ENGINEERING

USCIP Report 890

LEVEL II

10

UNIVERSITY OF SOUTHERN CALIFORNIA

ESTIMATION OF IMAGE SIGNALS WITH POISSON NOISE

by

Chun Moo Lo

June 1979

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

Sponsored by

Advanced Research Projects Agency
Contract No. F-33615-76-C-1203
ARPA Order No. 3119

IPI

IMAGE PROCESSING INSTITUTE

(14) USCIPI Report-890

(6) # ESTIMATION OF IMAGE SIGNALS WITH POISSON NOISE

by

(10) Chun Moo/Lo

(11) June 1979

(12) 206 p.

(1) Technical rept.

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

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Wright Patterson Air Force Base under Contract F-33615-76-C-1203, ARPA Order No. 3119.

(15) ARPA Order-3119

391 141

mt

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER USCIPI Report 890	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ESTIMATION OF IMAGE SIGNALS WITH POISSON NOISE		5. TYPE OF REPORT & PERIOD COVERED Technical Report, June 1979 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Chun Moo Lo		8. CONTRACT OR GRANT NUMBER(s) F-33615-76-C-1203
9. PERFORMING ORGANIZATION NAME AND ADDRESS Image Processing Institute University of Southern California University Park, Los Angeles, Calif. 90007		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS ARPA Order No. 3119
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209		12. REPORT DATE June 1979 13. NUMBER OF PAGES 205
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Wright Patterson Air Force Base U.S. Air Froce Air Force Avionics Laboratory Air Force Systems Command, Ohio 45433		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for release: distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image Processing, Image Restoration,
Maximum A Posteriori Estimation, Poisson Process,
Poisson Noise

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

An optimal filter in the sense of maximum a posteriori probability (MAP) is derived for image signals detected at low light levels. These signals suffer from Poisson noise and blurring degradations.

The low level photon resolved image signal is modeled as an inhomogeneous Poisson point process. The photon noise is inherent in any detected image, and is particularly serious at low

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

light levels. At these low light levels, the emission of photons is described by a Poisson point process, with the average rate of emission proportional to the integrated intensity. The blurring degradation model in the system includes space-variant and space-invariant effects such as atmospheric turbulence, linear motion, diffraction, and aberrations. The estimation is performed assuming that the photon events counted in each detector are independent, Poisson distributed random processes for the large time-bandwidth product case. Since the variance of the Poisson distribution is identical to its mean, the Poisson noise is neither multiplicative noise nor a linear additive Gaussian noise, and is generally signal-dependent. It has been demonstrated that MAP estimation with the Poisson noise model has improved performance because the MAP filter can be generalized to linear or nonlinear image models and to noise models different from additive Gaussian noise. In addition, the MAP filter can be a local adaptive processing filter and extended to the case of space-variant blurring. It also has been shown that image models with a nonstationary mean and stationary variance give useful a priori information for the MAP filter. The MAP estimation equations are nonlinear and have large dimensionality. A sectioning method with a Newton-Raphson solution has been adapted to cope with these problems. It has been shown that the strategy is an effective and fast way to solve nonlinear MAP estimation equations.

The Cramer-Rao lower bound (CRLB) on the mean-square estimation error of the MAP unbiased estimate is derived for the Poisson noise model. It is shown to be a very useful bound for finding the best suboptimal sectioning filters.

Finally, a comparison between the performance of the MAP filter and that of the linear minimum mean-square error (LMMSE) filter is made for Poisson noise models. The performance of the MAP filter is much better than that of the LMMSE filter. The LMMSE filter works very well for higher signal-to-noise ratios, but the MAP filter works better for low signal-to-noise ratios where Poisson noise dominates.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Professor Alexander A. Sawchuk, whose assistance and invaluable suggestions made this dissertation possible. I am indebted to him for his advice and guidance throughout my graduate career. I am deeply grateful to Professor Ta-Liang Teng for his countless assistance and encouragement. Without his help, my graduate education would still be a dream. I also would like to thank Professor Jerry M. Mendel for his helpful suggestions on this dissertation.

I would like to thank Professor Harry C. Andrews, who was the director of the Image Processing Institute, for providing financial support during the dissertation research period. I would also like to thank all the staff of the Image Processing Institute for their help, especially Hilda Marti for her preparation of seminar material, and Amy Yiu for her superb job in typing this completed work.

Finally, I would like to thank my wife Shu-Taur (Tammy). Without her constant sacrifice and encouragement the work would not have been completed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

An optimal filter in the sense of maximum a posteriori probability (MAP) is derived for image signals detected at low light levels. These signals suffer from Poisson noise and blurring degradations.

The low level photon resolved image signal is modeled as an inhomogeneous Poisson point process. The photon noise is inherent in any detected image, and is particularly serious at low light levels. At these low light levels, the emission of photons is described by a Poisson point process, with the average rate of emission proportional to the integrated intensity. The blurring degradation model in the system includes space-variant and space-invariant effects such as atmospheric turbulence, linear motion, diffraction, and aberrations. The estimation is performed assuming that the photon events counted in each detector are independent, Poisson distributed random processes for the large time-bandwidth product case. Since the variance of the Poisson distribution is identical to its mean, the Poisson noise is neither multiplicative noise nor a linear additive Gaussian noise, and is generally signal-dependent. It has been

demonstrated that MAP estimation with the Poisson noise model has improved performance because the MAP filter can be generalized to linear or nonlinear image models and to noise models different from additive Gaussian noise. In addition, the MAP filter can be a local adaptive processing filter and extended to the case of space-variant blurring. It also has been shown that image models with a nonstationary mean and stationary variance give useful a priori information for the MAP filter. The MAP estimation equations are nonlinear and have large dimensionality. A sectioning method with a Newton-Raphson solution has been adapted to cope with these problems. It has been shown that the strategy is an effective and fast way to solve nonlinear MAP estimation equations.

The Cramer-Rao lower bound (CRLB) on the mean-square estimation error of the MAP unbiased estimate is derived for the Poisson noise model. It is shown to be a very useful bound for finding the best suboptimal sectioning filters.

Finally, a comparison between the performance of the MAP filter and that of the linear minimum mean-square error (LMMSE) filter is made for Poisson noise models. The performance of the MAP filter is much better than that of the LMMSE filter. The LMMSE filter works very well for higher signal-to-noise ratios, but the MAP filter works

better for low signal-to-noise ratios where Poisson noise dominates.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Image restoration can be viewed as an estimation process in which operations are performed on observed or measured noisy data to estimate the object. More clearly, image restoration is the estimation of the original image signal by both blur removal and noise suppression. Image enhancement is the attempt to improve the appearance of the image for better human viewing or machine processing. Hence, image enhancement may not specifically need knowledge of the degrading phenomena. However, in order to effectively develop an optimal restoration filter with various criteria, it is necessary to characterize quantitatively the image degradation effects of the physical image system. Image restoration begins with a model of degradation effects, assumes given a priori information and then develops an optimal filter to obtain a restored image. Hence, accurate image modeling and more a priori information are often the key to effective image restoration.

In this dissertation, optimal restoration filters are developed in the sense of maximum a posteriori probability (MAP) and maximum likelihood (ML) for blurred image signals detected at low light levels. This low level photon resolved image signal is modeled as an inhomogeneous Poisson point process. The photon noise (which we call Poisson noise throughout this thesis) is inherent in any detected image signal particularly at low light levels. At these low light levels, the emission of photons is described by a Poisson process with the average rate of emission proportional to the integrated incident intensity. The estimation is performed assming that the number of photon events counted by the detectors are independent, Poisson distributed random processes for a given unknown object radiance. Since the variance of the Poisson distribution is identical to its mean, the Poisson noise is neither multiplicative noise nor linear additive Gaussian noise. It is, indeed a signal-dependent noise.

1.2 Organization and Contributions of the Dissertation

In next chapter, we discuss three image models and some system models for image noise; we also present an inhomogeneous Poisson process model which is a photon counting system. In Chapter 3, we review some important linear and nonlinear image restoration filters for Poisson noise models and their motivation for the work in this

dissertation. We also review other restoration filters for different image noise models. In Chapter 4, we develop and implement an MAP filter without blurring degradations for the Poisson noise model. In Chapter 5, we develop and implement MAP filters with blurring degradations for the Poisson noise model. In Chapter 6, we derive Cramer-Rao lower bounds (CRLB) on the estimation error for MAP filters and discuss the results. In Chapter 7, we compare the restored image performance of the MAP filter with that of the LMMSE filter. In Chapter 8, we conclude this dissertation and discuss future research on the problem.

The specific research contributions of this dissertation are now summarized. A model for photon resolved low light level image signals detected by a counting array is developed. These signals are impaired by signal dependent Poisson noise and linear blurring.

An optimal restoration filter based on maximizing the a posteriori probability density (MAP) is developed. A suboptimal overlap-save sectioning method using a Newton-Raphson iterative procedure is used for the solution of the high dimensionality, the nonlinear estimation equations for any type of space-variant and invariant linear blur. An accurate image model with a nonstationary mean and stationary variance is used to provide a priori information for the MAP restoration filter. The Cramer-Rao

lower bound (CRLB) of the unbiased MAP restoration filter is derived. Finally, a comparison between the MAP filter and a linear space-invariant minimum mean-square error (LMMSE) filter is made.

CHAPTER 2

SYSTEM MODELS FOR IMAGE NOISE AND IMAGE MODELS

2.1 Introduction

Using a statistical approach to develop an optimal restoration filter, the effectiveness of the algorithm depends on the completeness of the statistical description of image and noise. Hence, we investigate image and noise models in detail in the following two sections.

In section 2.4, we present an inhomogeneous Poisson process model (we refer this inhomogeneous Poisson process as the Poisson noise degradation throughout this thesis) which is a photon counting systtem. In section 2.4.1 we investigate a general photon counting system containing blurring and Poisson noise degradations. In section 2.4.2, we first discuss a single detector model and derive its statistical properties. In section 2.4.3, we extend the single counter to an array. In section 2.4.4, we discuss quantum limitations of photon resolved image signals, and in section 2.4.5, we discuss the simulation of images with Poisson noise at different ensemble mean signal-to-noise ratios. A comparison is made between Poisson noise degraded images and linear additive Gaussian noise degraded

images. Finally some conclusions of this chapter are presented.

2.2 System Models for Image Noise

2.2.1 Linear Additive Gaussian Noise Model

This model is most often used for image formation in the field of digital image processing. Its block diagram is illustrated in Fig. 2.1, where the image $g(x,y)$ and object $f(x,y)$ may be considered intensity functions of two spatial dimensions (x,y), and $h(x,y)$ is the point-spread function (PSF) or impulse response of the imaging system. Because the linear blurring degradation in all image noise models is the same as in this model, we investigate more details about the PSF in this section. The function $n(x,y)$ is additive noise which is signal-independent and Gaussian. To unify the notation, we denote continuous functions with (x,y) and discrete functions with (i,j) throughout this thesis. The mathematical representation of Fig. 2.1 is

$$g(x,y) = \int_{-\infty}^{\infty}\int h(\varepsilon,\eta;x,y)f(\varepsilon,\eta)d\varepsilon d\eta + n(x,y). \qquad (2.1)$$

This equation is a first order Fredholm integral equation plus a random noise component, where (ε,η) is the spatial coordinate of the object of interest and $h(\varepsilon,\eta;x,y)$ is a general space-variant point-spread function (SVPSF) describing the effects of the optical imaging system.

Ideally, it is desirable that the point-spread function for a Dirac delta function $\delta(x,y)$, in which case the image $g(x,y)$ equals the object $f(x,y)$ in the absence of noise. Furthermore, if $h(\varepsilon,\eta;x,y)$ is a function only of the differences between respective coordinates, that is

$$h(\varepsilon,\eta;x,y) = h(x-\varepsilon,y-\eta) \tag{2.2}$$

the PSF is said to be spatially invariant or isoplanar. In all other situations, the PSF is said to be spatially variant or anisoplanar. The physical meaning of a spatially invariant PSF is that the blurring degradation is unchanged across the image plane and the image and object are mathematically related via a two-dimensional convolution

$$g(x,y) = \int_{-\infty}^{\infty}\int h(x-\varepsilon,y-\eta)f(\varepsilon,\eta)d\varepsilon d\eta+n(x,y) \tag{2.3}$$

or equivalently

$$g(x,y) = \int_{-\infty}^{\infty}\int h(\varepsilon,\eta)f(x-\varepsilon,y-\eta)d\varepsilon d\eta+n(x,y). \tag{2.4}$$

These convolution integrals can be Fourier transformed to yield

$$G(u,v) = H(u,v)F(u,v)+N(u,v) \tag{2.5}$$

where the capital letters denote the Fourier transform of

the respective function represented by lower case letters and (u,v) is spatial frequency. In the discrete domain Eq. (2.5) can be expressed by a discrete Fourier transform (DFT) equivalent.

To process image signals on a digital computer we need a spatially discrete form of signal. Equation (2.4) can be represented as a discrete-discrete system [2-1,2-2] by a matrix. This matrix can be represented as a vector by lexicographically ordering the column of the matrix, i.e., the (i,j)th element of the M×N matrix is the [(j-1)m+1]th element of the vector. This ordering permits the use of a very simple matrix model

$$\underline{g} = H\underline{f} + \underline{n} \tag{2.6}$$

where $\underline{g}$ is an $N^2\times 1$ recorded or measured image data vector

$\underline{f}$ is an $N^2\times 1$ original object vector

$\underline{n}$ is an $N^2\times 1$ additive Gaussian noise vector

H is an $N^2\times N^2$ blurring matrix which is a transformation matrix representing the blurring degradation.

Thus the linear restoration matrix model is as shown in Fig. 2.2. The additive nature of the noise in Eq. (2.6) is a model for thermal noise and amplifier electronics noise in image sensors. This additive noise is often itself modeled as a Gaussian process. Since this model represents the physical reality well and is mathematically tractable,

it is the most conventional practical model.

2.2.2 Multiplicative Image Noise Model

A block diagram of a multiplicative noise model [2-6,2-7,2-8] and the associated restoration filter is illustrated in Fig. 2.3. Its mathematical expression is

$$g(x,y) = [f(x,y) \circledast h(x,y)] \cdot n(x,y) \tag{2.7}$$

or

$$g(x,y) = \left[\int_{\infty}^{\infty} \int h(\varepsilon,\eta;x,y) f(\varepsilon,\eta) d\varepsilon d\eta \right] \cdot n(x,y) \tag{2.8}$$

where $h(x,y)$ is the PSF of the linear system, $f(x,y)$ and $g(x,y)$ are the obejct and degraded image functions respectively. Here $n(x,y)$ denotes signal-independent Gaussian noise, and $\circledast$ denotes two-dimensional convolution. T. Yatagai [2-9] has used this model for speckle noise in the sense that standard deviation of the speckle is equal to its mean.

2.2.3 Additive Signal-Modulated Image Noise Model

The additive signal-modulated image noise model and the associated restoration filter is illustrated in Fig. 2.4. Its mathematical expression is

$$g(x,y) = f(x,y) \circledast h(x,y) + c[f(x,y) \circledast h(x,y)] \cdot n(x,y). \tag{2.9}$$

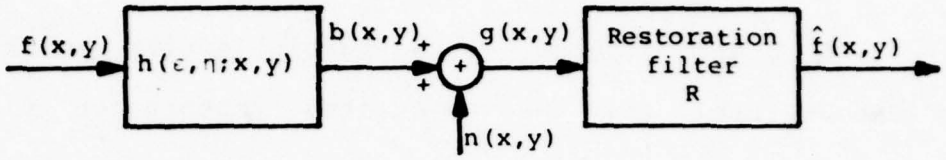


Figure 2.1. Block diagram of the linear additive Gaussian noise model and the associated restoration filter

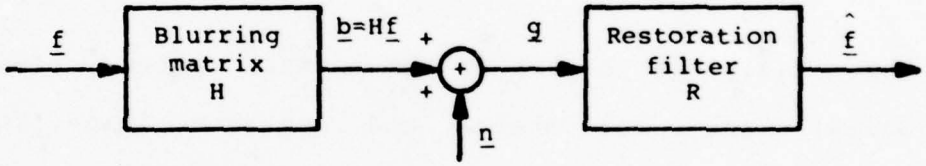


Figure 2.2. Block diagram of the linear restoration vector model with additive noise

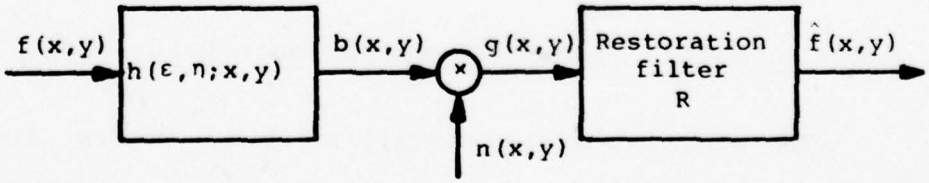


Figure 2.3. Block diagram of the multiplicative noise model and the associated restoration filter

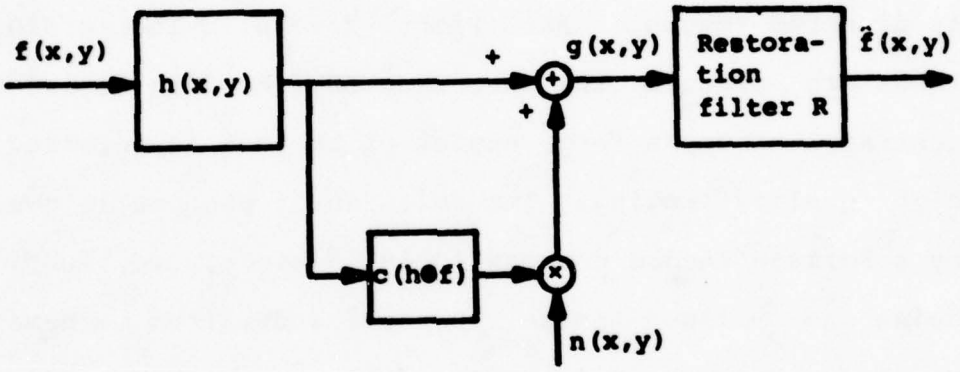


Figure 2.4. Block diagram of the additive, signal modulated image noise model and the associated restoration filter

Here C is generally a memoryless point nonlinear function, n(x,y) is the signal-independent noise and ⊛ denotes two-dimensional convolution. Film-grain noise and magnetic tape recording noise are accurately modeled as additive signal-modulated noise [2-10,2-11].

