USCIPI Report 610

UNIVERSITY OF SOUTHERN CALIFORNIA

FAST COMPUTATIONAL TECHNIQUES FOR PSEUDOINVERSE AND WIENER IMAGE RESTORATION

by

Faramarz Davarian

August 1975

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007

Sponsored by
Advanced Research Projects Agency
Contract No. F08606-72-C-0008
ARPA Order No. 1706

IPI

IMAGE PROCESSING INSTITUTE

USC'PI Report 610

FAST COMPUTATIONAL TECHNIQUES FOR PSEUDOINVERSE AND WIENER IMAGE RESTORATION

by

FARAMARZ DAVARIAN

August 1975

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Air Force Eastern Test Range under Contract No. F08606-72-C-0008, ARPA Order No. 1706.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)

Image Processing Institute
University of Southern California, University Park
Los Angeles, Calif. 90007

2a. REPORT SECURITY CLASSIFICATION

UNCLASSIFIED

2b. GROUP

3. REPORT TITLE

IMAGE PROCESSING RESEARCH.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Rep[redacted], August 1975

5. AUTHOR(S) (First name, middle initial, last name)

FARAMARZ DAVARIAN

6. REPORT DATE

Aug[redacted]75

7a. TOTAL NO. OF PAGES

163

7b. NO. OF REFS

63

8a. CONTRACT OR GRANT NO.

F08606-72-C-0008

b. PROJECT NO.

ARPA Order [redacted]-1706

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

USCIPI [redacted]-610

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Approved for release: distribution unlimited

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

13. ABSTRACT

A fast minimum mean-square error technique for restoring images degraded by blur is presented in this dissertation.

The phenomenon of linear shift-invariant degradation in an incoherent optical system can be described by means of a convolution integral. Since digital processing of pictorial data requires discretization of this integral by means of a quadrature technique, a theoretical study of a broad class of quadrature formulae is first presented.

The discrete image degradation phenomenon is modelled by two distinct vector space formulations: dark background objects correspond to a model possessing an overdetermined blur matrix; objects with unknown background, however, result in a system that is underdetermined. It is shown that these models become equivalent if the background of the object is artificially set to zero by processing the observed image. This fact results in introduction of a fast restoration technique in the absence of noise.

The noisy restoration problem is resolved by employing Wiener estimation. It is shown that with proper arrangement of the observed image data, the covariance matrix of the object becomes a circulant matrix. Hence, the Fourier domain properties of circulants gives rise to a computationally efficient Wiener restoration technique. A certain approximation imposed on this technique results in a suboptimal, but faster, restoration filter. It is shown that the computational saving gained by this approximation is significant, while the increase in the error variance is quite small.

* *

14. Key words: Digital Image Restoration, Quadrature Formulae, Discrete Image Degradation Modelling, Pseudoinverse Restoration, Fast Computational Techniques, Fast Wiener Filtering

DD FORM 1 NOV 65 1473

UNCLASSIFIED

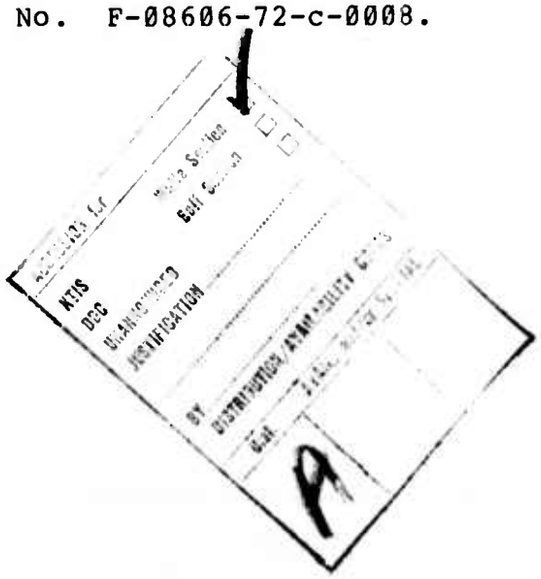
Security Classification

391 141

ACKNOWLEDGEMENT

The author wishes to express his sincere gratitude to the chairman of his committee, Professor William Pratt, for his guidance throughout the author's graduate career. Without Professor Pratt's encouragement and financial support, this graduate study could not have been possible. The author also feels deep appreciation toward Professors Alexander Sawchuk and Zdenek Vorel for their assistance, and to Michael Huhns, Javad Peyrovian, and Robert Wallis for their valuable discussions and aid.

This research was supported in part by the Advanced Research Projects Agency of the Department of Defense, and was monitored by the Air Force Test Range under Contract No. F-08606-72-c-0008.



ABSTRACT

A fast minimum mean-square error technique for restoring images degraded by blur is presented in this dissertation.

The phenomenon of linear shift-invariant degradation in an incoherent optical system can be described by means of a convolution integral. Since digital processing of pictorial data requires discretization of this integral by means of a quadrature technique, a theoretical study of a broad class of quadrature formulae is first presented.

The discrete image degradation phenomenon is modelled by two distinct vector space formulations: dark background objects correspond to a model possessing an overdetermined blur matrix; objects with unknown background, however, result in a system that is underdetermined. It is shown that these models become equivalent if the background of the object is artificially set to zero by processing the observed image. This fact results in introduction of a fast restoration technique in the absence of noise.

The noisy restoration problem is resolved by employing Wiener estimation. It is shown that with proper arrangement of the observed image data, the covariance matrix of the object becomes a circulant matrix. Hence, the Fourier domain properties of circulants gives rise to a

computationally efficient Wiener restoration technique. A certain approximation imposed on this technique results in a suboptimal, but faster, restoration filter. It is shown that the computational saving gained by this approximation is significant, while the increase in the error variance is quite small.

TABLE OF CONTENTS

1. INTRODUCTION

The concept of digital restoration in this text is interpreted as the reconstruction of an image to an original object by the removal of degradation phenomenon known a priori, possibly in the presence of noise [1-1]. Thus, restoration techniques require some form of knowledge concerning the degradation phenomenon, and this knowledge either comes from a deterministic assumption about the phenomenon or statistical models. Degradation systems are often quite complex in general. However, in many cases of practical importance, the degrading systems can be presented by a linear smoothing operation followed by the addition of noise known only in a statistical sense [1-2].

Early image restoration techniques, which were introduced by the pioneers of this field [1-3] [1-4], did not prove very successful because the proposed techniques did not acknowledge the existence of noise in the imaging systems. To be more specific, the failure of the early restoration methods was caused because of the amplification of a high frequency noise component by the inversion of some degrading process. Tsujiuchi [1-5], Harris [1-6], McGlamery [1-7], and Mueller [1-8] have all tried to overcome this hindrance by modifying the inversion technique. For example, Harris [1-6] has replaced the inverse filter by zero for the range of spatial frequencies

for which the noise power exceeds the power of the signal. Later, researchers in this field realized that a more successful recovery of the image could be achieved with a more realistic approach which utilized the characteristics of the noise in the imaging systems. Consequently, an optimum linear shift-invariant filter was introduced by Helstrom [1-9] which, when applied on a noisy degraded image, gives an estimate of the ideal image with the least mean-square error. This filter is indeed the same as the classical Wiener filter. Slepian [1-10] has also solved the same problem when the smoothing function itself is stochastic. It has later been illustrated that the minimum mean-square error technique gives better reconstructed images than the simple inversion method [1-11].

The Wiener filter technique unfortunately has the limitations of large storage requirements along with inefficient computational methods. Pratt [1-12] has introduced generalized Wiener filtering computation techniques which, by utilizing transform properties of imaging systems, have improved the computational efficiency. Furthermore, Pratt illustrated that a specific computational procedure could result in a significant reduction of the computational load, with only a small increase in estimation error. It has also been shown that lower-triangular transformations can give rise to an efficient suboptimal Wiener filter [1-13].

Constrained restoration is an alternative solution for the problem of noisy image reconstruction. Here the mean-square error is not necessarily minimized, but high frequency noise oscillations are dampened by observing the constraints governing the image forming systems. Hunt [1-14] has employed special properties of linear systems to introduce fast constrained image estimation techniques. A constrained restoration technique introduced by Mascarenhas [1-15] utilizes linear equality and inequality constraints. Linear inequality constraints involve solution of a quadratic programming problem, and require extensive computing when images of reasonable size are to be processed. A specific case of inequality constrained image restoration is when positiveness of image intensities is utilized for better image reconstruction purposes. A survey of positive image restoration techniques is given in reference [1-16].

Blind deconvolution, in which the point-spread function is assumed unknown [1-17], has attracted some attention. Ekstrom [1-18] has suggested means of estimating the unknown point-spread function by processing the blurred and noisy observation. Cole [1-19] has introduced the homomorphic filter which is the geometrical mean between the Wiener filter and the inverse filter. Similar methods have been applied to the problem of restoring old acoustic recordings as well as reconstruction

of blurry images [1-19]. (For a comparison of different restoration techniques, the reader is referred to reference [1-20]).

The fundamental purpose of this dissertation is to explore restoration techniques which are computationally fast and efficient, but can be applied to the restoration of large size images. This work starts with a brief, but sufficiently broad, discussion of the discretization methods of continuous linear shift-invariant degradation systems. Next, the problem of a fast pseudoinverse technique which can be applied to a general noise-free image is resolved; this part can be considered to be an application and extension of the Fourier transform properties of circulant matrices [1-21], [1-22]. In the presence of noise, a fast Wiener filtering technique is developed which overcomes many of the computational obstacles of the conventional Wiener filter. Finally, the problem of fast constrained filtering of degraded images is considered.

REFERENCES

1. B. Hunt, "Digital Image Processing," Proc. IEEE, April 1975, pp. 693-708.

2. H. Andrews, "Digital Images Restoration: A Survey," IEEE Computer, May 1974, pp. 36-45.

3. A. Marechal, P. Croce, and K. Dietzel, "Amelioration du Contraste des Details des Images Photographiques par Filtrage des Frequencies Spatiales" Opt. Acta, Vol. 5, 1958, pp. 256-262.

4. L. Cutrona et al., "Optical Data Processing and Filtering Systems," IRE Trans. Inform. Theory, June 1960, pp. 386-400.

5. J. Tsujiuchi, "Correction of Optical Images by Compensation of Aberrations and Spatial Frequency Fitering," Progress in Optics, Vol. 2, New York: Wiely, 1963, pp. 131-180.

6. J. Harris, "Image Evaluation and Restoration," J. Opt. Soc. Amer., May 1966, pp. 569-574.

7. B. McGlamery, "Restoration of Turbulece-Degraded Images," J. Opt. Soc. Amer., March 1967, pp. 293-297.

8. P. Mueller and G. Reynolds, "Image Restoration by Removal of Random Media Degradations," J. Opt. Soc. Amer., November 1967, pp. 1338-1344.

9. C. Helstrom, "Image Restoration by the Method of Least Squares," J. Opt. Soc. Amer., March 1967, pp. 297-303.

10. D. Slepian, "Linear Least Squares Filtering of Distorted Images," J. Opt. Soc. Amer., July 1967, pp. 918-922.

11. M. Sondhi, "Image Restoration: The Removal of Spatially Invariant Degradations," Proc. IEEE, July 1972, pp. 842-853.

12. W. Pratt, "Generalized Wiener Filter Computation Techniques," IEEE Trans. Comp., July 1972, pp. 636-641.

13. A. Habibi, "Fast Suboptimal Wiener Filtering of Markov Sequences," to appear in IEEE Trans. Comp.

14. B. Hunt, "The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer," IEEE Trans. Comp., September 1973, pp. 805-812.

15. N. Mascarenhas, Digital Image Restoration Under a Regression Model- Linear equality and Inequality Constrained Approaches, Ph. D. Dissertation, University of Southern California, January 1974.

16. H. Andrews, "Positive Digital Image Restoration Techniques: A Survey," ATR-73(8139)-2, Aerospace Corporation Technical Report, February 1973.

17. T. Stockham, T. Cannon, and R. Ingebretsen, "Blind Deconvolution Through Digital Signal Processing," Proc. IEEE, April 1975, pp. 678-692.

18. M. Ekstrom, "A Numerical Algorithm for Identifying Spread Functions of Shift-Invariant Imaging Systems," IEEE Trans. Comp., April 1973, pp. 322-328.

19. E. Cole, The Removal of Unknown Blur by Homomorphic Filtering, Ph. D. Dissertation, University of Utah, June 1973.

20. B. Hunt and H. Andrews, "Comparision of Different Filter Structures for Image Restoration," Proc. 6th Annual Hawaii Int. Conf. on Systems Sciences, January 1973.

21. B. Hunt, "A Matrix Theory Proof of the Discrete Convolution Theorem," IEEE Trans. Audio Electroacoust, December 1971, pp. 285-288.

22. W. Pratt, "Vector Space Formulation of Two Dimensional

Signal Processing Operations," Journal of Computer Graphics and Image Processing, Academic Press, March 1975.

2. THE RESTORATION PROBLEM IN ITS CONTINUOUS FORM

2. 1 The Image

Every visible object can be characterized by its radiant energy distribution which commonly is described by a two-dimensional function of spatial variables, f(x,y). The two-dimensional radiant energy distribution which enters the human eye is often referred to as the image. Even in the case of an observer with perfect vision, the image itself might be a distorted replica of the object function. This, inevitably, will cause an imperfect comprehension of the scene by the brain. A degraded image can result from many different phenomena. A turbulent atmosphere, for example, deforms the phase function while the light travels tnrough the air. When collected by this optical system aperture, a blurred image results. A photographic camera with a poor lens produces a low quality picture. Also a misfocused lens and movements of the objects in a scene both generate errors in recording the scene, and thus the image often differs from the original object in one way or another. The following section studies the mathematical model of the image-forming and image-degrading processes. Since this dissertation deals exclusively with linear phenomena, only degradations which can be represented by a linear system have been considered.

2. 2 Convolution as a Model for Linear Systems

A continuous space invariant (stationary) linear system is interpreted to be a convolution of two functions: a fixed function h, commonly referred to as the impulse response or the point spread function, and a function f which is the input to the system. The mathematical expression for convolution is

$$g(x)= \int_{-\infty}^{\infty} f(s)h(x-s)ds \qquad (2\text{-}1)$$

where g is the output of the system. In a more compact notation, eq. (2-1) can be written as

$$g(x)=f(x)\circledast h(x) \qquad (2\text{-}2)$$

The output of a linear system is equal to the impulse response, h, when the input to the system is an impulse.

In two dimensions the signals f, h, and g are functions of two variables as modelled by the relation

$$g(x,y)= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(r,s)h(x-r,y-s)drds \qquad (2\text{-}3)$$

The two dimensional convolution integral is the proper model for a spatially invariant degradation occurring under incoherent illumination [2-1], [2-2].

A linear system is fully defined by its impulse response. And, this is also true for a linear degradation where the impulse response of a particular process completely characterizes the degradation phenomenon.

A diffraction-limited rectangular optical system is characterized by a separable impulse response of the form [2-3]

$$h(r,s)= \left\{\frac{\sin(r)}{r}\right\}^2 \left\{\frac{\sin(s)}{s}\right\}^2 \qquad (2\text{-}4)$$

Blurring due to atmospheric turbulence has been modelled by a linear operator of the form [2-4]

$$h(r,s)=\exp[-(r^2+s^2)^{\frac{5}{6}}] \qquad (2\text{-}5)$$

Motion blur has a one dimensional point spread function defined as [2-5]

$$h(s)=\begin{cases} 1 & \text{if } -1/2< s <1/2 \\ 0 & \text{otherwise} \end{cases} \qquad (2\text{-}6)$$

Although the above is not an exhaustive list of sources of linear image degradation, the list provides a general intuition for some of the problems faced in image restoration. The next section describes the use of the

Fourier transform of linear systems for removal of spatially invariant degradations.

2. 3 Inverse Filtering

To avoid notational complexity and unnecessary formulations, only one-dimensional degradation is discussed at this point. This approach will be followed in most remaining sections of this dissertation, and except for the examples explicitly discussed, the extension of a one-dimensional problem to higher dimensions is assumed to be straight forward. Considering the model for image degradation formulated by eq. (2-1), the restoration task is phrased as follows: assuming the observation, g, and the point spread function, h, are both given, attempt to recover or estimate the image, f.

It has been shown that Fourier techniques can play an important role in attempting to obtain the object from the observation g(x) through the inversion of h(.) [2-6], [2-7]. For a function f(x) the Fourier transform, F(w), of f(x) is defined as

$$F(w) = \int_{-\infty}^{\infty} f(x) \exp[-iwx] dx \qquad (2\text{-}7)$$

The above integral does not, however, exist for every function f(x) [2-8], but the existence of this integral for

the class of functions encountered in this dissertation is certain.

An interesting utilization of the Fourier transform is its application to linear systems. Let G(w), F(w), and H(w) denote the Fourier transforms of g(x), f(x), and h(x), respectively. Then from equation (2-1) it is easily shown that

$$G(w)=F(w)H(w) \tag{2-8}$$

Observing the above equality, the concept of inverse filtering becomes clear. Inverse filtering simply consist of dividing both sides of eq. (2-8) by the blur transfer function H(w). If H(w) does not vanish at any point, the object, f(x), can be completely recovered by the operation

$$f(x)=\mathcal{F}^{-1}\frac{G(w)}{H(w)} \tag{2-9}$$

where the operation $\mathcal{F}^{-1}$ denotes the inverse Fourier transform

$$f(x)=\left(\frac{1}{2\Pi}\right)\int_{-\infty}^{\infty}F(w)\exp[iwx]dw \tag{2-10}$$

Equation (2-10) is the inverse of eq. (2-7), and the notation is the same in both equations.

Although the problem of noisy observations will be studied in later chapters, a brief comment on eq. (2-9) seems essential at this point. The impulse response H(w) almost always decreases rapidly with growth of w. On the other hand, any small amount of noise or uncertainty in the observation has a relatively flat spectral distribution. This means that the inverse filtering technique enhances high frequency noise so strongly that even for a small amount of noise the technique could not be applied, and thus must be modified.

REFERENCES

1. B. Hunt, "Digital Image Processing," Proc. IEEE, April 1975, pp. 693-708.

2. K. campbell, G. Wechksung, and C. Mansfield, "Spatial Filtering by Digital Holagraphy," Optical Engineering, vol. 13, May-June 1974, pp. 175-188.

3. J. Goodman, Introduction to Fourier Optics, New York: MacGraw-Hill, 1968.

4. R. Hufnagel and N. Stanly, "Modulation Transfer Function Associated with Image Transmission Through

Turbulent Media," J. Opt. Soc. Amer., January 1964, pp. 52-61.

5. T. Hung, W. Scheiber, and O. Tretiak, "Image Processing," Proc. IEEE, November 1971, pp. 1586-1609.

6. B. McGlamery, "Restoration of Turbulence-Degraded Images," J. Opt. Soc. Amer., March 1967, pp. 293-297.

7. P. Mueller and G. Reynolds, "Image Restoration by Removal of Random Media Degradations," J. Opt. Soc. Amer., November 1967, pp. 1338-1344.

8. A. Papoulis, Systems and Transforms with Applications in Optics, New York: MacGraw-Hill, 1968, pp. 61.

3. DISCRETIZATION OF THE CONTINUOUS MODEL

To treat degraded images by means of digital computers, the continuous model of eq. (2-1) must be discretized. The discretization process will, naturally, replace the integral by a discrete summation which takes advantage of values of the signals only at discrete points.

3. 1 Practical Considerations

In most practical situations, the physical sample image g(x) of eq. (2-1) is not available over the entire real line, and also the whole infinite extent of the object is not usually of particular interest to the image processer. In fact, in practice, only finite size of objects are to be restored by processing of finite size observations. The preceeding argument implies that, in reality, the limits on the definite integral of eq. (2-1) are not infinite. Another important feature in the model of eq. (2-1) is that the degrading function h usually vanishes beyond some point, and consequently the region for which h is nonzero has a finite length. In theory, of course, most point spread functions have infinite length, but invariably, h decreases rapidly for large values of x (the examples in Sec. (2-2), for instance, have this property). Considering this characteristic, a point spread function can be truncated to some lengh L without severe

modelling error providing that the length L is selected wisely. Figure (3-1) shows a truncated two-dimensional Gaussian point spread function.

In view of the above argument, eq. (2-1) is modified to the form

$$g(x)=\int_{u=x-\frac{L}{2}}^{v=x+\frac{L}{2}} f(s)h(x-s)ds \qquad (3\text{-}1)$$

where [u,v] is the region of integration.

3. 2 Quadrature Formulae

By definition, a quadrature formula (q.f.) is an approximation to a definite integral; the approximation is a linear combination of values of the integrand and its derivatives at certain points of the interval of integration called the nodes of the q.f. [3-1]. When the derivatives of the integrand are unknown, then the general form of a q.f. is expressed as

$$\int_u^v f(x)dx=\sum_{i=1}^{n} c_i f(x_i)+Rf \qquad (3\text{-}2)$$

where c_i and x_i are the coefficients and nodes of the q.f., respectively, and $u<x_1<x_2\ldots\ldots<x_n<v$. The term Rf is a functional which, for any given function f(*), equals the difference between the exact value of the integral and its

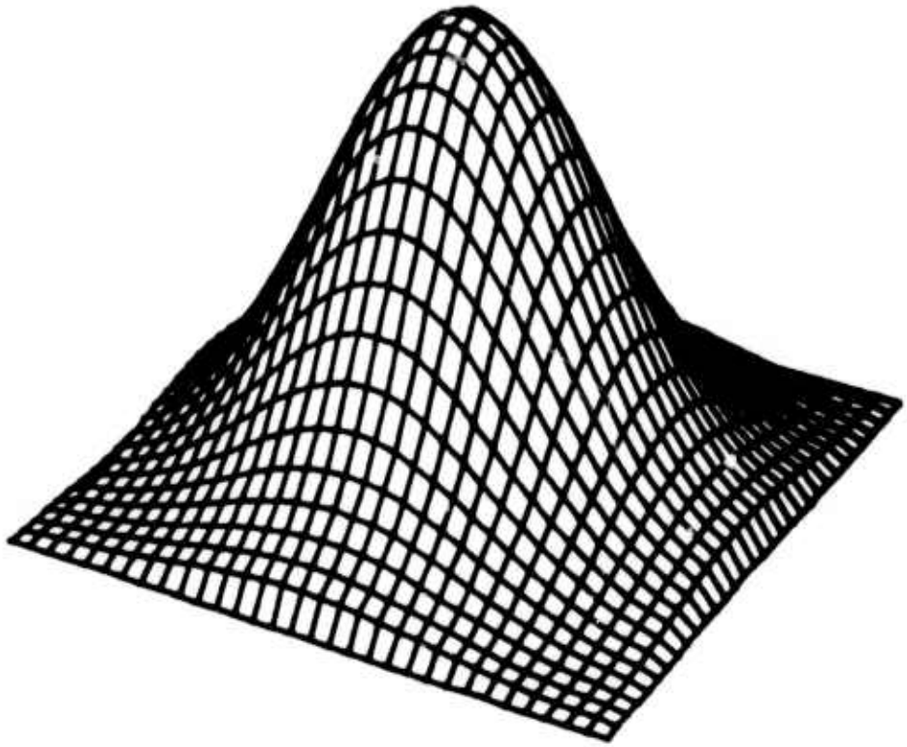


Figure (3-1) A Gaussian Point Spread Function

approximation, and Rf may vanish for some class of functions. If the nodes of the q.f. are selected in advance, the only available parameters to be treated are the coefficients c_i. Examples of this type (fixed node) are the method called pulse approximation (equal coefficient), Newton-Cotes, and the best q.f. in the sense of Sard. If the nodes of a q.f. are free, the best location of the nodes, in a certain sense, can be determined, and the q.f. is called optimal. Examples of the optimal type are Gauss-Legendre and the optimal q.f. in the sense of Sard. In image restoration the nodes x_i are usually preassigned, therefore, only fixed node quadrature formulae are discussed here.

3. 3 Spline Functions and Sard's Best Quadrature Formulae

Given a set of real numbers

$$x_0=u < x_1 < x_2 < \dots < x_n < v=x_{n+1} \tag{3-3}$$

a spline S(x) of degree m with the nodes $x_0, x_1, \dots, x_{n+1}$ is a function defined on the real line so that in each interval (x_i, x_{i+1}), for $i=0,1,\dots,n$, S(x) is represented by a polynomial of degree m or less, and the function and its derivatives of order m-1 or less are continuous on [u,v], thus S(x) is m-1 times continuously differentiable [3-2].

To represent splines, truncated power functions can be employed to construct a set of basis functions for the spline space. A truncated power function is defined as

$$x_+^m \triangleq \begin{cases} x^m & x>0 \\ 0 & x\leq 0 \end{cases} \tag{3-4}$$

A spline function of degree m and number of nodes n, $S_{m,n}(x)$, has a unique representation [3-2] of the form

$$S_{m,n}(x)=\sum_{i=0}^{m} a_i x^i + (m!)^{-1}\sum_{i=1}^{n} c_i (x-x_i)_+^m \tag{3-5}$$

where a_i and c_i are unknown coefficients to be determined.

