Operator Π_i

An N-dimensional signal can be projected onto an (N-1) dimensional manifold:

$$(\Pi_i\,\underset{m}{X})(\xi_1, \xi_2, \ldots \xi_{i-1}, \xi_{i+1} \ldots \xi_N)$$

$$= \sum_{\xi_i=-\infty}^{\xi_i=\infty} \underset{m}{X}(\xi_1, \xi_2, \ldots \xi_i \ldots \xi_N) \tag{46}$$

In terms of z-transform it is

$$\Pi_i = \sum_{\xi_i=-\infty}^{\xi_i=\infty} z_i^{-\xi_i} \tag{47}$$

As (47) cannot be written in closed form, an exponentially decaying projection $\Pi_{i\lambda}$ is used:

$$\Pi_{i\lambda} = \sum_{\xi_i=-\infty}^{\xi_i=\infty} \lambda^{|\xi_i|} z_i^{-\xi_i} = \frac{1-\lambda^2}{(1-\lambda z_i^{-1})(1-\lambda z_i)} \tag{48}$$

where $0< \lambda < 1$. Let $\underset{\sim}{\Pi}_\lambda$ denote $(\Pi_{1\lambda}, \Pi_{2\lambda}, \Pi_{3\lambda} \ldots)^T$.

The problem of reconstructing a scalor signal in N-dimensional space, from (N-1) dimensional projections can be formulated as follows:

$$\underset{\sim}{Y} = \underset{\sim}{\Pi}_\lambda X + \underset{\sim}{N} \tag{49}$$

Equation (44) gives the optimum processor as

$$\underset{\sim}{P}_o = \Phi_{xx} \underset{\sim}{\Pi}_\lambda^T \underset{\sim}{\Phi}_{yy}^{-1} = \underset{\sim}{\Pi}_\lambda^T (\Phi_{xx}^{-1} \underset{\sim}{\Phi}_{yy})^{-1} \tag{50}$$

where

$$\Phi_{xx}^{-1} \underset{\sim}{\Phi}_{yy} = \underset{\sim}{\Pi}_\lambda \underset{\sim}{\Pi}_\lambda^T + \Phi_{xx}^{-1} \underset{\sim}{\Phi}_{nn} \tag{51}$$

The Sampling Operator*

The sampling operator samples on one dimension only, z_1, and takes one frame out of every T samples:

$$\underset{\sim}{X}^* (\vec{z}) = \sum_{r=-\infty}^{r=\infty} \sum_{\vec{\xi}'} \underset{\sim}{x} (rT, \vec{\xi}') \, z_1^{rT} \vec{z}'^{-\vec{\xi}'} \tag{52}$$

where $\vec{z}'$ and $\vec{\xi}'$ denote $(z_2, z_3 \ldots z_N)$ and $(\xi_2, \xi_3, \ldots \xi_N)$. The sampled signal or processor can be treated as a transformed variable with $\underset{m}{x}(\vec{\xi}) = 0$ for all values of ξ_1 which are not integer multiples of T. Consequently

$$(\underset{m}{X}^*)^* = \underset{m}{X}^* \tag{53}$$

$$(\underset{m}{X}^* \underset{m}{F})^* = \underset{m}{X}^* \underset{m}{F}^* \tag{54}$$

$$(\underset{m}{F} \underset{m}{X}^*)^* = \underset{m}{F}^* \underset{m}{X}^* \tag{55}$$

Let $\underset{m}{X} = \underset{m}{F} \underset{m}{U}^*$, and $\underset{m}{Y} = \underset{m}{G} \underset{m}{V}^*$ where $\underset{m}{u}, \underset{m}{v}$ are correlated stationary random processes. Then for any $\vec{\xi}_a$ and $\vec{\tau}$

$$\underset{m}{x}(\vec{\xi}_a) = \sum_r \sum_{\vec{\xi}'} \underset{m}{f}(\xi_{1a} - rT, \vec{\xi}_a' - \vec{\xi}') \underset{m}{u}(rT, \vec{\xi}') \tag{56}$$

$$E\{\underset{m}{x}(\vec{\xi}_a) \underset{m}{v}^T (\vec{\xi}_a - \vec{\tau})\} = \sum_r \sum_{\vec{\xi}'} \underset{m}{f}(\xi_{1a} - rT, \vec{\xi}_a' - \vec{\xi}') \underset{m}{\phi}_{uv}(rT - \xi_{1a} + \tau_1, \vec{\xi}' - \vec{\xi}_a' + \vec{\tau}') \tag{57}$$

The correlation matrix $E\{\underset{m}{x} \underset{m}{v}^T\}$ is a function of $\vec{\tau}$ and position of ξ_{1a} in the sampling cycle. Averaging over ξ_{1a} gives

$$\underset{m}{\phi}_{xv}(\vec{\tau}) \triangleq \frac{1}{T} \sum_{\xi_{1a}=1}^{\xi_{1a}=T} E\{\underset{m}{x}(\vec{\xi}_a) \underset{m}{v}^T (\vec{\xi}_a - \vec{\tau})\}$$

$$= \frac{1}{T} \sum_{\vec{\xi} \varepsilon U_\infty} \underset{m}{f}(\vec{\xi}) \underset{m}{\phi}_{uv}(\vec{\tau} - \vec{\xi}) \tag{58}$$

The transform of (58) is

$$\underset{m}{\Phi}_{xv}(\vec{z}) = \frac{1}{T} \underset{m}{F}(\vec{z}) \underset{m}{\Phi}_{uv}(\vec{z}) \tag{59}$$

Equations (58) and (59) do not imply the invariance of the correlation matrix, but that of the averaged correlation matrix over one sampling cycle. The averaged correlation matrix is then defined as $\underset{\sim}{\phi}_{xv}$ with its transform $\underset{\sim}{\Phi}_{xv}$. As averaged mean square error is to be minimized in the data reconstruction problem, $\underset{\sim}{\phi}_{xv}$ and $\underset{\sim}{\Phi}_{xv}$ are the pertinent matrices to use.