2.2.4 Poisson Image Noise Model

Because of the quantum nature of light, photons arrive at random times and give rise to a fundamental graininess in detected images at low light levels. The graininess tends to obscure the detection of fine details and faint contrasts, thus, a large number of photons is required for high quality imaging. The emission of photons is governed by a Poisson random process [2-19], hence, we label the noise as Poisson noise. Because a detailed mathematical model for Poisson noise is given in section 2.4, only qualitative comments on Poisson noise are given here for the completeness of this section.

Poisson noise is another basic type of signal-dependent image noise which is quite different from film-grain noise and speckle noise. The signal-dependent nature of Poisson noise is associated with the fact that the variance of the Poisson probability distribution is equal to its mean. If the signal information received by an array of photo detector elements is contained in the mean number of events recorded by each element, then the

Poisson distribution of these events implies that Poisson noise is a form of signal-dependent noise. All low-level photon resolved signals are examples of signals corrupted by Poisson noise. These situations occur in scintillation camera imaging, medical imaging, astronomical imaging, and low light level television systems.

2.3 Image Models

Three of the more detailed image models are discussed in this section. Better estimates of statistical images should come from more accurate image models.

2.3.1 Gaussian Image Model

On the basis of physical arguments and mathematical tractability, Hunt [2-12,2-13] developed one of the most accurate image models. The image is modeled as a multivariate Gaussian process with nonstationary ensemble mean and with stationary covariance. The image vector *a priori* probability density function (pdf) is

$$p(\underline{f}) = ((2\pi)^{N/2}|R_f|^{1/2})^{-1}\exp\{-\frac{1}{2}(\underline{f}-\bar{\underline{f}})^T R_f^{-1}(\underline{f}-\bar{\underline{f}})\} \quad (2.10)$$

where $\bar{\underline{f}}=E[\underline{f}]$ is the nonstationary mean vector

$R_f=E[(\underline{f}-\bar{\underline{f}})(\underline{f}-\bar{\underline{f}})^T]$ is the stationary covariance matrix

$|R_f|$ is the determinant of R_f and T denotes transpose.

Lower case p denotes probability density function (pdf) and

capital case P denotes probability throughout this thesis. Equation (2.10) describes a random process of stationary fluctuations about a nonstationary mean value. Hunt derived this image model based on a heuristic analysis from following "thought" experiment. Suppose that several thousands of photographs with similar statistical properties (such as driver's license photographs) were used to calculate an ensemble mean image. Each face is positioned in approximately the same way in each image frame. Clearly such an ensemble mean image would not consist of a uniform shade of gray indicating spatial stationary. More likely, the mean image probably consists of an oval region where the face is expected to be and some dark spot where the eyes, nose, and other facial features are expected to be. Thus, images are generally nonstationary in first order statistics and are described by a spatially non-stationary mean vector $\underline{\bar{f}}$. The ensemble mean is strongly dependent upon the context which is established by the sample mean of the image to be modeled. The ensemble is called a context-dependent image ensemble. Wintz [2-16] has shown that images may have identical covariance statistics and the same constant mean intensity, but be completely unrelated in context. Therefore,the context-dependent ensemble properties are portrayed most strongly in the mean vector $\underline{\bar{f}}$, since this vector has the gross structure that represents the context of the ensemble

from which the sample vector $\underline{f}$ is drawn.

The spatial statistics $\underline{\bar{f}}$ and R_f of Eq. (2.12) do not represent a "second-order" stationary random process if the context-dependent ensemble mean is assumed, since $\underline{\bar{f}}$ is assumed to have nonstationary gross structure that depends upon the ensemble and its context. The covariance statistics can be described as spatially stationary fluctuations about a spatially nonstationary mean vector $\underline{\bar{f}}$. The random process associated with the image ensemble is not ergodic in the mean, since the ensemble average of the context-dependent ensemble is not equal to the spatial average of an ensemble member. However, the process can have a stationary autocovariance.

If an image f(j,k) can be described as being the sum of two components, a low frequency or blurred component $\bar{f}(j,k)$ and a high-frequency component s(j,k) of the fluctuations about $\bar{f}$, i.e.

$$f(j,k) = \bar{f}(j,k) + s(j,k) \qquad (2.11)$$

then $\bar{f}(j,k)$ represents the nonstationary mean and the variance of the difference image between the image f and the nonstationary mean image $\bar{f}$ is approximately equal to that of the high-frequency component s(j,k). Thus, the ensemble mean of the random process for an image is nonstationary and carries the low spatial frequency gross

features of an image, while the covariance properties about this ensemble mean represent random perturbations carrying the detailed image structure. This image model is one of the most sophisticated and will be used in deriving optimum nonlinear filters in later chapters.

2.3.2 Laplacian Image Model

Trussel et al. [2-13] using a nonlinear least-square fit technique found that a Laplacian pdf had a better fit to experimental image pdf's than the Gaussian pdf. Thus the model of Eq. (2.9) can be restated as

$$p_L(\underline{f}) = k_2 \exp\{-\sqrt{2}[(\underline{f}-\overline{\underline{f}})^T R_f^{-1} (\underline{f}-\overline{\underline{f}})]^{\frac{1}{2}}\} \qquad (2.12)$$

where

$$\overline{\underline{f}} = E[\underline{f}]$$
$$R_f = E[(\underline{f}-\overline{\underline{f}})(\underline{f}-\overline{\underline{f}})^T].$$

This process is similar to the Gaussian model because it contains a nonstationary mean and stationary covariance. The square root of the term in brackets in Eq. (2.12) makes the model Laplacian instead of Gaussian. The MAP filter and other restoration methods are later derived for both the Gaussian pdf and Laplacian pdf.

2.3.3 Lebedev's Composite Image Model

This image model repesents a completely different approach from Hunt's image model. Lebedev and Mirkin called their model a "composite model of an image fragment"

which includes the nonstationary statistical properties of an image [2-14,2-15]. They model an ensemble of images as a random field with an n-dimensional point pdf $p(\underline{f})$ where $\underline{f} = [f_1, f_2, \ldots, f_n]$. They decompose the image statistically into M classes of fragments, whose structure is distinguished by the type of correlational links between pixels. Some classes are formed by fragments with isotropic structure; others are found by fragments with some anisotropy.

Let $p(\underline{f})$ be the pdf of a fragment of image $\underline{f}$, on the condition that the fragment belongs to the class θ ($\theta = 1,2,\ldots,M$). Denoting the *a priori* density of the classes by $\pi(\theta)$, we have

$$\sum_{\theta=1}^{M} \pi(\theta) \doteq 1 \tag{2.13}$$

and

$$p(\underline{f}) = \sum_{\theta=1}^{M} \pi(\theta) p_{\theta}(\underline{f}). \tag{2.14}$$

Expression (2.14) is a matrix density decomposition of $p(\underline{f})$ in terms of $p_{\theta}(\underline{f})$, $\theta = 1,2,\ldots,M$ [2-18]. This representation is especially useful when $p(\underline{f})$ is approximated closely by a small collection of Gaussian distributions

$$p_\theta(\underline{f}) = |R_\theta|^{-1} \exp\{-\tfrac{1}{2}\underline{f}^T R_\theta^{-1} \underline{f}\} \qquad (2.15)$$

where R_θ is the covariance matrix corresponding to the class θ fragment image.

Using this image model for developing a spatial restoration filter with a maximum likelihood (ML), maximum a posteriori (MAP) or Wiener criterion leads to a multicategory filtering problem because of the decomposition properties of the image model. This composite image technique seems to be a good model for the local nonhomogeneous information in the image signals, hence an optimal filter can be a local adaptive filter. Although this model is not used in the results presented, it is believed that using this model with the MAP criterion may yield good results in future work.

2.4 Inhomogeneous Poisson Process Photon Counting System Model

2.4.1 Photon Counting System

In many practical situations a detected image can be modeled as a photon counting system illustrated in Fig. 2.5 with its corresponding block diagram shown in Fig. 2.6. For ease of notation, we use a lexicographic ordering vector notation in which

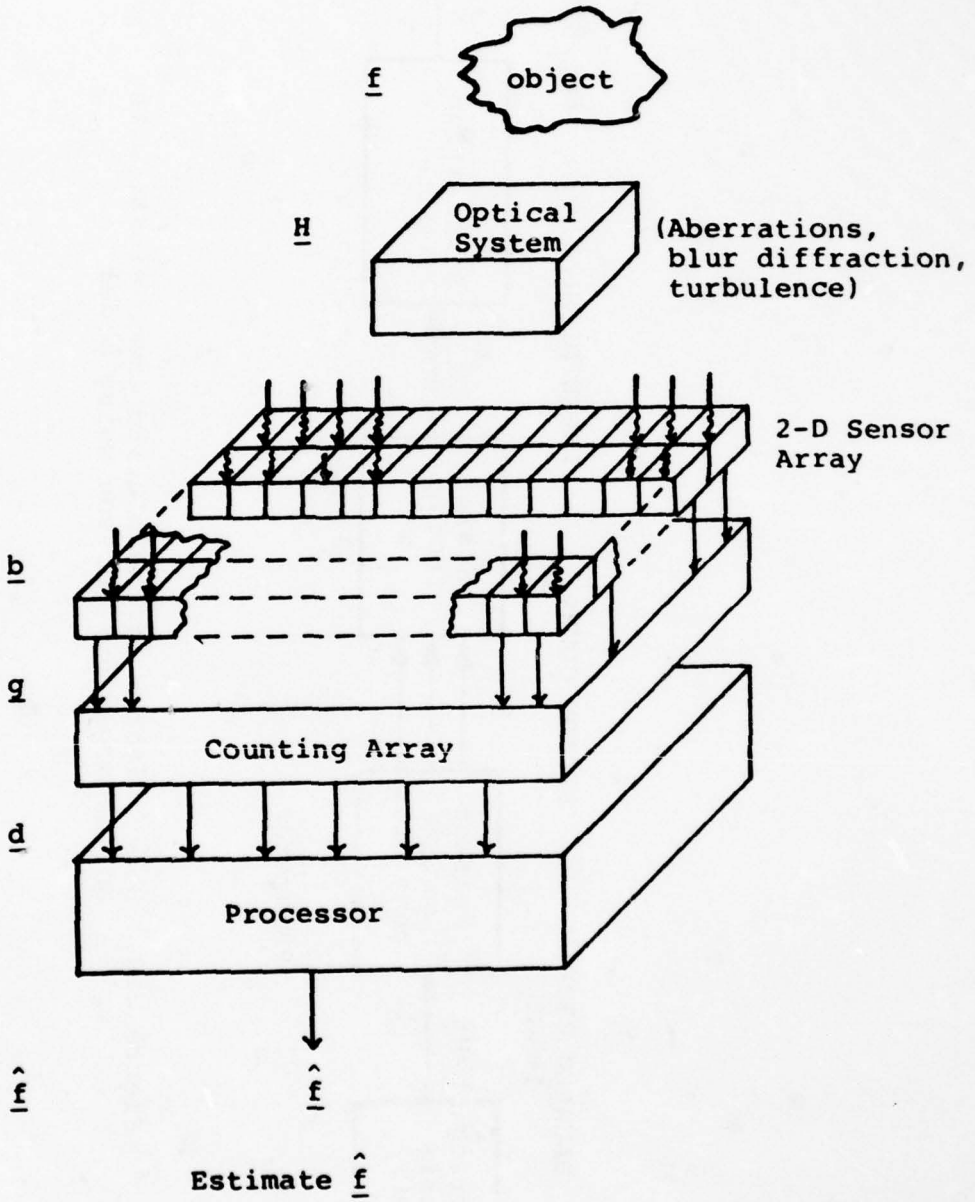


Figure 2.5. Two-dimensional photon counting system

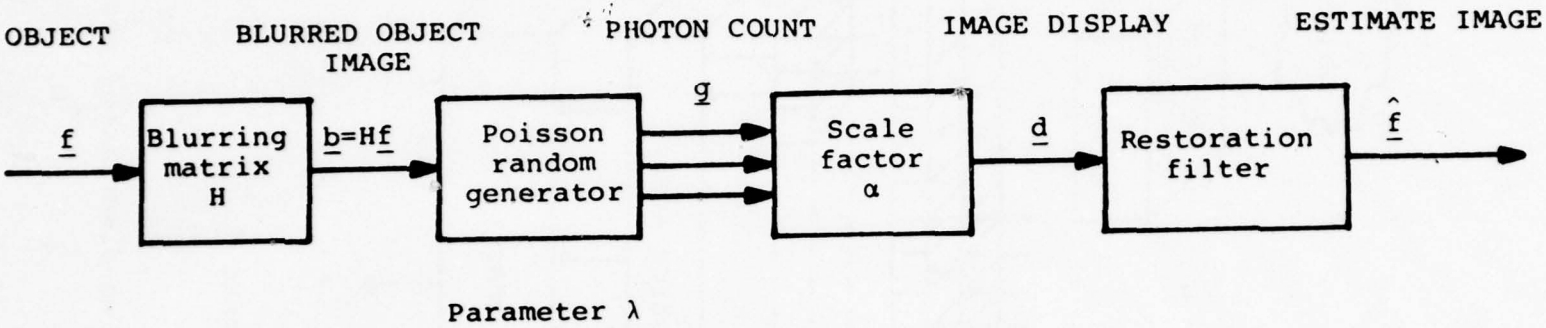


Figure 2.6. Block diagram of the blurred image signal with an inhomogeneous Poisson noise model

$\underline{f}$ is an N^2x1 object vector

$\underline{b}$ is an N^2x1 blurred image vector

$\underline{g}$ is an N^2x1 photon count vector

$\underline{d}$ is an N^2x1 display image vector

$\hat{\underline{f}}$ is an N^2x1 estimate object vector

and H is the N^2xN^2 blurring matrix.

Here, H lumps together all the linear blurring degradation effects such as motion blurring, turbulence, diffraction, and aberrations. The detailed derivation of the two-dimensional discrete operator H from a two-dimensional continuous superposition integral is discussed in [2-1,2-2].

The counting array counts photons incident on the two-dimensional sensor array. The mean of the photon count is proportional to the incident integrated image intensity and the counts themselves are random variables which have a Poisson distribution.

The photon count is a dimensionless number, so we include a scale factor α which provides an image intensity which is displayed. Thus the photon counting system is modeled as an inhomogeneous Poisson process. This is a key assumption of this thesis because the MAP estimate of the image in the non-blurring and blurring cases in the subsequent two chapters is based directly on it. For mathematical simplicity and ease of physical understanding,

we first derive the statistics of a single counter and later extend it to a large array.