To develop Sard's best q.f., let K(x) be a monospline of degree m [3-3] with n preassigned nodes. By definition, a monospline of degree m is a spline of degree m-1 plus a polynomial of degree m; thus, K(x) can be formulated as

$$K(x)=\frac{x^m}{m!}*S_{m-1,n}(x) \tag{3-6}$$

It is known that an arbitrary monospline can give rise to a q.f. [3-4]. To achieve this, set

$$K^{(i)}(u)=K^{(i)}(v)=0 \tag{3-7}$$

for i=0,1,...,m-1, and using K(x) as the kernel [3-4], then

$$\int_u^v f(x)\,dx = \sum_{i=1}^{n} c_i\, f(x_i) + Rf \qquad (3\text{-}8)$$

where

$$Rf = \int_u^v f^{(m)}(x)\,K(x)\,dx \qquad (3\text{-}9)$$

Note that Rf vanishes if f is a polynomial of degree m-1 or less. If K(x) has the least square deviation (minimum norm) among all kernels of the form (3-6), then the q.f. (3-8) is called best in the sense of Sard. Thus,

$$K = \int_u^v [K(x)]^2\,dx = \text{minimum} \qquad (3\text{-}10)$$

Schoenberg [3-3], [3-5] has shown that there exists a unique monospline H(x) of degree 2m

$$H(x) = \frac{x^{2m}}{(2m)!} + S_{2m-1,n}(x) \qquad (3\text{-}11)$$

with nodes $x_1, x_2, \ldots, x_n$, that satisfies the following three conditions:

$$H(x_i) = 0 \qquad i = 1,2,\ldots,n \qquad (3\text{-}12a)$$

$$H^{(m+i)}(u) = 0 \qquad i = 0,1,\ldots,m-1 \qquad (3\text{-}12b)$$

$$H^{(m+i)}(v)=0 \qquad i=0,1,\ldots,m-1 \tag{3-12c}$$

In terms of H(x), the kernel K(x) of Sard's best q.f. is given by

$$K(x)=H^{(m)}(x) \tag{3-13}$$

and the minimum norm of K(x) is obtained as

$$\|K\| = \int_u^v [K(X)]^2\,dx=(-1)^m\int_u^v H(x)\,dx \tag{3-14}$$

By normalizing [u,v] to [-1,1] and applying condition (3-12b), H(x) simplifies to

$$H(x)=\frac{(x+1)^{2m}}{2m} - \sum_{i=1}^{m-1} a_i x^i - \sum_{i=1}^{n} c_i \frac{(x-x_i)^{2m-1}}{(2m-1)!} \tag{3-15}$$

Conditions (3-12a) and (3-12c) produce m+n independent equations, whose solution gives the coefficients a_i and c_i. An upper bound can be derived for the error term Rf using eq. (3-9). Thus,

$$Rf=\int_{-1}^{1} K(x) f^{(m)}(x)\,dx \leq \|K\| \left\| f^{(m)} \right\| \tag{3-16}$$

or

$$Rf \leq (-1)^m \int_{-1}^{1} H(x)dx \left\| f^{(m)} \right\| \qquad (3\text{-}17)$$

The q.f. of form eq. (3-8) with coefficients c obtained from eq. (3-12) has some interesting properties. By varying m from 1 to n, eq. (3-8) presents a large family of quadrature formulae. The case m=1, if x are placed uniformly, is sometimes called tne pulse approximation method. An upper bound for the error term Rf can be derived when m<n. This property is an important one because it makes study of the error possible even in the simple case of pulse approximation. When m=n the technique is called Newton-Cotes method. Newton-Cotes q.f. results in zero error for the class of polynomials of degree n-1 or less, but no explicit error term is given if the integrand does not belong to this class. The following example is designed to aid the reader in better understanding of the best q.f. in the sense of Sard.

Let m=1 and assume x_i are uniformly placed on [-1,1]. Figure (3-2) illustrates the location of the nodes for the case of n=5. When the nodes are placed uniformly on [-1,1], the location of the nodes is obtained from

$$x_i = -1 - \frac{1}{n} + 2\frac{i}{n} \qquad (3\text{-}18)$$

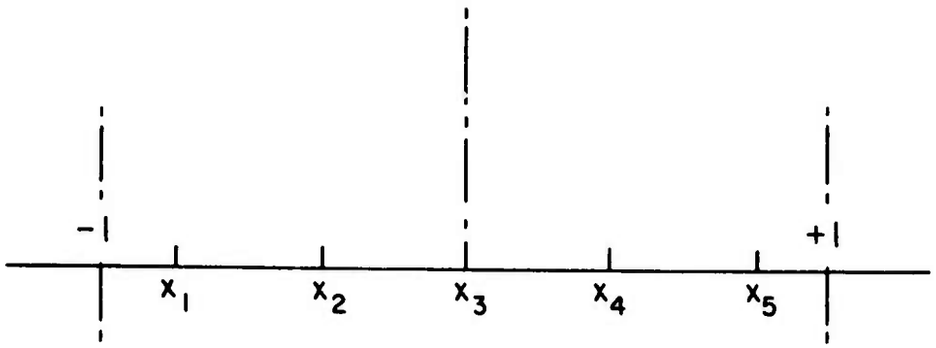


Figure (3-2) Uniform spacing of the nodes.

Here the expression for H(x) is given by

$$H(x)=\frac{(x+1)^2}{2} -a_0 -\sum_{}^{n} c_i (x-x_i)_+ \qquad (3\text{-}19)$$

Equation (3-12b) then reduces to

$$\sum_{i=1}^{n} c_i = 2 \qquad (3\text{-}20)$$

and eq. (3-12a) becomes

$$\sum_{j=1}^{n} c_j (x_i - x_j)_+ - a_0 = \frac{(x_i+1)^2}{2} \qquad (3\text{-}21)$$

For i=1,2,...,n. From equations (3-20) and (3-21), a_0 and c_i are obtained as

$$a_0 = \frac{1}{2n} \qquad (3\text{-}22a)$$

$$c_i = \frac{2}{n} \qquad (3\text{-}22b)$$

for i=1,2,...,n. Substituting (3-22b) in (3-8), the latter equation becomes

$$\int_{-1}^{1} f(x)\,dx = \frac{2}{n}\sum_{i=1}^{n} f(x_i) + Rf \qquad (3\text{-}23)$$

To estimate Rf, the norm of K(x) must be established.

Therefore,

$$\int_{-1}^{1} H(x)dx=\frac{(x+1)^3}{6}\Bigg|_{-1}^{1} \frac{1}{2n}x\Bigg|_{-1}^{1} -\frac{1}{n}\sum_{i=1}^{n}(1-x_i)^2 \qquad (3\text{-}24)$$

$$=+\frac{4}{3}-\frac{1}{n}-\frac{(4n-1)}{3n}=-\frac{2}{3n}$$

Thus the norm of K(x) can be obtained as

$$\left\| K \right\| =-\int_{-1}^{1} H(x)dx=\frac{2}{3n} \qquad (3\text{-}25)$$

Using the inequality of form eq. (3-16), the upper bound of Rf is then established as

$$Rf\leq \left\|K\right\| \left\| f''\right\| =(\frac{2}{3n}) \left\| f'\right\| \qquad (3\text{-}26)$$

And, as one would expect

$$\lim_{n\to\infty} Rf=\emptyset \qquad (3\text{-}27)$$

In the process of determining a q.f., one is faced with the task of selecting parameter m. Equation (3-9) requires that, for a given m, $f^{(m)}$ must exist, and this condition is met by most functions dealt with in this dissertation for any value of m. The stability of the

discrete system as well as the error term Rf are affected by the choice of m. Reference [3-6] studies the behavior of Rf for different classes of functions regarding specific values of m. In view of the conclusion at [3-6], m=1, 2, and 3 are considered to be good choices for the problems discussed in this dissertation.

3. 4 The Overdetermined Model

A discrete image degradation system which is of full column rank is characterized as overdetermined. In practice this situation arises from either a dark background in the object scene. or over-sampling of the observation [3-7]. Assuming that the object exists only on some given interval [a,b], the observation at a given point x is formulated as

$$g(x_i)=\int_a^b f(s)h(x_i-s)ds \qquad (3\text{-}28)$$

where x_i is in $[a-\frac{L}{2},b+\frac{L}{2}]$ and h(x) is a space limited function. Thus,

$$h(x_i-s)=\emptyset \qquad \text{if} \quad x_i-s>\frac{L}{2} \qquad (3\text{-}29)$$

Considering the above condition, eq. (3-28) becomes

$$g(x_i)=\int_{x_i-L/2}^{x_i+L/2} f(s)h(x_i-s)\,ds \qquad (3\text{-}30)$$

Using a q.f. on the integral of eq. (3-30) and taking the nodes at a discrete set of points, the above model is discretized as

$$g(x_i)=\sum_j c_j\ f(s_j)h(x_i-s_j) \qquad (3\text{-}31)$$

If the nodes s_j are placed uniformly (a valid assumption in regard to image restoration problems), s_j can be assumed to coincide with the integers on the real line without loss of generality. Thus

$$g(x_i)=\sum_j c_j\ f(j)h(x_i-j) \qquad (3\text{-}32)$$

Equation (3-32) formulates a general discrete image degradation process. If the observation g(x) is sampled uniformly at points $x_i=i$, then

$$g(i)=\sum_{j=i+\frac{L-1}{2}}^{i+\frac{L-1}{2}} c_j\ f(j)h(i-j) \qquad (3\text{-}33)$$

Employing vector space notation, eq. (3-33) can be presented in a more compact form. To construct the vector space model, assume the object is defined with N samples on [a,b]. Thus an object vector, $\underline{f}$, can be constructed as

$$\underline{f}_i = f(i) \tag{3-34}$$

for i=1,2,...,N. Assuming that the number of observed samples is M, an M-dimensional vector, $\underline{g}$, can be defined as

$$\underline{g}_i = g(i) \tag{3-35}$$

for i=1,2,...,M. The relationship between M and N is then given by

$$M = N+L-1 \tag{3-36}$$

The vector space formulation of eq. (3-33) is

$$\underline{g} = \underline{\underline{D}}\,\underline{f} \tag{3-37}$$

Where $\underline{\underline{D}}$ is the overdetermined blur matrix, defined by

$$\underline{D} = \begin{bmatrix} c_1 h_1 & 0 & 0 & 0 & \cdots & 0 \\ c_2 h_2 & c_1 h_1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & & & \\ c_L h_L & c_{L-1} h_{L-1} & 0 & \cdots & \cdots & 0 \\ 0 & c_L h_L & & & & \\ & & \ddots & & & \\ \vdots & \vdots & & & & c_1 h_1 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & c_L h_L \end{bmatrix} \qquad (3\text{-}38)$$

where

$$\underline{h}=[h_1, h_2, \ldots, h_L] \tag{3-39}$$

is the impulse response vector.

Equation (3-37) states the fundamental relationship which exists between an object vector and the corresponding observation. Under the present model there are more observed samples than unknown parameters, which is a direct consequence of the condition imposed on the object scene. This condition (a scene with dark setting) is equivalent to knowledge that the object is in the window. Figure (3-3) illustrates that equal rate sampling of the observation and the object results in more observed quantities than the unknown parameters. The stability of the system of equations defined by eq. (3-37) can be examined by studying the condition number of $\underline{D}$. This number depends on the shape and the variance of the blur function as well as the choice of the g.f. coefficients c_i. Reference [3-8] contains a study of the condition number of overdetermined systems versus the degrading function h.

3. 5 The Underdetermined Model

Continuing on the discussion of the previous section, a more realistic model evolves if no restrictions are

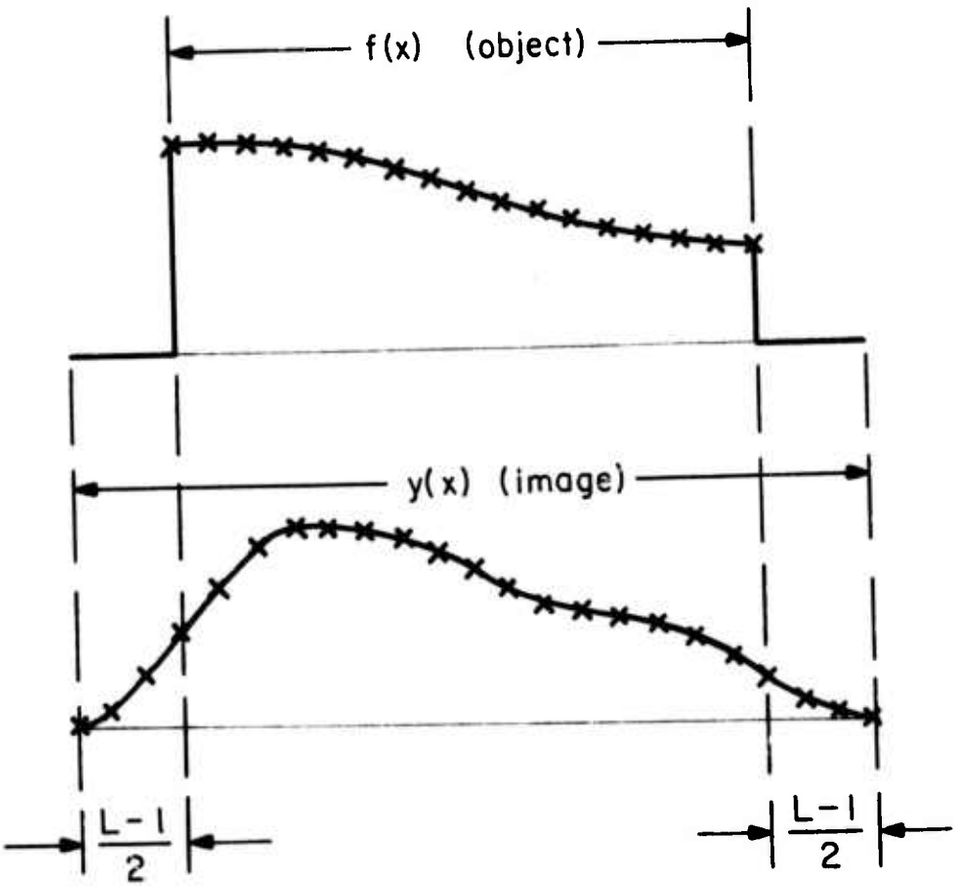


Figure (3-3) Imaging with dark background.

imposed upon the background of the scene. The model developed in section (3.4) is rarely very accurate because most scenes are not necessarily placed in a setting which has zero intensity, or even satisfy the less restrictive condition of constant intensity. Often, in the process of restoring an object, the observed image is partitioned into many smaller portions, and these small sections are processed separately. This partitioning process, by itself, contradicts any assumption made on the setting which surrounds the image because different portions of an image do not share the same background as the whole image does.

Assuming lack of information about the background of a scene, the object is assumed to extend very far in both directions, and consequently, so does the image. But, being able to handle only finite segments, one should be able to construct a model which relates portions of the object to corresponding segments of the observation. Figure (3-4) illustrates the concept.

Since the physical extent of the observation $g(x)$ is smaller than the physical extent of the object $f(x)$, the discrete observation $\underline{g}$ is represented by a smaller number of samples than the discrete object $\underline{f}$. This, of course, holds true when the sampling rate is kept the same for both the object and the observation. Thus, for an equal

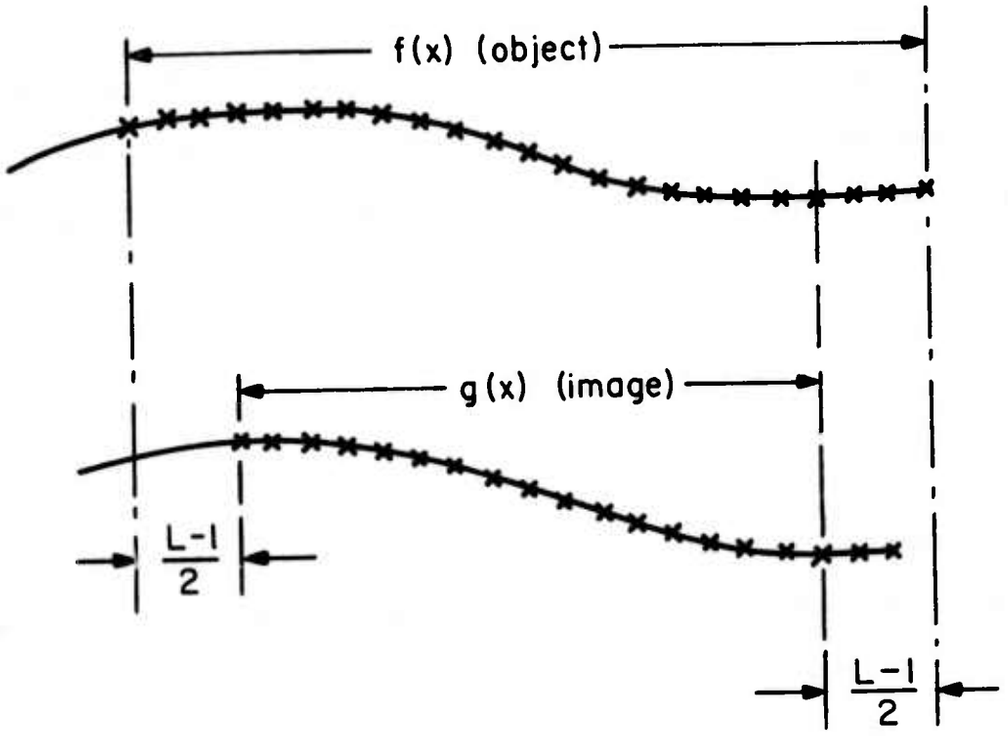


Figure (3-4) Imaging with unknown background.

sampling rate, the system defining the degradation phenomenon is equivalent to a set of linear equations which contains a greater number of unknown parameters than the number of equations available for solving for these parameters. The i-th equation is given by

$$g(x_i)=\sum_j c_j f(s)h(x_i-s) \qquad (3\text{-}40)$$

where x_i is in $[a+L/2,b-L/2]$. Following the approach adopted in section (3-4), an N-dimensional object vector $\underline{f}$, and an M-dimensional vector $\underline{g}$ are defined. Using notation $\underline{B}$ for the blur matrix, the model is expressed as

$$\underline{g}=\underline{B}\underline{f} \qquad (3\text{-}41)$$

where $\underline{B}$ is an MxN matrix, and

$$M=N-L+1 \qquad (3\text{-}42)$$

Since M< N, the blur matrix is not of full column rank, and for this reason $\underline{B}$ is called an underdetermined matrix. The blur matrix $\underline{B}$ is defined as

$$\underline{B} = \begin{bmatrix} c_L h_L & c_{L-1} h_{L-1} & \cdots & c_1 h_1 & 0 & \cdots & 0 \\ 0 & c_L h_L & \cdots & c_1 h_1 & 0 & & \\ \vdots & & \ddots & & \ddots & & \\ & & & c_L h_L & & c_1 h_1 & 0 \\ 0 & & & 0 & c_L h_L \cdots\cdots & & c_1 h_1 \end{bmatrix} \qquad (3\text{-}43)$$

Unlike the blur matrix $\underline{D}$ of the previous section, $\underline{B}$ does not possess a finite condition number. Thus, as will be discussed in the next chapter, the degradations introduced under the model of eq. (3-41) are impossible to remove completely. Another major difference between the two models is that although eq. (3-41) is a more realistic model for image degradation phenomena, it does not possess a structure which leads to the computational simplicity of the overdetermined model of eq. (3-37) for purposes of image restoration.

REFERENCES

1. I. Schoenberg, "Monosplines and Quadrature Formulae," Theory and Applications of Spline Functions, New York: Academic Press, 1969, pp. 158.

2. T. Greville, "Introduction to Spline Functions," Theory and Applications of Spline Functions, pp. 1.

3. I. Schoenberg, "On Monosplines of Least Square Deviation and best Quadrature Formulae,", J. SIAM Numer. Anal. Sec. B2, 1965, pp. 144-170.

4. I. Schoenberg, "Monosplines and Quadrature Formulae," pp. 163

5. I. Schoenberg, "On Monosplines of Least Square Deviation and best Quadrature Formulae II" SIAM J. Numer. Anal., 3, 1966, pp. 321-328.

6. F. Davarian and M. Peyrovian, "Quadrature Formulae Application in Image Restoration," University of Southern California Technical Report, USCIPI Report 560, March 1975, pp. 60-67.

7. F. Davarian, "Nonuniform Sampling of Observation

Space," USCIPI Report 540, September 1974, pp. 48-56.

8. N. Mascarenhas, Digital Image Restoration Under a Regression Model- The Unconstrained, Linear Equality and Inequality Constraint Approaches, Ph. D. Dissertation, University of Southern California, January 1974.

4. NOISE FREE RESTORATION

For the sake of simplicity, let $\underline{H}$ represent a general blur matrix of size M by N. The mathematical expression governing the discrete degradation phenomenon, with $\underline{H}$ as the degrading matrix, is given by

$$\underline{g}=\underline{H}\underline{f} \tag{4-1}$$

The above expression is a vector space equality, and in essence, it is a set of M linear equations with N unknown parameters. The most straightforward approach for recovering $\underline{f}$, is to effectively invert $\underline{H}$. If $\underline{H}$ is square and nonsingular, the estimate $\underline{f}$ is obtained as

$$\hat{\underline{f}}=\underline{H}^{-1}\underline{g} \tag{4-2}$$

where $\underline{H}^{-1}$ is the inverse of $\underline{H}$. Unfortunately, $\underline{H}$ is seldom a square matrix, and even if so, $\underline{H}$ may not be invertible. To define an inverse process for the system eq. (4-1) which will work in all circumstances, a different inverse for $\underline{H}$ is defined as

$$\underline{H}^{+} = \lim_{d\to 0} (\underline{H}^{T}\underline{H} + d^{2} I)^{-1} H^{T} \tag{4-3a}$$

for rank N or

$$\underline{H}^{+} = \lim_{d \to 0} \underline{H}^{T} (\underline{H}\underline{H}^{T} + d^{2} \underline{I})^{-1} \qquad (4\text{-}3b)$$

for rank M, where $\underline{H}^{+}$ is called the pseudoinverse of $\underline{H}$ [4-1], and $\underline{I}$ denotes the identity matrix. It has been proven that $\underline{H}^{+}$, as defined by eq. (4-3), always exists and is unique [4-2]. The most attractive property of $\underline{H}^{+}$ is stated as follows. Given the observation $\underline{g}$, the estimate

$$\hat{\underline{f}} = \underline{H}^{+} \underline{g} \qquad (4\text{-}4)$$

is the vector of minimum norm among those which minimize

$$\underline{\&} = \underline{g} - \underline{H}\underline{f} \qquad (4\text{-}5)$$

The above property of $\underline{H}^{+}$ introduces the key approach to the restoration techniques adopted for many image restoration applications. If $\underline{H}$ possesses some specific structure, computation of $\underline{H}^{+}$ may be simplified greatly. For example, if $\underline{H}$ is a square and singular matrix whose first r diagonal entries are nonzero

$$\underline{H} = \begin{bmatrix} H(1) & 0 & 0 & \cdot & \cdot & 0 \\ 0 & H(2) & 0 & & & 0 \\ \cdot & & \cdot & & & \cdot \\ \cdot & & & H(r) & & \\ \cdot & & & & 0 & \\ \cdot & & & & & \cdot \\ \cdot & & & & & 0 \end{bmatrix} \qquad (4\text{-}6)$$

the pseudoinverse of $\underline{H}$ becomes

$$\underline{H}^{+} = \begin{bmatrix} H(1)^{-1} & 0 & 0 & \cdot & \cdot & 0 \\ 0 & H(2)^{-1} & 0 & & & 0 \\ \cdot & & \cdot & & & \cdot \\ \cdot & & & H(r)^{-1} & & \\ \cdot & & & & 0 & \\ \cdot & & & & & \cdot \\ \cdot & & & & & 0 \end{bmatrix} \qquad (4\text{-}7)$$

If $\underline{H}$ is of full column rank or full row rank (two frequently encountered situations in image restoration tasks), simple forms for $\underline{H}^{+}$ can be obtained. The next section deals with the problem when the rank of a matrix is equal to its row size.

4. 1 Blur Matrix with Full Row Rank

Let $\underline{B}$ denote the blur matrix possessing full row rank. Then, the image degradation model is described by

$$\underline{g}=\underline{B}\underline{f} \tag{3-41}$$

The minimum norm estimate of an object blurred under eq. (3-41) is given as

$$\hat{\underline{f}}=\underline{B}^{+}\underline{g} \tag{4-8}$$

where $\underline{B}^{+}$ can be computed from

$$\underline{B}^{+} = \lim_{d\to 0} \underline{B}^{T}(\underline{B}\underline{B}^{T}+d^{2}\underline{I})^{-1} \tag{4-9}$$

Since $\underline{B}$ has full row rank, $\underline{B}\underline{B}^{T}$ is nonsingular, and the limit on the right hand side of eq. (4-9) can be carried out yielding

$$\underline{B}^{+}=\underline{B}^{T}\{\lim_{d\to 0}(\underline{B}\underline{B}^{T}+d^{2}\underline{I})\}^{-1} \qquad (4\text{-}10a)$$

or [4-3]

$$\underline{B}^{+}=\underline{B}^{T}(\underline{B}\underline{B}^{T})^{-1} \qquad (4\text{-}10b)$$

Although eq. (4-10b) suggests a much simpler method for computing $\underline{B}^{+}$ than eq. (4-9), an N by N matrix $\underline{B}\underline{B}^{T}$ must still be inverted.

There are two drawbacks to the estimation method described in this section. The first stems from the fact that the model does not allow full recovery of the object vector $\underline{f}$ since there are fewer equations than unknown parameters in the system of eq. (3-41). The second drawback is in the need for inverting a matrix of size N by N. For moderate sizes of the object, the ill conditioning of the matrix $\underline{B}\underline{B}^{T}$ could cause difficulties [4-4]. When relatively large images are to be processed, the limited size of available computers could put an intolerable restriction on the size of the object. Often in situations like this, the observation must be broken into smaller segments, each of which is used to estimate the corresponding object section. Notice that although the estimate, $\hat{\underline{f}}$, is not in general equal to the object, $\underline{f}$, the following equality always holds

$$\underline{B}\hat{\underline{f}}=\underline{B}\underline{f}=\underline{g} \tag{4-11}$$

4. 2 Blur Matrix with Full column Rank

Equation (3-37) describes the model, where $\underline{D}$ is the overdetermined blur matrix of form given by eq. (3-38)

$$\underline{g}=\underline{D}\underline{f} \tag{3-37}$$

Since $\underline{D}$ has full column rank, $\underline{D}^T\underline{D}$ is always invertible. Using eq. (4-3a), the pseudoinverse of $\underline{D}$ can be obtained as [4-3]

$$\underline{D}^{+}=(\underline{D}^T\underline{D})^{-1}\underline{D}^T \tag{4-12}$$

Thus, the minimum norm estimate is given by

$$\underline{f}=\underline{D}^{+}\underline{g} \tag{4-13}$$

Since system eq. (4-9) has more observed parameters than unknowns, the object $\underline{f}$ can be recovered with no error. Thus,

$$\hat{\underline{f}}=\underline{D}^{+}\underline{g} \tag{4-14a}$$

or

$$\hat{\underline{f}}=(\underline{D}^T\underline{D})^{-1}\underline{D}^T\underline{Df} \qquad (4\text{-}14b)$$

or

$$\hat{\underline{f}}=\underline{f} \qquad (4\text{-}15)$$

The above equality states the basic characteristic of overdetermined systems. Full recovery of the object is an advantage which only systems of full column rank enjoy. Another superiority of these systems is in the possiblity of introducing efficient methods which drastically reduce the computational complexity of the associated filters. The next section introduces a technique in which the overdetermined model is modified to pave the road for constructing computationally simple filters.