Similarly

$$\underset{\sim}{\phi}_{xy}(\vec{\tau}) \triangleq \frac{1}{T} \sum_{\xi_{1a}=1}^{\xi_{1a}=T} E\{\underset{\sim}{x}(\vec{\xi}_a)\, \underset{\sim}{y}^T(\vec{\xi}_a-\vec{\tau})\} \tag{60}$$

It is shown in the Appendix that the transformed matrix is

$$\underset{\sim}{\Phi}_{xy}(\vec{z}) = \frac{1}{T} \underset{\sim}{F}\, \underset{\sim}{\Phi}^*_{uv}\, \underset{\sim}{G}^\dagger \tag{61}$$

In microscopic study of microorganisms, various sections of the three dimensional object can be photographed by focusing the microscope at various depths and the three dimensional object is to be reconstructed from sectional photos. The sectional photos are represented by sampled planes in the following formulation:

$$\underset{\sim}{Y}^* = (\underset{\sim}{M}\, \underset{\sim}{X})^* + \underset{\sim}{N}^* \tag{62}$$

$$\underset{\sim}{\tilde{X}} = \underset{\sim}{X} - \underset{\sim}{P}\, \underset{\sim}{Y}^* = \underset{\sim}{X} - \underset{\sim}{P}\, [\, (\underset{\sim}{M}\, \underset{\sim}{X})^* + \underset{\sim}{N}^*\,] \tag{63}$$

From (59) and (61)

$$\underset{\sim}{\Phi}_{\tilde{x}\tilde{x}} = \underset{\sim}{\Phi}_{xx} - \frac{1}{T} \underset{\sim}{\Phi}_{xx}\, \underset{\sim}{M}^\dagger\, \underset{\sim}{P}^\dagger - \frac{1}{T} \underset{\sim}{P}\, \underset{\sim}{M}\, \underset{\sim}{\Phi}_{xx} + \frac{1}{T} \underset{\sim}{P}\, \underset{\sim}{\Phi}^*_{yy}\, \underset{\sim}{P}^\dagger \tag{64}$$

where

$$\underset{\sim}{\Phi}^*_{yy} = (\, \underset{\sim}{M}\, \underset{\sim}{\Phi}_{xx}\, \underset{\sim}{M}^\dagger)^* + \underset{\sim}{\Phi}^*_{nn} \tag{65}$$

The optimum processor and error matrix are given as:

$$P_o = \Phi_{xx} M^{\dagger} \Phi_{yy}^{*-1} \qquad (66)$$

$$\Phi_{\epsilon\epsilon} = \Phi_{xx} - \frac{1}{T}\Phi_{xx} M^{\dagger} \Phi_{yy}^{*-1} M \Phi_{xx} + (P - P_o)\Phi_{yy}^{*}(P - P_o) \qquad (67)$$

Example 3 In a three dimensional scalor signal reconstruction problem,

$$\phi_{xx}(\vec{\tau}) = S\rho^{(|\tau_1| + |\tau_2| + |\tau_3|)} \qquad (68)$$

and $\Phi_{nn} = 1$. The samples are taken at $\xi_1 = nT$ where $n = \ldots -3,-2,-1,0,1,2,3\ldots$, and $M = 1$. Determine the optimum processor P_o.

Solution: From (65) and (66), the solution can be written as

$$P_o = \left(\frac{\Phi_{xx}}{\Phi_{xx}^{*}}\right)\left[\frac{1}{1+\Phi_{xx}^{*-1}}\right] \qquad (69)$$

The second factor gives optimum processing at sampled planes while the first factor gives interpolation formula. From (68)

$$\Phi_{xx} = \frac{(1-\rho^2)^3 \; S}{\prod_{i=1}^{i=3}(1-\rho z_i^{-1})(1-\rho z_i)} \qquad (70)$$

$$\Phi_{xx}^{*} = \frac{(1-\rho^2)^2 \; (1-\rho^{2T})S}{(1-\rho^T z_1^{-T})(1-\rho^T z_1^{T})\prod_{i=2}^{i=3}(1-\rho z_i^{-1})(1-\rho z_i)} \qquad (71)$$

After some simplification, the interpolation formula is given as

$$\frac{\Phi_{xx}}{\Phi_{xx}^{*}} = 1 + \sum_{i=1}^{i=T-1}\rho^{i}\left[\frac{1-\rho^{2(T-i)}}{1-\rho^{2T}}\right](z_1^{-i} + z_1^{i}) \qquad (72)$$

V Real Time Processing of Motion Related Signals.