2.4.2 Single Photon Counter

From Fig. 2.5 we first assume the H matrix is the identity matrix, and assume only one counter instead of a vector array as in Fig. 2.7. According to the semi-classical theory of photon detection [2-19,2-26], the probability that g_i photon events occur for a given fixed intensity f_i is

$$p(g_i|f_i) = \frac{(\lambda f_i)^{g_i} e^{-\lambda f_i}}{g_i!} \quad , \tag{2.16}$$

where λ is a constant rate parameter. By direct summation, the conditional mean and variance of g_i for constant f_i is

$$\bar{g}_i = \lambda f_i, \tag{2.17}$$

$$\sigma^2_{g_i} = \lambda f_i, \tag{2.18}$$

and

$$\frac{\bar{g}_i}{f_i} = \lambda = \frac{\text{Average \# of photon counts}}{\text{Intensity unit}} \quad . \tag{2.19}$$

With the low light level image signals in which we are interested, we have [2-21,2-26]

$$WT >> 1, \tag{2.20}$$

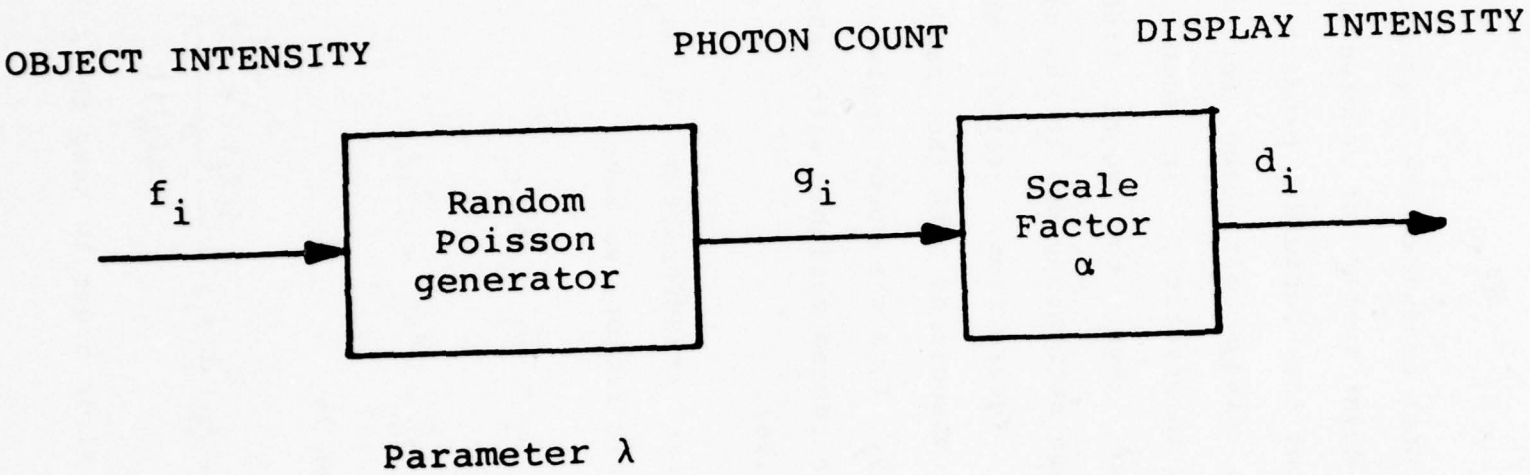


Figure 2.7. Block diagram of a single photon counter model

or equivalently,

$$\bar{g}_i << WT, \tag{2.21}$$

where W is the temporal bandwidth and T is the integration time, i.e. the mean number of counted photo events is small compared to the time-bandwidth product of the light. This condition is always satisfied for natural thermal sources encountered in practice. In such cases, Mandel [2-26] has shown that the count fluctuations are predominantly Poisson shot noise due to the discrete nature of wave-detector interaction, rather than classical "fluctuation noise" associated with the random nature of the image intensity. Thus the count registered on the ith counter is a Poisson random variable with mean $\bar{g}_i$ and pdf expressed by Eq. (2.16).

From the linear transformation $d_i=\alpha g_i$ with α a constant display scale factor, we have

$$\bar{d}_i = \alpha\bar{g}_i = \alpha\lambda f_i, \tag{2.22}$$

$$\sigma_{d_i}^2 = \alpha^2\sigma_{g_i}^2 = \alpha^2\lambda f_i. \tag{2.23}$$

and $p(d_i|f_i)$ is given by

$$p(d_i|f_i) = \frac{1}{\alpha}p(\frac{d_i}{\alpha}|f_i) = \frac{(\lambda f_i)^{\frac{d_i}{\alpha}} e^{-\lambda f_i}}{\alpha(\frac{d_i}{\alpha})!} \tag{2.24}$$

We usually choose $\alpha\lambda=1$ in order to keep the mean value of

the processed display image signal the same as the observed noisy image.

Before we define an r.m.s. signal-to-noise ratio, it is useful to discuss some details about signal-to-noise ratios (SNR's) in general. A study of SNR's at different points in a system enables us to pinpoint the significant contributions to the noise. It is also a simple criterion for the design of systems to minimize from the noise degradation and thus it provides a measure of the "noisiness" of a system. In most cases, the SNR criterion is applied with signal-independent zero mean additive noise. If the signal and noise are dependent, then the SNR is difficult to define because the cross correlation and other moments are nonzero. Poisson noise is a case of signal-dependent noise because the variance in Eqs. (2.18) and (2.23) depends on the signal. In order to compare the noisiness of images with Poisson noise, we will define an r.m.s. SNR denoted by $(SNR)_{rms}$. Because an image signal is a random process in space and time, we must define an ensemble mean $(\overline{SNR})_{rms}$ as the ensemble average of the $(SNR)_{rms}$ over the random image field.

In the case of Poisson noise, the $(SNR)_{rms}$ is

$$(SNR)_{rms} = \frac{\overline{d_i}}{\sigma_{d_i}} = (\lambda f_i)^{\frac{1}{2}} \tag{2.25}$$

and the ensemble mean $(SNR)_{rms}$ is

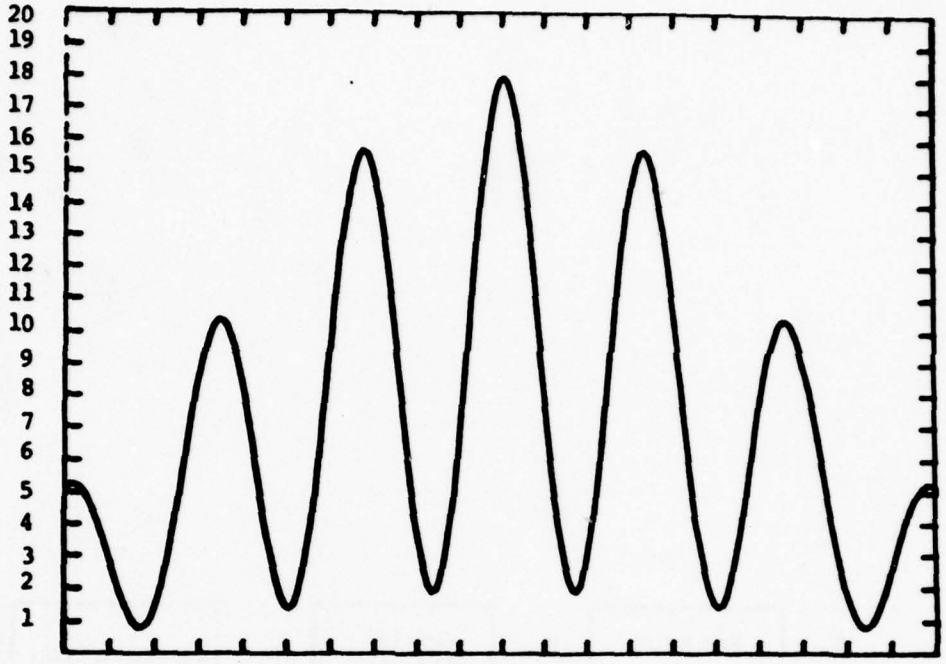
$$(\overline{SNR})_{rms} = \overline{(\lambda f_i)^{\frac{1}{2}}} = \overline{\lambda f_i^{\frac{1}{2}}} \qquad (2.26)$$

Thus, $(\overline{SNR})_{rms}$ is proportional to the square root of the signal, and is signal-dependent.

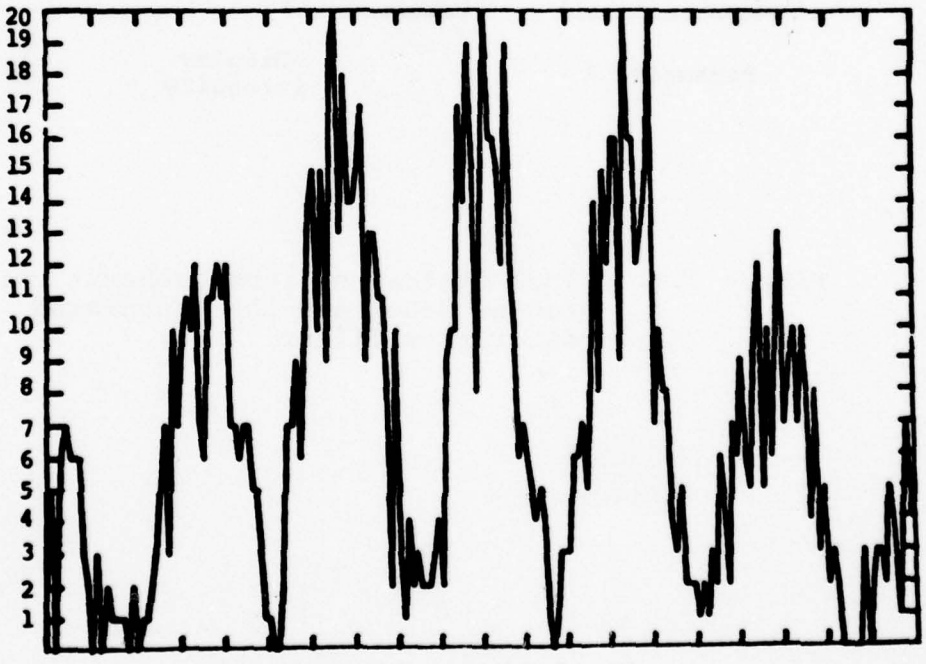
This behavior is quite different from additive noise, multiplicative noise, or film-grain noise. Figure 2.8 is a digital simulation of a low-level one-dimensional image signal whose peak SNR is approximately 18. It is clear from the illustration of Fig. 2.8, that at low signal levels the noise is statistically nonstationary and non-Gaussian. As the signal becomes photon resolved at very low intensities, there is little resemblance to the classical signal. However, at higher signal levels, the noise becomes more Gaussian [2-27].

2.4.3 Statistics of Array Counters

An array counter model for non-blurred image signals with Poisson noise is shown in Fig. 2.9. For one single counter the conditional density is given by Eq. (2.16). For an array of counters, we must find the joint ensemble statistics for a given object vector $\underline{f}$. Some assumptions are necessary to find these joint ensemble Poisson statistics. Walkup [2-20,2-21], Clark [2-22], and Wang [2-23] have shown that given $\underline{f}$, the joint ensemble photon



a) Original object signal



b) Detected image signal

Figure 2.8. Inhomogeneous Poisson model simulation of the image signal and noise

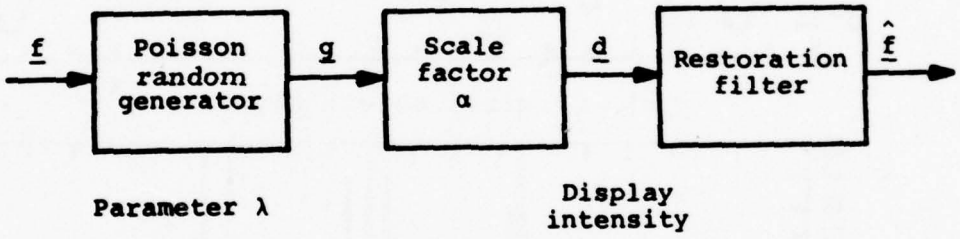


Figure 2.9. Block diagram of inhomogeneous Poisson process model and the associated restoration filter

counting statistics at the N detectors of Fig. 2.5 are independent Poisson variates whose mean is proportional to the mean intensities at each of the N disjoint detectors. For this to be true, the time-bandwidth product WT must satisfy WT>>1 where W is the temporal bandwidth of the optical image signal and T is integration time. This is the usual case in this thesis. From the WT >>1 assumption, Mandel [2-26] also has shown that the counts registered by the N detector/counters may be taken to be statistically independent, since classical fluctuation noise is negligible for light with a low degeneracy parameter, when Poisson shot noise predominates in the photon count fluctuations. The degeneracy parameter describes the variance characteristics of the photon counts, and it is defined as the ratio of classical fluctuation noise to Poisson shot noise. Thus, all g_i are independent for a given $\underline{f}$ (i.e. every Poisson generator is independent) and each g_i depends only on its corresponding f_i. Also, we assume that individual detectors have a smaller scale than the spatial intensity variations of the image so that no loss of information results from the sampling. We also initially assume that background intensities and thermal noise in the measurement system is negligible compared to the Poisson noise. Representing the array values by

$$\underline{g} = [g_1, g_2, \ldots, g_N]^T \quad \text{and} \quad \underline{f} = [f_1, f_2, \ldots, f_N]^T \qquad (2.27)$$

we have

$$p(\underline{g}|\underline{f}) = p(g_1|\underline{f})p(g_2|\underline{f})\ldots p(g_N|\underline{f}) \tag{2.28}$$

Now, each g_i depends only on its corresponding f_i, thus

$$p(\underline{g}|\underline{f}) = \prod_i \frac{(\lambda f_i)^{g_i} e^{-\lambda f_i}}{g_i!} \tag{2.29}$$

Now from Eqs. (2.22)-(2.24) we have

$$\begin{aligned} p(\underline{d}|\underline{f}) &= p(d_1|\underline{f})p(d_2|\underline{f})\ldots p(d_N|\underline{f}) \\ &= p(d_1|f_1)p(d_2|f_2)\ldots p(d_N|f_N) \end{aligned} \tag{2.30}$$

$$p(\underline{d}|\underline{f}) = \prod_i \frac{(\lambda f_i)^{\frac{d_i}{\alpha}} e^{-\lambda f_i}}{\alpha(\frac{d_i}{\alpha})!} \tag{2.31}$$

From this conditional density an MAP estimate is derived in Chapter 4.

2.4.4 The Quantum Limitations of Photon Resolved Image Signals

The information content of a finite amount of light is limited by the finite number of photons, by the random character of their distribution, and by the need to avoid false alarms. These limitations mean that a considerable number of photons is required to delineate the fine detail of images. Low light level image signals conspicuously suffer from these quantum limitations [2-24].

(i) Discreteness of Light Quanta

The number of picture elements N required for a well resolved image signal often lies in the range of 10^6-10^7 [2-24]. Hence, 10^8-10^9 photons are needed to delineate the location and brightness of pixels, assuming that 100 photons per pixel are arranged in a precise array. Unfortunately, nature does not work in so orderly a fashion, and photons arrive at random times and places and give rise to a fundamental graininess in any detected image. This grain noise obscures the detection of fine detail in low contrast images and is called Poisson noise at these low light levels.

(ii) Random Character of the Photon Distribution

Suppose that the average number of photons arriving in a given area is n_0, and that a number $n_1, n_2, \ldots, n_i$ are distributed around n_0 in such a fashion that the average value of $(n_i - n_0)$ is also n_0. Then the mean of the photon count is the same as the variance and the $(SNR)_{rms}$ is $n_0^{1/2}$.

If we need to detect a 1% contrast variation in a signal, we require that the noise level defined as the r.m.s. deviation of the mean number of photons also not exceed 1%. This can be achieved by having an average of 10^4 photons falling on each pixel of the image. The r.m.s. deviation (noise) is then the square root of 10^4, or 10^2,

and the ratio of this random deviation to the average will be 10^{-2}, or 1%. In general, if we want to detect an image signal with contrast C we must increase the number of photon counts proportional to $\frac{1}{C^2}$. In addition, we must guard against false alarms, that is, the mistaking of any particularly random fluctuation for the real signal to be detected.

(iii) False Alarms

False alarms and spurious visual patterns may arise from the random character of the photon distribution and not from the original scene itself. If we define the mean photon count as the signal, and the r.m.s. deviation as noise, then we can use detection theory to find the false alarm probability. Figure 2.10 shows the distribution of noise fluctuations around the mean value of a parameter. The ordinate is the probability density and the abscissa k is plotted in units of the r.m.s. deviation. The second abscissa scale n is a particular example for which the average number of photons is 900 and its r.m.s. deviation is 30. From Fig. 2.10, we can calculate the false alarm probability given a signal which is k units above its mean value. Table 2.1 gives the probability that noise fluctuations will exceed the mean value of the background by an integer number of units of the r.m.s. value of the noise.

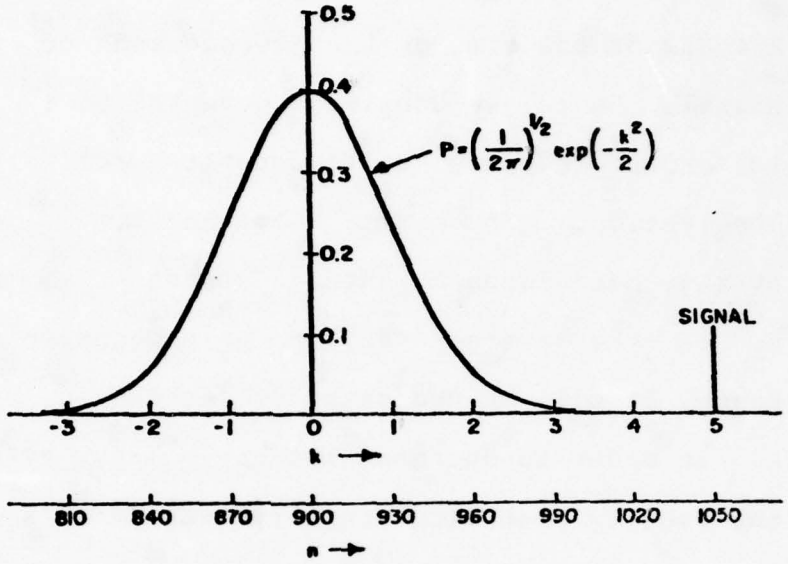


Figure 2.10. Probability distribution of a RMS deviation of photon counts about its mean

Table 2.1

False alarm probability of exceeding various values of k

k	prob. of exceeding k
1	0.15
2	0.023
3	1.3×10^{-3}
4	3×10^{-5}
5	3×10^{-7}
6	2×10^{-9}

If we locate the signal at k=5, then we find that only 15% of the time will the signal appear $k \leq 4$. It will exceed k=4 85% of the time on the average and be judged a real signal. Hence, we usually choose k=5 to avoid false alarms in order to give a reasonable reliability to our observation. The ratio of the r.m.s. deviation to the average background brightness varies as $n_0^{\frac{1}{2}}/n_0 = 1/n_0^{\frac{1}{2}}$, where n_0 is the average number of photons in the background. Hence, it will be necessary to increase n by a factor of k in order to decrease the ratio $1/n_0^{\frac{1}{2}}$ by k. In summary, the density of photons required varies as k^2 in order to avoid false alarms. In general, the expression for the total number of photons required to detect a contrast C where $C \triangleq \Delta B/B$ and $0 \leq C \leq 1$ is

$$\text{Total number of photons} = N \frac{1}{C^2} k^2 \tag{2.32}$$

Here, N is the total number of pixels in the picture and reflects the discrete nature of the photons. The factor $\frac{1}{C^2}$ is a consequence of contrast the C and the random character of photon distributions, the factor k^2 reflects both the random character of photon distribution and the need to avoid false alarms. The expression (2.32) is only for the case in which we do not have any <u>a priori</u> information about the image. However, we usually assume that the image is a Markovian random field with correlation coefficient ρ in image restoration work. Thus, we should be able to obtain

better results than obtained in Eq. (2.32)

2.4.5 Comparison Between Linear Additive Gaussian Noise Degraded Image Signals and Poisson Noise Degraded Image Signals

Figure 2.11 illustrates experimental results of pictures with Poisson noise and Gaussian noise for comparable $(SNR)_{rms}$.