4. 3 A Circulant Model

The objective in this section is to establish an image degradation vector space model equivalent to the one stated by eq.(3-37), in which the blur matrix $\underline{D}$ is replaced by a circulant blur matrix $\underline{C}$. A circulant matrix can best be explained by illustration: a K by K circulant matrix has the following particular structure.

$$\underline{C} = \begin{bmatrix} c_1 & c_2 & \cdot & & c_K \\ c_K & c_1 & \cdot & \cdot & c_{K-1} \\ \cdot & & & & \cdot \\ c_2 & \cdot & & & c_1 \end{bmatrix} \qquad (4\text{-}16)$$

Each row, or column, of the above matrix is a circular right shift of the row, or column, immediately preceeding. This property extends from the last row (column) to the first row (column) since the first row is a circulant shift of the last row (column).

To set up a degradation model with a blur matrix of structure defined by eq. (4-15), two auxiliary vectors are defined as follows. Let K be an integer, where $K \geq M$ and define an extended object vector of size K, $\underline{f}_c$, where

$$\begin{aligned} \underline{f}_c(i) &= f(i) && \text{for } i=1,2,\ldots,N \\ &= \emptyset && \text{for } i=N+1,\ldots,K \end{aligned} \qquad (4\text{-}17)$$

Likewise form an extended observation vector $\underline{g}_c$ of size K

$$\underline{g}_c(i)=\underline{g}(i) \qquad \text{for } i=1,2,\ldots,M$$

$$=\emptyset \qquad \text{for } i=M+1,\ldots,K \tag{4-18}$$

Next, placing the impulse response vector in the first column, construct a circulant matrix

$$\underline{C}=\begin{bmatrix} h(1) & \emptyset & \cdots & \emptyset & h(L-1) & \cdots & h(2) \\ \vdots & & & & & & \vdots \\ & & & & & \emptyset & h(L-1) \\ h(L) & & & & & & \\ \emptyset & & & & & & \\ \vdots & & & & & & \\ \emptyset & \cdots & \cdots & \emptyset & h(L) & \cdots & h(1) \end{bmatrix} \tag{4-19}$$

The discrete convolution summation of eq. (3-30) is equivalent to the vector space equality [4-5]

$$\underline{g}_c = \underline{C}\underline{f}_c \tag{4-20}$$

Equation (4-20) which presents the desired model, is similar to eq. (3-37), where $\underline{g}$, $\underline{D}$, and $\underline{f}$ are replaced by

$\underline{g}_c$, $\underline{C}$, and $\underline{f}_c$.

It is known that a K by K matrix which possesses K independent eigenvectors can be diagonalized through a similarity transformation [4-6]. Hunt [4-7] demonstrates that matrices of the form of eq. (4-15) have K independent eigenvectors, and that the Fourier transform diagonalizes circulant matrices. In fact, the eigenvectors of circulants are the Fourier basis vectors. Let $\underline{A}$ represent the Fourier transform matrix, thus

$$A(i,k)=\exp\{-(\frac{2\Pi j}{K})ik\} \tag{4-21}$$

The similarity transform which diagonalizes the blur matrix $\underline{C}$ of eq. (4-19) is

$$\underline{C}=\underline{A}^{-1}\underline{\Lambda}\,\underline{A} \tag{4-22}$$

where $\underline{\Lambda}$ is the diagonal matrix of the eigenvalues of $\underline{C}$

$$\underline{\Lambda} \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \bullet & & \\ & & & \bullet & \\ & & & & \lambda_K \end{bmatrix} \tag{4-23}$$

The λ_i are obtained by Fourier transforming the first column of $\underline{C}$. Thus,

$$\lambda_i = \sum_{k=0}^{L-1} \exp\{-(i-1)k\frac{2\Pi j}{K}\}h(k+1) \qquad (4\text{-}24)$$

for i=1,...,K. To study eq. (4-20) in Fourier space, substitute eq. (4-22) in eq. (4-20) to obtain

$$\underline{g}_c = \underline{A}^{-1}\,\underline{\Lambda}\,\underline{A}\underline{f}_c \qquad (4\text{-}25)$$

Next, eq. (4-27) is rearranged as the following

$$\underline{A}\underline{g}_c = \underline{\Lambda}\underline{A}\underline{f}_c \qquad (4\text{-}26)$$

By definition $\underline{A}\underline{g}_c$ and $\underline{A}\underline{f}_c$ are the discrete Fourier transforms of the vectors $\underline{g}_c$ and $\underline{f}_c$, respectively.

$$\underline{F}_c = \underline{A}\underline{f}_c \qquad (4\text{-}27a)$$

$$\underline{G}_c = \underline{A}\underline{g}_c \qquad (4\text{-}27b)$$

where

$$F_c(i) = (\frac{1}{K})\sum_{k=0}^{K-1} f_c(k+1)\exp\{-(i-1)k\frac{2\Pi j}{K}\}$$
$$G_c(i) = (\frac{1}{K})\sum_{k=0}^{K-1} g_c(k+1)\exp\{-(i-1)k\frac{2\Pi j}{K}\} \qquad (4\text{-}28)$$

for i=1,...,K. Thus, in the transform domain, eq. (4-20)

simplifies to

$$\underline{G}_c = \underline{\Lambda F}_c \tag{4-29}$$

Since $\underline{\Lambda}$ is a diagonal matrix, eq. (4-29) is actually a scalar equation

$$G_c(i) = \lambda_i F_c(i) \tag{4-30}$$

for i=1,...,K. To restore $\underline{F}_c$, multiply both sides of eq. (4-29) by Λ^{-1}. Thus,

$$\hat{\underline{F}}_c = \underline{\Lambda^{-1} G}_c \tag{4-31}$$

or in the scalar form

$$\hat{F}_c(i) = \frac{1}{\lambda_i} G_c(i) \qquad i=1,..,K \tag{4-32}$$

Inverse Fourier transforming of $\hat{\underline{F}}_c$ results in the estimate

$$\hat{\underline{f}}_c = \underline{A}^{-1} \hat{\underline{F}}_c \tag{4-33}$$

The object estimate, $\hat{\underline{f}}$, can be obtained by extracting out the first N entries of $\hat{\underline{f}}_c$. That is,

$$\hat{\underline{f}} = \underline{S1}_K^N \hat{\underline{f}}_c \tag{4-34}$$

where $\underline{S1}_K^N$ is a selection matrix of the following form

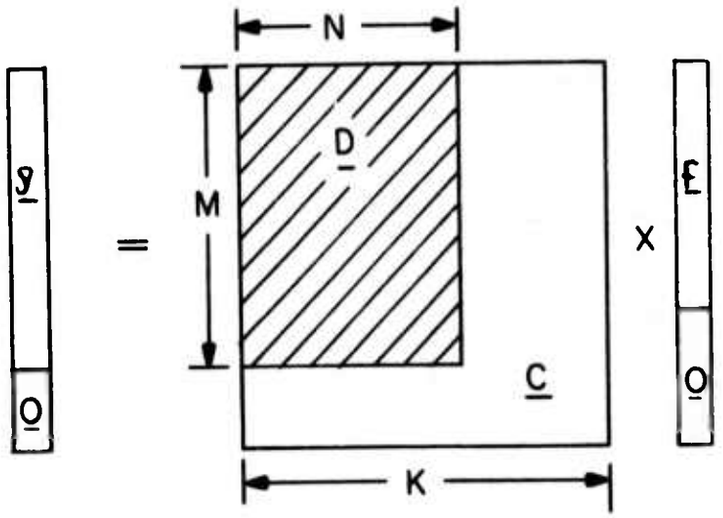
$$\underline{S1}_K^N = \overbrace{\left[\begin{array}{c|c} \underline{I} & \underline{\emptyset} \end{array}\right]}^{K} \Bigg\} N \qquad (4\text{-}35)$$

Figure (4-1) illustrates the relationship between the models expressed by eq. (3-38) and eq. (4-20).

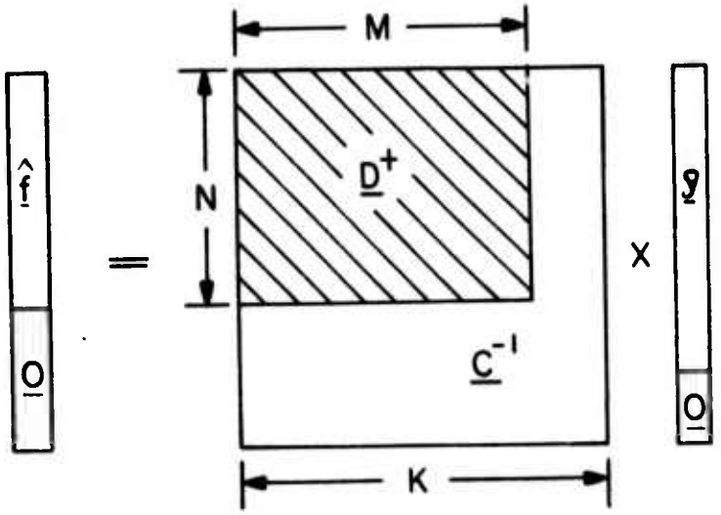
As noted earlier, the model developed in this section is equivalent to the overdetermined model established in section (3. 4); thus, the restriction of objects with dark background still exists. The incentive for taking a detour by introducing the circulant model is in the present system's computational superiority over the previous model. The inverse filter, $\underline{\Lambda}^{-1}$, used in estimating $\underline{f}_c$ does not require a matrix inversion. Thus, the ill-conditioning, or the large size of the system is not a major obstacle in the process of computing the circulant filter.

4. 4 Experimental Results

Figure (4-2) illustrates an object with zero



a) $\underline{g}_c = \underline{C}\ \underline{f}_c$



b) $\hat{\underline{f}}_c = \underline{C}^{-1}\underline{g}_c$

Figure (4-1) The relationship between the overdetermined model and the circulant model.

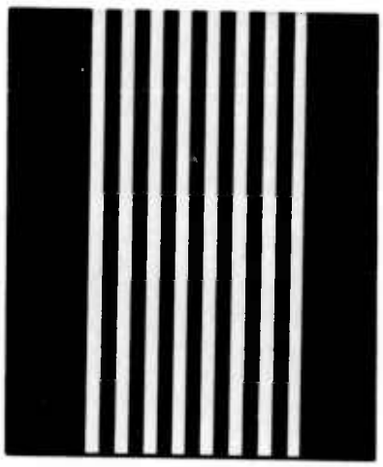
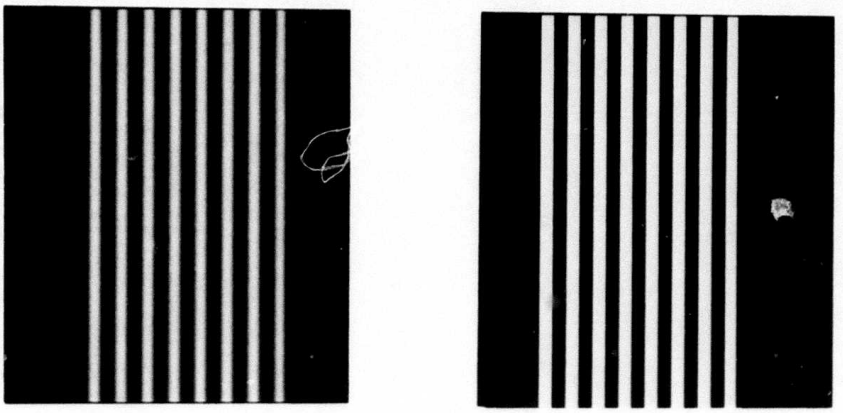


Figure (4-2) The test object.

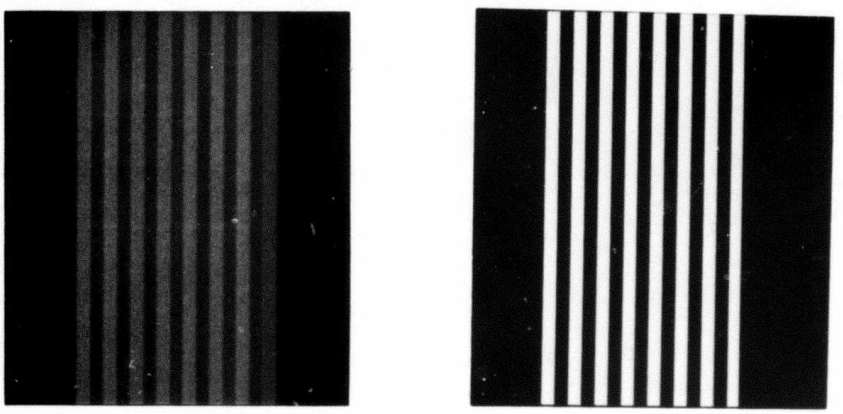
background. The object in this case is a set of bars of constant intensity separated from each other by a set of bars with zero intensity. The object and its background form a square picture which is sampled at 256x256 uniformly spaced points. Figure (4-3a) shows this object after undergoing a one-dimensional blur degradation of Gaussian shape with standard deviation of 2.5 pixels. Figure (4-3b) shows the estimate. This estimate has been obtained using the method explained in section 4. 3. Since the restoration has been performed in the Fourier domain (a scalar operation), the relatively large size of the image is of no concern. The estimate is identical to the object itself. This interesting achievement holds for all examples in which the object possesses a black background, and the environment is noise free. Figure (4-4a) illustrates the object after undergoing an extreme amount of blur. The degradation models a motion blur for which L=15. Figure (4-4b) is the estimate which is error free.

If the images discussed above were to be restored using an underdetermined model (section 4. 1), the large size of the image would require that the observation be partitioned into smaller segments. Note that an estimate of the form given by eq. (4-8) cannot be obtained by direct Fourier techniques [4-5], [4-8]. Figure (4-5) plots the error for the underdetermined model estimate. Each observed segment has 25 pixels and is assumed to estimate



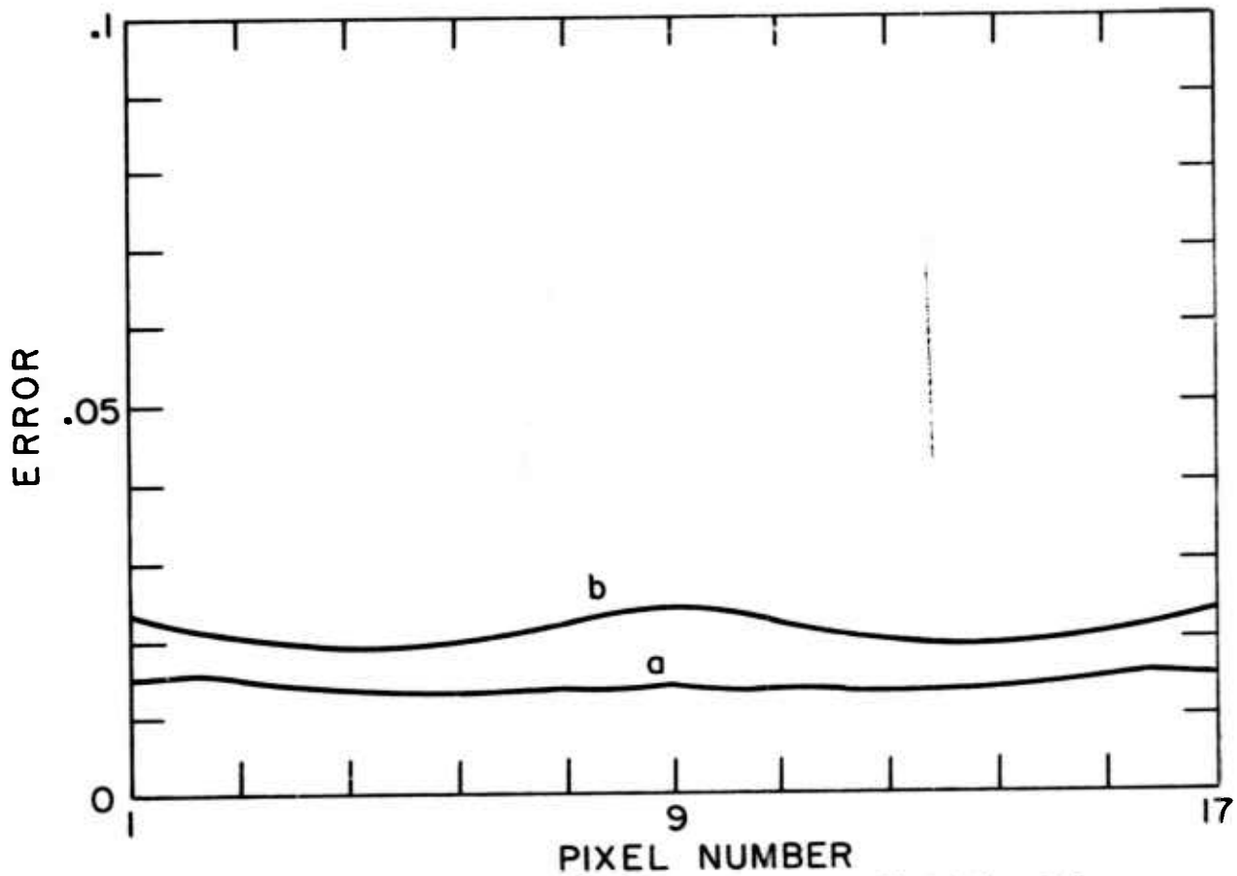
(a) Blurred (b) Restored

Figure (4-3) Image restoration with Gaussian shape blur.



(a) Blurred (b) Restored

Figure (4-4) Image restoration with motion blur.



a) Gaussian shape blur b) Motion blur

Figure (4-5) Error for underdetermined estimation.

17 pixels of the object. The error is plotted for two different blur models: a Gaussian shaped blur of standard deviation 2.5; and motion blur. In both cases L is assumed to be 9.

REFERENCES

1. A. Albert, Regression and the Moore Penrose pseudoinverse, New York: Academic Press, 1972.

2. A. Albert, Reg. Moore-Penrose, pp. 19.

3. A Albert, Reg. Moore-Penrose, pp. 21.

4. N. Mascarenhas, Digital Image Restoration Under a Regression Model- The Unconstrained, Linear Equality and Inequality Constrained Approaches, Ph. D. Dissertation, University of Southern California, January 1974.

5. W. Pratt, "Vector Space Formulation of Two Dimensional Signal Processing Operations," Journal of Computer Graphics and Image Processing, March 1975.

6. J. Moore, Elements of Linear Algebra and Matrix Theory, New York: MaGraw-Hill, 1968, pp. 291.

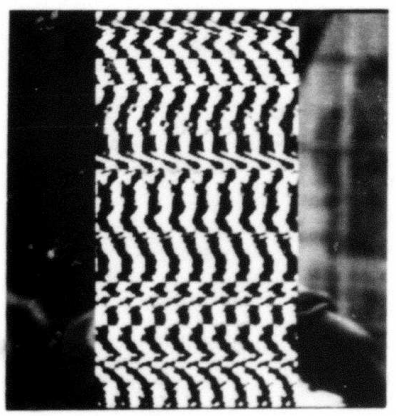
7. B. Hunt, "A Matrix Theory Proof of the Discrete Convolution Theorem," IEEE Trans. Audio Electroacoust, December 1971, pp. 285-288.

8. W. Pratt, "Transform Domain Signal Processing Techniques," Proc. National Electronics Conference, Chicago, IL, 1974.

5. WINDOWING OPERATOR

The previous section introduced a circular image degradation model which resulted in a computationally efficient restoration technique. Unfortunately, this technique can only be applied to a scene which possesses a dark background. Hence, portions of a given image, or an object placed in a nonzero intensity setting cannot be recovered by brute forcing the circulant filter on the observed image. In fact, since this kind of imaging can not be expressed by an overdetermined model, employment of a system of the form of eq. (4-20) for representing the degradation phenomenon, when objects of unknown background are involved, can cause a modelling error at the boundaries of the image. This error can be looked at as signal correlated noise; hence, considering the ill-conditioning inherent in imaging models, the implementation of the circulant filter becomes catastrophic. To illustrate this claim, figure (5-1a) is selected as a test image. This image is obtained by slightly blurring the original object. Figure (5-1b) shows the same image after an attempt is made to restore the center portion of the object by utilizing the fast Fourier technique of eq. (4-31). Observing the high frequency noise component in figure (5-1b), it is evident that an unwise choice of a degradation model not only does not improve the observation, but, on the contrary, it could be quite destructive. To approach this

(a) Observed



(b) Processed

Figure (5-1) Image restoration using wrong model.

problem in a more systematic manner, the relationship between the background of the object, the system model, and the observed image must be studied. The next section discusses this subject and suggests means of treating a given observation in order to modify the physical samples to suit a desired model.

5. 1 Overdetermined System with Unrestricted Observation

The primary difference between the two sets of equations (3-41) and (3-37) is stated as follows. There are only M equations (or observed parameters) in the system of eq. (3-41) to solve for M+L-1 unknown parameters, thus, since L>1, an exact restoration of the object is impossible. On the other hand, the system described by eq. (3-37) has M observations for only M-L+1 unknowns. This means that number of equations is actually larger than the minimum number necessary for a complete restoration of the object. Now, considering the assumption made on the data of the latter model, it appears that the overdetermined model is in essence equivalent to its underdetermined counterpart if some of the unknown parameters of the underdetermined system are set to be equal to zero. An M by N underdetermined system is represented by

$$\underline{g}=\underline{\underline{B}}\underline{f} \qquad (3\text{-}41)$$

The following lemma states the compatiblity of system eq. (3-41) with an overdetermined system of form eq. (3-37), where the corresponding blur matrix is of size M by M-L+1.

Lemma 5-1. The underdetermined system of eq. (3-41) becomes equivalent to an overdetermined system of the form eq. (3-37) if the first and the last L-1 entries of the object vector $\underline{f}$ are set equal to zero.

Proof. Let $\underline{e}$ represent the object vector after the first and the last L-1 entries are set to zero.

$$\underline{e}(i)=f(i) \qquad \text{if } L<i<N-L+1$$
$$\underline{e}(i)=\emptyset \qquad \text{otherwise} \qquad (5\text{-}1)$$

and assume $\underline{d}$, a vector of size M-L+1, represents the nonzero center portion of $\underline{e}$. The observation corresponding to object $\underline{e}$ is given as follows

$$\underline{g}=\underline{\underline{B}}\underline{e} \qquad (5\text{-}2)$$

Notice that

$$\underline{B}\underline{e}=\underline{D}\underline{d} \tag{5-3}$$

where $\underline{D}$ is the overdetermined matrix of dimensions M by M-L+1 and is given by eq. (3-34). Equality (5-3) holds because of the particular structure that both $\underline{D}$ and $\underline{B}$ have. If eq. (5-3) is substituted in eq. (5-2) then

$$\underline{q}=\underline{D}\underline{d} \tag{5-4}$$

Equation (5-4) is the desired result. To find $\underline{q}$, subtract eq. (3-41) from eq. (5-2) giving

$$\underline{q}-\underline{g}=\underline{B}(\underline{e}-\underline{f}) \tag{5-5}$$

or

$$\underline{q}=\underline{g}-\underline{B}(\underline{f}-\underline{e}) \tag{5-6}$$

Introducirg an N by N selection matrix $\underline{S}_N^L$,

$$\underline{S}_N^L = \begin{bmatrix} \underline{I} & & \\ & \underline{0} & \\ & & \underline{I} \end{bmatrix} \begin{matrix} \\ \\ \} K \\ \\ \underbrace{\quad}_{L} \end{matrix} \tag{5-7}$$

eq. (5-6) simplifies to

$$\underline{g}=\underline{g}-\underline{B}\underline{S}_N^L\underline{f} \tag{5-8}$$

Equation (5-8) holds true since

$$\underline{f}-\underline{e}=\underline{S}_N^L\underline{f} \tag{5-9}$$

The role of $\underline{S}_N^L$ is to select the first and the last L-1 entries of $\underline{f}$. Vector $\underline{f}-\underline{e}$, in fact, represents the background of the object, and vector $\underline{B}(\underline{f}-\underline{e})$ in eq.(5-6) represents the contribution of this background to the image.

What lemma (5-1) suggests is restated as follows. Any observed image can be suitably processed for an overdetermined model provided that the intensity function describing the setting of the object is obtainable. The major drawback of the above statement is that the intensity function of the surrounding of the object is not usually known a priori. But often this function can be estimated with an acceptable accuracy. This is the subject of tne following section

5. 2 Windowing of the observation

Since the only available source of information

concerning the scene is the observed vector, the estimator which estimates the object background must take use of the the physical sample vector $\underline{g}$. Assume that matrix $\underline{W}$ represents the combination of a background estimator plus the system which removes the effect of the estimated background from the observation. The product of the image vector, $\underline{g}$, by matrix $\underline{W}$ results in the desired observation, $\underline{q}$, which can, successfully, be used in an overdetermined system. The structure of matrix $\underline{W}$ (the windowing matrix) is explained in the following paragraph.