The signals and observed signals are represented by $\underset{m}{X}(\vec{z},t)$ and $\underset{m}{Y}(\vec{z},t)$:

$$\underset{m}{X}(\vec{z},t+1) = \underset{m}{A}(\vec{z},t))\ \underset{m}{X}(\vec{z},t) + \underset{m}{U}(\vec{z},t) \tag{73}$$

$$\underset{m}{Y}(\vec{z},t) = \underset{m}{M}(\vec{z},t)\ \underset{m}{X}(\vec{z},t) + \underset{m}{V}(\vec{z},t) \tag{74}$$

Equation (73) gives the kinetic relations between signals of successive frames (different t). There is no correlation between $\underset{m}{U}$, $\underset{m}{V}$; between $\underset{m}{U}(\vec{z},t')$ and $\underset{m}{V}(\vec{z},t)$, $\underset{m}{V}(\vec{z},t')$ for $t \neq t'$. At $t = 0$, the best estimate of $\underset{m}{X}$ is

$$\underset{m}{\hat{X}}(\vec{z},0) = \underset{m}{P_0}(\vec{z})\ \underset{m}{Y}(\vec{z},0) \tag{75}$$

where $\underset{m}{P_0}(\vec{z})$ is given by (44) and the initial error matrix $\underset{m}{\Phi_{XX}}(\vec{z},0)$ is given by the first term of (45).

A reccursive processing procedure is developed as follows:

$$\underset{m}{\hat{X}}(\vec{z},t+1) = \underset{m}{A}(\vec{z},t)\ \underset{m}{\hat{X}}(\vec{z},t) + \underset{m}{K}(\vec{z},t+1)[\ \underset{m}{Y}(\vec{z},t+1) - \underset{m}{\hat{Y}}(\vec{z},t+1)\ ] \tag{76}$$

where

$$\underset{m}{\hat{Y}}(\vec{z},t+1) = \underset{m}{M}(\vec{z},t+1)\ \underset{m}{A}(\vec{z},t)\ \underset{m}{\hat{X}}(\vec{z},t) \tag{77}$$

is the expected $\underset{m}{Y}(\vec{z},t+1)$ <u>before</u> it is observed, and $\underset{m}{K}(\vec{z},t+1)$ is a generalized version of Kalman gain matrix. Subtracting (76) from (73) gives

$$\underset{m}{\tilde{X}}(\vec{z},t+1) = \underset{m}{A}(\vec{z},t)\ \underset{m}{\tilde{X}}(\vec{z},t) + \underset{m}{U}(\vec{z},t) - \underset{m}{K}(\vec{z},t)\ \{\ \underset{m}{M}(\vec{z},t+1)[\underset{m}{A}(\vec{z},t)\underset{m}{\tilde{X}}(\vec{z},t) + \underset{m}{U}(\vec{z},t)] + \underset{m}{V}(\vec{z},t+1)\} \tag{78}$$

As the three random variables $\hat{\underset{\sim}{X}}\,(\vec{z},t)$, $\underset{\sim}{U}\,(\vec{z},t)$, and $\underset{\sim}{V}\,(\vec{z},t+1)$ are mutually independent, the error matrix is given as

$$\underset{\sim}{\Phi}_{\tilde{x}\tilde{x}}(\vec{z},t+1) = (\underset{\sim}{1} - \underset{\sim}{K}\,\underset{\sim}{M})\,\underset{\sim}{A}\,\underset{\sim}{\Phi}_{\tilde{x}\tilde{x}}\,\underset{\sim}{A}^{\dagger}\,(\underset{\sim}{1} - \underset{\sim}{K}\,\underset{\sim}{M})^{\dagger}$$

$$+ (\underset{\sim}{1} - \underset{\sim}{K}\,\underset{\sim}{M})\,\underset{\sim}{\Phi}_{uu}\,(\underset{\sim}{1} - \underset{\sim}{K}\,\underset{\sim}{M})^{\dagger} + \underset{\sim}{K}\,\underset{\sim}{\Phi}_{vv}\,\underset{\sim}{K}^{\dagger} \tag{79}$$

In the R.H.S. of (78), all matrices are at t or t+1 as specified in (78). The optimum gain is

$$\underset{\sim}{K}_{o}\,(\vec{z},t+1) = \underset{\sim}{\Psi}_{x}(\vec{z},t+1)\,\underset{\sim}{M}^{\dagger}\,(\vec{z},t+1)\,\underset{\sim}{\Psi}_{y}\,(\vec{z},t+1)^{-1} \tag{80}$$

$$\underset{\sim}{\Psi}_{x}\,(\vec{z},t+1) = \underset{\sim}{A}\,(\vec{z},t)\,\underset{\sim}{\Phi}_{\tilde{x}\tilde{x}}(\vec{z},t)\,\underset{\sim}{A}\,(\vec{z},t)^{\dagger} + \underset{\sim}{\Phi}_{uu}(z,t) \tag{81}$$

$$\underset{\sim}{\Psi}_{y}\,(z,t) = \underset{\sim}{M}\,(z,t)\,\underset{\sim}{\Psi}_{x}(z,t)\,\underset{\sim}{M}\,(z,t)^{\dagger} + \underset{\sim}{\Phi}_{vv}\,(z,t) \tag{82}$$

The $\underset{\sim}{\Psi}$'s are variances of $\underset{\sim}{x} - \hat{\underset{\sim}{x}}$ and $\underset{\sim}{y} - \hat{\underset{\sim}{y}}$ <u>before</u> the observation $\underset{\sim}{y}$ is made.