The upper left image (A) in Fig. 2.11 is the original ideal object. The upper right image (B) contains Poisson noise with some constant amplification and $\overline{(SNR)_{rms}}$=6 db. The lower left image (C) has additive Gaussian noise with $(SNR)_{rms}$ also 6 db. The lower right one (D) also has linear additive Gaussian noise with $(SNR)_{rms}$ approximately 1 db. From this experimental result, we have demonstrated that images with Poisson noise are more severely degraded than images with linear additive Gaussian noise for comparable $(SNR)_{rms}$ even though the $(SNR)_{rms}$ is 6 db lower than that of an image with Poisson noise. The Poisson noise has obliterated almost completely the detailed edge and contrast of the face. The image takes on a mottled appearance which depends on brightness whereas the image degraded by Gaussian noise appears uniformly degraded with some edge contrast still discernible at comparable (SNR).

Figure 2.11 Poisson noise image vs linear additive Gaussian noise image

(A) Original object image
(B) Poisson noise image for $\overline{(SNR)}_{rms}$ =6 db
(C) Linear additive Gaussian noise image for $\overline{(SNR)}_{rms}$=6 db
(D) Linear additive Gaussian noise image for $\overline{(SNR)}_{rms}$=1 db

2.5 Conclusions

We have presented some system models for image noise and three image models. We have discussed the difference between the Poisson image noise model and other important image noise models from a system point of view. The model which we have developed in the section 2.4 is used in later chapters for developing and implementing an MAP filter for the Poisson noise model. Appendix A describes a computer algorithm which is used to generate Poisson noise for all simulations throughout this thesis. The quantum limitations of photon resolved image signals are presented in order to further understand the physical meaning and causes of Poisson noise in low light level image signals. By modeling the low level signals as an inhomogeneous Poisson process, a very accurate, complete model for many physical systems including medical imaging, astronomical imaging is developed.

CHAPTER 3

REVIEW OF LINEAR AND NONLINEAR OPTIMAL IMAGE RESTORATION FILTERS FOR THE IMAGE NOISE MODELS

3.1 Introduction

In this chapter, we review previous work on linear and nonlinear restoration filters with Poisson image noise models and other important image noise models. These image spatial restoration fillers are developed based upon the system models for image noise given in Chaper 2. The relationship of these fillers to the work presented in later chaptes is also discussed.

3.2 Spatial Restoration Filters for the Signal-Independent Noise Model and Signal-Dependent Noise Model

3.2.1 Spatial Restoration Filter for Signal-Independent Noise---Linear Additive Gaussian Noise Model

Most previous work in image restoration is based on the model of section 2.2.1. Different filters such as the inverse filter, constrained least-squares filter, parametric Wiener filter, homomorphic filter, maximum entropy filter, and pseudo-inverse filter have been developed under various criteria [3-1,3-2]. Hunt and

Trussell [3-3,3-4,3-5] have also developed a nonlinear MAP filter based on the direct maximization of the posterior density function with a linear additive image noise model.

3.2.2 Wiener Spatial Filters for Signal-Dependent Noise Model---Multiplicative Noise, Additive Signal-Modulated Noise

The models in sections 2.2.2 and 2.2.3 are signal-depdnent noise models. Walkup et al. [3-7], Kondo et al. [3-8] and Yatagai [3-9] developed an optimal spatial filter in the sense of minimizing the mean of the squared error $[f(x,y)-\hat{f}(x,y)]^2$ in the manner of a Wiener filter. From the minimum mean-square error criterion, the orthogonality principle is developed, leading to a spatial Wiener filter $W=R_{fg}[R_{gg}]^{-1}$, where R_{fg} and R_{gg} are the cross-covariance matrices between object and image and the covariance matrices of the image respectively [3-12]. If $g(x,y)$ and $f(x,y)$ are spatially wide-sense stationary random processes, then the two-dimensional Wiener filter has a spatial frequency domain transfer function

$$\omega(u,v) = \frac{\Phi_{fg}(u,v)}{\Phi_{gg}(u,v)} \tag{3.1}$$

where Φ_{fg} and Φ_{gg} represent the cross-spectral densities of the image $g(x,y)$ with the object $f(x,y)$ and the spectral density of the image $g(x,y)$, respectively.

Naderi and Sawchuk [3-10] also developed a better spatial adaptive Wiener filter for signal-dependent film-grain noise using a more accurate, complex noise model. Walkup et al. [3-11] developed an MAP spatial filter for signal-dependent noise models such as film-grain noise and magnetic recording noise. Their filter involved scalar rather than vector processing and did not include blurring degradations.

3.3 Linear, Invariant Least-Squares Restoration Filter for the Poisson Image Noise Model

Goodman and Belsher [3-13] have modeled a low light level imaging system with blurring and Poisson noise by the continuous system shown in Fig. 3.1. This system is a special case of the discrete restoration model discussed in section 2.4. The detected data is represented as

$$g(x,y) = \sum_{n=1}^{N} \delta(x-x_n, y-y_n) \qquad (3.2)$$

where $\delta(\cdot,\cdot)$ is a two-dimensional Dirac delta function

(x_n, y_n) represents the location of the nth photo event

N is the total number of photo events produced by the blurred image b(x,y)

In the expression (3.2), N, x_n and y_n are all regarded as random variables.

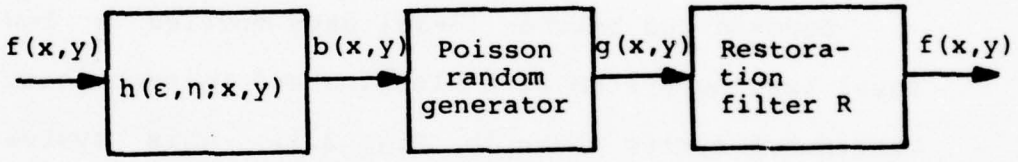


Figure 3.1. Block diagram of the continuous restoration filter with Poisson noise model

From the semi-classical theory of photo detection [3-14,3-21], the probability that N events occur in an area A on the detector is taken to be Poisson,

$$p_A(N) = \frac{\left[\iint_A \lambda(x,y)\,dxdy\right]^N}{N!} \exp\left\{-\iint_A \lambda(x,y)\,dxdy\right\} \qquad (3.3)$$

where $\lambda(x,y)$ is a rate function

$$\lambda(x,y) = \frac{\eta b(x,y)}{h\,\bar{\nu}}\,\tau \qquad (3.4)$$

η is the quantum efficiency

h is plank's constant = 6.624×10^{-34} W-sec/HZ

$\bar{\nu}$ is the mean optical frequency

τ is the detector integration time

and

$$b(x,y) = \int_{-\infty}^{\infty}\int h(\varepsilon,\eta;x,y)\,f(\varepsilon,\eta)\,d\varepsilon d\eta \quad . \qquad (3.5)$$

Because the image f(x,y) is a random process, the rate function $\lambda(x,y)$ is a random process. Thus the whole process is called a compound Poisson process or doubly Poisson process [3-14,3-15]. The event locations (x_n,y_n) are independent variables for different n's for $\lambda(x,y)$ given, and it has pdf.

$$p(x_n,y_n) = \frac{\lambda(x_n,y_n)}{\int_{-\infty}^{\infty}\int \lambda(x,y)\,dxdy} \quad . \qquad (3.6)$$

In Fig. 3.2, we illustrate a one-dimensional typical object intensity distribution and a corresponding typical detected image.

Based on the model of Eq. (3.2) and Fig. 3.1, Goodman and Belsher [3-13] developed an linear, space-invariant, least-square restoration filter (LMMSE filter). The filter is derived in the Fourier transform domain using the orthogonality principle. The form of the filter is

$$W(u,v) = \frac{\bar{N}\mathcal{H}^*(u,v)\Phi_f(u,v)}{1+\bar{N}|\mathcal{H}(u,v)|^2\Phi_f(u,v)} \tag{3.7}$$

where $\mathcal{H}$ is the two-dimensional Fourier transform of h(x,y)

Φ_f is the spectral density of the object f(x,y)

$\bar{N}$ is the mean number of photon counts

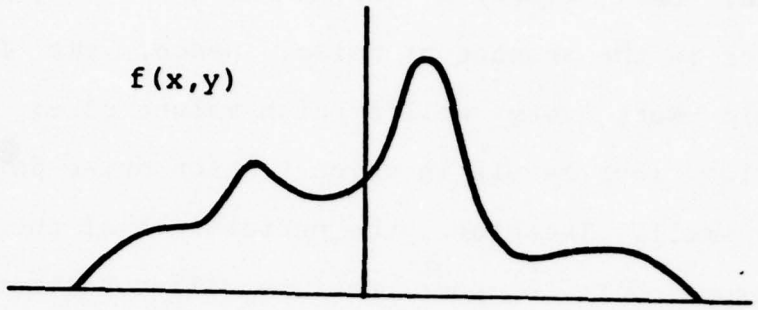
* denotes complex conjugate

(u,v) are spatial frequency variables

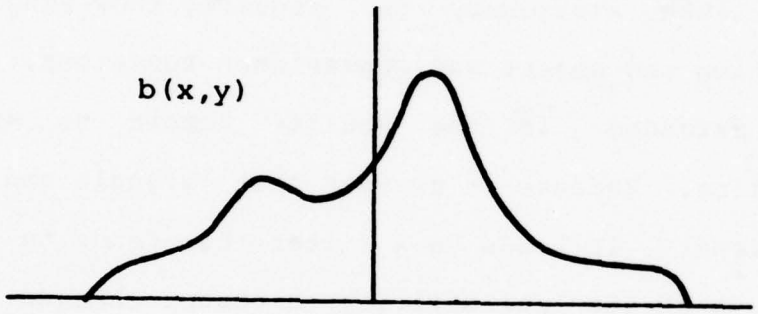
Equation (3.7) can be rewritten as

$$W(u,v) = \frac{\mathcal{H}^*(u,v)}{|\mathcal{H}(u,v)|^2+\frac{1}{\alpha}} \tag{3.8}$$

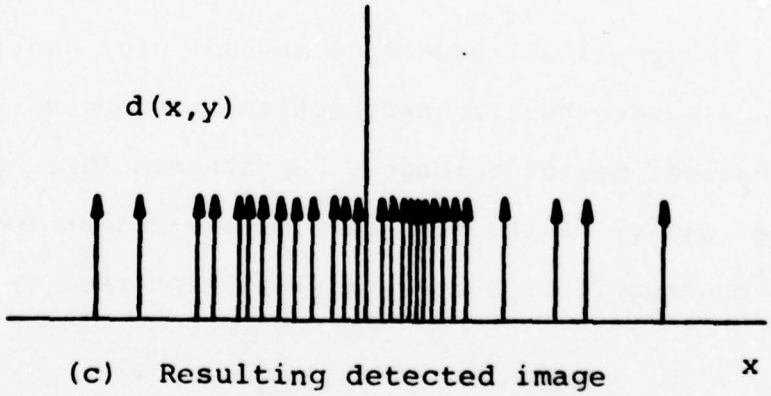
where $\alpha=\bar{N}\Phi_f(u,v)$. This filter can be implemented in the Fourier domain using the FFT. Goodman and Belsher did not apply this filter to two-dimensional picture data, so it is implemented in Chapter 7 and compared with the MAP filter described in Chapter 4 and 5. From Eq. (3.8), if α is very



(a) Object intensity



(b) Blurred object intensity



(c) Resulting detected image

Figure 3.2. Model of photon-resolved imagery
(a) object intensity
(b) blurred object intensity
(c) resulting detected image

large, then $W(u,v)=H^{-1}(u,v)$ which is the ideal inverse filter in the absence of noise. Hence, the LMMSE filter should work very well at high values of α. However, for the low light levels in which Poisson noise dominates, α is very small. Therefore, the performance of the LMMSE filter is expected to be poor due to ill-conditioning in the deconvolution process. Furthermore, the LMMSE filter assumes that the object signal and noise must be at least wide-sense stationary and requires knowledge of the blur function and object and convariance functions. It cannot be extended in the Fourier domain to space-variant blurring. Because it assumes that signal and noise are stationary, it leads to a filter that tends to smooth edges because of its insensitivity to abrupt changes.

3.4 Nonlinear Filtering with the Square-Root Transform for the Poisson Image Noise Model

Inouye [3-16] developed an ad hoc nonlinear filter with a square-root transformation to suppress Poisson noise in nuclear medicine images. He assumed that the observed data $g(x,y)$ is the summation of the object function $f(x,y)$ and quantum fluctuation $n(x,y)$ as expressed by

$$g(x,y) = f(x,y)+n(x,y) \tag{3.9}$$

where $n(x,y)$ depends on $f(x,y)$ according to

$$\overline{n(x,y)} = 0 \quad , \tag{3.10}$$

$$\overline{n^2(x,y)} = \alpha f(x,y) \quad , \tag{3.11}$$

where α is a constant of proportionality. Taking the square root of both sides of Eq. (3.9) we get

$$[g(x,y)]^{\frac{1}{2}} = [f(x,y)+n(x,y)]^{\frac{1}{2}},$$

$$[g(x,y)]^{\frac{1}{2}} = f^{\frac{1}{2}}(x,y)\left(1+\frac{n(x,y)}{f(x,y)}\right)^{\frac{1}{2}} . \tag{3.12}$$

Taking the first two terms of a Taylor series expansion of the above and assuming $f(x,y)>>n(x,y)$ yields an approximation

$$[g(x,y)]^{\frac{1}{2}} \stackrel{\cdot}{\approx} f^{\frac{1}{2}}(x,y)\left(1+\frac{n(x,y)}{2f(x,y)}\right) . \tag{3.13}$$

We now rewrite Eq. (3.13) as follows:

$$[g(x,y)]^{\frac{1}{2}} = \tilde{f}(x,y) + \tilde{n}(x,y), \tag{3.14}$$

where

$$\tilde{f}(x,y) = f^{\frac{1}{2}}(x,y),$$
$$\tilde{n}(x,y) = \frac{1}{2}n(x,y)/f^{\frac{1}{2}}(x,y).$$

From Eqs. (3.10) and (3.11), we have

$$\overline{\tilde{n}(x,y)} = 0, \tag{3.15}$$

$$\overline{[\tilde{n}(x,y)]^2} = \frac{1}{4}\,\alpha \qquad (3.16)$$

and the noise statistics of the square-root transform are stationary [3-17]. Thus this transform of g(x,y) gives a signal-independent fluctuating component $n(x,y)/2f^{\frac{1}{2}}(x,y)$ around an average value $(f(x,y))^{\frac{1}{2}}$. Equation (3.13) becomes an approximately linear additive noise model and the usual linear filtering technique is applied. Inouye did not include any burring effects of the imaging system and it is difficult to judge the level of improvement from the line printer pictures in his paper. In addition, this square-root transform works only at higher (SNR) image signals, because he assumes $f(x,y) >> n(x,y)$.

3.5 Nonlinear Optimal Filter for the Poisson Image Noise Model

From Bayes' law we have the posteriori conditional density

$$p(\underline{f}|\underline{d}) = \frac{p(\underline{d}|\underline{f})p(\underline{f})}{p(\underline{d})} \quad , \qquad (3.17)$$

where $\underline{f}$ is the original image (object) to be estimated and $\underline{d}$ is the observed data. The use of the posterior density for estimation is well known [3-18]. The minimum mean-square error estimate (MMSE) is the mean of the posterior density given by $E[\underline{f}|\underline{d}]$, the maximum *a posteriori*

estimate (MAP) is the mode of the posteriori density as expressed by $\max_{\underline{f}} p(\underline{f}|\underline{d})$. The maximum likelihood estimate (ML) given by $\max_{\underline{f}} p(\underline{d}|\underline{f})$ can be viewed as a special case of the MAP estimate when the a posteriori density is equal to a priori density (i.e. when $p(\underline{f})$ has a uniform distribution). The MMSE, MAP, and ML estimates are generally nonlinear, depending on the form of the probability density functions. The MMSE estimate also needs the density of the observed data $p(\underline{d})$, but this is often difficult to obtain in practice. Thus, a linear minimum mean-square error (LMMSE) estimate is commonly used as described in section 3.3. The MAP estimate tries to find the value of object $\underline{f}$ which maximizes the posterior density $p(\underline{f}|\underline{d})$. Thus, it does not need $p(\underline{d})$ at all but does need $p(\underline{d}|\underline{f})$ and $p(\underline{f})$, while the ML estimate only needs the a priori density $p(\underline{d}|\underline{f})$.