Let $\underline{q}$ represent the observation vector if the first and the last L-1 entries of $\underline{f}$ are zero. The objective here is to express $\underline{q}$ in terms of the elements of the observation $\underline{g}$ according to the relation

$$q(i)=\begin{cases} g(i)-\sum_{j=1}^{L-1} h(L+1-j)\,f(j+i-1) & \text{if } i<L \\ g(i)_{L-1-M+i} & L<i<M-L \\ g(i)-\sum_{j=1} h(j)\,f(N+1-j+i-M) & \text{if } i>M-L+1 \end{cases} \qquad (5\text{-}10)$$

Since the entries of $\underline{f}$ are not known, the correspondence of eq. (5-11) cannot be made directly. However by making an assumption on the continuity of the original image vector that

$$f(i)=f(L) \qquad \text{for } i<L \qquad (5\text{-}11)$$

and

$$f(N-i+1)=f(N-L) \qquad i>N-L \tag{5-12}$$

then the vector $\underline{g}$ can be estimated as

$$\underline{g}=\underline{W}\underline{g} \tag{5-13}$$

where $\underline{W}$ is an M by M matrix of the form given by eq. (5-14).

A zero order predictor is inherent in the structure of matrix $\underline{W}$ expressed by eq. (5-14). The prediction of the image background is the main idea in expression (5-13). Therefore, the success of the operation defined by matrix $\underline{W}$ depends on the validity of the prediction method used to obtain $\underline{W}$. There are, of course, other prediction methods which can be employed. For example a first order predictor results in a smaller (overall) mean square error. Figure

$$\underline{W} = \begin{bmatrix} h(1) & 0 & 0 & \cdots & 0 \\ \sum_{i=3}^{L} h(i) & 1 & \cdot & & \cdot \\ \cdot & & 1 & & \\ \sum_{i=L}^{L} h(i) & 0 & & & \\ 0 & & & & \\ \cdot & & & & 0 \\ & & & & -\sum_{i=1}^{1} h(i) \\ & & & & \cdot \\ & & & & 1-\sum_{i=1}^{L-2} h(i) \\ 0 & & & & 0 \quad h(1) \end{bmatrix} \qquad (5\text{-}14)$$

(5-2) illustrates the expected mean square restoration error of the object-estimate for two prediction algorithms, as a function of element correlation p. Fortunately the zero order predictor, in practice, produces sufficiently accurate results. Since this predictor has a simple form, it has been employed in the remaining material of this text.

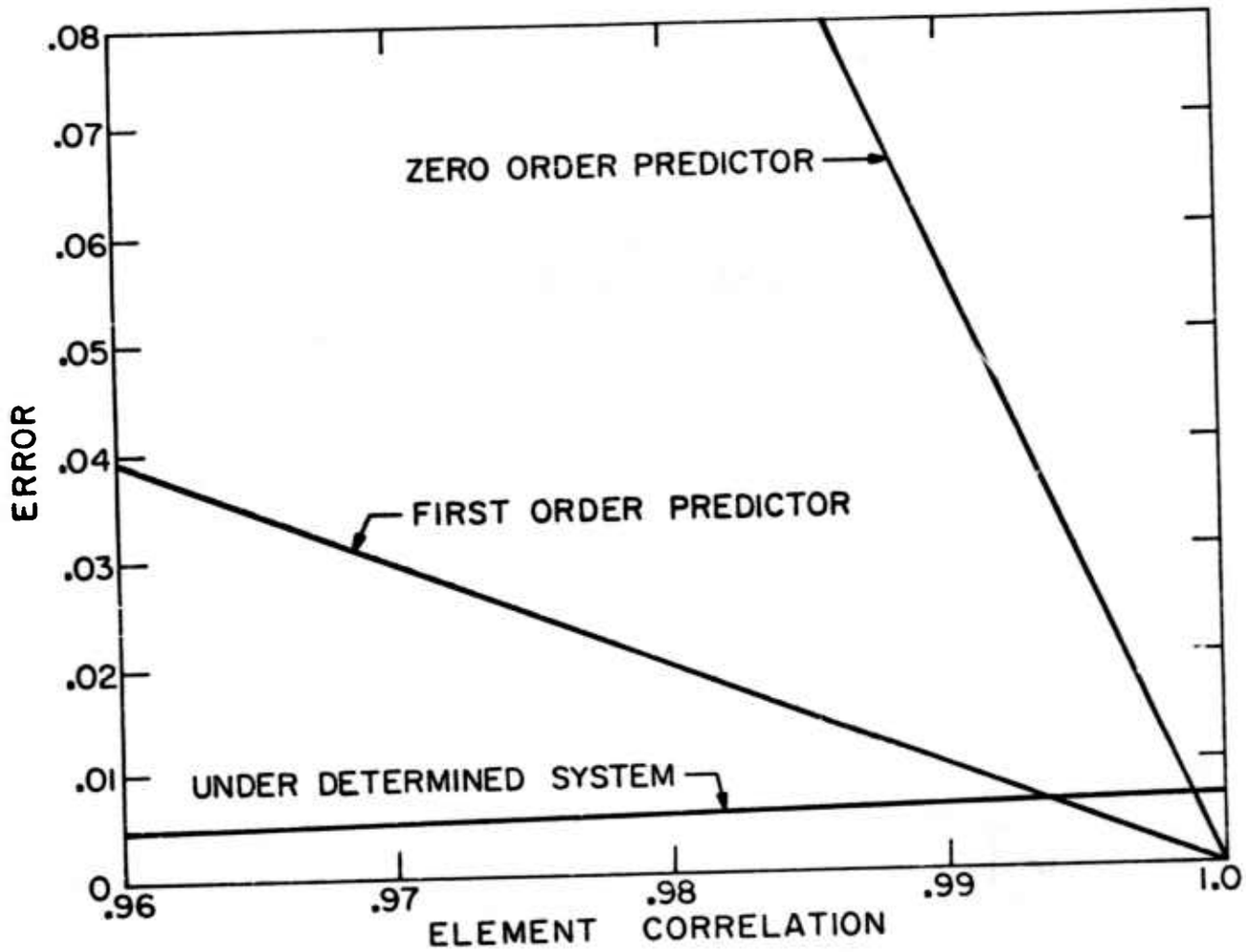


Figure (5-2) Restoration error for fast pseudoinverse restoration

It is interesting to note that the physical contribution of multiplying an image vector by $\underline{W}$ is an attenuation of the first and the last L-1 entries of the corresponding vector. This is expected since the observation resulting from the class of objects encircled by a dark region illustrates dim object boundaries (boundary of the object with its background). Figure (5-3a) represents the general form of a typical observed image line while figure (5-3b) shows the same image function after undergoing a windowing operation.

5. 3 Error Analysis

Assume $\underline{x}$, a vector of size N, represents the object, and let

$$\underline{g}=\underline{B}\underline{x} \qquad (5\text{-}15)$$

symbolize the observed image after $\underline{x}$ has undergone a degradation of known impulse response. The size of $\underline{g}$ is given by

$$M=N-L+1 \qquad (3\text{-}38)$$

The objective here is to estimate the center elements of $\underline{x}$ using the overdetermined system model. Since the physical sample vector $\underline{g}$ and the system equation (3-41) are not

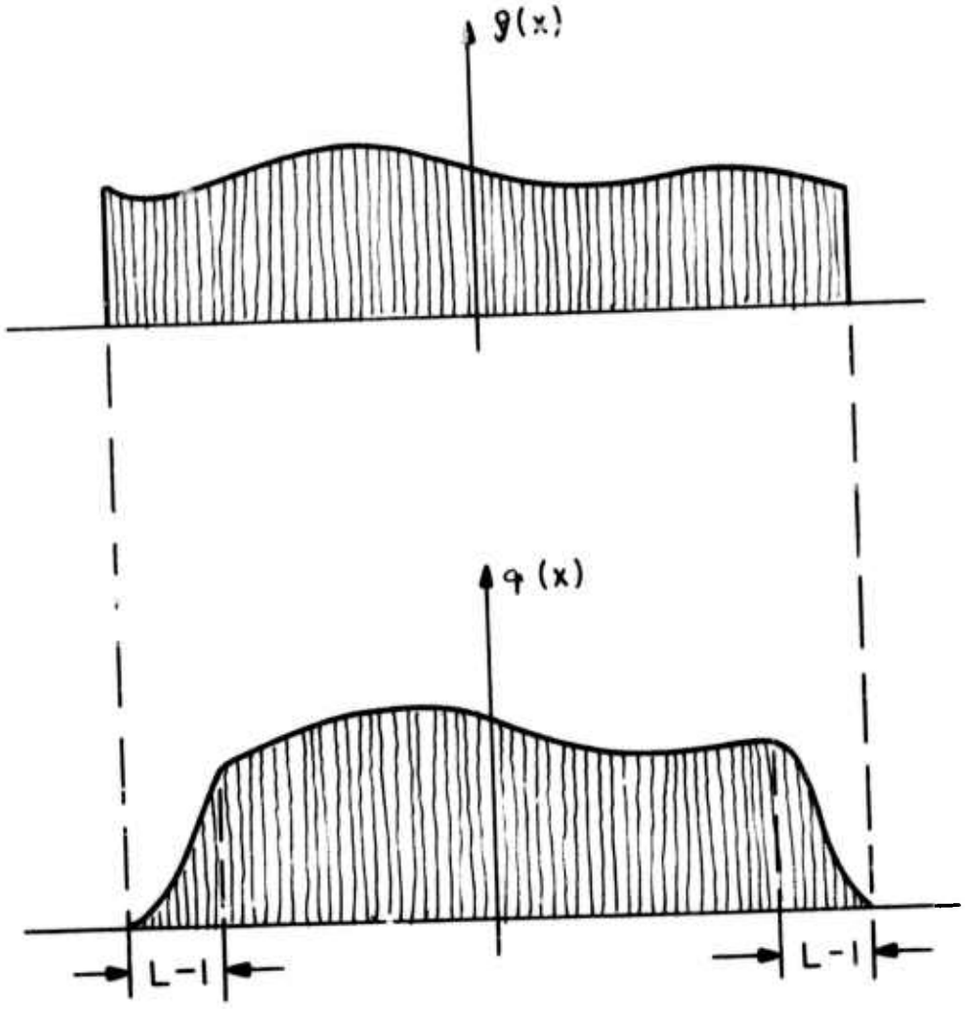


Figure (5-3) A blurred image line.

compatible, the observed vector $\underline{g}$ must undergo the windowing operation described in the previous section. The modified observation, q, is obtained as

$$\underline{q}=\underline{W}\underline{g} \tag{5-16}$$

or

$$\underline{q}=\underline{W}\underline{B}\underline{x} \tag{5-17}$$

At this stage, the filter derived for the overdetermined system, Sec. 4-2, can be employed to estimate the center part of $\underline{x}$. Therefore,

$$\hat{\underline{f}}=\underline{D}^{+}\underline{q} \tag{5-18}$$

Where, $\underline{D}^{+}$ is given by eq. (4-12), and $\hat{\underline{f}}$ is the estimated center part of $\underline{x}$. The length of $\hat{\underline{f}}$, K, is given by

$$K=M-L+1 \tag{5-19}$$

If expression (5-17) is used to substitute for $\underline{q}$, eq. (5-18) becomes

$$\hat{\underline{f}}=\underline{D}^{+}\underline{W}\underline{B}\underline{x} \tag{5-20}$$

Let $\underline{f}$, a vector of size K, denote the center portion of the ideal image vector $\underline{x}$. This vector can be extracted from $\underline{x}$ using the selection matrix

$$\underline{S2}_N^K = \overbrace{\left[\begin{array}{c|c|c} \underline{0} & \underline{I} & \underline{0} \end{array}\right]}^{N} \Bigg\} K \qquad (5\text{-}21)$$

Thus,

$$\underline{f} = \underline{S2}_N^K \underline{x} \qquad (5\text{-}22)$$

Figure (5-4) illustrates the correspondence between the vectors $\underline{x}$, $\underline{g}$, $\underline{g}$, and $\underline{f}$.

The estimation error, $\underline{e}$ is defined as

$$\underline{e} = \underline{f} - \hat{\underline{f}} \qquad (5\text{-}23)$$

To analyze the error vector from a statistical point of view, the error vector $\underline{e}$ is assumed to be mean zero. this

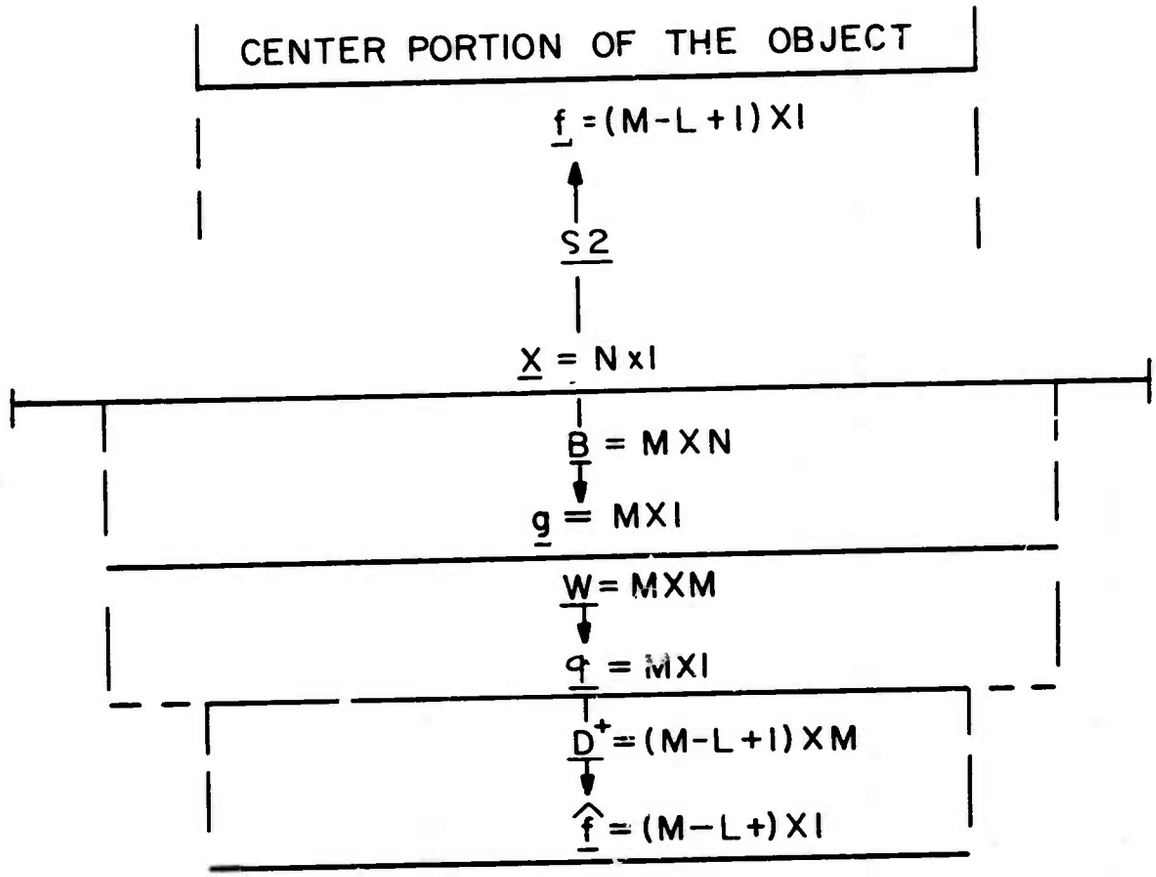


Figure (5-4) The relationship between the physical observed samples and the object.

assumption is merely made for the sake of convenience. It is always possible to subtract the mean value of the object from $\underline{x}$. This will automatically modify all the corresponding vectors to become mean zero. The error covariance matrix, $\underline{R}_e$, is defined by

$$\underline{R}_e = E\{\underline{e}\underline{e}^T\} \tag{5-24}$$

Where, the notation E denotes the ensemble average. Using the expression for $\underline{e}$ from eq. (5-23), $\underline{R}_e$ becomes

$$\underline{R}_e = E\{(\underline{f}-\hat{\underline{f}})(\underline{f}-\hat{\underline{f}})^T\} \tag{5-25}$$

or

$$\underline{R} = E\{\underline{f}\underline{f}^T\} + E\{\hat{\underline{f}}\hat{\underline{f}}^T\} - E\{\underline{f}\hat{\underline{f}}^T\} - E\{\hat{\underline{f}}\underline{f}^T\} \tag{5-26}$$

Let

$$\underline{R}_x = E\{\underline{x}\underline{x}^T\} \tag{5-27}$$

represent the correlation matrix of the object. The right hand side of eq. (5-26) can then be evaluated element by element as

$$E\{\underline{f}\underline{f}^T\} = \underline{S2}_N^K \, E\{\underline{x}\underline{x}^T\} (\underline{S2}_N^K)^T \tag{5-28}$$

or

$$E\{\underline{\hat{f}}\underline{\hat{f}}^T\}=\underline{S2}_N^K \underline{R}_x (\underline{S2}_N^K)^T \tag{5-29}$$

The second term can be evaluated as

$$E\{\underline{\hat{f}}\underline{\hat{f}}^T\}=\underline{D}^+ \underline{WB}[E\{\underline{x}\underline{x}^T\}]\underline{B}^T \underline{W}^T (\underline{D}^+)^T \tag{5-30}$$

or

$$E\{\underline{\hat{f}}\underline{\hat{f}}^T\}=\underline{D}^+ \underline{WBR}_x \underline{B}^T \underline{W}^T (\underline{D}^+)^T \tag{5-31}$$

The expected value of $\underline{f}\underline{f}^T$ can be obtained as

$$E\{\underline{f}\underline{\hat{f}}^T\}=\underline{D}^+ \underline{WB}[E\{\underline{x}\underline{x}^T\}](\underline{S2}_N^K)^T \tag{5-32}$$

or

$$E\{\underline{f}\underline{\hat{f}}^T\}=\underline{D}^+ \underline{WBR}_x (\underline{S2}_N^K)^T \tag{5-33}$$

Likewise $E\{\underline{\hat{f}}\underline{\hat{f}}^T\}$ can be derived as

$$E\{\underline{\hat{f}}\underline{\hat{f}}^T\}=\underline{S2}_N^K \underline{R}_x \underline{B}^T \underline{W}^T (\underline{D}^+)^T \tag{5-34}$$

Now the total error covariance can be expressed as

$$\underline{R}_e=\underline{D}^+\underline{WBR}_x\underline{B}^T\underline{W}^T\underline{D}^{+T}+\underline{S2}_N^K\underline{R}_x(\underline{S2}_N^K)^T-\underline{D}^+\underline{WBR}_x(\underline{S2}_N^K)^T$$
$$-\underline{S2}_N^K\underline{R}_x\underline{B}^T\underline{W}^T(\underline{D}^+)^T \qquad (5\text{-}35)$$

It is possible to process the physical samples of the blurred image, $\underline{g}$, with the filter derived under the underdetermined system model assumption. In this case the filter is given by eq. (3-41). Therefore, the estimate, $\hat{\underline{x}}$, is given by

$$\hat{\underline{x}}=\underline{B}^+\underline{g} \qquad (5\text{-}36)$$

And the center portion of the estimate, $\hat{\underline{f}}$, is simply obtained by premultiplying eq. (5-36) by $\underline{S2}_N^K$ to yield

$$\hat{\underline{f}}=\underline{S2}_N^K\hat{\underline{x}} \qquad (5\text{-}37)$$

or

$$\hat{\underline{f}}=\underline{S2}_N^K\underline{B}^+\underline{g} \qquad (5\text{-}38)$$

The error term

$$\underline{e}=\underline{f}-\hat{\underline{f}} \qquad (5\text{-}39)$$

has the following covariance

$$\underline{R}_e = E\{(\underline{f}-\hat{\underline{f}})(\underline{f}-\hat{\underline{f}})^{\hat{T}}\} \qquad (5\text{-}40)$$

Going through the steps similar to those described in the previous case, the error covariance can be derived as

$$\underline{R}_e = \underline{S2}_N^K [\underline{I}-\underline{B}^+\underline{B}]\underline{R}_x[\underline{I}-\underline{B}^+\underline{B}](\underline{S2}_N^K)^T \qquad (5\text{-}41)$$

where, $\underline{I}$ is the M by M identity matrix. Figure (5-2) of the previous section illustrates the expected mean square restoration error of $\underline{f}$ as a function of the correlation of elements in $\underline{f}$ under the assumption that $\underline{f}$ is a sample of a Markov process with correlation factor p.

Figure (5-2) illustrates that, except when the object element correlation coefficient is near unity, the error resulting from the estimate given by eq. (5-20) is larger than the one resulting from eq. (5-38). But, considering that most images observed by a human viewer possess strong correlation among their sampled pixels, this extra contribution of error is not, usually, unreasonably high in practice. Also, with the $\underline{D}$ operator, contained in eq. (5-20), it is possible to perform the restoration by Fourier domain processing quite efficiently.

5. 4 Experimental results

Figure (5-5a) illustrates an image corrupted by a Gaussian blur modelling a blur degradation caused by imaging through a turbulent atmosphere. The center portion of the image has been filtered to produce the corresponding section of the object. A zero order predictor was employed for estimating the object background, and the restoration is performed in the Fourier domain as indicated in Sec. 5. 3. Figure (5-5b) shows the image after the central portion of the object has been restored. Unfortunately, since the background predictor is not error free, the filter is not applicable for severe amounts of blur. Since the system modelling the degradation process is basically ill-conditioned, the restoration technique greatly amplifies any uncertainty in the observation. Figure (5-6a) illustrates a test object after undergoing a severe amount of blur. Figure (5-6b) is an attempt to restore the object, which clearly has been unsuccessful. If the background of the object is of constant intensity, the zero-order predictor inherent in the windowing matrix can make an exact estimate of the background. This would insure an error free observation. Figure (5-7a) illustrates this case. The center part of the object has been processed to artificially generate a constant intensity background. Figure (5-7b) is the restored version of fig. (5-7a). As before, only the center part of

(a) Observed

(b) Restored

Figure (5-5) Examples of fast pseudoinverse image restoration.

(a) Observed

(b) Restored

Figure (5-6) Image restoration with severe blur

(a) Observed

(b) Restored

Figure (5-7) Image restoration when the object possesses a constant background.

the object is restored.

6. NOISY RESTORATION

The models formulating the image degradation phenomenon in the preceeding chapters have ignored the presence of noise. In practice, however, imaging systems are seldom totally noise free. In image forming systems, noise or uncertainty may arise from a variety of sources; probably the most common sources of noise are measurement and recording errors. Scanners, the devices which measure images, invariably add an element of uncertainty to the measured (or scanned) signal [6-1]. Usually, after an image is scanned, it is operated upon in a computer of finite precision, creating truncation errors. Coding and channel errors are caused if the particular image is to be transmitted through a noisy channel [6-1]. Lastly, film noise may be added to the image when the signal is recorded [6-2]. It is, of course, impractical to make an exhaustive list of all noise producing elements in an image forming process, but the noise sources listed above are the most significant.

The next section studies the continuous image degradation problem in the presence of additive white noise.

6. 1 The Continuous Model

The continuous image degradation model of eq. (2-1) is modified in the presence of additive white noise to

$$g(x)=\int_{-\infty}^{\infty} f(s)h(x-s)ds \quad +n(x) \qquad (6\text{-}1)$$

where n(x) represents the noise component. As a result of the existence of noise in the system, the inverse filtering technique cannot be employed to recover the object. If any attempt is made to force the inverse filter to process the physical sample image g(x), the high frequency component of the noise will be greatly amplified ruining the restoration. To avoid high frequency amplification of the noise, the inverse filter can be truncated at a certain point. The point can be selected so that beyond this point the noise power exceeds the signal energy [6-3]. Figure (6-1) illustrates this method.

A more intelligent technique to recover the object in the presence of noise is the classical Wiener filter. This filter controls the noise component by keeping into account the ratio between the signal power and the noise variance at each point of the Fourier space. The Wiener filter output is a minimum mean-square error estimate of the original object, provided that the statistics used in the filter are carefully obtained.

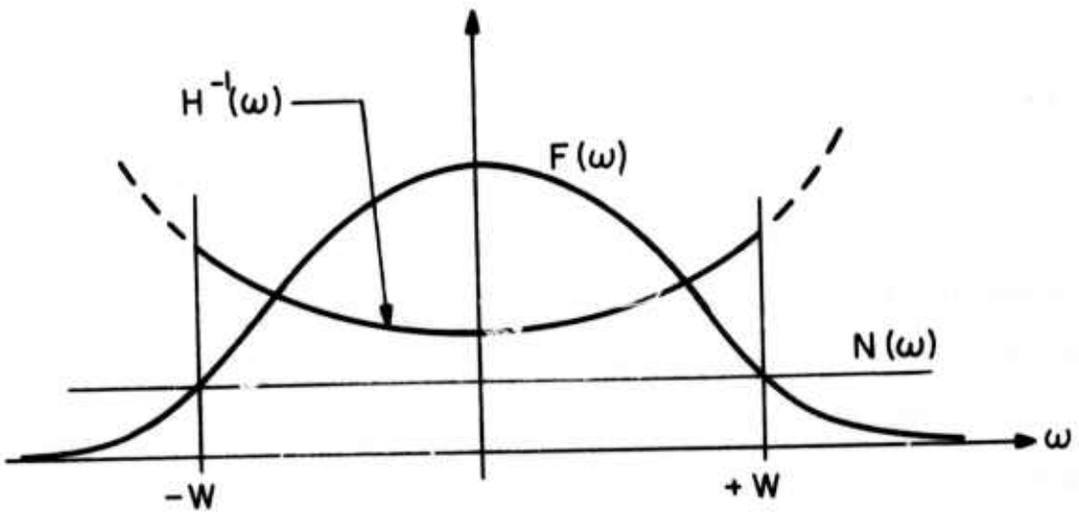


Figure (6-1) Truncated inverse filter.