Substituting (80) into (79) gives

$$\underset{\sim}{\Phi}_{\tilde{x}\tilde{x}}(\vec{z},t) = \underset{\sim}{\Psi}_{x}(\vec{z},t) - \underset{\sim}{\Psi}_{x}(\vec{z},t)\,\underset{\sim}{M}\,(\vec{z},t)^{\dagger}\,\underset{\sim}{\Psi}_{y}\,(\vec{z},t)^{-1}\,\underset{\sim}{M}(\vec{z},t)\,\underset{\sim}{\Psi}_{x}(\vec{z},t) \tag{83}$$

<u>Example 4</u> The pictures are reconstructed with the purpose of tracking an object which moves 5 units in the ξ_1 direction every unit of t. Determine the optimum processor.

<u>Solution</u> The desired relation is

$$\underset{\sim}{x}\,(\xi_1,\vec{\xi}\,', t+1) = \underset{\sim}{x}\,(\xi_1-5, \vec{\xi}\,',t) \quad \text{Therefore } \underset{\sim}{A}\,(\vec{z}.t) = z_1^{-5} \text{ and } \underset{\sim}{U} = 0.$$

The matrices $\underset{\sim}{\Psi}_{x}$, $\underset{\sim}{\Psi}_{y}$, and $\underset{\sim}{\Phi}_{\tilde{x}\tilde{x}}$ at various t are obtained by repeated use of (81), (82), and (83). The Kalman gain $\underset{\sim}{K}_{o}$ is then determined from (80).

Equation (76) gives the optimum processing equation:

$$\hat{X}_m(\vec{z},t+1) = z_1^{-5}\,\hat{X}_m(\vec{z},t) + K_{m0}(\vec{z},t+1)\,[\,Y_m(\vec{z},t+1) - M_m(\vec{z},t+1)z_1^{-5}\,\hat{X}_m(\vec{z},t)] \qquad (84)$$

The optimum processed signal is the sum of two parts, a translated signal from an earlier time, and an optimally processed part on the unexplained portion of the observed signal.

VI. Conclusion

Linear processors are noncausal filters which are easy to design and easy to implement. In designing optimum processor to be used in a given application, the problem is the same as optimum filtering without realizability constraint on its poles and zeros. The processor is realizable if its denominator polynomial is nonzero on the unit circle product subspace C which has no interior point. When it fails to satisfy this condition, it can be approximated by a stable processor which satisfy this condition. However, a truly optimum stable processor then does not exist for the problem.

Algebraic expansion of the optimum processor function leads to a recursive processing scheme with a localized frame processor. The speed of convergence to optimum is given theoretically, and can be observed by comparing two successively modified frames. From the standpoint of practical implementation, frame recursive scheme has two important advantages: (i) It is ideal for parallel processing, as the same program can be executed simultaneously on all the points in the same frame. (ii) In an adaptive application, only the frame processor needs to be modified, and its coefficients are readily calculated from the statistical parameters.

Theoretically, it is shown that every rational processor function can be approximated by repeated operations with a finite impulse response frame processor with exponentially decaying remainder if its denominator polynomial is not equal to zero on the unit circle product subspace C. It is also shown that the convolution product of two stable processors is a stable processor. In the frequency domain, the processor function takes the place of filter transfer function in the new context of a discrete N-dimensional system.

Because of its theoretical simplicity, optimum linear processors have been obtained herein for a variety of applications, namely: two dimensional picture processing, three dimensional signal restoration from two dimensional projections and from two dimensional slices, and real time processing of multi-dimensional motion signals. In each case the optimum processor gives least mean square error while utilizing all available information. The analysis also gives an expression for the additional mean square error, if a different processor is used.

VII <u>Appendix: Mathematical Development of Linear Processor Theory</u>

Definition 1: A γ - modulus funtion $M(\vec{z})$ is defined by

$$M(\vec{z}) \triangleq \sum_{\vec{\xi}\epsilon U_\infty} m(\vec{\xi})\ \vec{z}^{-\vec{\xi}} \tag{A1}$$

where $m(\vec{\xi}) = A_1\ \gamma^{|\xi|}$. $\quad |\xi| = |\xi_1| + |\xi_2| + \cdots |\xi_N|$,

$A_1 > 0$, and $0 < \gamma < 1$.

<u>Theorem 1</u>: The infinite series $M(z)$ is absolutely, unconditionally, uniformly, convergent in N_c.

Proof: Let $M(|\vec{z}|, L)$ denote the sum

$$\sum_{\vec{\xi}\epsilon U_L} m(\vec{\xi})\,|\vec{z}|^{-\vec{\xi}} = A_1 \prod_{i=1}^{i=N} \sum_{\xi_i=-L}^{\xi_i=L} \gamma^{|\xi_i|}\ |z_i|^{-\xi_i}$$

$$= A_1 \prod_{i=1}^{i=N} \left\{ \frac{1-|\gamma z_i^{-1}|^{L+1}}{1-\gamma|z_i|^{-1}} + \frac{\gamma|z_i|-|\gamma z_i|^{L+1}}{1-\gamma|z_i|} \right\}$$

In $N_c(\varepsilon)$ with $0 < \varepsilon < 1 - \gamma$: $\quad |\gamma z_i^{-1}| < 1, \quad |\gamma z_i| < 1$.