Burke [3-19] has developed a ML spatial filter based on the Poisson noise model in Fig. 3.1. Its ML estimate is obtained recursively according to the iteration

$$f_k^{(n+1)} = f_k^{(n)} \exp\left\{\beta \sum_j \left[\frac{g_j}{\beta \sum_\ell H_{j\ell} f_\ell^{(n)}} - 1\right] H_{jk}\right\} \quad (3.18)$$

where f_k is the kth component of the object to be estimated

g_j is the observed data $\underline{g}$ of the jth component

H_{ij} is ijth component of the blurring matrix H

β is the quantum efficiency constant,

ans the superscript denotes the iterative step.

Burke's experimental results are very impressive compared to the Wiener filter results. Unfortunately, only simple images of very small size (32x32 pixels) were processed and the recursive Eq. (3.18) converges very slowly. The ML estimate assumes that no information can be extracted from the a priori term $p(\underline{f})$. Modeling the object as a nonstationary random field with a probability density $p(\underline{f})$ to perform an MAP estimate can be thought of an extension of ML estimation.

3.6 Conclusions

It is well known that linear least squares image restoration is not optimal in the sense of maximum likelihood or maximum a posteriori probability when the image statistics are Poisson. Rather, nonlinear filtering is required for true optimality. In addition, Fourier techniques cannot treat space-variant imaging. Thus, it seems reasonable that a nonlinear filters should perform better than linear space-invariant filters in the presence of signal-dependent Poisson noise. In subsequent chapters, we try to formulate and implement MAP filters in order to perform nonlinear MAP filter for the low SNR's image signals with Poisson noise.

CHAPTER 4

RESTORATION OF IMAGE SIGNALS WITH POISSON NOISE

4.1 Introduction

In this chapter, we develop an optimal MAP filter for non-blurred image signals with Poisson noise based on the model developed in previous chapters. There are many practical situations in which the degradation due to blurring is negligible or zero. More importantly, we develop a framework for MAP estimation with a Poisson noise model to evaluate the concept for future application to more complex problems. In section 4.2 we present the formulation and solution to the MAP estimate equations. In section 4.3, 4.4, and 4.5 we address the implementation of an MAP estimate with different a priori knowledge. In section 4.6, we describe a recursive MAP filter for the Poisson noise model. In section 4.7, we discuss a local adaptive MAP filter and finally some conclusions are presented.

4.2 MAP Estimate Formulation [4-10]

Previously we derived a conditional density $p(\underline{d}|\underline{f})$ for displayed image data $\underline{d}$ given an ideal object $\underline{f}$ with a

Poisson noise model. Also from chapter 2, we have the pdf of the object expressed as Eq. (2.10).

From the MAP estimate definition, we have

$$\underset{\underline{f}}{\text{Max}}\ p(\underline{f}|\underline{d}) = \frac{p(\underline{d}|\underline{f})p(\underline{f})}{p(\underline{d})} \tag{4.1}$$

It is often easier to maximize a monotonic function of the posterior density, such as the logarithm. Taking the log of Eq. (4.1) we have

$$\underset{\underline{f}}{\text{Max}}\ [\ell np(\underline{f}|\underline{d}) = \ell np(\underline{d}|\underline{f}) + \ell np(\underline{f}) - \ell np(\underline{d})]. \tag{4.2}$$

Since the last term $\ell n\ p(\underline{d})$ on the right does not depend on $\underline{f}$, we neglect it in maximization with respect to $\underline{f}$. Thus the MAP estimate equation is given by

$$\frac{\partial \ell np(\underline{f}|\underline{d})}{\partial \underline{f}} = \frac{\partial \ell np(\underline{d}|\underline{f})}{\partial \underline{f}} + \frac{\partial \ell np(\underline{f})}{\partial \underline{f}} = \underline{0}^T \tag{4.3}$$

From Eq. (2.31), we have

$$\begin{aligned} \ell np(\underline{d}|\underline{f}) &= \sum_i \ell n\left[\frac{(\lambda f_i)^{d_i/\alpha} e^{-\lambda f_i}}{\alpha(\frac{d_i}{\alpha})!}\right] \\ &= \sum_i \left\{\frac{d_i}{\alpha}\ \ell n(\lambda f_i) - \lambda f_i - \ell n\alpha - \ell n[(\frac{d_i}{\alpha})!]\right\}. \end{aligned} \tag{4.4}$$

From Eq. (2.10), we have

$$\ell np(\underline{f}) = \ell n K_f - \frac{1}{2}(\underline{f}-\bar{\underline{f}})^T R_f^{-1}(\underline{f}-\bar{\underline{f}}). \tag{4.5}$$

Differentiating the above two equations individually with respect to $\underline{f}$ we get

$$\frac{\partial \ell np(\underline{d}|\underline{f})}{\partial \underline{f}} = \left[\frac{d_1}{\alpha f_1} - \lambda, \frac{d_2}{\alpha f_2} - \lambda, \ldots, \frac{d_N}{\alpha f_N} - \lambda\right], \quad (4.6)$$

and

$$\frac{\partial \ell np(\underline{f})}{\partial \underline{f}} = -\frac{1}{2} \cdot 2(\underline{f}-\bar{\underline{f}})^T R_f^{-1} = -(\underline{f}-\bar{\underline{f}})^T R_f^{-1}. \quad (4.7)$$

Substituting Eqs. (4.6) and (4.7) into Eq. (4.3), we get

$$\left[\frac{d_1}{\alpha f_1} - \lambda, \frac{d_2}{\alpha f_2} - \lambda, \ldots, \frac{d_N}{\alpha f_N} - \lambda\right] - (\underline{f}-\bar{\underline{f}})^T R_f^{-1} = \underline{0}^T. \quad (4.8)$$

Taking the transpose on both sides of the above equation and assuming that the covariance R_f^{-1} is a symmetric matrix, (i.e. $R_f^{-1}=(R_f^{-1})^T$) we then get

$$\begin{bmatrix} \frac{d_1}{\alpha f_1} - \lambda \\ \frac{d_2}{\alpha f_2} - \lambda \\ \vdots \\ \frac{d_N}{\alpha f_N} - \lambda \end{bmatrix} - R_f^{-1}(\underline{f}-\bar{\underline{f}}) = \underline{0} \quad (4.9)$$

From Eq. (4.9), we know if the norm of R_f denoted by $\|R_f\|$ is very large, then

$$\hat{\underline{f}}_{MAP} \doteq \underline{d} = \hat{\underline{f}}_{ML}$$

where $\underline{d}$ is the observation data vector and $\hat{\underline{f}}_{ML}$ is the

maximum likelihood estimate vector. In the blurred image case, the $\hat{\underline{f}}_{ML}$ is the inverse solution instead of the observation data.

On the other hand, if the norm of R_f is very small then

$$\hat{\underline{f}}_{MAP} \approx \underline{\bar{f}}$$

where $\underline{\bar{f}}$ is the a priori nonstationary mean of the image.

Therefore, solving these equations for $\hat{\underline{f}}_{MAP}$ tries to move the solution of $\underline{f}$ from the a priori nonstationary mean $\underline{\bar{f}}$ to a maximum likelihood estimate $\hat{\underline{f}}_{ML}$. Here R_f is a measure of our confidence in the nonstationary mean $\underline{\bar{f}}$ and maximum likelihood estimate $\hat{\underline{f}}_{ML}$ as a solution to the restoration problem. Equation (4.9) appears very simple, but the complexity of the estimate implementation depends heavily on the structure of the R_f covariance matrix. Thus we will discuss in the two following sections methods of implementing Eq. (4.9). One assumes R_f is an identity matrix, and the other assumes that R_f is a Markovian matrix.

4.3 MAP Estimate Implementatin with an A Priori Image Covariance Matrix - Identity Covariance Matrix

FoR simplicity, we assume $R_f=\sigma_f^2 I$, thus each pixel of the picture is uncorrelated. In real picture data, each pixel is highly correlated with its neighbors [4-11], and

this assumption is treated in the next section. From Eq. (4.9) and $R_f=\sigma_f^2 I$ we then have

$$\begin{bmatrix} \frac{d_1}{\alpha f_1} - \lambda \\ \frac{d_2}{\alpha f_2} - \lambda \\ \frac{d_N}{\alpha f_N} - \lambda \end{bmatrix} - \frac{1}{\sigma_f^2} \begin{bmatrix} f_1 - \bar{f}_1 \\ f_2 - \bar{f}_2 \\ \vdots \\ f_N - \bar{f}_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{4.10}$$

From Eq. (4.10), we see that the MAP estimate becomes a very simple point process instead of a vector process because equations are decoupled. Hence we can get a closed form solution

$$f_i = \frac{(\bar{f}_i - \lambda\sigma_f^2) \pm \sqrt{(f_i - \lambda\sigma_f^2)+4\lambda\sigma_f^2 d_i}}{2} \tag{4.11}$$

where the positive root is taken because intensity is always non-negative.

4.3.1 Implementation and Experimental Results

The observation data are photon counts with some amplification gain α. The photon count is simulated from an original picture (256x256) through a Poisson random noise generator wich is described in detail in Appendix A.

In order to implement Eq. (4.11), we must estimate the variance σ_f^2 and nonstationary mean $\bar{f}$ from the available data. Hunt [4-13] has shown that an estimate of the nonstationary ensemble mean of a context-dependent ensemble

can be made by blurring the given ensemble member with a point-spread function as expressed by

$$\langle f_K(x,y) \rangle = \bar{f}(x,y) = h(x,y) \circledast f(x,y) \qquad (4.12)$$

where $\langle f_K(x,y) \rangle$ denotes an ensemble average at (x,y) over all k ensemble images

$\bar{f}(x,y)$ is the nonstationary mean of the object

f(x,y) is the object intensity

$\circledast$ denotes two-dimensional convolution

h(x,y) is a point-spread function related to the probabilistic process which generates the ensemble.

Equation (4.12) states that an estimate of the ensemble mean can be made by convolving the original image with a point-spread function h(x,y). In the experimental work shown in this section, an estimate of the nonstationary mean along each line is made by blurring the noisy measurement data with a point-spread function h(x,y) chosen to be square blurring function. We have chosen an 11 pixel linear moving average window along each line in order to obtain nonstationary mean estimate. Because of this large moving window, the noise is averaged out while the low frequency components of the object remain. The image variance is estimated by an unbiased estimate of the population variance.

The restored pictures produced by the solution of Eq. (4.10) are shown in Fig. 4.1 for different mean rms signal to noise ratio denoted by $\overline{(SNR)_{rms}}$. From Fig. 4.1 it can be seen that the restored pictures are improved compared to the noisy pictures and that the nonstationary mean carries the important low frequency image information.

4.4 MAP Estimate Implementation With An A Priori Image Covariance Matrix - Markovian Covariance Matrix

In this section, we assume R_f is a Markovian covariance matrix with correlation coefficient ρ. The Markovian covariance matrix is very good approximation for real image signals. The Markovian covariance matrix for a one-dimensional image model is

$$R_f = \sigma_f^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ \rho & 1 & \rho & \cdots & \rho^{N-2} \\ \vdots & & & & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & \cdots & 1 \end{bmatrix} \qquad (4.13a)$$

where ρ is the correlation coefficient between pixels and $|\rho|<1$. It can be shown that the inverse of R_f is [4-12]

$$R_f^{-1} = r \begin{bmatrix} \frac{1}{1+\rho^2} & -\beta & & & 0 \\ -\beta & 1 & -\beta & & \\ & -\beta & 1 & -\beta & \\ & & -\beta & 1 & \\ 0 & & & -\beta & \frac{1}{1+\rho^2} \end{bmatrix} \qquad (4.13b)$$

where

Figure 4.1a Restored image with a MAP filter with $R_f = \sigma^2 I$ and $\overline{(SNR)}_{rms} = \sqrt{5}$

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 4.1b Restored image with a MAP filter with $R_f=\sigma^2 I$ and $\overline{(SNR)}_{rms}=\sqrt{10}$

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 4.1c Restored image with a MAP filter with $R_f=\sigma^2 I$ and $\overline{(SNR)}_{rms}=\sqrt{20}$

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

$$r = \frac{1+\rho^2}{1-\rho^2} \cdot \frac{1}{\sigma_f^2} \qquad \beta = \frac{\rho}{1+\rho^2}, \quad |\rho| < 1 \text{ and } |\beta| < \frac{1}{2}.$$

Substituting Eq. (4.14) into Eq. (4.9), we obtain a set of three types of nonlinear equations with N unknowns

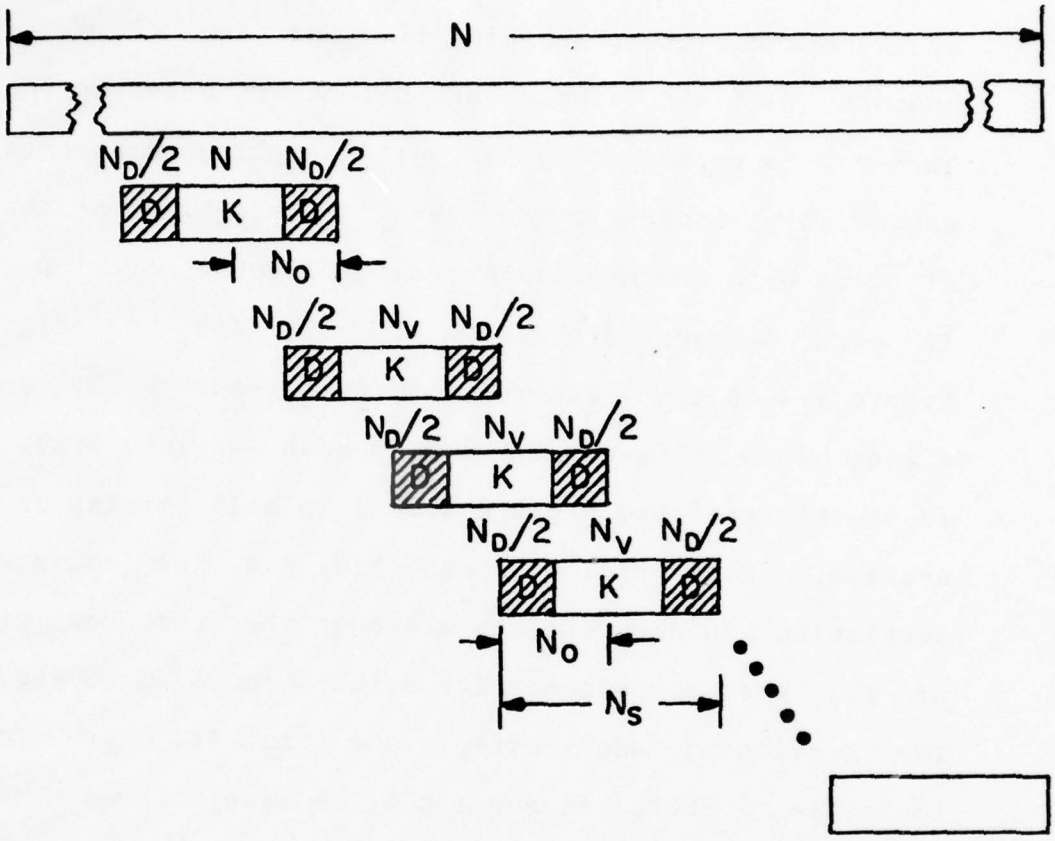
$$\begin{aligned}
&\left(\frac{d_1}{\alpha f_1}-\lambda\right)-\frac{r}{1+\rho^2}(f_1-\bar{f}_1)+\beta r(f_2-\bar{f}_2) = 0 \\
&\left(\frac{d_i}{\alpha f_i}-\lambda\right)+\beta r(f_{i-1}-\bar{f}_{i-1})-r(f_i-\bar{f}_i)+\beta r(f_{i+1}-\bar{f}_{i+1}) = 0 \qquad (4.14) \\
&\qquad i = 2,3,\ldots,N-1 \\
&\left(\frac{d_N}{\alpha f_N}-\lambda\right)+\beta r(f_{N-1}-\bar{f}_{N-1})-\frac{r}{1+\rho^2}(f_N-\bar{f}_N) = 0 \quad .
\end{aligned}$$

Due to the large dimensionality and nonlinear nature of the above system equations, ordinary linear signal processing matrix operations and Fourier methods are of no use. Thus, it is impossible to directly solve those equations in order to obtain optimal solutions. Instead a suboptimal method with sectioning to reduce the dimensionality of the equations is used. After trying several techniques for numerical solution of the nonlinear equations, we have found that a Newton-Raphson iterative method [4-4] is the best. This method is very easily implemented and converges rapidly with a good starting guess. In this application, convergence is generally reached in about three to four iterative steps. The details of applying this method to the MAP estimate will be discussed in Appendix B.

4.5 Implementation and Experimental Results

The estimate of image variance σ_f^2 and the nonstationary mean are done by the same technique as in the last section. The initial value of $\underline{f}$ in the Newton-Raphson method might be assumed to be the nonstationary mean $\hat{\underline{f}}$ or the raw observation data, but the final estimates must converge to the same values. The convergence criterion is based on the numerical closeness to the ideal image (object). The convergence rate is also controlled by the estimated variance $\hat{\sigma}_f^2$ of the image. An accurate estimate of the local variance is very important to the convergence of the algorithm. The iterative Newton-Raphson method employs the gradient of the function to obtain the increment value for iteration, thus it converges much more rapidly than other numerical methods.