6. 2 The Discrete Wiener Filter

In the presence of noise, the system governing the image degradation phenomenon is formulated as [6-3]

$$\underline{g}=\underline{\underline{H}}\underline{f}+\underline{n} \tag{6-2}$$

where $\underline{\underline{H}}$ is a blur matrix and $\underline{n}$ is the noise component vector. To derive the minimum mean-square error filter for the system described by eq. (6-2) the well known orthogonality principle [6-4], [5-6] can be employed. Letting $\underline{\underline{U}}$ represent the desired filter; the estimate $\underline{f}$ is obtained as

$$\hat{\underline{f}}=\underline{\underline{U}}\underline{g} \tag{6-3}$$

According to the orthogonality principle the following equality must hold

$$E(\underline{f}-\hat{\underline{f}})\underline{g}^T=\underline{\underline{\emptyset}} \tag{6-4}$$

After substituting for $\underline{f}$ from eq. (6-3) and carrying out the appropriate steps, the solution is given as [6-3]

$$\underline{\underline{U}}=\underline{\underline{R}}_f\,\underline{\underline{H}}^T(\underline{\underline{H}}\underline{\underline{R}}_f\,\underline{\underline{H}}^T+\underline{\underline{V}})^{-1} \tag{6-5}$$

where $\underline{\underline{R}}_f$ is the correlation matrix of the object and $\underline{\underline{V}}$ is

the noise correlation matrix. The object and the noise are assumed to be uncorrelated. The following matrix identity [6-6]

$$\underline{A}\underline{B}^T(\underline{C}+\underline{B}\underline{A}\underline{B}^T)^{-1}=(\underline{A}^{-1}+\underline{B}^T\underline{C}^{-1}\underline{B})^{-1}\underline{B}^T\underline{C}^{-1} \tag{6-6}$$

can be used to obtain a different formulation for filter $\underline{U}$ given by

$$\underline{U}=(\underline{R}_f^{-1}+\underline{H}^T\underline{V}^{-1}\underline{H})^{-1}\underline{H}^T\underline{V}^{-1} \tag{6-7}$$

The error variance matrix $\underline{C}_e$ is given as [6-7]

$$\underline{C}_e=\underline{R}_f-(\underline{R}_f\underline{H}^T)(\underline{H}\underline{R}_f\underline{H}^T+\underline{V})^{-1}(\underline{R}_f\underline{H}^T)^{-1} \tag{6-8}$$

And by using the matrix identity [6-6]

$$(\underline{C}^{-1}+\underline{B}^T\underline{A}\underline{B})^{-1}=\underline{C}-\underline{C}\underline{B}^T(\underline{B}\underline{C}\underline{B}^T+\underline{A})^{-1}\underline{B}\underline{C} \tag{6-9}$$

the error term simplifies to

$$\underline{C}_e=(\underline{R}^{-1}+\underline{H}^T\underline{V}^{-1}\underline{H})^{-1} \tag{6-10}$$

For a white noise process, the noise correlation matrix has the following simple form

$$\underline{V}=d^2\,\underline{I} \tag{6-11}$$

where d^2 is the noise power, and $\underline{I}$ denotes the identity matrix. Representing the image power by z^2, expression (6-7) becomes

$$\underline{U}=(z^{-2}\,\underline{C}_f+d^{-2}\,\underline{H}^T\underline{H})^{-1}\,\underline{H}^T d^{-2} \tag{6-12}$$

or

$$\underline{U}=(S^{-1}\underline{C}_f+\underline{H}^T\underline{H})^{-1}\,\underline{H}^T \tag{6-13}$$

where $\underline{C}_f$ represents the normalized object covariance matrix, and S represents the signal to noise ratio

$$S=\frac{z^2}{d^2} \tag{6-14}$$

Likewise, expression (6-5) simplifies to

$$\underline{U}=\underline{C}_f\,\underline{H}^T(\underline{H}\underline{C}_f\,\underline{H}^T+S^{-1}\underline{I})^{-1} \tag{6-15}$$

Although expressions (6-13) and (6-15) are equivalent, the first expression is employed when blur matrix $\underline{H}$ is of full column rank (type $\underline{D}$), and the second one is used if $\underline{H}$ has full row rank (type $\underline{B}$). The reason for this is simply because of computational savings. Note that for an

overdetermined matrix the number of columns is less than the number of rows (Sec. 3.4). Hence, the dimensionality of $\underline{H}^T\underline{H}$ (NxN) is less than the dimensionality of $\underline{H}\underline{H}^T$ (MxM). And, for an underdetermined matrix, the order is reversed; the dimensionality of $\underline{H}^T\underline{H}$ (NxN) is less than the dimensionality of $\underline{H}\underline{H}^T$ (MxM). Thus, for the case of overdetermined matrices an N by N matrix (N<M) has to be inverted, and in the other case (underdetermined blur matrix) the inversion of an M by M matrix (M<N) is required.

The necessity of inverting matrices of relatively large size is an unattractive property of the discrete Wiener filter. To prevent this, the next section introduces a technique which eliminates the matrix inversion requirement.

6. 3 The Fast Wiener Filter

Section (4. 3) introduced an image degradation model which, by utilizing the Fourier domain properties of circulant matrices, successfully eliminated the matrix inversion requirement of the pseudoinverse filter. Although the circular model provides attractive computational savings in a noise-free environment, it cannot be applied to a noisy observation. In fact, since the windowing operation of Sec. (5.2) introduces an element

of uncertainty in the observation space, the fast pseudoinverse technique is unable to recover the object after it has undergone a severe amount of degradation. To overcome this shortcoming, it is necessary to consider the existence of noise in the system. Thus, this section is devoted to the derivation of a system model which retains computational simplicity and, at the same time, tolerates noise-corrupted observation samples.

In the presence of additive noise, the image degradation model with a circulant blur matrix $\underline{C}$ is formulated as

$$\underline{g}_c = \underline{C}\underline{f}_c + \underline{n} \qquad (6\text{-}16)$$

According to the developments and the definitions established in Sec. (4. 3), both vectors $\underline{g}$ and $\underline{f}$ are of size K ($K \geq M$), and represent the hypothetical observation and object, respectively. Since the actual physical sample image, $\underline{g}$, is of size M and the actual object vector, $\underline{f}$, is of size N, the last K-M entries of $\underline{g}_c$ are pure noise, and the last K-N entries of $\underline{f}_c$ are deterministic and equal to zero. In searching for a minimum mean-square error filter corresponding to the observation in eq. (6-16), the Wiener filter may seem the logical candidate; however, this is not true. To illustrate this fact, consider the Wiener filter solution for the white noise case as given by

$$\underline{U}=(\underline{S}^{-1}\underline{C}_{f_c}^{-1}+\underline{C}^T\underline{C})^{-1}\underline{C}^T \tag{6-17}$$

where C_{f_c} is the covariance matrix of the object. Since the covariance matrix $\underline{C}_{f_c}$ is not circulant, the undesirable KxK matrix inverse operation cannot be avoided by simply Fourier transforming the filter. Furthermore, since some of the entries of $\underline{f}_c$ are deterministic, $\underline{C}_{f_c}$ is a singular matrix, and thus $\underline{C}_{f_c}^{-1}$ does not exist. The preceeding argument implies that a circulant Wiener filter is not achieveable because of the particular statistical properties of the vector $\underline{f}_c$. It is conceivable that a circulant and nonsingular covariance matrix could be obtained if the postulated object vector, $\underline{f}_c$, had a proper statistical background. To obtain a nonsingular covariance matrix, none of the entries of $\underline{f}_c$ can be deterministic. To achieve this result, form a new vector $\underline{f}_c$ by augmenting the actual object $\underline{f}$ with a vector $\underline{y}$ which has certain desired statistical properties to be described later. The augmented vector is

$$\underline{f}_c = \begin{bmatrix} \underline{f} \\ \\ \underline{y} \end{bmatrix} \tag{6-18}$$

The covariance matrix of the above vector is

$$E(\underline{f}_c \underline{f}_c^T) = \begin{bmatrix} \underline{C}_f & \underline{C}_{fy} \\ \underline{C}_{yf} & \underline{C}_y \end{bmatrix} \tag{6-19}$$

where

$$\underline{C}_f = E(\underline{f}\underline{f}^T) \tag{6-20}$$

is the actual object NxN covariance matrix, and

$$\underline{C}_{fy} = E(\underline{f}\underline{y}^T) \tag{6-21}$$

and

$$\underline{C}_y = E(\underline{y}\underline{y}^T) \tag{6-22}$$

is the (K-N)x(K-N) dimensional covariance matrix of $\underline{y}$. At this stage, the aim is to illustrate that a certain statistical assumption on $\underline{y}$ plus a specific value of K can result in a circulant and positive definite matrix $\underline{C}_{f_c}$.

As an example let the object vector, $\underline{f}$, consist of exactly four samples given by

$$\underline{f} = [f_1, f_2, f_3, f_4]^T \tag{6-23}$$

Assuming that $\underline{f}$ arises from a Markovian random process, the covariance matrix of $\underline{f}$ has the following form

$$\underline{C} = \begin{bmatrix} 1 & p & p^2 & p^3 \\ p & 1 & p & p^2 \\ p^2 & p & 1 & p \\ p^3 & p^2 & p & 1 \end{bmatrix} \tag{6-24}$$

where p is the element correlation coefficient. Since the size of $\underline{f}$, N, is four, $\underline{C}_f$ is of size 4x4. Selecting K to be equal to 6, a 6 by 6 circulant matrix is introduced

$$\underline{C}_{f_c} = \begin{bmatrix} 1 & p & p^2 & p^3 & p^2 & p \\ p & 1 & p & p^2 & p^3 & p^2 \\ p^2 & p & 1 & p & p^2 & p^3 \\ p^3 & p^2 & p & 1 & p & p^2 \\ p^2 & p^3 & p^2 & p & 1 & p \\ p & p^2 & p^3 & p^2 & p & 1 \end{bmatrix} \tag{6-25}$$

Since the matrix in eq. (6-25) is symmetric, it must be proven that it is positive definite in order to conclude that $\underline{C}_{f_c}$ is indeed a covariance matrix. Since $\underline{C}_{f_c}$ is circulant, Fourier transformation of the first column can generate the eigenvalues of this matrix [6-8]. Figure (6-2) lists all the eigenvalues of $\underline{C}_{f_c}$ according to their number. Since the element correlation is always less than

k	λ_k
1	$(1+p)(1+p+p^2)$
2	$(1+p)^2(1-p)$
3	$(1-p)^2(1+p)$
4	$(1-p)(1-p+p^2)$
5	$(1-p)^2(1+p)$
6	$(1+p)^2(1-p)$

(a) List of eigenvalues

k	λ_k
1	5.5623
2	.1901
3	.0048
4	.0476
5	.0048
6	.1901

(b) List of eigenvalues for the case of p=.95

Figure (6-2) Eigenvalues of a 6 by 6 extended object covariance matrix.

unity, all the eigenvalues are positive. Thus, $\underline{C}_{f_c}$ is a positive definite and circulant matrix. Since K=6 in this example, vector $\underline{y}$ is exactly of size 2

$$\underline{y}=[y_1,y_2]^T \tag{6-26}$$

According to eq. (6-25), $\underline{y}$ must have the following statistics ($\underline{y}$ is mean-zero)

$$\underline{C}_y = \begin{bmatrix} 1 & p \\ p & 1 \end{bmatrix} \tag{6-27}$$

and

$$\underline{C}_{fy} = \begin{bmatrix} p^2 & p \\ p^3 & p^2 \\ p^2 & p^3 \\ p & p^2 \end{bmatrix} \tag{6-28}$$

Equations (6-27) and (6-28) can be used to generate the random process samples $\underline{y}$.

To generalize the above example, the case of an N-dimensional object is considered here. Parameter K is selected according to the following rule

$$K=N+N-2 \tag{6-29}$$

And similar to the covariance matrix of eq. (6-25), a KxK covariance matrix is constructed using the NxN Markovian matrix. To acheive this, N-2 center entries of the first row are folded over and attached to the first row itself to produce the first line of the KxK circulant covariance matrix. Since the resulting matrix is circulant, only the first row is needed to construct the remainder of the matrix. Figure (6-3a) illustrates the first row of an NxN Markovian matrix, and fig. (6-3b) shows the resulting first line of the circulant matrix. Figure (6-4) illustrates the resulting KxK circulant matrix. The eigenvalues of this matrix can be obtained by Fourier transforming the first row of the matrix itself.

Let a K-dimensional vector $\underline{r}$, which represents the first row of the circulant matrix, be defined as

$$\underline{r}=[\ 1 \quad p \quad p^2 \ \ldots \ p^{N-2}\ p^{N-1}\ p^{N-2} \ldots \ p^2 \quad p] \qquad (6\text{-}30)$$

The Fourier transform of $\underline{r}$, $\underline{R}$, is a vector of size K which is obtained from

$$\underline{R}=\underline{A}\underline{r} \qquad (6\text{-}31)$$

where A represents the discrete Fourier transform matrix. The vector $\underline{R}$ contains the eigenvalues of the matrix in fig. (6-4). Since $\underline{r}$ has a particular structure, it is

$$[\ 1 \quad p \quad p^2 \ \ldots \ p^{N-3}\ p^{N-2}\ p^{N-1}\ ]$$

a) First line of an N by N Markovian covariance matrix

$$[\ 1 \quad p \quad p^2 \ \ldots \ p^{N-3}\ p^{N-2}\ p^{N-1}\ p^{N-2}\ p^{N-3} \ \ldots \ p^2 \quad p\ ]$$

b) First row of the resulting Markovian circulant matrix

Figure (6-3) The first line of an NxN Markovian matrix gives rise to the first line of a circulant matrix.

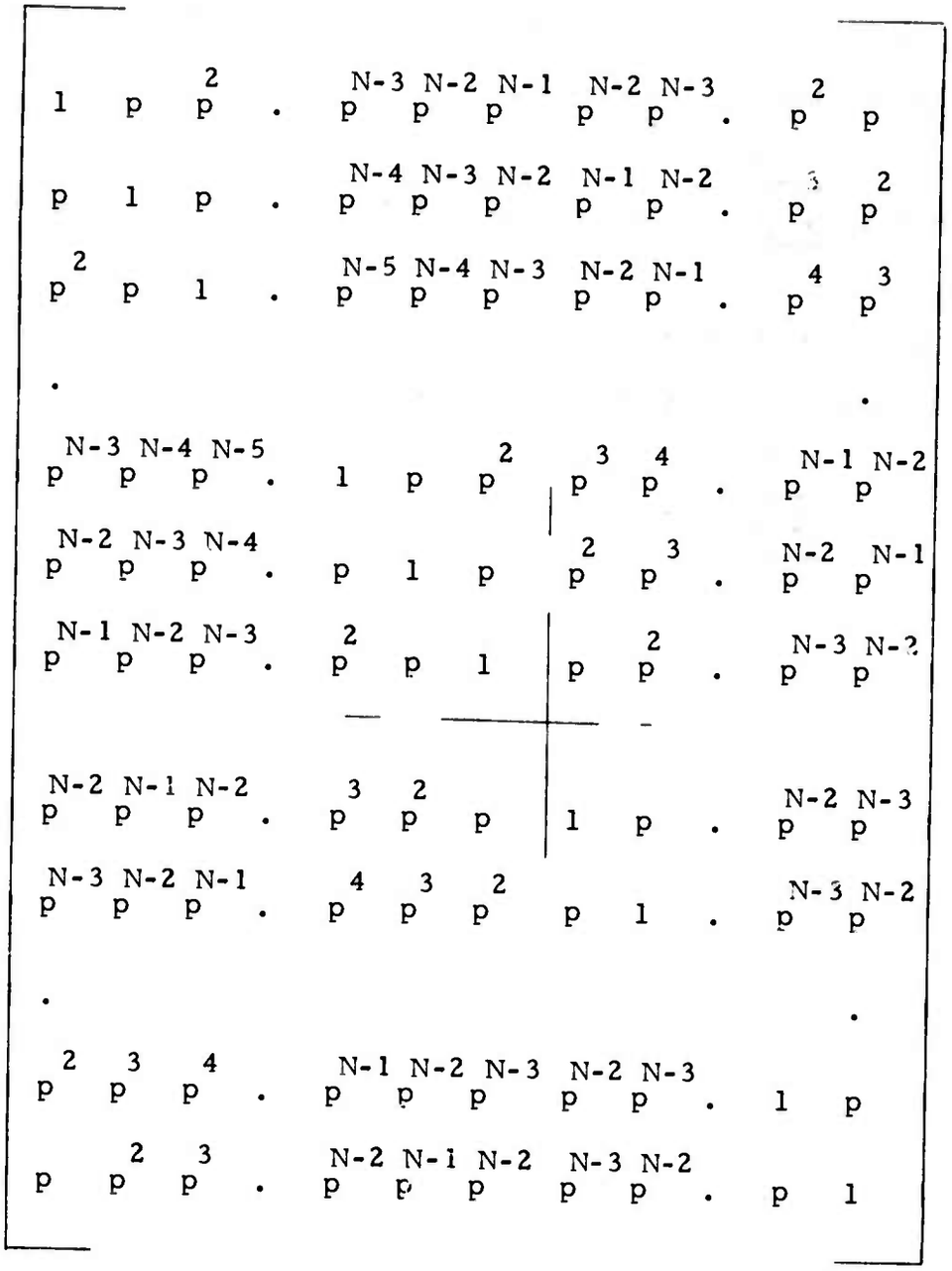


Figure (6-4) Circulant covariance matrix

possible to find a closed form solution for the entries of $\underline{R}$. By definition of the Fourier transform, the (k+1)th entry of $\underline{R}$, R(k+1), is given as

$$R(k+1)=\sum_{j=0}^{2N-3} r(j+1)\exp\{-[\frac{2\Pi i}{(2N-2)}]jk\} \qquad (6-32)$$

where k varies from zero to K=2N-3, and i is the square root of unity. Equation (6-31) can be partitioned as

$$R(k+1)=\sum_{j=0}^{N-1} r(j+1)\exp\{-[\frac{2\Pi i}{(2N-2)}]jk\}+$$
$$r(j+1)\exp\{-[\frac{2\Pi i}{(2N-2)}]jk\} \qquad (6-33)$$

The exact value of r(j+1) can be obtained from eq. (6-30). If eq. (6-30) is used in eq. (6-33), then along with further simplifications the following equation can be obtained

$$R(k+1)=1+(-1)^k p^{N-1}$$
$$+\sum_{j=1}^{N-2} p^j \{\exp[-\frac{2\Pi ijk}{2N-2}]+\exp[\frac{2\Pi ijk}{2N-2}]\} \qquad (6-34)$$

or

$$R(k+1)=1+(-1)^k p^{N-1}+2\sum_{j=1}^{N-2} p^j \cos[-\frac{2\Pi jk}{2N-2}] \qquad (6-35)$$

Observing the following identity [6-9]

$$\sum_{j=0}^{N-1} p^j \cos(jt) = \frac{[1-p\cos(t)][1-p^N\cos(Nt)]+p^{N+1}\sin(t)\sin(Nt)}{1-2p\cos(t)+p^2} \quad (6\text{-}36)$$

or

$$= \frac{1-p\cos(t)+p^{N+1}\cos(N-1)t-p\cos(Nt)}{1-2p\cos(t)+p^2} \quad (6\text{-}37)$$

eq. (6-35) further simplifies to

$$R(k+1) = -1-(-1)^k p^{N-1} + 2\frac{1-p\cos(t)+p^{N+1}\cos(N-1)t-p^N\cos(Nt)}{1-2p\cos(t)+p^2} \quad (6\text{-}38)$$

where $t=\frac{\pi}{N-1}k$. Equation (6-38) is the desired result which provides a closed form solution for the eigenvalues of the matrix in fig. (6-4). In order to use the matrix in fig. (6-4) in the fast Weiner filter equation, fig. (6-4) must represent a covariance matrix for real data. Thus, the matrix of fig. (6-4) is required to have nonnegative eigenvalues, which is the subject of the following lemma.

Lemma (6-1). Figure (6-4) illustrates a nonnegative definite matrix which becomes singular only when p=1. Hence, this matrix is positive definite for all the values of p such that $0 \leq p < 1$.

Proof. It must be illustrated here that R(k+1) of

eq. (6-38) is nonnegative for all the increments of k. Since $t=\frac{\Pi}{N-1}k$, then

$$(N-1)t=\Pi k \tag{6-39}$$

and

$$\mathrm{Cos}[(N-1)t]=(-1)^k \tag{6-40}$$

Also

$$Nt=\Pi k+\frac{\Pi}{N-1}k \tag{6-41}$$

thus

$$\mathrm{Cos}(Nt)=(-1)^k\,\mathrm{Cos}(t) \tag{6-42}$$

Equations (6-40) and (6-42) can be substituted in eq. (6-38) to give

$$R(k+1)=-1+(-1)^k p^{N-1}+2\frac{1-p\mathrm{Cos}(t)+(-1)^k p^{N+1}-(-1)^k p^N \mathrm{Cos}(t)}{1-2p+\mathrm{Cos}(t)+p^2} \tag{6-43}$$

or

$$R(k+1)=\frac{(1-p^2)[1-(-1)^k p^{N-1}]}{1-2p\mathrm{Cos}(t)+p^2} \tag{6-44}$$

Since $p<1$, the numerator of the above equation is always positive, and also since $\text{Cos}(t)\leq 1$ the denominator is also always positive. thus $R(k+1)$ is larger than zero for all values of k and $0\leq p<1$. If $p=1$, $R(k+1)$ becomes zero with the exception of the case when t is the zero angle ($k=0$). If $t=0$, then $\text{Cos}(t)=1$, and

$$R(1)=\frac{(1-p)(1+p)(1-p^{N-1})}{(1-p)^2} \tag{6-45}$$

or

$$R(1)=\frac{(1+p)(1-p^{N-1})}{1-p} \tag{6-46}$$

Now if p approaches unity, then the following nonzero valu for $R(1)$ is obtained

$$R(1)=1+p^{N-1}+2\sum_{j=1}^{N-2}p=2N-2 \tag{6-47}$$

Since the assumption of $p<1$ is always valid, lemma (6-1) has illustrated that matrix of fig. (6-4) is a nonsingular covariance matrix.

The developments of the fast Wiener estimator are summarized in the following steps

1. The circular image degradation model is defined by equation (6-16).

2. An extended object vector is defined by augmenting a vector $\underline{y}$ of certain statistical characteristics to the actual object vector $\underline{f}$. Equation (6-18) illustrates this vector.

3. A circulant nonsingular covariance matrix of form fig. (6-4) is constructed to be used in the Fast Weiner filter of eq. (6-17)

4. The augmented vector $\underline{y}$ is generated - this step is yet to be established.

5. The observation $\underline{g}_c$ is modified to correspond to the extended object $\underline{f}_c$ -this step is yet to be established.

Continuing from step 4 of above, the mean zero random process vector $\underline{y}$ has the following statistics

$$E(\underline{y}\underline{y}^T)=\underline{C}_y \tag{6-48}$$

and

$$E(\underline{f}\underline{y}^T)=\underline{C}_{fy} \tag{6-49}$$

where $\underline{C}_y$ is the lower right (N-2)x(N-2) portion of matrix fig. (6-4) and $\underline{C}_{fy}$ is the upper right Nx(N-2) portion of

the same matrix. Thus, $\underline{C}_y$ is a Markovian covariance matrix of size (N-2)x(N-2), and $\underline{C}_{fy}$ is given as

$$\underline{C} = \begin{bmatrix} p^{N-2} & p^{N-3} & \cdots\cdots & p^2 & p \\ p^{N-1} & p^{N-2} & \cdots\cdots & p^3 & p^2 \\ p^{N-2} & p^{N-1} & \cdots\cdots & p^4 & p^3 \\ \vdots & \vdots & & & \\ p^3 & p^4 & \cdots\cdots & p^{N-1} & p^{N-2} \\ p^2 & p^3 & \cdots\cdots & p^{N-2} & p^{N-1} \\ p & p^2 & \cdots\cdots & p^{N-3} & p^{N-1} \end{bmatrix} \qquad (6\text{-}50)$$

The statistical information on $\underline{y}$ can be employed to generate this vector. To proceed, let the entries of $\underline{y}$ be equal to a linear combination of the entries of the observed physical image samples, $\underline{g}$, plus a noise term. Assuming $\underline{Q}$ represents the linear operation, vector $\underline{y}$ can be expressed as

$$\underline{y}=\underline{Q}\underline{g}+\underline{u} \qquad (6\text{-}51)$$

where $\underline{u}$ is an independent noise term. The observation vector $\underline{g}$ is given by eq. (3-41). Thus,

$$\underline{y}=\underline{Q}\underline{B}\underline{f}+\underline{Q}\underline{n}+\underline{u} \qquad (6\text{-}52)$$

The desired covariance of $\underline{y}$, $\underline{C}_y$, is known. Hence,

$$E[(\underline{QBf}+\underline{Qn}+\underline{u})(\underline{QBf}+\underline{Qn}+\underline{u})^T]=\underline{C}_y \tag{6-53}$$

or

$$\underline{QBC}_f\underline{B}^T\underline{Q}^T+\underline{QVQ}^T+\underline{C}_u=\underline{C}_y \tag{6-54}$$

where $\underline{V}$ represents the covariance of $\underline{n}$, and $\underline{C}_u$ represents the (unknown) covariance of $\underline{u}$. Cross correlation of $\underline{f}$ with $\underline{y}$ gives rise to

$$E[\underline{f}(\underline{QBf}+\underline{Qn}+\underline{u})^T]=\underline{C}_{fy} \tag{6-55}$$

or

$$\underline{C}_f\underline{B}^T\underline{Q}^T=\underline{C}_{fy} \tag{6-56}$$

Equation (6-56) can be solved for matrix $\underline{Q}$ to give the following result

$$\underline{Q}=\underline{C}_{fy}^T\underline{C}_f^{-1}\underline{B}^T(\underline{BB}^T)^{-1} \tag{6-57}$$

And the above equation can be used in eq. (6-54) to solve for $\underline{C}_u$

$$\underline{C}_u = \underline{C}_y - \underline{QBC}_f \underline{B}^T \underline{Q}^T - \underline{QVQ}^T \qquad (6\text{-}58)$$

Equation (6-58) is used to generate the mean-zero independent noise term $\underline{u}$. The numerical value of $\underline{u}$ must be summed with vector $\underline{Qg}$ to generate vector $\underline{y}$ in order to complete step 4. In the last step, the extended observation vector $\underline{g}$ must be established. To proceed at this point, two hypothetical K-dimensional observation vectors $\underline{g}_1$ and $\underline{g}_2$ are defined as follows: let $\underline{g}_1$ be formed by augmenting $\underline{g}$ with a vector of zeros. The resulting vector is

$$\underline{g}_1 = \begin{bmatrix} \underline{g} \\ \\ \underline{0} \end{bmatrix} \qquad (6\text{-}59)$$

And $\underline{g}_2$ is formed by artificially degrading (blurring) a vector of zeros augmented by $\underline{y}$. Vector $\underline{g}_2$ has the following form

$$\underline{g}_2 = \underline{C} \begin{bmatrix} \underline{0} \\ \\ \underline{y} \end{bmatrix} \qquad (6\text{-}60)$$

The vector $\underline{g}_c$ is found as the sum of $\underline{g}_1$ and $\underline{g}_2$

$$\underline{g}_c = \underline{g}_1 + \underline{g}_2 \tag{6-61a}$$

Hence,

$$\underline{g}_c = \underline{C}\begin{bmatrix} \underline{f} \\ \\ \underline{\emptyset} \end{bmatrix} + \underline{C}\begin{bmatrix} \underline{\emptyset} \\ \\ \underline{Y} \end{bmatrix} + \underline{n} \tag{6-61b}$$

or

$$\underline{g}_c = \underline{C}\underline{f}_c + n \tag{6-61c}$$

where $\underline{n}$ is a noise term which incorporates the original observed noise plus the uncertainty introduced by a possible windowing operation applied on the observation. Equation (6-61c) represents a circular degradation phenomenon, whose object vector $\underline{f}_c$ has a circulant nonsingular covariance matrix. The minimum mean-square error estimate of the object, $\underline{f}_c$, is given as

$$\hat{\underline{f}}_c = (S\,\underline{C}_{f_c}^{-1} + \underline{C}^T\underline{C})^{-1}\,\underline{C}^T\,\underline{g}_c \tag{6-62}$$

where S is the signal to noise ratio. Extraction of the first N entries of $\hat{\underline{f}}_c$ results in the true object estimate $\hat{\underline{f}}$. Thus

$$\hat{\underline{f}} = \underline{\underline{Sl}}_K^N \underline{f} \qquad (6\text{-}63)$$

where $\underline{\underline{Sl}}_K^N$ is the proper selection matrix.