Therefore

$$\lim_{L\to\infty} M(|\vec{z}|, L) = A_1 \prod_{i=1}^{i=N} \frac{1-\gamma^2}{(1-\gamma|z_i|^{-1})(1-\gamma|z_i|)}$$

$M(\vec{z})$ converges absolutely, then it converges unconditionally, uniformly in N_c. [15]

<u>Definition 2</u>: A linear processor $F(\vec{z})$ is said to be <u>stable</u> if it has a γ-modulus:

$$|f(\vec{\xi})| < A_1\gamma^{|\xi|}$$

Theorem 2: If $F(\vec{z})$ is stable, then the series

$$F(\vec{z}) = \sum_{\vec{\xi}\varepsilon U_\infty} f(\vec{\xi})\ \vec{z}^{-\vec{\xi}}$$

converges absolutely, unconditionally, uniformly in N_c.

Proof: Theorem 2 follows from Theorem 1 and Weierstrass M-test. [15].

Definition 3: A polynomial of z is a sum of finite number of terms, and each term is of the form $C_i\ z_1^{\xi_1} z_2^{\xi_2} \cdots z_N^{\xi_N}$ where $\xi_1, \xi_2, \cdots \xi_n$ are real integers (positive, negative, or zero). The power of each term is defined as $|\xi| = |\xi_1| + |\xi_2| + \cdots |\xi_N|$. The power of a polynomial is the largest power of its terms.

Theorem 3 Let $D(\vec{z})$ be a polynomial of $\vec{z}$, and $D(\vec{z})$ is real and positive (non zero) on C. Then the following:

(i) There exist D_u, and $D_\ell > 0$:

$$D_u = \underset{\vec{z}\varepsilon C}{\mathrm{Max}}\ D(\vec{z}),\ D_\ell = \underset{\vec{z}\varepsilon C}{\mathrm{Min}}\ D(\vec{z})$$

(ii) The following expansion converges on C if $A > \frac{D_u}{2}$;

$$\frac{1}{D(\vec{z})} = \frac{1}{A-[A-D(\vec{z})]} = \frac{1}{A}\sum_{n=0}^{n=\infty} (1-\frac{D(\vec{z})}{A})^n \qquad (A2)$$

(iii) Let $f(\vec{\xi})$ and $f_L(\vec{\xi})$ be defined by

$$f(\vec{\xi}) = \frac{1}{(2\pi j)^N}\oint_{\vec{z}\ \in\ C}\cdots\oint \frac{\vec{z}^{\vec{\xi}-1}}{D(\vec{z})}\ d\vec{z}$$

$$f_L(\vec{\xi}) = \text{Coefficient of } \vec{z} \text{ in } \frac{1}{A}\sum_{n=0}^{n=L} (1-\frac{D(\vec{z})}{A})^n$$

Then

$$\lim_{L\to\infty} f_L(\vec{\xi}) \to f(\vec{\xi})$$

(iv) Let $F(\vec{z})$ be defined in terms of $f(\vec{\xi})$ following the convention expressed in (A1), $F(\vec{z})$ is the unique, stable, expansion of $1/D(\vec{z})$ in N_c.

Proof: (i) Since C is a closed set and D(z) is a continuous mapping of $\vec{z}$ onto the complex plane, its mapping of C, D(C), is also closed. D(C) is a closed subset on the real line and contains its superior and inferior limits. $D_\ell > 0$ follows from $0 \notin D(c)$.

(ii) Let a denote the larger of the two positive numbers: $|1 - \frac{D_u}{A}|$, $|1 - \frac{D_\ell}{A}|$

Then

$$|1 - \frac{D(\vec{z})}{A}| \leq a < 1$$

and the expansion (A2) converges absolutely uniformly in N_c.

(iii) Let $J(L,\vec{\xi})$ be defined by

$$J(L,\vec{\xi}) = \frac{1}{(2\pi j)^N} \oint \cdots \oint_{\vec{z}\,\varepsilon\, C} \left\{ F(\vec{z}) - \frac{1}{A} \sum_{n=0}^{n=L} (1 - \frac{D(\vec{z})}{A})^n \right\} \vec{z}^{(\vec{\xi}-\vec{1})} \, d\vec{z} \qquad \text{(A-3)}$$

Then

$$|J(L,\vec{\xi})| \leq \frac{1}{(2\pi j)^N A} \oint \cdots \oint_{\vec{z}\,\varepsilon\, C} \sum_{n=L+1}^{n=\infty} \left| (1 - \frac{D(z)}{A})^n \vec{z}^{\vec{\xi}} \right| \vec{z}^{-1} \, d\vec{z}$$

$$\leq \frac{a^{L+1}}{A(1-a)} \qquad \text{(A-4)}$$

Let N_D denote the power of polynomial $D(\vec{z})$. Let L denote the integer which satisfies

$$N_D L < |\xi| \leq N_D(L + 1) \qquad \text{(A-5)}$$

There is no term of the form $\vec{z}^{-\vec{\xi}}$ in the sum $\sum_{n=0}^{n=L} (1 - \frac{D(\vec{z})}{A})^n$ since the latter is a polynomial of power $N_D L$. Equation (A-3) gives

$$J(L,\vec{\xi}) = f(\vec{\xi}) \qquad \text{(A-6)}$$

Condition (A-4) gives

$$|f(\vec{\xi})| \leq \frac{1}{A(1-a)} \cdot a^{\frac{|\xi|}{N_D}} < A' \gamma^{|\xi|}$$

where $A' > \frac{1}{A(1-a)}$, and $\gamma = a^{\frac{1}{N_D}} < 1$. Therefore $F(\vec{z})$ is stable.