The boundaries between sections must be carefully considered when the MAP equations are solved by Newton-Raphson techniques. The overlap sectioning method may minimize the boundary effects, but more computing time is required because the number of arithmetic operations in the Newton-Raphson solution goes up roughly as the cube of the section size. Thus, there is a compromise between section size and processing speed. The sectioning technique is best explained with the use of diagram shown in Fig. 4.2. In implementing the one-dimensional MAP



(1) N_S measured image data points are processed in each section with an overlap of N_0 data points between sections.

(2) Each section keeps N_V valid processed data points and discards $N_D/2$ erroneous processed data points at the two ends.

(3) N_S, N_0, N_V and N_D are the section size, overlap size, valid processed data size and discarded data size respectively.

Figure 4.2 One-dimensional sectioning method diagram

filter we have found that a section size of 32 pixels and overlap of 16 pixels in each section works well. This choice reduces the processing time and also minimizes the boundary effects. These assumptions are based on the fact that pixels separated by 16 unit sampling distances are essentially uncorrelated even if ρ=0.95. The restored pictures with the MAP filter using different overlap sizes in each section are shown in Fig. 4.3 and Fig. 4.4. Figure 4.3 is the restored picture produced by an MAP filter with 8 pixels overlap in each section. Figure 4.4 is the MAP restored picture with 16 pixels overlap in each section. From Fig. 4.3 and Fig. 4.4, we can see some sectioning boundary effects although the restored picture of Fig. 4.4 has reduced edge effects compared to Fig. 4.3. The sectioning edge effects of Fig. 4.4 are almost invisible. Thus, an overlap of 16 pixels in each section is a good practical choice for minimizing the sectioning boundary effects. Therefore, all the following one-dimensional processed pictures have used the overlap sectioning method with a section size of 32 pixels and an overlap of 16 pixels in each section.

We have also found that the cpu processing time of the MAP filter with ρ=0.95 is lesser than that of the MAP filter with ρ=0.0 because of the *a priori* knowledge of pixel correlation. The restored pictures with MAP filters are shown in Fig. 4.5 for different $\overline{(SNR)_{rms}}$ with ρ =0 and

Figure 4.3 Restored image with a MAP filter 8 pixels overlap in each section

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 4.4 Restored image with a MAP filter 16 pixels overlap in each section

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 4.5a $\overline{(SNR)}_{rms}=\sqrt{5}$ and $\rho=0$

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 4.5b Restored image with a MAP filter at $\overline{(SNR)}_{rms}=\sqrt{5}$ and $\rho=0.95$

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 4.5c Restored image with a MAP filter at $\overline{(SNR)}_{rms}=\sqrt{10}$ and $\rho=0$

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 4.5d Restored image with a MAP filter at $\overline{(SNR)}_{rms}=\sqrt{10}$ and $\rho=0.95$

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 4.5e Restored image with a MAP filter at $\overline{(SNR)}_{rms}=\sqrt{20}$ and $\rho=0$

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 4.5f Restored image with a MAP filter at $\overline{(SNR)}_{rms}=\sqrt{20}$ and $\rho=0.95$

(A) Original object image
(B) Poisson noisy image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

ρ=0.95. The restored images are optimal solutions between the maximum likelihood (ML) solution and the a priori nonstationary mean solution. The MAP solution smooths out the Poisson noise degradations and has also extracted some higher frequency information from the noisy images. The images restored with the MAP filter have more detail information particularly at higher $\overline{(SNR)}_{rms}$.

4.6 Recursive MAP Estimate for the Poisson Noise Model

From Eq. (4.14) and simple algebraic manipulations, we get

$$f_2 = \bar{f}_2 + \frac{1}{\beta(1+\rho^2)}(f_1 - \bar{f}_1) - \frac{\lambda}{\beta r}\left(\frac{d_1}{\bar{f}_1} - 1\right) \tag{4.15a}$$

$$f_{i+1} = \bar{f}_{i+1} + \frac{1}{\beta}(f_i - \bar{f}_i) - \frac{\lambda}{\beta r}\left(\frac{d_i}{\bar{f}_i} - 1\right) - (f_{i-1} - \bar{f}_{i-1}), \tag{4.15b}$$

$$i = 3,4,\ldots,N-1$$

and

$$f_N = \frac{B + \sqrt{B^2 - 4C}}{2} \tag{4.15c}$$

where

$$B = -\frac{\lambda(1+\rho^2)}{r} + \beta(1+\rho^2)(f_{N-1} - \bar{f}_{N-1}) + \bar{f}_N$$

$$C = \frac{(1+\rho^2)}{r} d_N.$$

If we can estimate f_1 and $\bar{f}$, then Eq. (4.15) is a

recursive relation

$$\hat{f}_{i+1} = \hat{f}_i + kg(\hat{f}_i, d_i, \hat{\underline{f}}) \tag{4.16}$$

where $\hat{f}_i$ is the previous estimate, d_i is the observation data, k is a residual gain, $\hat{\underline{f}}$ is the estimate of the nonstationary mean vector and g is a function of $\hat{f}_i$, d_i, $\hat{\underline{f}}$. Because images with Poisson noise have a very low $\overline{(SNR)}_{rms}$, the nonstationary mean can only be estimated approximately. Also, the estimation error is propagated through all the estimates $\hat{f}_i$. Thus, Eq. (4.16) is very unstable and it is impossible to obtain an accurate recursive solution. Some simulations have been performed with different estimates of $\hat{f}_i$ and $\hat{\underline{f}}$ for different $\overline{(SNR)}_{rms}$. All experimental results quickly blow up, obliterating all the image information.

4.7 A Local Adaptive Processing Filter

As discussed in Ch. 2, image signals are a nonstationary random field whose statistical properties vary in a local region of the image. Hence, a local adaptive processing filter should have many advantages compared to global processing filters which are defined over the entire image field. Global processing filters generally average over detail information in local regions of the images. The local adaptive filter should be particularly useful when the image noise is signal-dependent as with Poisson noise. The sectioned MAP

filter as described can be made to operate as a local adaptive processor. Because the MAP filter contains an ML (maximum likelihood) term and an *a priori* term (nonstationary mean), the filter can be implemented by adaptively weighting the terms as a function of local properties such as the first and second moments of the image or the nonlinearity of the human eye. The local adaptive MAP filter can be expressed as

$$Q_i * (\text{ML term}) + (1-Q_i) * (\text{a priori term}) = 0$$
$$i = 1,2,\ldots,L \tag{4.17}$$

where Q_i is the weight of the ith section and L is the total number of sections in the entire image. The local adaptive filter can also be extended to space-variant degradations. For simplicity in the experimental results that follow, we have simulated the global adaptive processing filter with $Q_i=Q_j=Q$ for all sections. The equations to be solved are:

$$Q*\left(\frac{d_1}{\bar{f}_1}-1\right)-(1-Q)*\left(\frac{r}{(1+\rho^2)\lambda}\right)(f_1-\bar{f}_1)+(1-Q)*\frac{\beta r}{\lambda}(f_2-\bar{f}_2) = 0 \tag{4.18a}$$

$$Q*\left(\frac{d_i}{\bar{f}_i}-1\right)-(1-Q)*\frac{\beta r}{\lambda}(f_{i-1}-\bar{f}_{i-1})-(1-Q)*\frac{r}{\lambda}(f_i-\bar{f}_i) +(1-Q)*\frac{\beta r}{\lambda}(f_{i+1},\bar{f}_{i+1}) = 0 \tag{4.18b}$$

$$Q*\left(\frac{d_N}{\bar{f}_N}-1\right)+(1-Q)*\frac{\beta r}{\lambda}(f_{i-1},\bar{f}_{i+1})-(1-Q)*\frac{r}{(1+\rho^2)} (f_N-\bar{f}_N) = 0 \tag{4.18c}$$

A Newton-Raphson iterative technique with sectioning was

used to solve Eq. (4.18). All the Newton-Raphson iterative techniques are described in Appendix B. The only factor that changes are some constant coefficients in the equations.

The resulting pictures of this global adaptive filter are shown in Figs. 4.6 through 4.8 for different $\overline{(SNR)}_{rms}$ and different weights Q=0.3, 0.5, and 0.8. The Q=0.5 weights both terms equally. The region of higher $\overline{(SNR)}_{rms}$ can be given a higher weight and then can extract higher frequency components from the ML term, while the region of lower $\overline{(SNR)}_{rms}$ can be weighted more towards the *a priori* term. These facts can be seen in Figs. 4.6 through 4.8 where we can see clearly that detail information is more visible and noise suppression is decreased as the ML term is increased. However, if Q is too large, then the resulting image is the same as the unprocessed data.

4.8 Conclusions

From this chapter, we have found that the estimated nonstationary mean carries most of the gross information in MAP estimation, and that the covariance matrix carries important detail information of an image. Also, the variance affects the convergence rate of the algorithm. The variance can act as a weighting factor in sectioned suboptimal MAP estimation because this method is very dependent on the local nonstationary variance. The

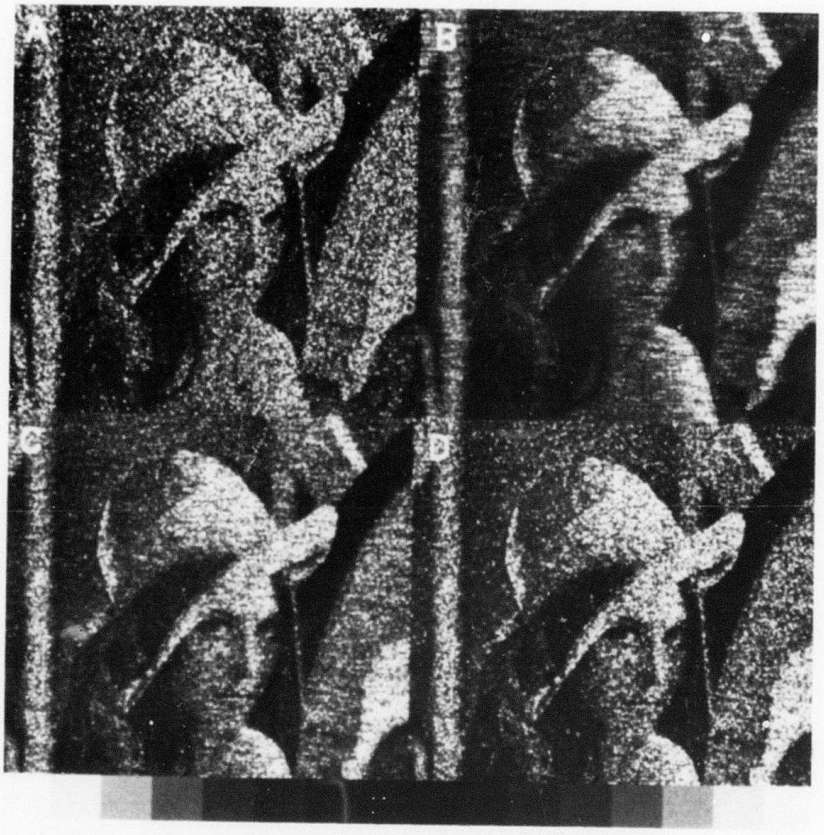


Figure 4.6 Restored images with an adaptive MAP filter at $\overline{(SNR)}_{rms}=\sqrt{5}$

(A) Poisson noisy image
(B) Restored image with $w_i=0.3$
(C) Restored image with $w_i=0.5$
(D) Restored image with $w_i=0.8$

Figure 4.7 Restored images with an adaptive MAP filter at $\overline{(SNR)}_{rms}=\sqrt{10}$

(A) Poisson noisy image
(B) Restored image with $w_i=0.3$
(C) Restored image with $w_i=0.5$
(D) Restored image with $w_i=0.8$

Figure 4.8 Restored images with an adaptive MAP filter at $\overline{(SNR)}_{rms}=\sqrt{20}$

(A) Poisson noisy image
(B) Restored image with $w_i=0.3$
(C) Restored image with $w_i=0.5$
(D) Restored image with $w_i=0.8$

sectioning method with a Newton-Raphson solution has been effective at coping with large dimension nonlinear MAP estimation equations and has produced good results.

CHAPTER 5

RESTORATION OF BLURRED IMAGE WITH POISSON NOISE

5.1 Introduction

This chapter we extend the previous results to the more general imaging model including blurring degradations and Poisson noise degradations as discussed in section 2.4. In many practical situations of interest, the detected image data arises from a linearly blurred image of the object. Although the blurring may arise from many different sources, we simply lump them together as a blurring matrix H. This system and its block diagram are shown in Fig. 2.5 and Fig. 2.6 respectively. In section 5.2, we formulate the MAP estimation equations and its solution. In section 5.3, we briefly review the sampled infinite area superposition operator model for image blur. In section 5.4, we discuss sectioning methods and in section 5.5, we discuss implementation of an MAP filter with one-dimensional blurring and its experimental results. In section 5.6, we discuss implementation of the MAP filter with two-dimensional blurring and present some experimental results. In the final section, we present some conclusions of this chapter.

5.2 MAP Derivation with A Blurring Matrix H

From Bayes' Law we have the posterior density as follows:

$$p(\underline{b}|\underline{d}) = \frac{p(\underline{d}|\underline{b})p(\underline{b})}{p(\underline{d})} \tag{5.1}$$

where $\underline{b}=H\underline{f}$ and $b_i = \sum_j H_{ij}f_j$.

From Fig. 2.6, we know that $\underline{g}$ is the Poisson noise degraded version of $(H\underline{f})$. As before we have the conditional density

$$p(g_i|b_i) = \frac{(\lambda b_i)^{g_i} e^{-\lambda b_i}}{g_i!} \tag{5.2}$$

for the measured counts g_i as a function of incident intensity b_i. We also assume that counters i and j are independent for a given $\underline{b}$. Hence

$$p(\underline{g}|\underline{b}) = p(g_1|\underline{b})p(g_2|\underline{b})\ldots p(g_N|\underline{b}) \tag{5.3}$$

Because each g_i only depends on its corresponding b_i,

$$\begin{aligned} p(\underline{g}|\underline{b}) &= p(g_1|b_1)p(g_2|b_2)\ldots p(g_N|b_N) \\ &= \prod_i \frac{(\lambda b_i)^{g_i} e^{-\lambda b_i}}{g_i!} \end{aligned} \tag{5.4}$$

and setting $d_i=\alpha g_i$ we have

$$p(\underline{d}|\underline{b}) = \prod_i \frac{(\lambda b_i)^{d_i/\alpha} e^{-\lambda b_i}}{\alpha(\frac{d_i}{\alpha})!} \tag{5.5}$$

5.2.1 The Probability Density of the Blurred Image

The *a priori* density $p(\underline{f})$ is multivariate normal with nonstationary mean $\overline{\underline{f}}$ and stationary covariance matrix R_f given by [5-1]

$$p(\underline{f}) = k_f \exp\{-\tfrac{1}{2}(\underline{f}-\overline{\underline{f}})^T R_f^{-1}(\underline{f}-\overline{\underline{f}})\} \quad . \tag{5.6}$$

The blurred data $\underline{b}$ is given by the linear equation,

$$\underline{b} = H\,\underline{f} \tag{5.7}$$

and the probability density for $\underline{b}$ can be shown to be

$$p(\underline{b}) = K_b \exp\{-\tfrac{1}{2}(\underline{b}-\overline{\underline{b}})^T R_b^{-1}(\underline{b}-\overline{\underline{b}})\} \tag{5.8}$$

where

$$R_b = H R_f H^T, \tag{5.9}$$

$$\overline{\underline{b}} = H\,\overline{\underline{f}} \tag{5.10}$$

Substituting Eqs. (5.9) and (5.10) into Eq. (5.8), we get

$$p(H\underline{f}) = K_b \exp\{-\tfrac{1}{2}(\underline{f}-\overline{\underline{f}})^T R_f^{-1}(\underline{f}-\overline{\underline{f}})\} \quad . \tag{5.11}$$

5.2.2 Estimation Equations

From chapter 4, we have the estimation equations as follows:

$$\frac{\partial}{\partial \underline{f}} \ell np(\underline{d}|\underline{b}) + \frac{\partial}{\partial \underline{f}} \ell np(\underline{b}) = \underline{0}^T, \tag{5.12}$$

where

$$\underline{b} = H\underline{f} \quad \text{and} \quad b_i = \sum_j H_{ij} f_j \tag{5.13}$$

From Eq. (5.5), we have

$$\ell np(\underline{d}|\underline{b}) = \sum_i \ell n \frac{(\lambda b_i)^{d_i/\alpha} e^{-\lambda b_i}}{\alpha(\frac{d_i}{\alpha})!} \tag{5.14}$$

$$= \sum_i \{\frac{d_i}{\alpha} \ell n(\lambda b_i) - \lambda b_i - \ell n\alpha - \ell n[(\frac{d_i}{\alpha})!]\} \tag{5.15}$$

$$\frac{\partial}{\partial f_K} \ell np(\underline{d}|\underline{b}) = \sum_i (\frac{d_i}{\alpha} \frac{H_{iK}}{b_i} - \lambda H_{iK}) \tag{5.16}$$

$$= \sum_i \lambda (\frac{d_i}{b_i} - H_{iK}) \tag{5.17}$$

Thus

$$\frac{\partial}{\partial \underline{f}} \ell n p(\underline{d}|H\underline{f}) = \left[\sum_i \lambda \left(\frac{d_i}{b_i}-1\right) H_{i1}, \sum_i \lambda \left(\frac{d_i}{b_i}-1\right) H_{i2}, \ldots, \sum_i \lambda \left(\frac{d_i}{b_i}-1\right) H_{iN} \right] \tag{5.18}$$

and from before

$$\frac{\partial}{\partial \underline{f}} p(H\underline{f}) = -(\underline{f}-\bar{\underline{f}})^T R_f^{-1} \tag{5.19}$$