6.4 A Comment on the Optimality of the Fast Filter

The previous section developed a fast image restoration technique which is optimal in a minimum mean-square error sense. It should be noted, however, that the optimality of this filter depends greatly on the validity of the statistical assumptions imposed upon the hypothetical object vector $\underline{f}_c$. Although the proper selection of the physical samples of the auxiliary random process vector $\underline{y}$ is essential for recovery of the object with the least mean-square error, the error analysis of the next section illustrates that the increase in the error can be extremely small.

Because of the circularity of the covariance matrix $\underline{\underline{C}}_{f_c}$, the few end pixels of the true object $\underline{f}$ happen to be highly correlated with certain entries of $\underline{y}$. To illustrate this claim, fig. (6-5a) shows the pixels of $\underline{f}_c$ arranged around a circle. Under a Markovian assumption, it is noted that

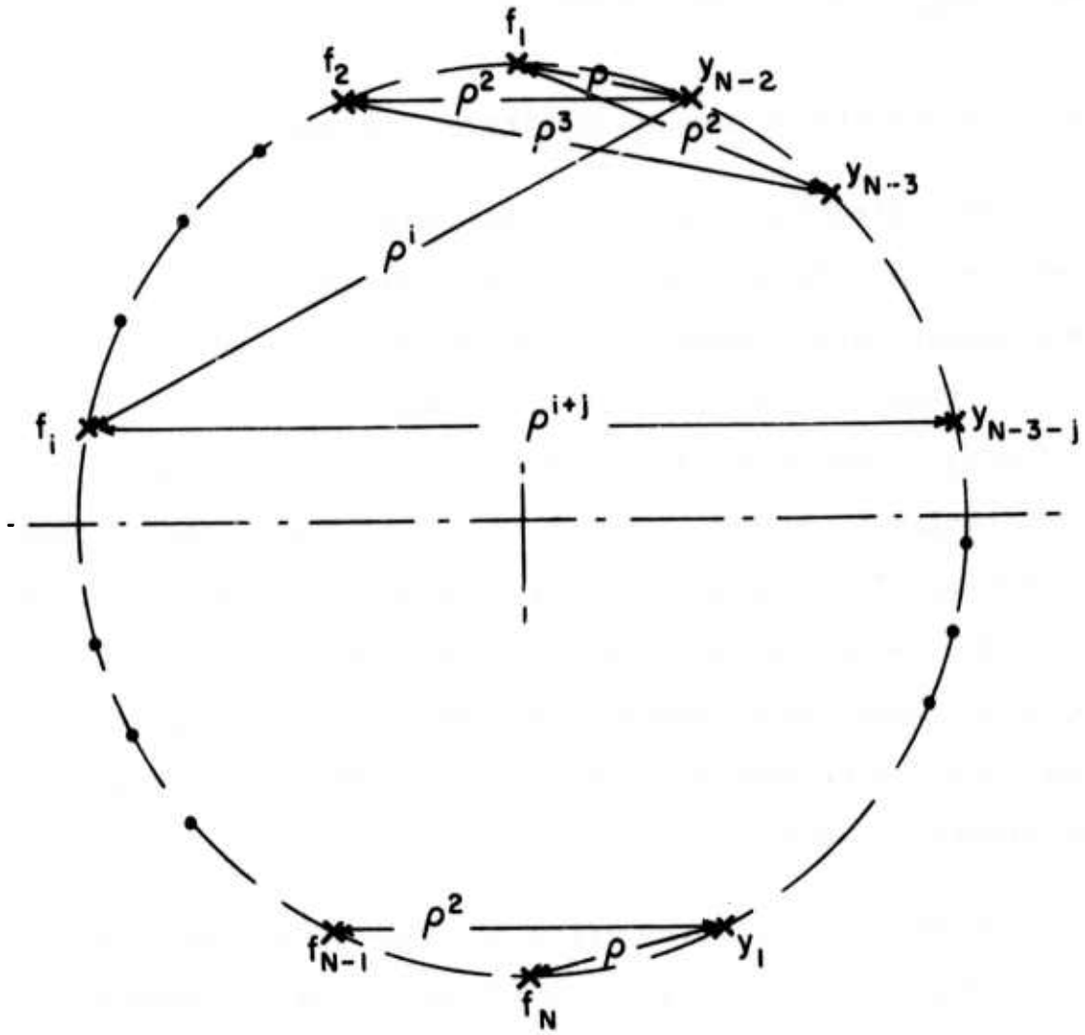


Figure (6-5) Circular statistical property of the vector $\underline{y}_c$.

$$E[f_1 y_{N-2}] = p^i \qquad i \leq \frac{N}{2}$$

$$E[f_{N+1-i} y_1] = p^i \qquad i \geq \frac{N}{2} \tag{6-64}$$

Equation (6-64) illustrates that the central pixels of $\underline{f}$, where the subscript i is not close to either one or N, are just slightly correlated with any pixel in $\underline{y}$. On the other hand, the pixels of $\underline{f}$ lying close to the two ends, where i is either close to one or N, are strongly correlated with y_1 or y_{N-2}.

In view of the above argument, a poor estimation of the physical samples of $\underline{y}$ would mostly affect the few pixels of the object estimate at its two ends. Therefore, it can be expected that a reasonably small amount of error in the vector $\underline{y}$ would not influence the center portion of the object estimate $\hat{\underline{f}}$. Thus it appears that computation of the physical samples of the random process $\underline{y}$ can be avoided if this vector is replaced by its mean. In this case, since $\underline{y}$ is mean zero, the hypothetical object vector $\underline{f}_c$ has the following form

$$\underline{f}_c = \begin{bmatrix} \underline{f} \\ \underline{0} \end{bmatrix} \tag{6-65a}$$

and the observation is given by

$$\underline{g}_c = \begin{bmatrix} \underline{g} \\ \underline{\emptyset} \end{bmatrix} \tag{6-65b}$$

Representing the fast filter by $\underline{U}$, the object estimate $\hat{\underline{f}}_c$ is obtained as

$$\hat{\underline{f}}_c = \underline{\underline{U}}\underline{g}_c \tag{6-66}$$

The first N entries of $\hat{\underline{f}}_c$ result in the true object estimate $\hat{\underline{f}}$. Thus,

$$\hat{\underline{f}} = \underline{\underline{Sl}}_K^N \hat{\underline{f}}_c = \underline{\underline{Sl}}_K^N \underline{\underline{U}}\underline{g}_c \tag{6-67}$$

It should be noted that the above equation represents a suboptimal estimate of the object $\underline{f}$, but the estimation technique associated with eq. (6-67) illustrates extreme amount of computational efficiency. The efficiency of this technique is due to the fact that only Fourier transform operation is employed for obtaining the estimate $\underline{f}$ of eq. (6-67). The next section illustrates that replacing vector $\underline{y}$ by its mean results only in an slight increase of the error in the estimation of the few pixels lying at the two ends of the object vector.

6. 5 Error Analysis

In order to analyze the error for both overdetermined

and underdetermined systems using a uniform terminology, the following approach has been adopted. Assume $\underline{x}$ is an object of size Q and that the N middle pixels of this object, $\underline{f}$, are to be estimated. Vector $\underline{x}$ is blurred to generate the observation $\underline{g}$. In fact, $\underline{g}$ consists of the blurred object plus an additive noise term $\underline{n}$. Next, the observation $\underline{g}$ undergoes the windowing operation $\underline{W}$. This modified observation is referred to as $\underline{\bar{g}}$. Figure (6-5b) illustrates the vectors and the corresponding operations. Both $\underline{g}$ and $\underline{\bar{g}}$ are of size M, where

$$M=Q-L+1 \tag{6-68a}$$

and

$$N=M-L+1 \tag{6-68b}$$

The observation $\underline{g}$ is obtained from the following expression (see Sec. 3. 5)

$$\underline{g}=\underline{\underline{B}}\underline{x}+\underline{n} \tag{6-69}$$

where $\underline{\underline{B}}$ is an M by Q blur matrix (underdetermined), and $\underline{n}$ is the white noise term. The output of the windowing operation is given by

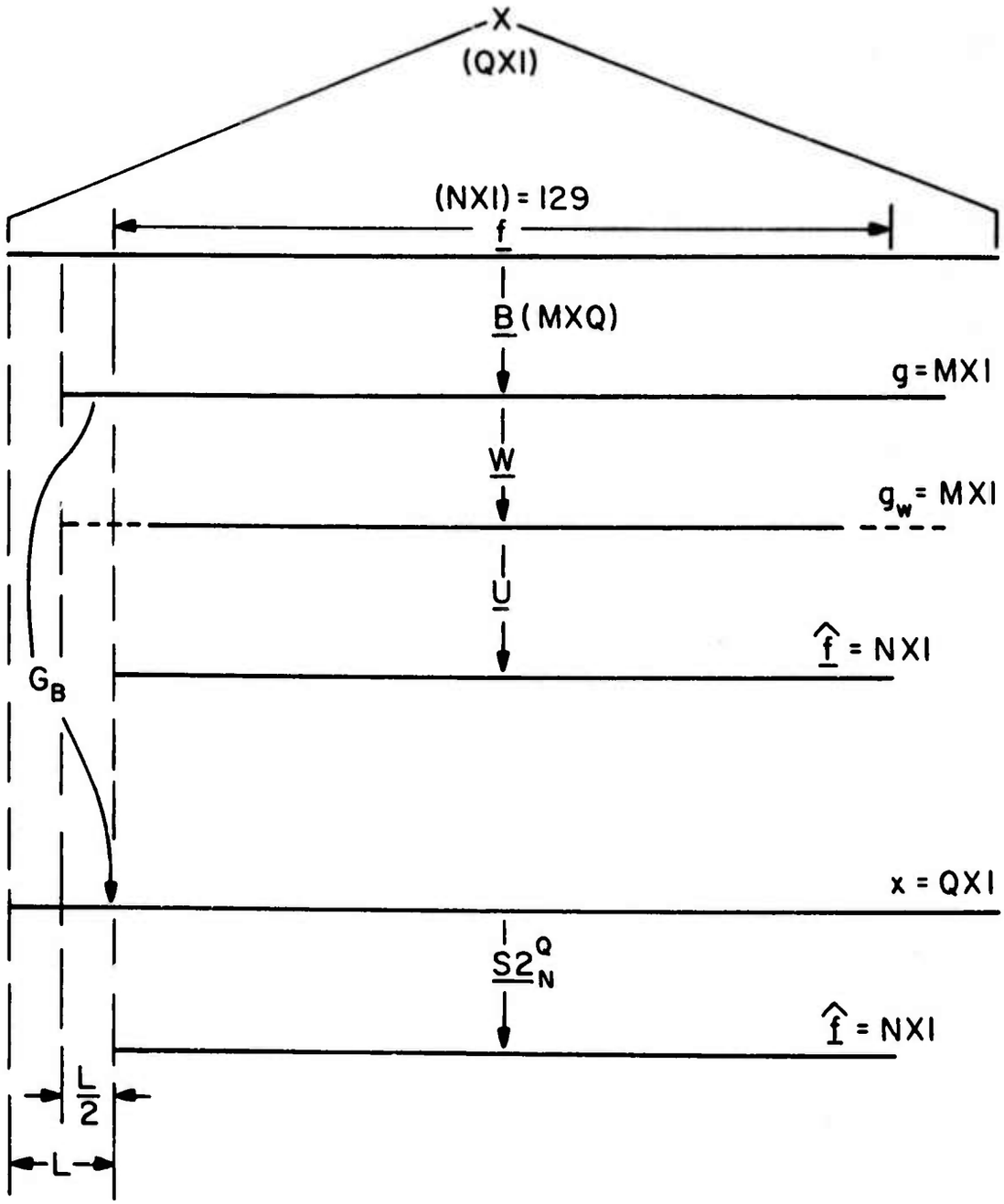


Figure (6-5b) Vectors and matrices used for error analysis

$$\underline{g}=\underline{W}\underline{B}\underline{x}+\underline{n} \qquad (6\text{-}70)$$

To estimate the object $\underline{f}$, by means of the fast Wiener technique and by use of the observation $\underline{g}$, the random process vector $\underline{y}$ must be obtained. For the sake of computational simplicity, this vector can be replaced by its mean value which is zero. It must be realized that this approximation and application of the windowing operator are the only approximations made in computing the classical Wiener filter. In other words, if the vector $\underline{y}$ is generated by executing the process explained in Sec. (6. 3), the error term for the fast filter must be identical to the error function derived for the classical Wiener filter (assuming a black background for the scene, or absence of blur). As will be illustrated later, the approximation of $\underline{y}$ by its mean results in a small increase in the variance of the error. Nevertheless, considering the computational savings the assumption yields, this approximation is worthwhile. Thus the estimate $\hat{\underline{f}}$ is obtained by

$$\begin{bmatrix} \underline{f} \\ \underline{0} \end{bmatrix} = \underline{U} \begin{bmatrix} \underline{g} \\ \underline{0} \end{bmatrix} \qquad (6\text{-}71a)$$

or

$$\hat{\underline{f}}=\underline{Sl}^{N}_{K}\,\underline{U}(\underline{Sl}^{N}_{K})^{T}\underline{g} \tag{6-71b}$$

where $\underline{U}$ represents the fast filter. For the sake of shortening the length of the preceeding equations, let

$$\underline{T}=\underline{Sl}^{N}_{K}\,\underline{U}(\underline{Sl}^{N}_{K})^{T} \tag{6-72}$$

where $\underline{T}$ is an N by M matrix (filter). Using eq. (6-72) in eq. (6-71b) results in

$$\hat{\underline{f}}=\underline{Tg} \tag{6-73}$$

Next, $\underline{g}$ can be substituted for as follows

$$\hat{\underline{f}}=\underline{TWBx}+\underline{TWn} \tag{6-74}$$

The error covariance matrix is defined as

$$\underline{C}_{e}=E[(\underline{f}-\hat{\underline{f}})(\underline{f}-\hat{\underline{f}})^{T}] \tag{6-75a}$$

or

$$\underline{C}_{e}=E(\underline{ff}^{T})+E(\hat{\underline{f}}\hat{\underline{f}}^{T})-E(\underline{f}\hat{\underline{f}}^{T})-E(\hat{\underline{f}}\underline{f}^{T}) \tag{6-75b}$$

To obtain $\underline{C}_{e}$, the four individual terms in the right hand side of eq. (6-75b) must be obtained. Starting from the

simplest term $E(\underline{f}\underline{f}^T)$, and assuming that the object is a sample from a Markov process, this term can be represented by an N by N Markovian covariance matrix. The term $E(\underline{\hat{f}}\underline{\hat{f}}^T)$ is obtained as

$$E(\underline{\hat{f}}\underline{\hat{f}}^T)=E[(\underline{TWBx}+\underline{TWn})\,\underline{f}^T] \tag{6-76a}$$

or

$$E(\underline{\hat{f}}\underline{\hat{f}}^T)=\underline{TWB}[E(\underline{x}\underline{f}^T)] \tag{6-76b}$$

where $E(\underline{x}\underline{f}^T)$ is the cross covariance of a Q dimensional object with its own N middle pixels. Presenting this term by $\underline{C}_{xf}$

$$E(\underline{\hat{f}}\underline{\hat{f}}^T)=\underline{TWBC}_{xf} \tag{6-77}$$

where $\underline{C}_{xf}$ is the Q by N center portion of a Q by Q Markovian covariance matrix. The term $E(\underline{x}\underline{f}^T)$ is simply obtained by transposing $E(\underline{f}\underline{x}^T)$. Thus,

$$E(\underline{x}\underline{f}^T)=(\underline{TWBC}_{xf})^T \tag{6-78}$$

And the term $E(\underline{\hat{f}}\underline{\hat{f}}^T)$ is obtained as

$$E(\underline{\hat{f}}\underline{\hat{f}}^T)=E(\underline{TWBx}+\underline{TWn})\,(\underline{TWBx}+\underline{TWn})^T \tag{6-79}$$

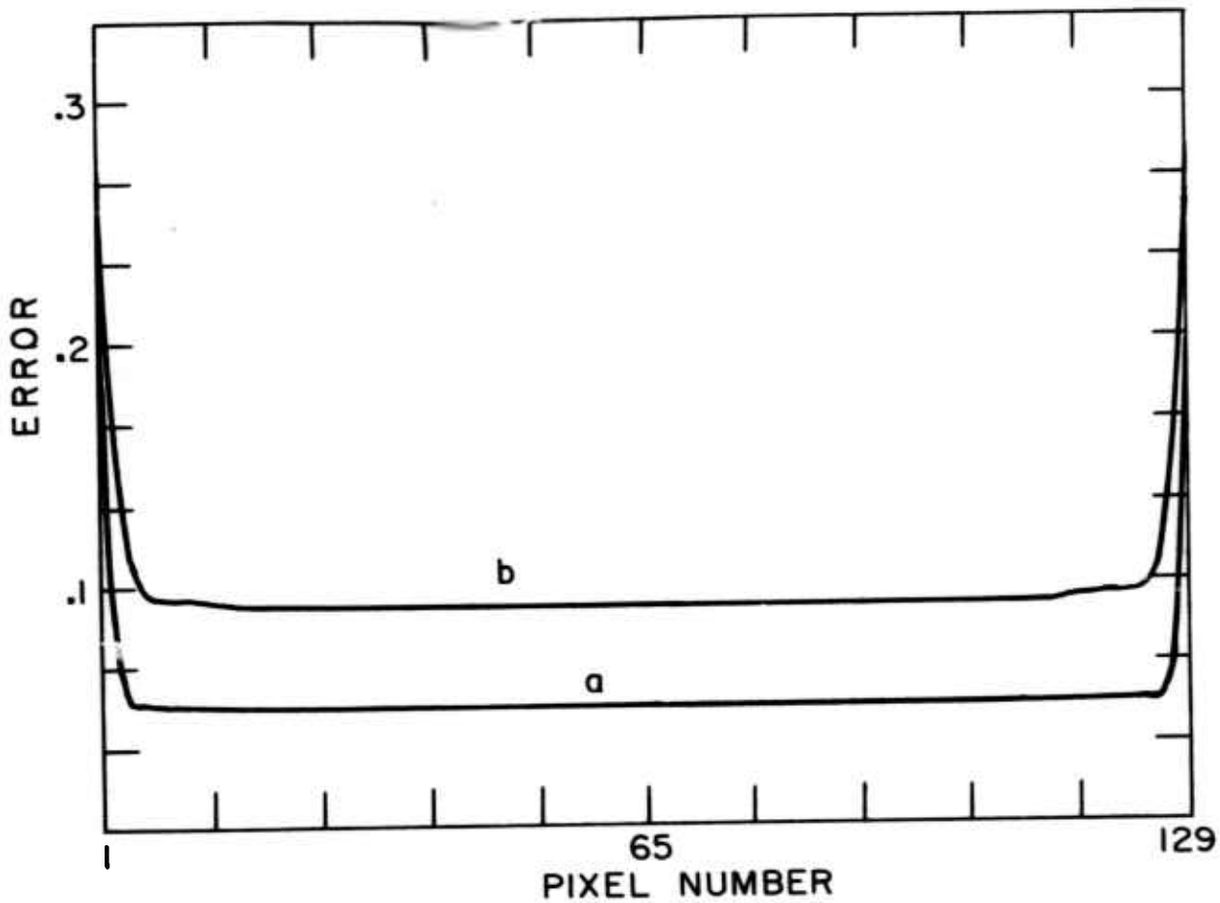
or

$$E(\underline{\hat{f}}\underline{\hat{f}}^{T}) = \underline{TWBC}_{x}\underline{B}^{T}\underline{W}^{T}\underline{T}^{T} + \underline{TWVW}^{T}\underline{T}^{T} \qquad (6\text{-}80)$$

where $\underline{C}_x$ is a Q by Q Markovian covariance matrix and $\underline{V}$ is the noise covariance. Finally, the error covariance is given by the following equation

$$\underline{C}_e = \underline{TWBC}_{x}\underline{B}^{T}\underline{W}^{T}\underline{T}^{T} + \underline{C}_f + \underline{TWVW}^{T}\underline{T}^{T} - [\underline{TWBC}_{xf} + (\underline{TWBC}_{xf})^{T}] \qquad (6\text{-}81)$$

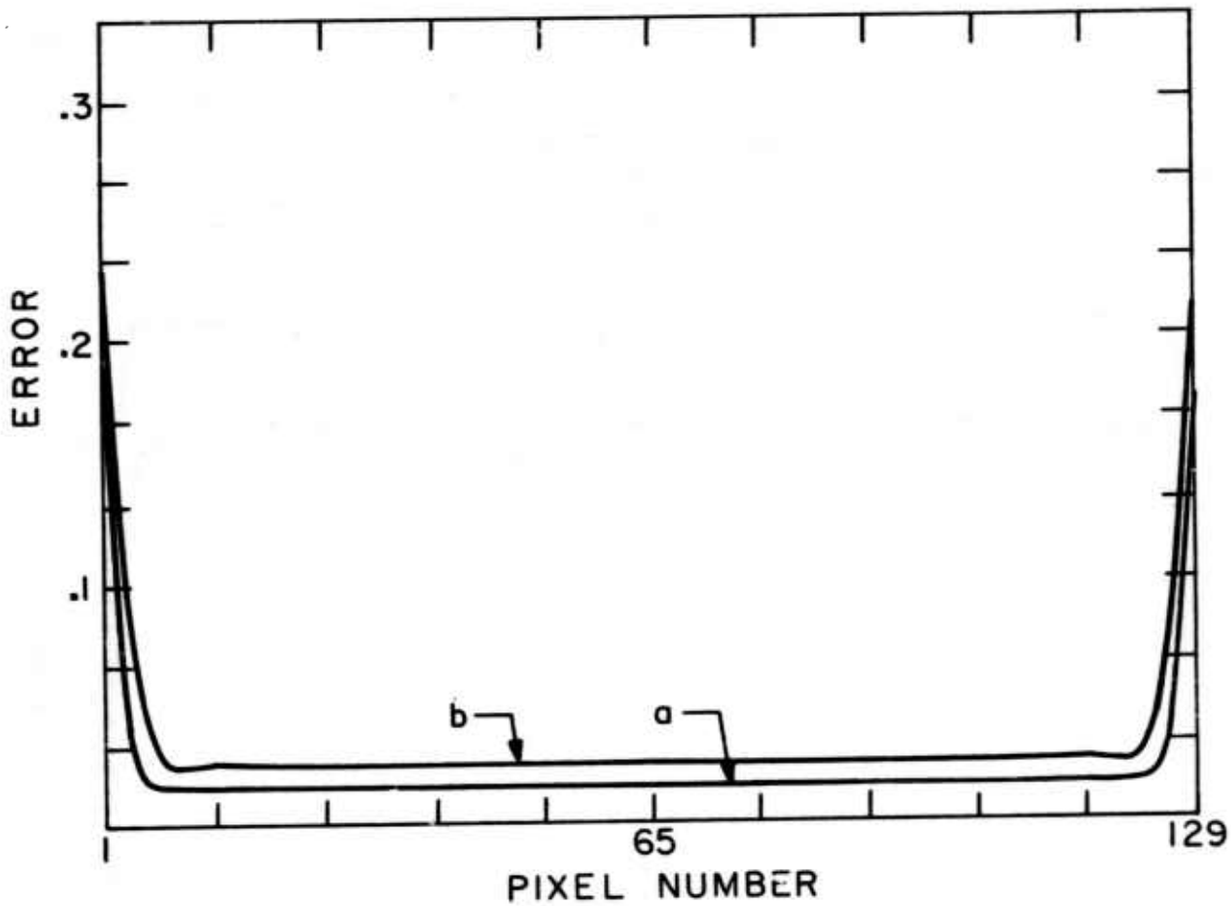
Figure (6-6) illustrates the error variance for an object estimate containing 129 pixels. The error is plotted for two different blurring effects. Plot (a) represents the error for Gaussian-shape blur of standard deviation 2. Plot (b) shows the error for motion blur. In both cases L is 15, the signal-to-noise ratio is 10, and the element correlation coefficient is 0.95. Notice that the error for the first and the last few pixels is considerably higher than the remainder. This can be explained by the approximation made on vector $\underline{y}$. The circularity property of the current model assumes a strong correlation between $\underline{y}$ and the first and the last few pixels of the object. As a consequence of approximating $\underline{y}$ by a vector of zeros, these pixels become correlated with the wrong data —the zeros— which results in a higher error variance. Figure (6-7)



a) Gaussian shape blur

b) Motion blur

Figure (6-6) Mean-square for an object of size 129 pixels and element correlation 0.95



a) Gaussian shape blur b) Motion blur

Figure (6-7) Mean-square error for an object of size 129 pixels and element correlation 0.99

illustrates the error for a higher element correlation of Ø.99. As expected, both error curves decrease considerably. Figure (6-8) contains two plots for a Gaussian-shape blur of the same standard deviation but having different element correlations. Plot (a) is obtained for a correlation coefficient of Ø.95, and (b) is obtained by assuming the coefficient is Ø.99. Figure (6-9) is the counterpart of the previous figure for motion blur.