Let $f(\vec{\xi}, n)$ denote the coefficient of the term $\vec{z}^{-\vec{\xi}}$ in the polynomial

$$\frac{1}{A}\left(1 - \frac{D(\vec{z})}{A}\right)^n$$

With values of L no longer restricted by (A-5), equation (A-3) gives

$$J(L,\vec{\xi}) = f(\vec{\xi}) - \sum_{n=0}^{n=L} f(\vec{\xi}, n)$$

Condition(A4) gives

$$\sum_{n=0}^{n=L} f(\vec{\xi}, n) \xrightarrow[L \to \infty]{} f(\vec{\xi})$$

(iv) Let

$$\frac{1}{D(\vec{z})} = \sum_{\vec{\xi}} f'(\vec{\xi})\, \vec{z}^{-\vec{\xi}}$$

be a stable expansion of $\frac{1}{D(\vec{z})}$. Then

$$f'(\vec{\xi}) = \frac{1}{(2\pi j)^N} \oint \cdots \oint \frac{\vec{z}^{(\vec{\xi}-\vec{1})}}{D(\vec{z})}\, d\vec{z} = f(\vec{\xi})$$

<u>Theorem 4</u> A linear processor of finite terms is stable. Let $F(\vec{z})$ and $G(\vec{z})$ be stable linear processors. The linear processor $F(\vec{z})\, G(\vec{z})$ is stable.

<u>Proof</u> A linear processor $P(\vec{z})$ of N_p terms can be represented as

$$P(\vec{z}) = \sum_{i=1}^{i=N_p} p(\vec{\xi}_i)\, \vec{z}^{\vec{\xi}_i}$$

Given any γ, the constant A can be selected to satisfy

$$A > \max_{i=1,2\ldots N_p} \left(\frac{|p(\vec{\xi}_i)|}{\gamma^{|\xi_i|}} \right)$$

To prove the second part of the theorem: Stability of $F(\vec{z})$ and $G(\vec{z})$ implies existence of a γ-modulus function for each:

$$f(\vec{\xi}) < A_1 \gamma_1^{|\xi|}$$

$$g(\vec{\xi}) < A_2 \gamma_2^{|\xi|}$$

Let $M_1(\vec{z})$ and $M_2(\vec{z})$ denote the two γ-modulus functions

$$H(\vec{z}) = M_1(\vec{z})\, M_2(\vec{z})$$

Then

$$h(\vec{\xi}) > \left|(f*g)(\vec{\xi})\right|$$

However, $H(\vec{z})$ in closed form can be expressed as $1/D(\vec{z})$ satisfying condition of Theorem 3. A γ-modulus function exists for $1/D(\vec{z})$ and it is also a γ-modulus for $F(\vec{z})\ G(\vec{z})$.

<u>Theorem 5</u> Let $N(\vec{z})$ and $D(\vec{z})$ denote polynomials of $\vec{z}$ such that $D(\vec{z}) \neq 0$ on C. Then the rational function $N(\vec{z})/D(\vec{z})$ has a unique, stable expansion.

<u>Proof</u>

$$\frac{N(\vec{z})}{D(\vec{z})} = \frac{N(\vec{z})\bar{D}(\vec{z})}{D(\vec{z})\bar{D}(\vec{z})}$$

The polynomial $D(\vec{z})\ \bar{D}(\vec{z})$ is positive real on C. From Theorem 3, its reciprocal

$$F(\vec{z}) = \frac{1}{D(\vec{z})\bar{D}(\vec{z})}$$

has a unique stable expansion. From Theorem 4, $N(\vec{z})\bar{D}(\vec{z})F(\vec{z})$ has a unique stable expansion.

<u>Theorem 6</u> Let x, y denote processed output of stationary signals u,v with invariant stable L.P.'s f and g:

$$X(\vec{z}) = F(\vec{z})\ U(\vec{z}) \tag{A7}$$

$$Y(\vec{z}) = G(\vec{z})\ V(\vec{z}) \tag{A8}$$

Then $\Phi_{xy}(\vec{z}) = F(\vec{z})\,\Phi_{uv}(\vec{z})\,G(\vec{z}^{-1})$ (A9)

<u>Proof</u> Equations (A7) and (A8) can be rewritten as

$$x(\vec{\xi}_a) = \sum_{\vec{\sigma}\varepsilon U_\infty} f(\vec{\xi}_a-\vec{\sigma})\,u(\vec{\sigma}) \tag{A10}$$

$$y(\vec{\xi}_b) = \sum_{\vec{\tau}\varepsilon U_\infty} g(\vec{\xi}_b-\vec{\tau})\,v(\vec{\tau}) \tag{A11}$$

Multiplying (A10) and (A11) and $\vec{z}^{-(\vec{\xi}_a-\vec{\xi}_b)}$ and summing over $\vec{\xi}_a$ gives

$$\Phi_{xy}(\vec{z}) = \sum_{\vec{\xi}_a}\sum_{\vec{\sigma}}\sum_{\vec{\tau}} f(\vec{\xi}_a-\vec{\sigma})\vec{z}^{-(\vec{\xi}_a-\vec{\sigma})}\; g(\vec{\xi}_b-\vec{\tau})\vec{z}^{(\vec{\xi}_b-\vec{\tau})}\;\Phi_{uv}(\vec{\sigma}-\vec{\tau})z^{-(\vec{\sigma}-\vec{\tau})}$$

Equation (A9) follows from the above equation.