Substituting Eqs. (5.18) and (5.19) into Eq. (5.13), we get

$$\left[\sum_i \lambda \left(\frac{d_i}{b_i}-1\right) H_{i1}, \sum_i \lambda \left(\frac{d_i}{b_i}-1\right) H_{i2}, \ldots, \sum_i \lambda \left(\frac{d_i}{b_i}-1\right) H_{iN} \right] + [-(\underline{f}-\bar{\underline{f}})^T R_f^{-1}] = \underline{O}^T \tag{5.20}$$

Taking the transpose of Eq. (5.20) and assuming $R_f = R_f^T$ then, we get

$$\begin{bmatrix} \sum_i \lambda \left(\frac{d_i}{b_i}-1\right) H_{i1} \\ \sum_i \lambda \left(\frac{d_i}{b_i}-1\right) H_{i2} \\ \vdots \\ \sum_i \lambda \left(\frac{d_i}{b_i}-1\right) H_{iN} \end{bmatrix} - R_f^{-1}(\underline{f}-\bar{\underline{f}}) = \underline{O} \tag{5.21}$$

We expand the first term of the left side of Eq. (5.21) to obtain

$$\lambda \begin{bmatrix} (\frac{d_1}{b_1}-1)H_{11}+(\frac{d_2}{b_2}-1)H_{21}+\ldots+(\frac{d_N}{b_N}-1)H_{N1} \\ (\frac{d_1}{b_1}-1)H_{12}+(\frac{d_2}{b_2}-1)H_{22}+\ldots+(\frac{d_N}{b_N}-1)H_{N2} \\ \vdots \\ (\frac{d_1}{b_1}-1)H_{1N}+(\frac{d_2}{b_2}-1)H_{2N}+\ldots+(\frac{d_N}{b_N}-1)H_{NN} \end{bmatrix} \tag{5.22}$$

$$= \lambda \begin{bmatrix} H_{11} & H_{21} & H_{31} \cdots H_{N1} \\ H_{12} & H_{22} & H_{32} \cdots H_{N2} \\ \vdots & & \vdots \\ H_{1N} & H_{N2} & H_{N3} \cdots H_{NN} \end{bmatrix} \begin{bmatrix} \frac{d_1}{b_1}-1 \\ \frac{d_2}{b_2}-1 \\ \vdots \\ \frac{d_N}{b_N}-1 \end{bmatrix} \tag{5.23}$$

$$= \lambda \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1N} \\ H_{21} & H_{22} & \cdots & H_{2N} \\ \vdots & & & \vdots \\ H_{N1} & H_{N2} & \cdots & H_{NN} \end{bmatrix}^T \begin{bmatrix} \frac{d_1}{b_1}-1 \\ \frac{d_2}{b_2}-1 \\ \vdots \\ \frac{d_N}{b_N}-1 \end{bmatrix} \tag{5.24}$$

$$= \lambda H^T(\underline{q}-1) \tag{5.25}$$

Here the H matrix is not necessarily a square matrix but depends on the model of the blurring degradation where

$$\underline{q} = [q_1, q_2, \ldots, q_N]^T$$

and

$$q_i = \frac{d_i}{b_i} = \frac{d_i}{\sum_j H_{ij} f_j}$$

For the time-being, we assume that H is a square matrix

$$H \triangleq \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1N} \\ H_{21} & H_{22} & \cdots & H_{2N} \\ \vdots & & & \vdots \\ H_{N1} & H_{N2} & \cdots & H_{NN} \end{bmatrix} . \qquad (5.26)$$

Therefore, the MAP estimate equations reduce to

$$\lambda H^T(\underline{q}-\underline{1}) - R_f^{-1}(\underline{f}-\bar{\underline{f}}) = \underline{0} \qquad (5.27)$$

This equation is the key equation of MAP estimation with a blurring matrix.

The complexity of solving this equation is determined by the structure of the blurring matrix H as well as the covariance matrix of image denoted by R_f. If the matrix R_f is a separable matrix it can be expressed as $R_f = R_{fR} \circledast R_{fc}$ where $\circledast$ denotes the direct product [5-14] and R_{fR} and R_{fc} are NxN matrices of the form

$$R_{fR} = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ \rho & 1 & \rho & \cdots & \rho^{N-2} \\ \vdots & & & & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & \cdots & 1 \end{bmatrix} . \qquad (5.28)$$

The Markovian covariance matrix R_f is an accurate approximation to the statistics of many images [5-2]. The

H matrix is an N^2xN^2 blurring matrix which may be nonseparable or separable. If H is a separable matrix, then $H=H_R \otimes H_C$ where H_R or H_C is an NxN matrix. The detail of modeling continuous superposition integrals for blurring by a discrete operator H is discussed in the next section.

Solving the MAP estimation Eq. (5.27) is heavily dependent on the separability of the H matrix and R_f matrix. If we assume that both are separable, then the MAP estimate Eq. (5.27) can be obtained by separate row and column processing. Thus, the computation time of a separable two-dimensional MAP filter is twice that of a one-dimensional MAP filter. However, the computation time of a nonseparable two-dimensional MAP filter is approximately the square of the time for a one-dimensional MAP filter. Thus, computing $H(\underline{g}-\underline{l})$ takes approximately $2N^4$ operations (each operation includes one multiplication and one addition) for the nonseparable case and $4N^3$ operations for the separable case where N is the picture size, e.g. N=256. The tremendous amount of computing needed for the MAP estimate makes solution impossible even with the separable case. Thus, we adopt a suboptimal solution involving sectioning with a Newton-Raphson solution technique. This method has been developed in Chapter 4 for solving nonlinear MAP estimate equations of larger dimensionality. Using the sectioning method [5-4], the MAP estimate Eq. (5.27) becomes

$$\lambda H^{T(m)} (\underline{g}-\underline{l})^{(m)} - R_f^{-1(m)} (\underline{f}-\overline{\underline{f}})^{(m)} = \underline{0} \quad (5.29)$$

where the superscript m denotes the mth section of the image. If each section size is N_s^2, then $\underline{g}^{(m)}$, $\underline{f}^{(m)}$, and $\overline{\underline{f}}^{(m)}$ are N_s^2 x1 vectors and $H^{(m)}$ is an $N_s^2 \times N_s^2$ matrix. For clarity, we omit the superscript (m) in the discussion which follows. The H matrix is constructed from the known point spread function (PSF) h(x,y). The PSF h(x,y) is obtained either from a priori knowledge or a posteriori knowledge. The former assumes that the PSF h(x,y) is known a priori. The latter assumes that the PSF h(x,y) is not available a priori. It must be estimated from noisy observation data using techniques such as blind deconvolution. References [5-5,5-6,5-7,5-8] describe several methods for estimating the amplitude response of h(x,y), and the Knox-Thompson method can be used to estimate the phase response of h(x,y) [5-9,5-15]. These methods determine the degradation parameters by a posteriori methods [5-3]. This thesis assumes that PSF h(x,y) of the system is given as a priori information. We must know the detailed structure of the H blurring matrix in order to implement the MAP estimate Eq. (5.29). Hence, we now discuss how to construct the H matrix from a given PSF h(x,y) of the system.

5.3 Construction of the Blurring Matrix H

All blurring degradation effects are lumped together as a two-dimensional *point spread function* $h(x,y;\xi,\eta)$ as shown in Fig. 5.1. From linear system theory, we can describe Fig. 5.1 by the superposition integral

$$G(x,y) = \int_{-\infty}^{\infty}\int F(\alpha,\beta)h(x,y;\alpha,\beta)d\alpha d\beta \qquad (5.30)$$

If this linear system is spatially invariant, then

$$h(x,y;\alpha,\beta) = h(x-\alpha;y-\beta). \qquad (5.31)$$

In order to discretize the continuous Eq. (5.30), we must sample G(x,y) over a grid at spacing $(\Delta x,\Delta y)$ satisfying the Nyquist criterion [5-2]. For notational simplicity, the continuous object function $F(\alpha,\beta)$ and the continuous PSF $h(x,y;\xi,\eta)$ will also be assumed to be sampled over the same grid spacing. Thus, Eq. (5.30) can be expressed as a double summation over infinite limits by invoking the sampling theorem and a quadrature formula. In order to express the infinite area superpositon operator as a lexicographic ordering for vector processing, it is necessary to truncate the PSF to some spatial limit, say $(L\Delta x,L\Delta y)$ and restrict G(x,y) to an area $(M\Delta x,M\Delta y)$. Then the truncated superposition operator is

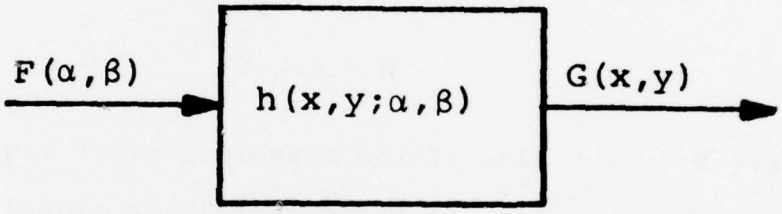


Figure 5.1 Two-dimensional linear system

$$G(m_1,m_2) = \sum_{n_1=m_1}^{L+m_1-1} \sum_{n_2=m_2}^{L+m_2-1} F(n_1,n_2)\,p(m_1-n_1+L,\ m_2-n_2+L;m_1,m_2) \tag{5.32}$$

where for simplicity, the grid spacing $\Delta x,\Delta y$ is dropped and p, assumed to be zero outside its range of indices, represents the sampled point-spread function which incorporates the quadrature integration. The detailed derivation of Eq. (5.32) is in [5-2]. In order to prevent serious approximate error at the boundaries of G, we must choose N such that

$$N \geq M+L-1 \tag{5.33}$$

where N is the size of the convolved image signal G

M is the size of the object image signal F

L is the size of PSf P

These boundary problems are important and closely related to the sectioning method used. In the next section we discuss the use of an overlap-save sectioning method to minimize the boundary edge effects.

If the arrays F and G are represented as lexicographic ordered vectors by $\underline{f}$ and $\underline{g}$ respectively, then the superposition operator can be written as

$$\underline{g} = H\,\underline{f} \tag{5.34}$$

where $\underline{f}$ and $\underline{g}$ are M^2x1 and N^2x1 vectors respectively. H is the M^2xN^2 matrix

$$H = \begin{bmatrix} \underline{H}_{11} & \underline{H}_{12} & \cdots & \underline{H}_{1L} & & 0 \\ & \underline{H}_{22} & & \underline{H}_{2,L} & \underline{H}_{2,L+1} & \\ 0 & & \underline{H}_{M,N-L+1} & \cdots & \underline{H}_{M,N} & \end{bmatrix} \tag{5.35}$$

where $\underline{H}_{ij}$ are MxN matrices with entries

$$H_{m_2,n_2}(m_1,n_1) = p(m_1-n_1+L, m_2-n_2+L) \tag{5.36}$$

for

$$1 \leq m_1 \leq M, \qquad 1 \leq m_2 \leq M$$

$$m_1 \leq n_1 \leq L+m_1-1, \quad m_2 \leq n_2 \leq L+m_2-1 \tag{5.37}$$

and P is a discretized truncated point spread function h(x,y). If the PSF is spatially invariant (SIPSF), then

$$\underline{H}_{m_2,n_2} = \underline{H}_{m_2+1,n_2+1} \cdot \tag{5.38}$$

When the PSF is spatially invariant and orthogonally separable

$$H = H_C \circledast H_R \tag{5.39}$$

and the two-dimensional convolution operation becomes

$$G = H_C F H_R^T \tag{5.40}$$

where $\circledast$ denotes direct product, G and F are the image and object arrays, respectively, and H_R and H_C are MxN matrices of the form

$$H_R = \begin{bmatrix} P_R(L) & P_R(L-1) & \cdots & P_R(1) & 0 \\ & \ddots & & \ddots & \\ 0 & & P_R(L) & \cdots & P_R(1) \end{bmatrix} . \quad (5.41)$$

Another discrete operator models the blurring degradation by a finite area superposition operator expressed as a D matrix. The form of this model is

$$Q(m_1,m_2) = \sum_{n_1=1}^{m_1} \sum_{n_2=1}^{m_2} F(n_1,n_2)p(m_1-n_1+1,m_2-n_2+1) \quad (5.42)$$

where M=N+L-1. Hence the processed array Q is of larger dimension than object data array F. Its vector form is

$$\underline{q} = D\,\underline{f} \quad (5.43)$$

where D is an M^2xN^2 matrix of the form

$$D = \begin{bmatrix} \underline{D}_{1,1} & & & 0 \\ \underline{D}_{2,1} & \underline{D}_{2,2} & \ddots & \\ \vdots & \vdots & & \\ \underline{D}_{L,1} & \underline{D}_{L-1,2} & \cdots & \underline{D}_{M-L+1,N} \\ & \underline{D}_{L,2} & & \vdots \\ 0 & & & \underline{D}_{M,N} \end{bmatrix} \quad (5.44)$$

where $\underline{D}_{ij}$ is an MxN matrix.

All the special cases of the D matrix are the same as those of the H matrix except for the matrix structure which has the form of Eq. (5.44). The difference between the H operator matrix and the D matrix for modeling a continuous superpositon are that the processed array for the finite area H computation is equivalent to the processed array for

the finite area H computation, surrounded by a boundary of L-1 superfluous elements. Conversely, if the processed array size is the same for two superposition operators, the L-1 boundary elements for the array obtained by the D operator will be in error. The resulting implementation of the MAP filter using the D operator for the blurring matrix is shown in Fig. 5.2. Figure 5.3 shows the results for the H operator. We can see the checkerboard noise appearing in the restored image of Fig. 5.2. The checkerboard noise results from using the D operator giving rise to the erroneous data surrounding each section. Thus, the two-dimensional MAP filter is implemented with an H operator matrix for the nonseparable blurring cases rather than the D operator. We can either add zero elements to the rectangular H matrix or truncate the boundary elements which wrap around, making it a square matrix as discussed in previous sections. However, the discrete blurring matrix H should be a rectangular matrix physically. Since m finite data points are convolved with n finite data points, the resulting processed data has m+n-1 points.

5.4 Sectioning Method

Due to the large dimensionality and nonlinearity of the MAP estimate equations, a sectioning method [5-4,5-10] is used with the Newton-Raphson solution to obtain a suboptimal solution. There are two sectioning methods:

Figure 5.2 Restored image with the MAP filter with a D operator for modeling the blurring degradation

(A) Original object image
(B) Degraded image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 5.3 Restored image with the MAP filter with an H operator for modeling the blurring degradation

(A) Original object image
(B) Degraded image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

overlap-add sectioning and overlap-save sectioning. The sectioning method chosen will give rise to various boundary edge effects. Hence, we will investigate which method is the best for MAP estimation. From Eq. (5.29), we have

$$H^T(\underline{g}-\underline{l})-R_f^{-1}(\underline{f}-\bar{\underline{f}}) = \underline{0} \tag{5.29}$$

where

$$\underline{g}_i = \frac{d_i}{\sum_K H_{iK} f_K}$$

In the overlap-add method of sectioning, filtered outputs of the mth and m+1th section are added together in the region of overlap to create the final output. This method is suitable only for linear estimation.

In our case however, the convolution with H is imbedded inside the function $\underline{g}$ of Eq. (5.29). Since $\underline{g}$ is a nonlinear function of $(H\underline{f})$, we have

$$\frac{d_i^{(m)}+d_i^{(m+1)}}{\sum_K H_{iK}(f_k^{(m)}+f_k^{(m+1)})} \neq \frac{d_i^{(m)}}{\sum_K H_{iK} f_k^{(m)}} + \frac{d_i^{(m+1)}}{\sum_K H_{iK} f_K^{(m+1)}} \tag{5.45}$$

which reduces to

$$\underline{g}(H\underline{f}^{(m)}+H\underline{f}^{(m+1)}) \neq \underline{g}(H\underline{f}^{(m)})+\underline{g}(H\underline{f}^{(m+1)}) \tag{5.46}$$

Thus, if $\underline{f}^{(m)}$ belongs to the overlapped portion of a section m, and $\underline{f}^{(m+1)}$ belongs to the overlapped portion of

an adjacent section m+1, then the *overlap-add sectioning* method is not valid because of the nonlinear function in the MAP estimate equation. Fortunately, the overlap-save method remains valid for the nonlinear case and can be used in our MAP estimate equation because incorrect boundary points in the overlap region are discarded, rather than being corrected by addition. Thus the overlap-save method can reduce the boundary edge effects because it discards the erroneous processed data of the overlapped region. Since Eq. (5.29) contains two H operators and assumes that the truncated point spread function matrix is LxL, the amount of overlap required is 2(L-1)x2(L-1). If we must correct NxN points of processed data at each section, then we must use a working section of [N+2(L-1)]x[N+2(L-1)]. The necessary overlap area constitutes the major overhead in the sectioned filtering process. It is clear that smaller sections have a larger percentage of overhead computation. It is also clear that the computation will be more inefficient for larger point spread functions. However, it should be kept in mind that the number of arithmetic operations in the Newton-Raphson solution for each section is the key computing burden of the sectioning filter. In order to find the updated incremental estimate vector in each iterative step of the Newton-Raphson method, we must solve a set of linear system equations. The dimension of the linear system equations is N_s^2 where N_s is

the number of pixels in each section and the structure of the gradient matrix depends on the structure of the blurring matrix H when the covariance matrix R_f is assumed first order Markovian. Hence, the number of arithmetic operations for solving linear system equations is heavily dependent on the form of the blurring matrix H. In general, the number of arithmetic operations in solving a set of linear system equations goes up roughly as the cube of the order of the system when the gradient matrix is a general square matrix. Thus, the smaller the section size, the less the computing time of sectioned filtering with the Newton-Raphson solution.