If an underdetermined system is used to model the degradation process, the object estimate $\hat{\underline{x}}$ can be obtained using the classical Wiener filter. The estimate $\hat{\underline{x}}$ is given by

$$\hat{\underline{x}}=\underline{U}\underline{g} \tag{6-82}$$

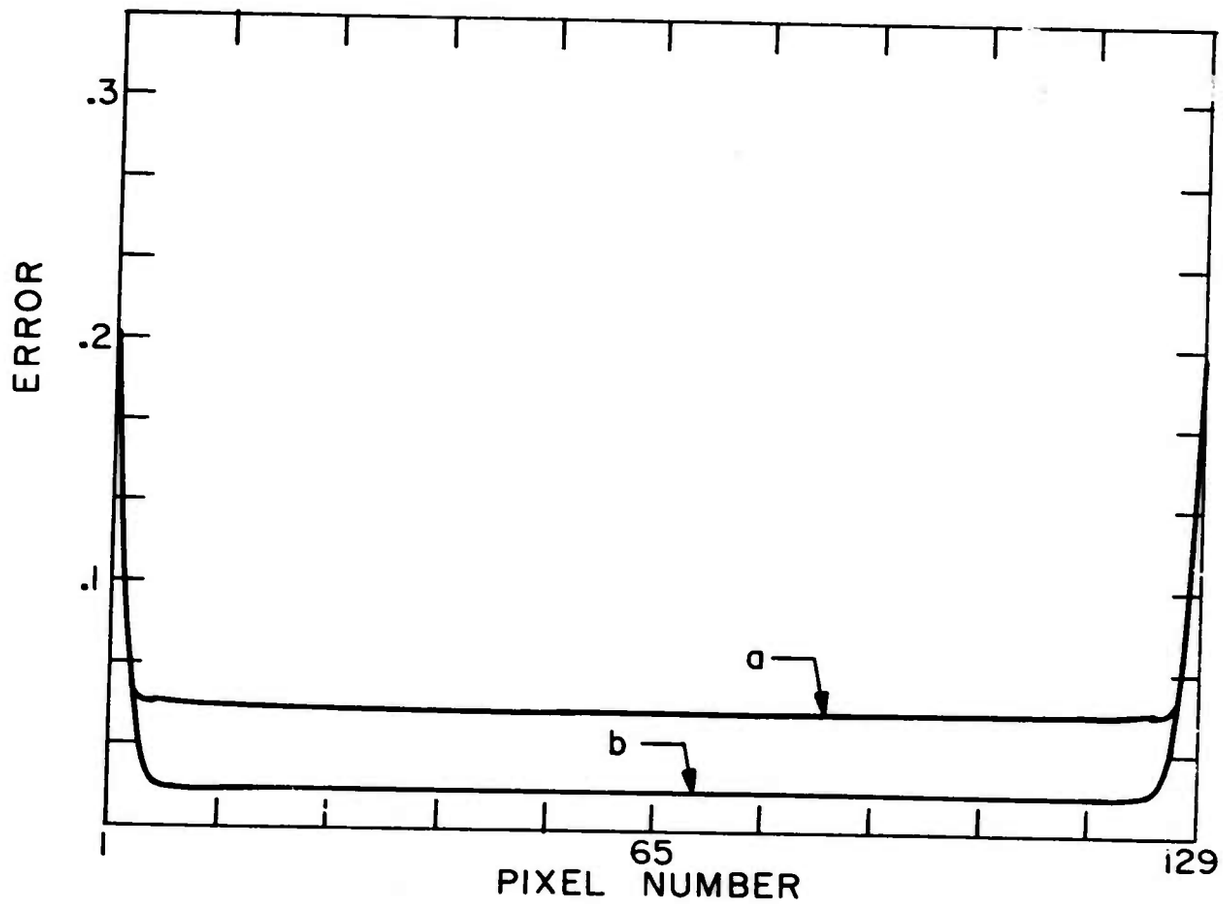
where $\underline{U}$ represents the Wiener filter. The error term is defined as

$$\underline{e}=\underline{x}-\hat{\underline{x}} \tag{6-83}$$

and the error vector for the N middle pixels of $\underline{x}$ is given by

$$\underline{f}-\hat{\underline{f}}=\underline{S2}_{Q}^{N}\underline{e} \tag{6-84}$$

where $\underline{S2}_{Q}^{N}$ is the appropriate selection matrix. The



a) Correlation Coeffi cient=0.95 b) Correlation coefficient=0.99

Figure (6-8) Mean-square error for Gaussian shape blur.

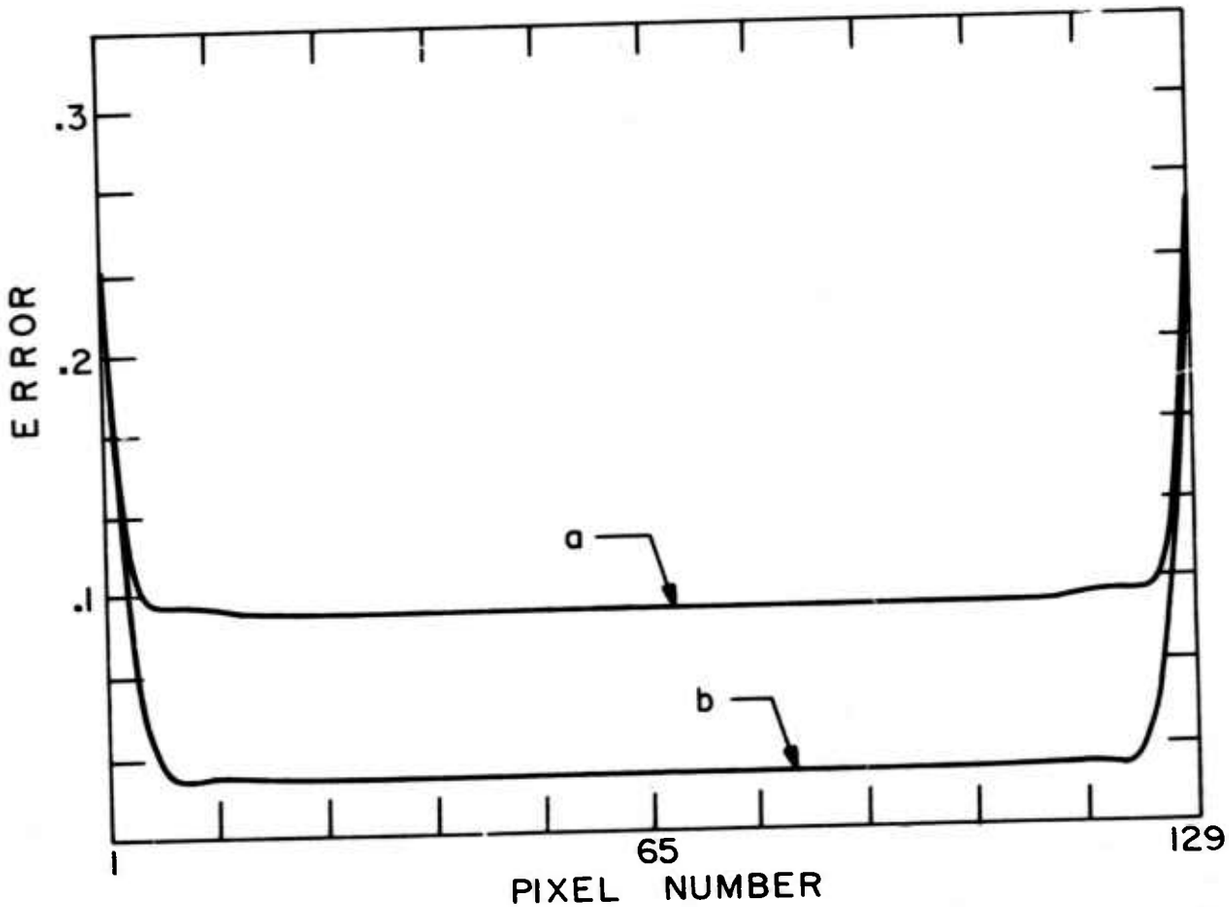


Figure (6-9) Error for motion blur.

a) Correlation coefficient=0.95

b) Correlation coefficient=0.99

covariance matrix of $\underline{e}$ is obtained as follows. Let $\underline{C}_e$ represent this covariance matrix, so that

$$\underline{C}_e = E[(\underline{x}-\hat{\underline{x}})(\underline{x}-\hat{\underline{x}})^T] \tag{6-85}$$

or

$$\underline{C}_e = E(\underline{xx}^T) + E(\hat{\underline{xx}}^T) - E(\underline{x}\hat{\underline{x}}^T) - E(\hat{\underline{x}}\underline{x}^T) \tag{6-86}$$

where $E(\underline{xx}^T)$ is a Q by Q Markovian matrix, and $E(\hat{\underline{xx}}^T)$ is obtained from

$$E(\hat{\underline{xx}}^T) = \underline{U}(E(\underline{gg}^T))\underline{U}^T \tag{6-87}$$

Recall that

$$E(\underline{gg}^T) = E(\underline{Bx}+\underline{n})(\underline{Bx}+\underline{n})^T \tag{6-88}$$

or

$$E(\underline{gg}^T) = \underline{BC}_x\underline{B}^T + \underline{V} \tag{6-89}$$

Equation (6-89) can then be used in equation (6-87) to yield the following equation

$$E(\hat{\underline{xx}}^T) = \underline{UBC}_x\underline{B}^T\underline{U}^T + \underline{UVU}^T \tag{6-90}$$

In a similar manner, $E(\underline{x}\hat{\underline{x}}^T)$ can be derived as

$$E(\underline{x}\hat{\underline{x}}^T)=\underline{C}_x\underline{B}^T\underline{U}^T \tag{6-91}$$

and

$$E(\hat{\underline{x}}\underline{x}^T)=\underline{UBC}_x \tag{6-92}$$

Equations (6-90), (6-91), and (6-92) can be used in eq. (6-86) to give the final expression for the error variance

$$\underline{C}_e=\underline{C}_x+\underline{UBC}_x\underline{B}\ \underline{U}\ +\underline{UVU}\ -\underline{C}_x\underline{B}\ \underline{U}\ -\underline{UBC}_x \tag{6-93}$$

The error covariance matrix for the N middle pixels of the object can be obtained by observing eq. (6-84)

$$E(\underline{f}-\hat{\underline{f}})(\underline{f}-\hat{\underline{f}})^T=\underline{S2}_Q^N\underline{C}_e(\underline{S2}_Q^N)^T \tag{6-94}$$

Equation (6-94) is the error equation for the N-dimensional object $\underline{f}$. Notice that, unlike the previous case (overdetermined system), the whole object, $\underline{x}$, is estimated first and then, using a selection matrix, $\underline{S2}_Q^N$, the N-dimensional object estimate, $\hat{\underline{f}}$, is obtained. Figure (6-10) illustrates the error for the case when N is 17, L

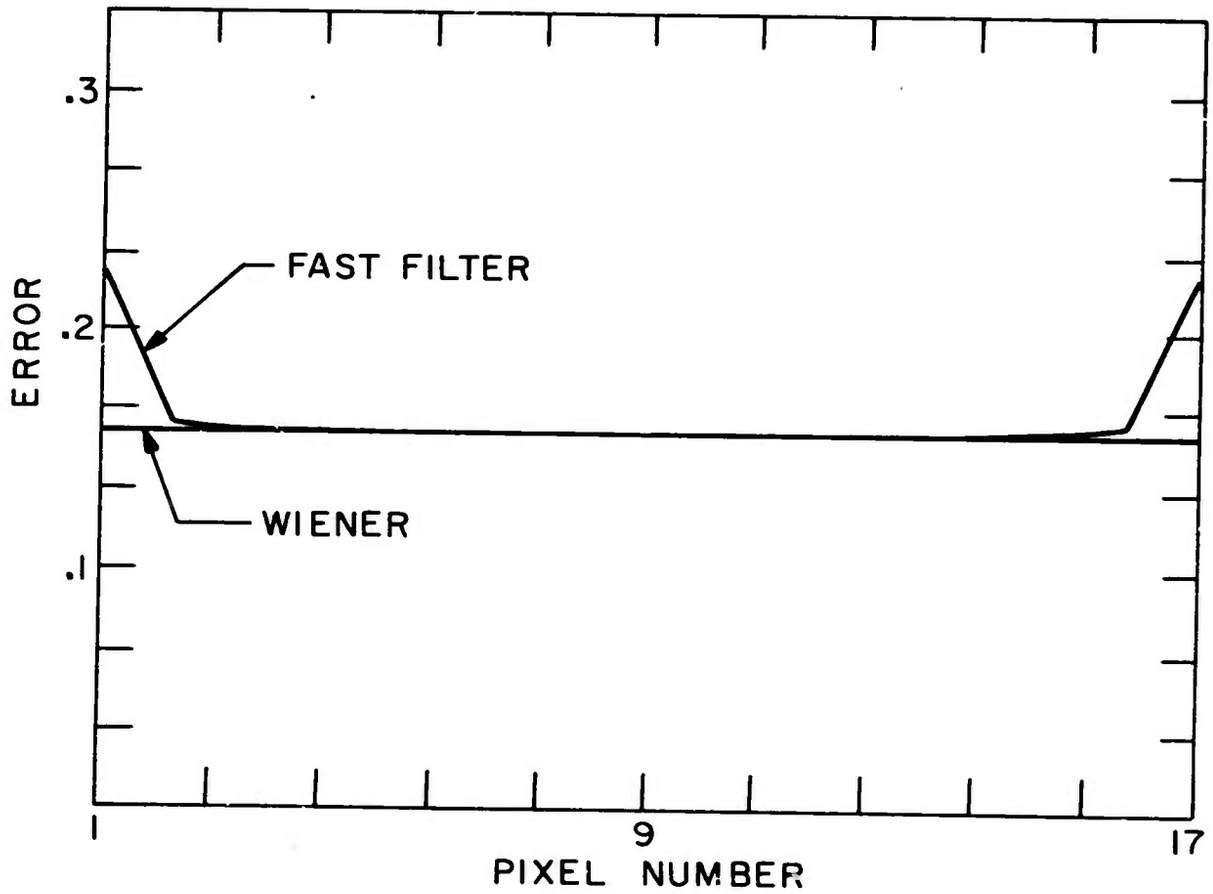


Figure (6-10) Mean-square error comparison of Wiener filter and fast computational Wiener filter.

equals 9, and the number of observed pixels, $\underline{g}$, is 25. Plot (a) shows the error for the classical Wiener filter, and (b) is the error plot for the fast filter. In this figure, except for the first and the last few pixels, the error is the same for both the Wiener and the fast Wiener filters. As was pointed out earlier, this phenomenon is caused by the approximation made on the vector $\underline{y}$. In both plots the blur is of Gaussian shape with a standard deviation of 1, and the correlation coefficient is assumed to be 0.7. The signal-to-noise ratio is 5. Figure (6-11) illustrates the error when there is no blur and the image degradation is due only to additive white noise. The signal-to-noise ratio is assumed to be 10, and the element correlation to be zero. In this example, both error functions are equal. Since no correlation between the elements is assumed, the approximation made on vector $\underline{y}$ does not affect the error plots. Also, in the absence of any blur, the windowing operation does not introduce any uncertainty. Figure (6-12) illustrates the error curves for a Gaussian-shape blur having standard deviation of 2, and an element correlation of 0.95. The signal-to-noise ratio is assumed to be 100.

6. 6 Experimental Results

To illustrate the function of the fast filter, Fig. (6-13a) is selected as a test scene. This scene is

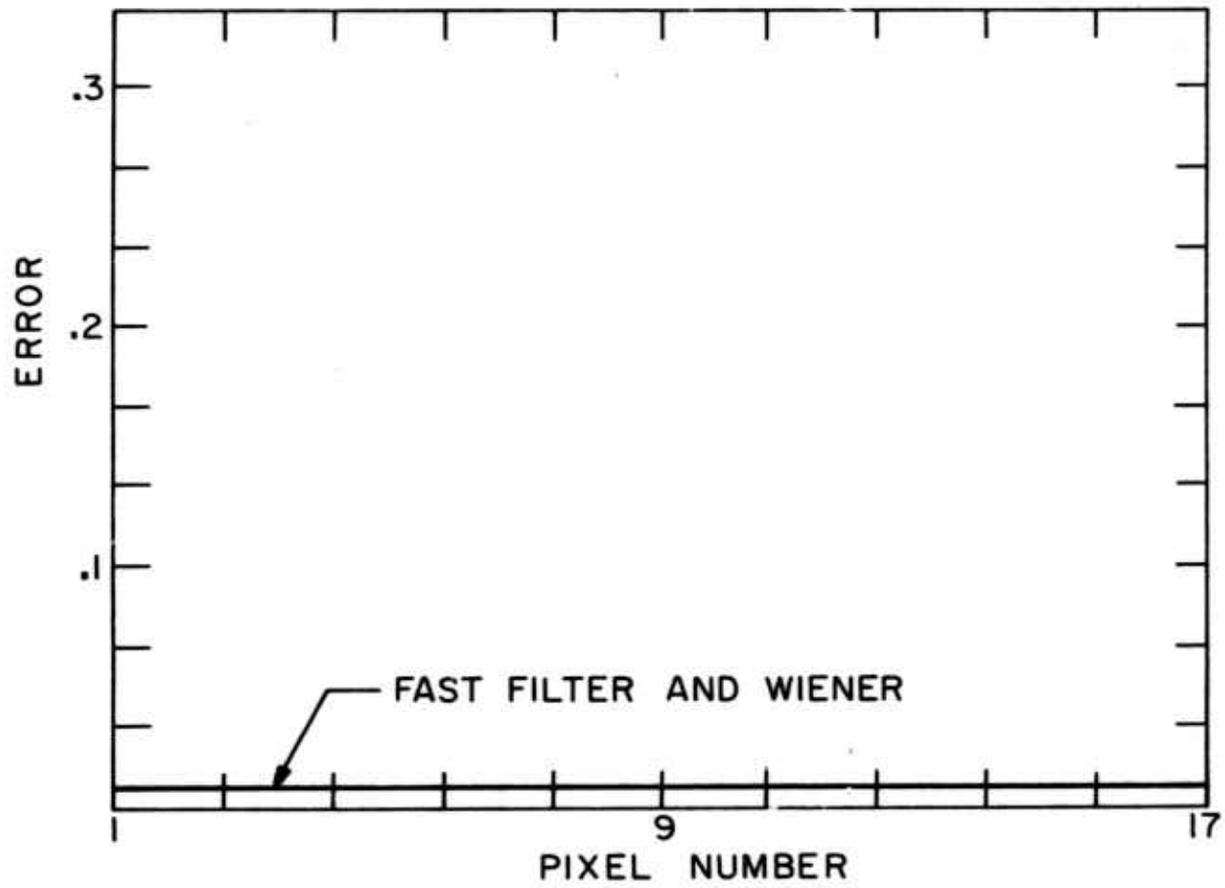


Figure (6-11) Error comparison of Wiener filter and fast computational Wiener filter.

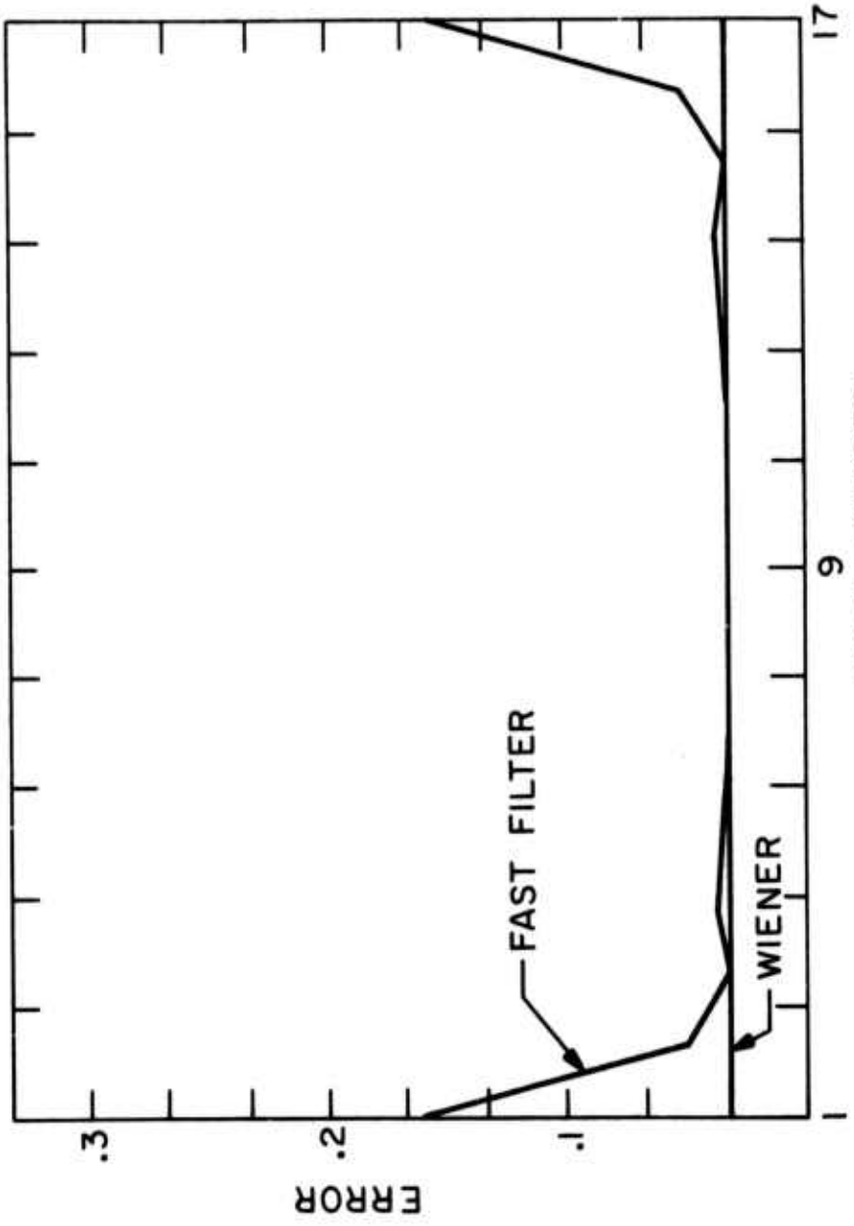


Figure (6-12) Error comparison of Wiener filter and computationaly fast Wiener filter.

(a) Original

(b) Observed

(c) Restored

Figure (6-13) Restoration of an image degraded by motion blur.

represented by an array of 256 by 256 pixels, with each pixel value uniformly quantized to 256 levels. This image was displayed on a flying spot scanner cathode ray tube display and photographed with polaroid type 52 film. Figure (6-13b) represents the image after it undergoesing a motion-blur degradation having an impulse response length of 15 pixels (L=15). To observe the work of the filter on an object with unknown background, only the middle N (N <256) pixels have been restored. The restored object, in this example, contains 129 (N=129) pixels, and it is placed in the blurred background to make the comparison between the observation and the restored image simple, as fig. (6-13c) shows. In this figure, the background of the restored object is obtained by carefully extracting the appropriate section of the blurry observation. This kind of object and background combination enables the observer to easily view the improvement made on the center region of the scene. Hence, the image in fig. (6-13c) can be interpreted as the blurry observation with an enhancing aperture placed in front of the scene. In this particular example the hypothetical aperture contains 256 by 129 pixels. Since motion blur is a one dimensional degradation, each line of the observed image has been processed separately. At the first step, 143 (143=129+15-1) pixels of a line of the observed image have been extracted. Next, these pixels, after undergoing a

windowing operation, have been augmented by a vector of zeros to form a hypothetical observation vector of size 256 (Note that the size of the hypothetical observed vector $\underline{g}$ is determined by eq. (6-29). After the hypothetical observation is formed, it is Fourier transformed, and then this transformed vector is multiplied by the corresponding filter coefficients in the Fourier domain (a scalar operation). Inverse Fourier transforming the filtered vector generates the hypothetical object vector $\underline{f}$. The first 129 (N=129) pixels of this vector form the true object vector $\underline{f}$.