<u>Proof of Equation (61)</u>

Let U_∞^* denote the set $\{rT, \vec{\xi}'\}$ with r and each component of $\vec{\xi}'$ ranging over all integers. Equation (56) can be rewritten as

$$\underset{m}{X}(\vec{\xi}_a) = \sum_{\vec{\sigma}\varepsilon U_\infty^*} \underset{m}{f}(\vec{\xi}_a-\vec{\sigma})\,\underset{m}{u}(\vec{\sigma})$$

Multiplying (56) with a similar expression for $\underset{m}{y}^T(\vec{\xi}_a-\vec{\tau})$ gives

$$E\{\underset{m}{x}(\vec{\xi}_a)\,\underset{m}{y}^T(\vec{\xi}_a-\vec{\tau})\} = \sum_{\vec{\sigma}\varepsilon U_\infty^*}\sum_{\vec{\lambda}\varepsilon U_\infty^*} \underset{m}{f}(\vec{\xi}_a-\vec{\sigma})\,\underset{m}{\phi}_{uv}(\vec{\sigma}-\vec{\lambda})\;g^T(\vec{\xi}_a-\vec{\tau}-\vec{\lambda}) \tag{A12}$$

One way to prove (61) is to check its correctness by expanding the RHS product:

$$\underset{m}{\Phi}_{xy}(z) = \frac{1}{T}\left(\sum_{\xi_{1a}=1}^{\xi_{1a}=T}\sum_{\vec{\sigma}\varepsilon U_\infty^*} \underset{m}{f}(\vec{\xi}_a-\vec{\sigma})\,\vec{z}^{-(\vec{\xi}_a-\vec{\sigma})}\right)$$

$$\left(\sum_{\vec{\lambda}\varepsilon U_\infty^*}\underset{m}{\Phi}_{uv}(\vec{\sigma}\,\vec{\lambda})\vec{z}^{-(\vec{\sigma}\,\vec{\lambda})}\right)\left(\sum_{\vec{\xi}_b\varepsilon U_\infty}\underset{m}{g}^T(\vec{\xi}_b-\vec{\lambda})\,\vec{z}^{-(\vec{\xi}_b-\vec{\lambda})}\right) \tag{A13}$$

Since $\underset{m}{F}$, $\underset{m}{\phi}_{uv}$, and $\underset{m}{G}^{\dagger}$ are stable, their product is also stable by Theorem 4. Theorem 2 states that the infinite series on the RHS of (13) converges absolutely, unconditionally and uniformly in N_C. All the summation signs are then moved to the front:

$$\underset{m}{\phi}_{xy}(z) = \frac{1}{T} \sum_{\xi_{1a}=1}^{\xi_{1a}=T} \sum_{\vec{\sigma}\varepsilon U_\infty^*} \sum_{\vec{\lambda}\varepsilon U_\infty^*} \sum_{\vec{\xi}_b \varepsilon U_\infty} \underset{m}{f}(\vec{\xi}_a-\vec{\sigma})\underset{m}{\phi}_{uv}(\vec{\sigma}\ \vec{\lambda})\underset{m}{g}^T(\vec{\xi}_b-\vec{\lambda})\vec{z}^{-(\vec{\xi}_a - \vec{\xi}_b)} \tag{A14}$$

The $\vec{z}^{-\vec{\tau}}$ term on the RHS of (A14) is given by $\vec{\xi}_b = \vec{\xi}_a-\vec{\tau}$. Therefore

$$\underset{m}{\phi}_{xy}(\vec{\tau}) = \frac{1}{T} \sum_{\xi_{1a}=1}^{\xi_{1a}=T} \sum_{\vec{\sigma}\varepsilon U_\infty^*} \sum_{\vec{\lambda}\varepsilon U_\infty^*} \underset{m}{f}(\vec{\xi}_a-\vec{\sigma})\underset{m}{\phi}_{uv}(\vec{\sigma}-\vec{\lambda})\underset{m}{g}^T(\vec{\xi}_a-\vec{\tau}-\vec{\lambda}) \tag{A15}$$

Equation (A15) checks with (60) and (A12).

References

1. T.S. Huang, W.F. Schreiber, and O.J. Tretiak, "Image Processing", Proceedings of IEEE, Vol. 59, pp 1586-1609, Nov. 1971.

2. N.E. Nahi, "Role of Recursive Estimation in Statistical Image Enhancement", Proceedings of the IEEE, Vol. 60, No. 7, pp. 872-877, July 1972.

3. A. Habibi, "Two Dimensional Bayesian Estimates of Images", Proceedings of the IEEE, Vol. 60, No. 7, pp 878-883, July 1972.

4. M.M. Sondhi, "Image Restoration: The Removal of Spatially Invariant Degradations", Proceedings of the IEEE, Vol. 60, No. 7, pp 842-853, July 1972.

5. W.K. Pratt, "Generalized Wiener Filtering Computation Techniques", IEEE Trans. on Computers, Vol. C-21, No. 7, pp 636-641, July, 1972.

6. R.M. Mersereau, and D.E. Dudgeon, "Two Dimensional Digital Filtering", Proceedings of IEEE, Vol. 63, No. 4, pp 610-623, April 1975.

7. B.R. Hunt, "Digital Image Processing", Proceedings of IEEE, Vol. 63, No. 4, pp 693-708, April 1975

8. R.P. Roesser, "A Discrete State-Space Model for Linear Image Processing", IEEE Trans. on Auto., Cont., Vol. AC-20, No. 1, pp 1-10, February 1975.