6. 7 The Problem of Unknown Point Spread Function

The process of restoring images when the point spread function of the degrading phenomenon is not known a priori is usually referred to as blind deconvolution [6-10]. This kind of problem arises when the characteristics of the imaging system are not known to the observer, and thus the impulse response must be directly measured from the observed image. In theory, the point spread function can be simply obtained by a direct measurement of the image that resuls from a point source of light. Such an experimental computation of the point spread function is severely limited in practice because of the lack of real point sources in the original scene. A similar technique known as edge measurement is an alternative choice when the

point spread function is isotropic. Unlike isolated point sources of light which rarely occur in a scene, edges are abundant in most images. To illustrate how an edge measurement can help to determine the point spread function, assume U(x) represents an object function of the following form

$$U(x)=\begin{cases}1 & x>0\\ 0 & x=0\\ -1 & x<0\end{cases} \qquad (6\text{-}95)$$

The function U(x) is known as the unit step function. It is a widely accepted discipline to represent a point object by the derivative of the unit step function [6-11], [6-12]. Thus

$$\Delta(x)=\frac{dU(x)}{dx} \qquad (6\text{-}96)$$

where $\Delta(x)$ is now the mathematical notation for a point object. Assume that the object U(x) has undergone a degradation with impulse response h(x). Note that this function, h(x), is yet to be determined, of course. The observation g(x) is given by the following equation

$$g(x)=\int U(x-t)h(t)dt \qquad (6\text{-}97)$$

The limits of the above definite integral are omitted since they play no rule in this analysis. Differentiating the above equation gives rise to the interesting resultation

$$\frac{dg(x)}{dx} = \int [\frac{d}{dx} U(x-t)] h(t) dt \qquad (6\text{-}98)$$

or

$$\frac{dg(x)}{dx} = \int \Delta(x-t) h(t) dt \qquad (6\text{-}99)$$

notice that the right hand side of the above equation (by definition) equals the impulse response h(x). Hence

$$h(x) = \frac{dg(x)}{dx} \qquad (6\text{-}100)$$

Equation (6-100) implies that differentiating the image pattern associated with an edge is equivalent to determination of the impulse response of a linear shift-invariant optical system. Figure (6-14a) contains an image photographed by a SEM Cambridge stereo scan type S4-10 electron microscope. The image represents a Ferrite (iron) particle taken from the record side of an audio magnetic tape. The magnification ratio is 130,000 to 1. Since this amount of magnification equals the limiting power of the system, the resulting image is blurry. To assure a nondegraded image, it has been experienced that

(a) Observed

(b) Restored

(c) Restored

Figure (6-14) Image restoration when the point spread function is not known a priori.

the magnification ratio on this system must stay below 100,000. Fortunately the image contains many edges which can help to determine the impulse response of the electron microscope. The estimated impulse response in this example was approximated by a separable two-dimensional Gaussian function of standard deviation 2. Figure (6-14b) illustrates the same image after the center section has been restored. The size of the restored region is 129x129, and the observed image, fig. (6-14a), contains 256x256 pixels which have been uniformely quantized to 8 bits. The restoration technique was the fast Wiener filtering of Sec. (6. 3). Figure (6-14c) shows the crystal after its upper left corner was filtered using the fast filter.

REFRENCES

1. W. Pratt, Digital Image Processing, to be published.

2. B. Hunt, "Digital Image Processing," Proc. IEEE, april 1975, pp. 693-708.

3. M. Sondhi, "Image Restoration: The Problem of Spatially Invariant Degradation," Proc. IEEE, July 1972, pp. 842,853.

4. A Papoulis, Probablity, Random Variabls, and Stochastic

Processes, New York: McGraw-Hill, 1965, pp. 390.

5. N. Nahi, Estimation Theory and Application, New York: Wiley, 1969, pp. 109.

6. P. Liebelt, An Introduction to Optimal Estimation, Menlo Park: Wesley, 1967, pp. 30.

7. P. Liebelt, pp. 146.

8. B. Hunt, "Matrix Theory Proof of the Discrete Convolution Theory," IEEE Trans. Audio and Electroacoustics, December 1971, pp. 285-288.

9. L. Jolley, Summation of Series, New York: Dover Publications, 1961, pp. 84.

10. T. Stockham, T. Cannon, and R. Ingebretsen, "Blind Deconvolution Through Digital Processing," Proc. IEEE, April 1975, pp. 678-692.

11. A. Papoulis, systems and Transforms with Applications in Optics, New York: McGraw-Hill, 1968, pp. 20.

12. A. Papoulis, Systems and Transforms, pp. 37.

7. THE PROBLEM OF POSITIVE RESTORATION, SUGGESTIONS FOR FURTHER RESEARCH, AND CONCLUSIONS

The Wiener restoration approach introduced in the previous chapter neglects certain a priori information concerning the pictorial data which, if utilized properly, could improve the quality of the restored images. Non-negativeness of an optical scene, for instance, is a restriction which can be utilized to reduce the estimation error variance. Also, it has been shown that the visual quality of a restored image can be enhanced if the human visual system response is utilized in the restoration process [7-1]. The following section summarizes certain non-negative restoration approaches which are adaptable to the fast restoration technique.

7. 1 Constrained Restoration

Positive restoration refers to image filtering techniques which employ a positiveness constraint to improve the restoration of degraded images. It has been illustrated that, in general, linear inequality constraints reduce the error covariance of the object estimate, and also improve the stablity of the system representing the degradation phenomenon. Reference [7-2] illustrates this claim by adopting a numerical analysis approach to the problem, although the increased computational requirements

of this approach impose a severe limitation on the allowable size of the image.

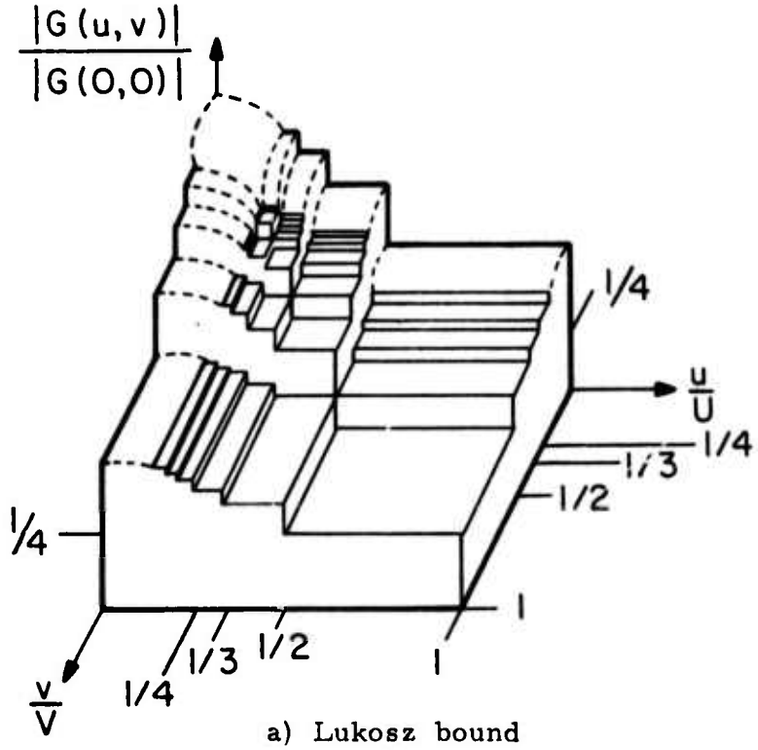
An alternative approach to this problem is to utilize the Fourier domain properties of positive signals. Lukosz [7-3] has determined upper bounds on the Fourier pattern of a non-negative signal and has shown that the amplitude of the Fourier transform of a non-negative, band-limited signal satisfies the constraint

$$|G(u,v)| \leq G(0,0)\left\{1-\frac{|u|}{U}\right\}\left\{1-\frac{|v|}{V}\right\} \qquad (7\text{-}1)$$

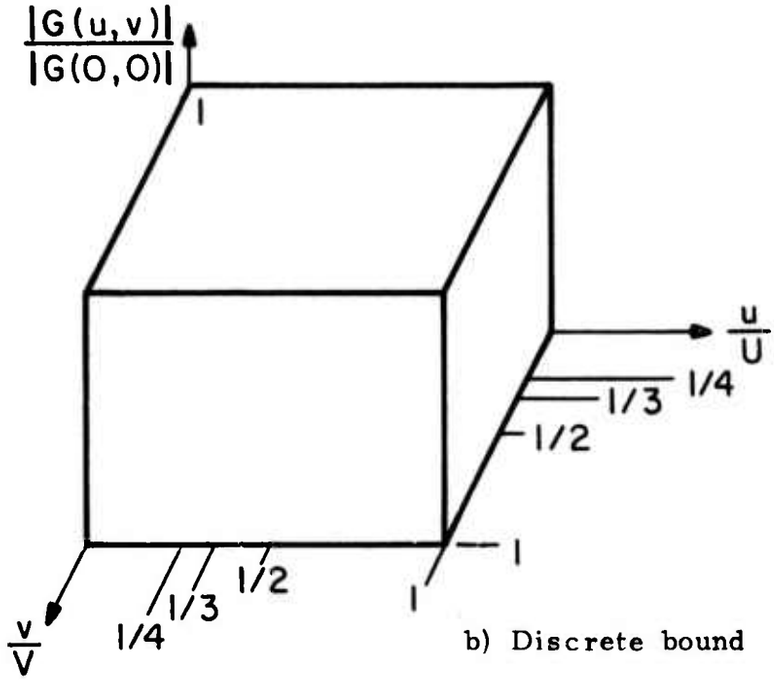
where G(u,v) represents the Fourier transform of the non-negative signal g(x,y), and U and V are the cutoff frequencies. Figure (7-1a) illustrates this bound. It should be noted that eq. (7-1) is a necessary condition, but not necessarily a sufficient restriction. In other words, there are many signals which satisfy inequality (7-1), but are not necessarily non-negative. The Lukosz bound in its present form does not apply to discrete signals, and the only constraint which appears to hold for discrete signals can be expressed as

$$|G(i,j)| \leq G(0,0) \qquad (7\text{-}2)$$

where the G(i,j) are the discrete values of the Fourier



a) Lukosz bound



b) Discrete bound

Figure (7-1) Fourier domain bounds.

transform of the discrete signal g(l,m). Figure (7-1b) illustrates this bound.

Constraints having the form shown in eq. (7-1) or eq. (7-2) are very simple to implement, particularly, if in the filtering process, the Fourier transforms can be computed in advance. For example, implementation of the inequality in eq. (7-2), for images restored by the fast Wiener filter of Sec. (6-3), is trivial in nature. This is because the Wiener filtering takes place in the Fourier domain, and all that remains is to check every coefficient of this domain against the d.c. value of the object, G(Ø,Ø). If any coefficient violates this bound, the coefficient will be automatically decreased until the bound is satisfied.

The major shortcoming of the Fourier domain inequality restrictions for positive image restoration stems from the fact that most images contain small quantities of high frequency components, and thus rarely violate an inequality of the form of eq. (7-2). For instance, although the restored object of fig. (6-14b) occasionally contains negative quantities, it never violates the bound described by eq. (7-2). Hence, before displaying the image, the negative entries were cliped to zero.

A brute force approach for assuring the positiveness of the images restored by the fast filter can be

implemented as follows. Considering a one-dimensional case, let F(i) represent the ith Fourier coefficient of a K-dimensional image estimate $\underline{\hat{f}}$. Starting from the coefficient associated with the lowest frequency, F(0), an image is constructed using the Fourier basis function corresponding to the selected coefficient, its complex conjugate, and all the lower order basis vectors. Note that since the image is real, every Fourier entry is accompanied by another coefficient which is its complex conjugate, and that the d.c. term F(0) is always positive. Thus, at the very first stage, the image is represented by a gray level which has a numerical value equal to F(0); at the second step, the two next lowest frequency vectors are added. This process is continued until a negative quantity is detected in the constructed image when, in this case, the two coefficients corresponding to the very last two vectors added to the image are modified to retain the image positiveness. Then, the procedure is carried on until all the coefficients are exhausted and the image is totally created. In some cases, however, the reconstructed image may never become negative. To avoid the lengthy requirements of positive restoration in such a case, the image can be first constructed by simply inverse Fourier transforming, and then if negative quantities are found, the positive restoration technique described above can be applied. The only drawback of the brute force method is in

its computational inefficiency. It takes a considerable amount of computing time to check if every single Fourier coefficient retains the image positiveness. One suggestion requiring further research in this area is to attempt to establish methods to predict the coefficients which are likely to produce negative quantities. This type of coefficient selection is likely to be a function of the size of the quantity itself. For instance, the quantities which are very small need not be checked out, for they usually cannot give rise to negative entries in the image vector; even if they produce negative quantities, their modification cannot improve the pictorial data substantially.

7. 2 The Fast Wiener Filter and the Eye Model

It is known that a subjectively optimal image estimate can be obtained if the human visual system characteristics are employed to constrain the restoration technique. Unfortunately, because of the inaccessiblity and complexity of the visual system, the true nature of this system is not fully known. However, indirect measurements and repetitive experiments have unveiled some of the mysteries of the process of vision. For instance, for low contrast images, the frequency response of the visual system is believed to be of form of fig. (7-2), and, it is known that the human visual system responds nonlinearly to incident light

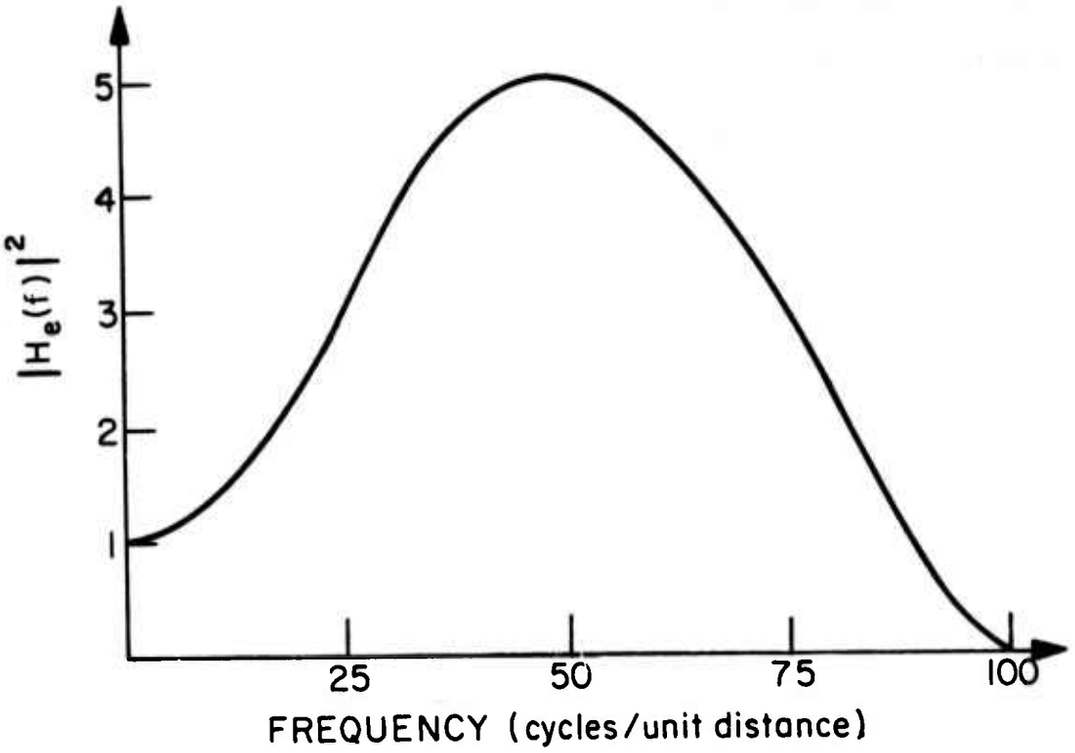


Figure (7-2) Frequency response of the eye.

intensity [7-4].

To simplify the notation let S* denote an operation modelling the human visual system. For instance S* can represent a simple logarithm operation, although, in reality S* can only be expressed by much more complicated functions. The complexity of the eye model makes it impractical to design a minimum mean-square restoration technique which satisfies the constraint defined by S*. A good approach for avoiding this hindrance is to process the image signal output of the visual system. Let g represent the observed signal, and let q denote the same observation after going through the system S*. It should be noted that q is actually the image which is perceived by the brain. To improve q, the Wiener filter technique can be applied to minimize the observation error in the mean-square sense. Figure (7-3) illustrates this approach. Figure (7-3a) represents the eye model, where f denotes the object signal and d represents the signal observed by the brain. Figure (7-3b) illustrates the filtering process. In this figure S* denotes the inverse function of the visual system. It should be noted that S* is followed immediately by the observers eye system S*. Thus, the final result is shown in fig. (7-3c). This figure illustrates that, by Wiener filtering the signal entering the brain q, the complicated eye model constrait can be avoided.

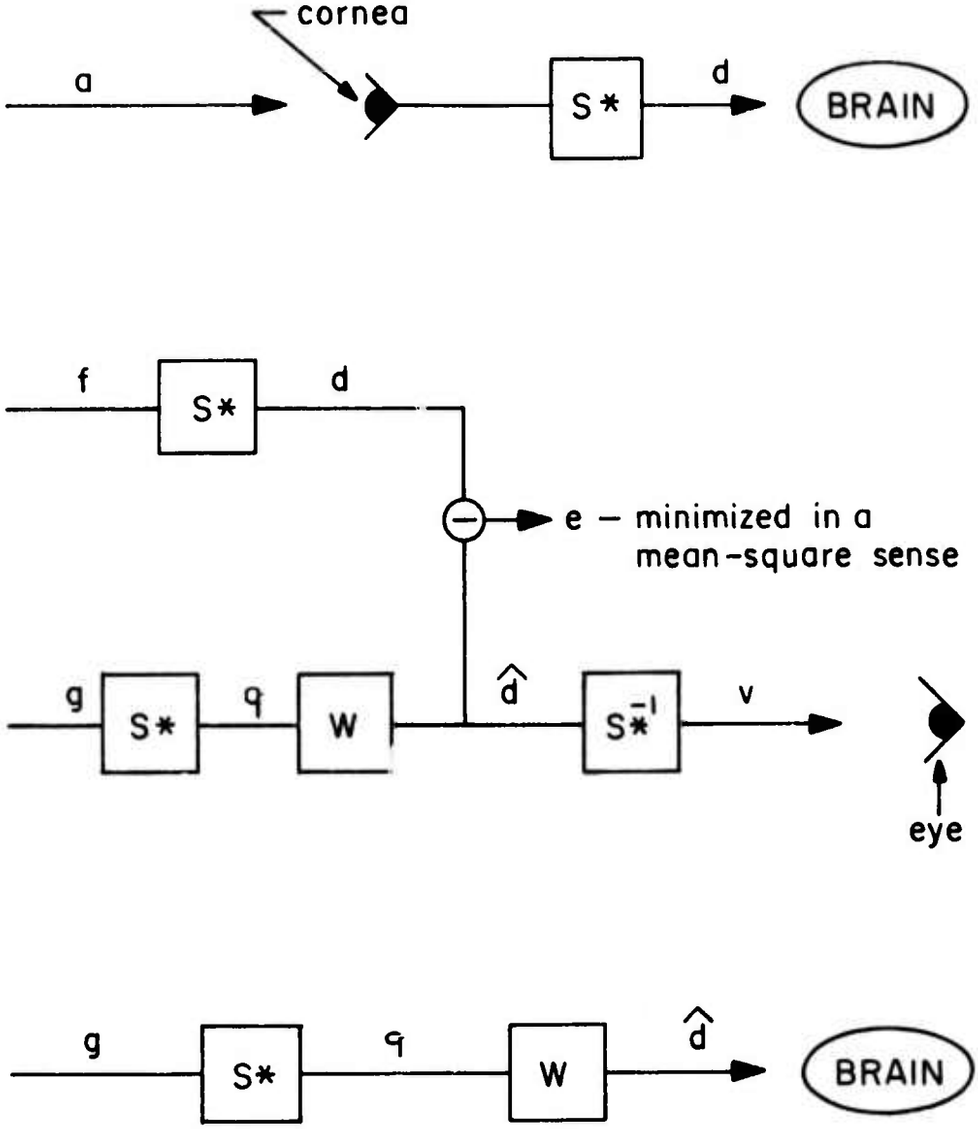


Figure (7-3) Wiener filter and the eye model.

To summarize this dissertation, the next section briefly reviews this work and suggests means of generalizing some of the techniques introduced in the previous chapters.

7. 3 Extensions to the fast Wiener filter

It was shown earlier in this dissertation that the fast Wiener filter in the frequency domain can be described by the following scalar equation

$$W(n) = \frac{a(n)}{|a(n)|^2 + S^{-1} b^{-1}(n)} \qquad (7\text{-}3)$$

where a(n) represents the n-th eigenvalue of the corresponding blur matrix, b(n) represents the n-th eigenvalue of the circulant covariance matrix, and S is the signal-to-noise ratio. If the noise power approaches zero, eq. (7-3) represents the inverse filter

$$I(n) = \frac{1}{a(n)} \qquad (7\text{-}4)$$

where I(n) represents the n-th entry of the inverse filter in the frequency domain.

It has been shown that the minimum mean-square filter is not usually the best technique for image restoration

applications. This is merely because this filter does not emphasize high frequency components of an image, and these components are very important for visual perception [7-4]. Figure (7-2) shows the frequency response of the human visual system, illustrating the importance of high frequency pictorial information for a human observer. Since the human visual system attenuates lower spatial frequencies, the higher frequency components must be of greater utility. Considering this argument, it may seem that the inverse filter approximates the human visual system closer than the Wiener filter. Unfortunately, in the presence of noise, the inverse filter is not feasible. However, a compromise can be achieved if the two filters are averaged in a geometrical sense [7-5]. A Geometrical Mean filter is formulated as

$$G(n)=\left(\frac{1}{a(n)}\right)^{s}\left(\frac{a^{*}(n)}{|a(n)|^{2}+S^{-1}b^{-1}(n)}\right)^{1-s} \qquad (7\text{-}5)$$

where $0 \leq s \leq 1$, and G(n) is the n-th entry of the filter in the frequency domain. Clearly, the Wiener and the inverse filters are both special cases of the Geometrical Mean filter for s=0 and s=1, respectively. Since eq. (7-5) defines a scalar operation, this filter is already in a computationally efficient (fast) form. This holds true since b(n) is an eigenvalue of the circulant covariance matrix. To apply a fast Geometrical Mean filter to a

blurry observation, the same steps which are involved in the application of the fast Wiener filter must be observed. For instance, here too the M-dimensional observation $\underline{g}$ is operated upon by the windowing matrix to reduce the wrap-around error and then is placed in a longer vector of zeros of size K, where K is given by eq. (6-29). And, after the K-dimensional observation is processed by the fast flter, a K-dimensional object estimate resulrs in which the first N entries of this vector are the true object estimate f. Since the Geometrical Mean filter represents a class of restoration techniques to which the Wiener and the Inverse filters are special cases, only one general software (filter) is necessary to be designed to give a wide selection of restoration techniques.

7. 4 Summary and Conclus ons

This dissertation has developed computationally efficient image restoration techniques, and pictorial examples have been presented to illustrate the efficiency of these techniques.

The continuous convolution integral has been utilized to represent linear shift-invariant degradation phenomena. It has been shown that the discrete image degradation problem can be modelled by an overdetermined system, if the object background is not known. To avoid this dual model,

a simple operation has been introduced which, by predicting the object background, almost eliminates the contribution of the background to the pictorial data. Hence, only the overdetermined model is employed to describe the degradation problem. This operation is, in essence, designed to control the modelling error. In the absence of noise, the inverse filter for the continuous case or the matrix pseudoinverse for the discrete data can be employed to restore degraded images. The computational shortcomings of the pseudoinverse techniques can be overcome by utilizing the Fourier domain properties of circulant matrices. Simulated pictorial examples were used to illustrate this point.

In a noisy environment, the statistics of the image and the noise can be employed to control high frequency noise oscillation in the restored data. Based on this fact, a minimum mean-square error filter can be constructed for image restoration. Since a filter of this kind is computationally unattractive, a fast Wiener filter was introduced for noisy image restoration. This filter was obtained by imposing certain modifications on the observed image. It was shown that the observed image can be operated upon to modify its statistical characteristics. Thus, certain operations were introduced which, when applied on pictorial data, allow the modified data to be characterized statistically by a circulant covariance

matrix. Later, these operations were approximated to gain more computational speed at the cost of a slight increase in the mean-square error. An error study then illustrated that for most of the length of the object vector, the error variance associated to the fast Wiener filter is equal to the one associated with the Wiener filter itself. The slight error increase occurs only at the beginning and the end of the object vector, and almost vanishes when the element correlation becomes small. Unlike the classical Wiener filter, it was shown that the fast filter is capable of operating upon large images without the need to break down the observation into small blocks.

Statistically speaking, it was assumed that a scene is a sample of a Markovian random process. This, of course, is not an essential restriction, and can be removed. It appears that the fast filter can be constructed for any image which possesses a monotonically decreasing correlation function. In fact, any covariance matrix which can be extended into a larger, circulant, and positive definite matrix characterizes a random image that can be operated by the fast filter. Thus, lemma (6-1) can be proved for certain non-Markovian sources as well.

Since the main objective in this dissertation has been to present computationally efficient restoration techniques, it seems proper at this stage to point out that

the classical Wiener approach is not the only restoration technique which can be modified for computational efficiency. In fact, as previously shown, there is a class of filters which can be efficiently computed through a technique similar to the Wiener approach. Notice that the phrase "computational efficiency" corresponds to both the time and storage requirements of a certain restoration technique. Hence, a fast counterpart of a restoration technique is computationally advantageous, in the sense that the fast filter permits processing of large images in a relatively small amount of time.

REFERENCES

1. B. Hunt, "Digital Image Processing," Proc. IEEE, April 1975, pp. 693-708.

2. N. Mascarenhas, Digital Image Restoration Under a Regression Model- The Unconstrained, Linear equality and Inequality Constrained Approaches, Ph. D. Dissertation, University of Southern California.

3. W. Lukosz, "Ubertragung Nicht-Negativer Signale durch Lineare Filter," Optica Octa, 9, 1962, pp. 335-364.

4. T. Stockham, "Image Processing in The Context of a Visual Model," Proc. IEEE, July 1972, pp.828-842.

5. H. Andrews, "Digital Image Processing: A Survey," IEEE Computer, May 1974, pp. 36-45.