9. E.N. Leith, "Complex Spatial Filters for Image Deconvolution", Proceedings of IEEE, Vol. 65, No. 1, pp 18-28, January 1977.

10. G.W. Stroke, M. Halioua, F. Thon, and D.H. Willasch, "Image Improvement and Three-Dimensional Reconstruction Using Holographic Image Processing", Proceedings of IEEE, Vol. 65, No. 1, pp 39-61, January 1977.

11. A.O. Aboutalib, M.S. Murphy, and L.M. Silverman, "Digital Restoration of Images Degraded by General Motion Blurs", IEEE Trans. on Auto. Cont., Vol. AC-22 No. 3, pp 294-301, June 1977

12. A. Lev, S.W. Zucker, and A. Rosenfeld", Iterative Enhancement of Noisy Images", IEEE Trans. on Systems, Man, & Cybernetics, Vol. SMC-7 No. 6, pp 435-442, June 1977.

13. E.L. Hall, "Comparison of Computations for Spatial Frequency Filtering", Proceedings of IEEE, Vol. 60, No. 7, pp 884-891, July 1972.

14. J.H. Justice, and J.L. Shanks, "Stability Criterion for N-Dimensional Digital Filters" IEEE Trans. on Auto. Cont., Vol. AC-18, No. 3, pp 284-286, June 1973.

15. L.M. Graves, "The Theory of Functions of Real Variables", McGraw Hill, 1956, pp 107-111.

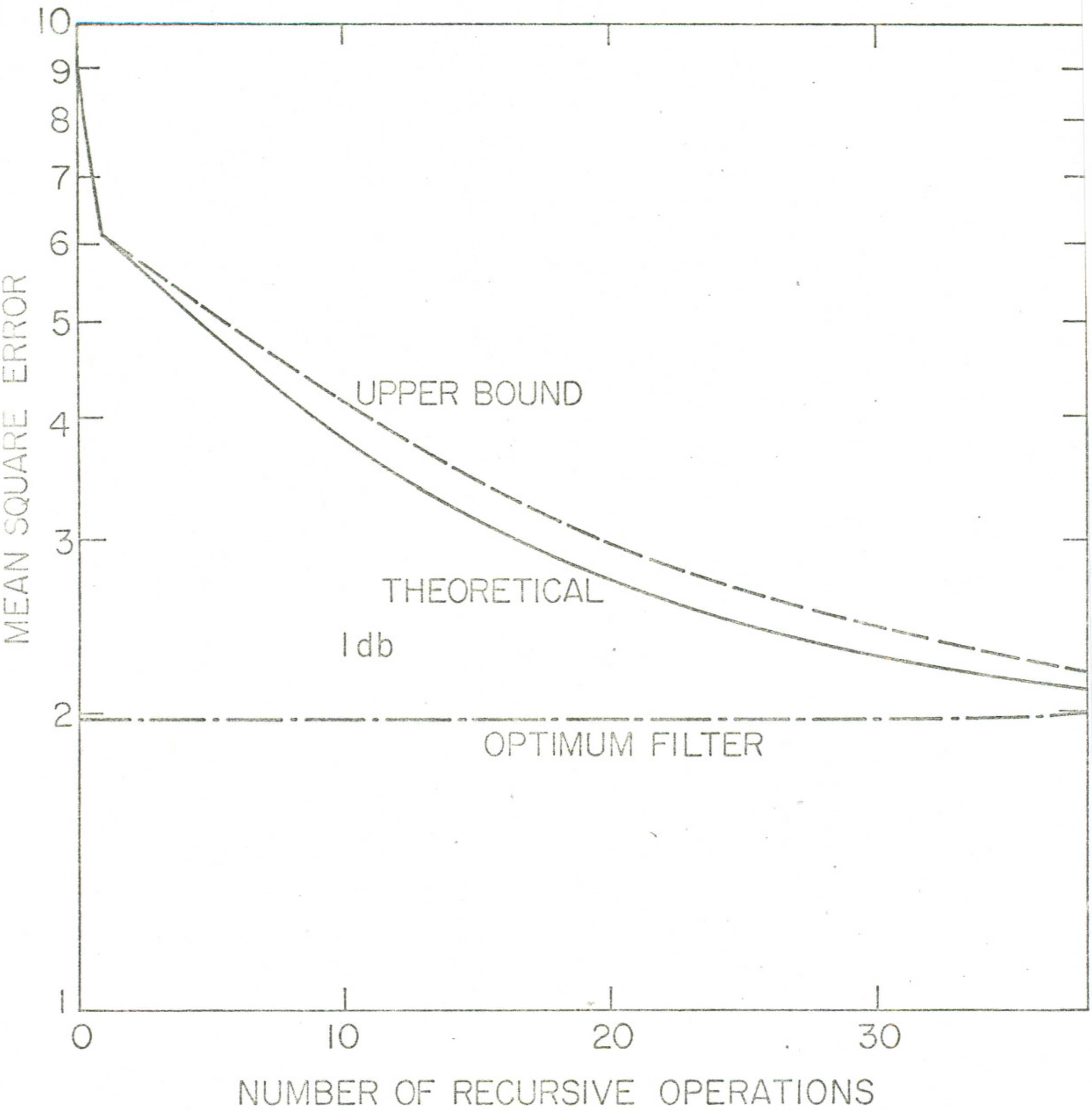

FIGURE 1: Mean square error of improved image versus the number of recursive operations.