5.5 Implementation of the MAP Filter with One-dimensional Blurring Degradation

As stated earlier, two of the most interesting sources of blur are atmospheric turbulence and linear motion [5-11,5-12]. In this section, we assume that the image is degraded by one dimensional linear motion blur with a rectangular point-spread function. The rectangular blurring degradation is troublesome because its amplitude response has a singularity and phase reversals. The blurring matrix H is the form of Eq. (5.41), where $p_R(L)$ is the discrete truncated point spread function of $h(x,y)$. We assume that R_f is a first order Markovian covariance matrix, and following Eq. (5.29), we can write the

equations to be solved as

$$\lambda\left[\sum_{j=1}^{N} H_{j1}\left(\frac{d_j}{\sum_K H_{1k}f_k} - 1\right)\right] - \frac{r}{1+\rho^2}(f_1-\bar{f}_1)+\beta r(f_2-\bar{f}_2) = 0 \quad (5.47a)$$

$$\lambda\left[\sum_{j=1}^{N} H_{ji}\left(\frac{d_j}{\sum_K H_{iK}f_K} - 1\right)\right] +\beta r(f_{i-1}-\bar{f}_{i-1}) - r(f_i-\bar{f}_i) + \beta r(f_{i+1}-\bar{f}_{i+1}) = 0 \quad (5.47b)$$

$$i = 2,3,\ldots,N-1$$

$$\lambda\left[\sum_{j=1}^{N} H_{jN}\left(\frac{d_j}{\sum_K H_{NK}f_K} - 1\right)\right] +\beta r(f_{N-1}-\bar{f}_{N-1}) - \frac{r}{1+\rho^2}(f_N-\bar{f}_N) = 0 \quad (5.47c)$$

where

$$r = \frac{1+\rho^2}{1-\rho^2} \cdot \frac{1}{\sigma_f^2}$$

$$C_2 = \frac{r}{(1+\rho^2)} = \frac{1}{(1-\rho^2)} \cdot \frac{1}{\sigma_f^2}$$

$$C_1 = \frac{\beta r}{\lambda} = \rho C_2$$

N is the number of pixels of each section,

ρ is the correlation coefficiency between pixel,

σ_f^2 is the variance of the object,

H_{ij} is the ijth element of the H matrix,

$\bar{f}_i$ is the nonstationary mean of the ith pixel of the section,

d_j is the observation measurement.

Using the overlap-save sectioning method with an

iterative Newton-Raphson solution, the MAP solution to Eq. (5.47) is obtained. The convergence is very rapid, generally requiring about two to three iterative steps as described in detail in Appendix B. The discrete point spread function of $h(x,y)$ is a rectangular blurring degradation with a width of 5 pixels. The nonstationary mean is estimated by a one-dimensional moving average on 11 pixels of observation data and its variance is obtained from all the picture data by an unbiased estimate. The linear system of equations for the gradient function of Eq. (5.47) which determines the increment value for the iterations is heavily dependent on the structure of the blurring matrix H. When the H matrix is symmetrical, the computing time of Eq. (5.47) with the Newton-Raphson solution can be reduced. A simulation is done for one directional linear motion blurring (5 pixels) and various $\overline{(SNR)}_{rms}$. The sectioned MAP filter has a section size of 36 pixels with 8 pixels of overlap. The restored images of the MAP filter are shown in Fig. 5.4 for different $\overline{(SNR)}_{rms}$.

From Fig. 5.4, we can see that the ill-conditioning of the restored image with $\rho=0$ is more severe for the higher $\overline{(SNR)}_{rms}$ image signals. A possible explanation is that the restoration filter performs more smoothing with a priori knowledge of high object correlation.

Figure 5.4a Restored image with a MAP filter for one-dimensional blurring degradation at $\overline{(SNR)}_{rms}=\sqrt{5}$

(A) Original object image
(B) Degraded image
(C) Nonstationary mean image
(D) Restored image with MAP filter

Figure 5.4b Restored image with a MAP filter for one-dimensional blurring degradation at $\frac{}{(SNR)_{rms}} = \sqrt{7.5}$

(A) Original object image
(B) Degraded image
(C) Nonstationary mean image
(D) Restored image with MAP filter

Figure 5.4c Restored image with a MAP filter for one-dimensional blurring degradation at $\overline{(SNR)}_{rms}=\sqrt{10}$

(A) Original object image
(B) Degraded image
(C) Nonstationary mean image
(D) Restored image with MAP filter

For a local adaptive MAP filter, the equation to be solved is

$$W\cdot[\lambda H^{T}(\underline{g}-\underline{l})]+(I-W)\cdot[-R_{f}^{-1}(\underline{f}-\bar{\underline{f}})] = \underline{0}, \qquad (5.48)$$

where $W= \text{diag}\{W_i\}$. Here, W_i is the weight of the ith secton which varies with the nonstationary mean and the second moment of the local properties of the image. The local adaptive MAP filter also can be used for the restoration of images degraded by spatially variant point spread functions. The point spread function at each photon detector may not be identical over the whole array. The image can be divided into sectioned images each with its own space invariant PSF. For simplicity, we have simulated a global adaptive MAP sectioning filter in which each W is fixed. The simulation is performed with $\rho=0.95$ for different section weights and different $\overline{(SNR)}_{rms}$. The experimental results are shown in Figs. 5.5 through 5.7 with different $\overline{(SNR)}_{rms}$. The $W=0.5$ gives equal weight to the (ML) solution and the a priori solution.

From these experimental results, we observe that more high frequencies are extracted if the weight on the ML term is increased. Also we see that large weight on the ML solution results in ill-conditioning of some of the solutions with the MAP estimate. With a large weight on the ML part of the solution, the MAP estimate asymptotically approaches the ML estimate. This is

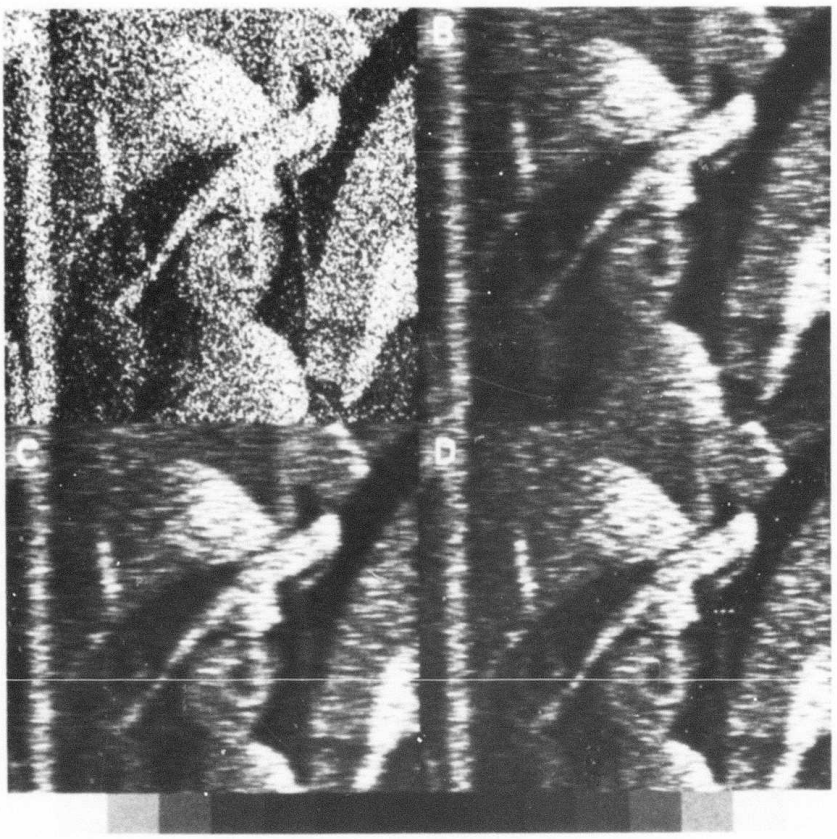


Figure 5.5 Restored images with a global adaptive MAP filter for $\overline{(SNR)}_{rms}=\sqrt{5}$

(A) Degraded image
(B) Restored image with $w_i=0.3$
(C) Restored image with $w_i=0.5$
(D) Restored image with $w_i=0.8$

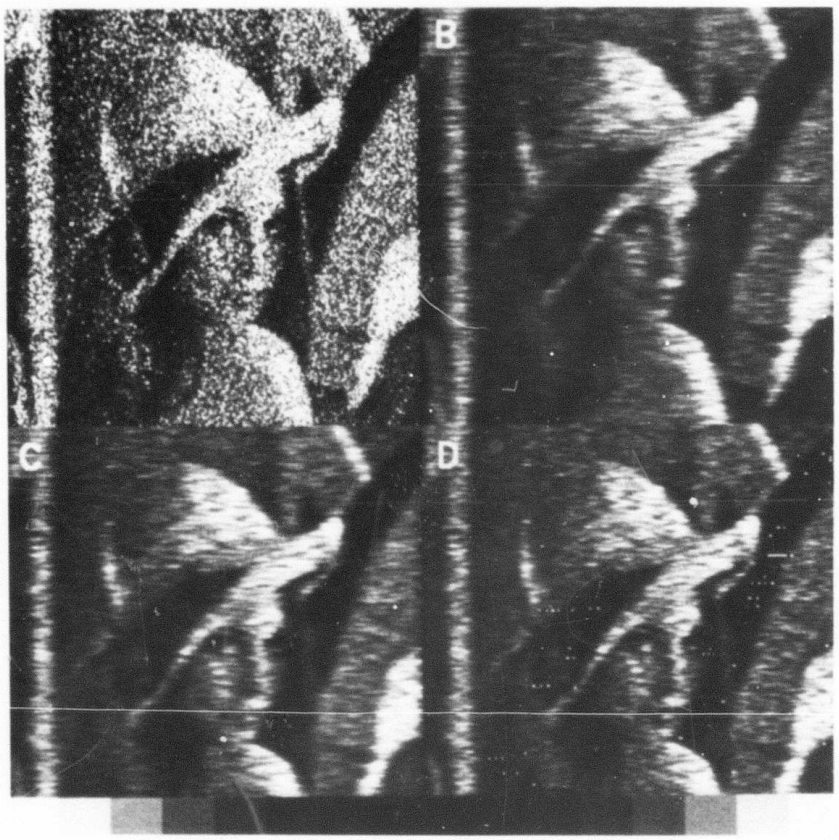


Figure 5.6 Restored images with a global adaptive MAP filter for $\overline{(SNR)}_{rms}=\sqrt{7.5}$

(A) Degraded image
(B) Restored image with $w_i=0.3$
(C) Restored image with $w_i=0.5$
(D) Restored image with $w_i=0.8$

Figure 5.7 Restored images with a global adaptive MAP filter for $\overline{(SNR)}_{rms}=\sqrt{10}$

(A) Degraded image
(B) Restored image with $w_i=0.3$
(C) Restored image with $w_i=0.5$
(D) Restored image with $w_i=0.8$

consistent with fact that the ML estimate is the inverse filter for image restoration with blurring degradations. (For the non blurring case, the inverse solution of image restoration just is the observation data as discussed in the last chapter). From those observations, we can conclude that there is an optimal section weight over the global adaptive filter as well as the local adaptive filter.

5.6 Implementation of the MAP Filter with a Two-dimensional Blurring Degradation

This section is divided into two subsections to separately discuss the assumptions of separability and non-separability. The blurring degradation is simulated by a 3x3 pixel moving window blurring for different $\overline{(SNR)_{rms}}$. The nonstationary mean is estimated by a 7x7 window moving average over the measured photon counts. This size of moving wndow was found to give a reasonably good estimate with a minimum amount of computing.

5.6.1 Separable Case

In this section references to separability means that the PSF is a separable space-invariant function (SSIPSF) in the sectioned MAP filter and that the covarance matrix R_f is Markovian and separable [5-3]. This can be expressed in vector notation as

$$H = H_R \circledast H_C \tag{5.49}$$

and $R_f = R_R \circledast R_C$, where

$$H_R = H_C = \begin{bmatrix} p_R(L) & p_R(L-1) & \cdots & p_R(1) & 0 \\ & \ddots & & \ddots & \\ 0 & & p_R(L) & & p_R(1) \end{bmatrix} \tag{5.41}$$

$$p_R = p_C = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ \rho & 1 & \rho & & \rho^{N-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & \cdots \rho & 1 \end{bmatrix} \tag{5.28}$$

and $\circledast$ denotes the direct product.

From direct product identities [5-14], we have

$$R_f^{-1} = (R_R \circledast R_C)^{-1} = R_R^{-1} \circledast R_C^{-1} \tag{5.50}$$

Therefore, the two-dimensional Eq. (5.29) can be implemented using Eq. (5.47) as a row processor and then using Eq. (5.47) as a column processor. The solution to Eq. (5.47) uses the same sectioning method with a Newton-Raphson iterative solution as before. The processing time of this two-dimensional MAP filter is twice that of the one-dimensional case. Without describing the details of the MAP implementaton, the experimental results are shown in Fig. 5.8 for different $\overline{(SNR)}_{rms}$ and $\rho = 0.95$.

From Fig. 5.8, we can see that the restored image with a two-dimensional separable filter performs considerable

Figure 5.8a Restored images with a separable MAP filter for $\overline{(SNR)}_{rms}=\sqrt{5}$

(A) Original object image
(B) Degraded image
(C) Restored image by one-dimensional MAP filter
(D) Restored image by two-dimensional MAP filter

Figure 5.8b Restored images with a separable MAP filter for $\overline{(SNR)}_{rms}=\sqrt{7.5}$

(A) Original object image
(B) Degraded image
(C) Restored image by one-dimensional MAP filter
(D) Restored image by two-dimensional MAP filter

Figure 5.8c Restored images with a separable MAP filter for $\overline{(SNR)}_{rms}=\sqrt{10}$

(A) Original object image
(B) Degraded image
(C) Restored image by one-dimensional MAP filter
(D) Restored image by two-dimensional MAP filter

smoothing of the Poisson noise. The restored images show some improvement over the noisy originals. Although the separable assumption is an accurate first-order approximation for the system, the image field itself is not separable. Thus, we implement a two-dimensional non-separable MAP sectioned filter next.

5.6.2 Non-separable Case

We now assume that the PSF is a non-separable space-invariant function (NSIPSF) and that R_f is an identity matrix to simplify the simulation. The two-dimensional non-separable sampled infinite area superposition model is used to reduce the two-dimensional blurring degradation to a matrix H. This matrix is an M^2xN^2 matrix, where M is the observed data size and N is the processed data size for the sectioned MAP filter. Thus, we need to solve a linear system of equations of order N^2xN^2 (Eq. (5.29)) in order to find the updated increment value of the root in each iterative step. Even though the sections are small, a large amount of computing is needed for this sectioned MAP filter. The details of computing time of the sectioned MAP filter are described in the last section. Since the sectioned MAP filter uses a lexicographic ordered vector representation and the observed image uses a matrix representation, some conversion between them is needed. The conversion relation

from matrix to vector is

$$\underline{D} = ((J-1)*M)+I \tag{5.51}$$

and from vector to matrix they are

$$I = \text{Mod}(P-1,M)+1 \tag{5.52}$$

$$J = [(P-1)\text{Mod}(P-1,M)]/M+1 \tag{5.53}$$

where P is the location of the vector element with lexicograhic ordering, M is the size of the section to be processed, (J,I) is the (row, column) location of the image pixel and Mod is the modulo operator. The overlap-save sectioned MAP filter is used to minimize the boundary edge effects of sectioning. The simulation is done with section size 9x9 and an overlap of 4x4. Since the blurring degradation is an unweighted average over 3x3 pixels, from the discussion of the previous section, the wrap around data is 2(L-1)x2(L-1) pixels which is 4x4 pixels. The nonstationary mean is estimated by a rolling window moving average method which efficiently keeps only the data required in high speed memory. Becuase of the large estimated CPU time for this filter on the DEC KL-10 only the right side of the noisy picture is processed. Two original images with different $\overline{(SNR)}_{rms}$ were filtered and the experimental results are shown in Fig. 5.9. From Fig. 5.9, we can see that the two-dimensional filter has

Figure 5.9a Restored image with a non-separable two-dimensional MAP filter for $\overline{(SNR)}_{rms}=\sqrt{7.5}$

(A) Original object image
(B) Degraded image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

Figure 5.9b Restored image with a non-separable two-dimensional MAP filter for $\overline{(SNR)}_{rms}=\sqrt{10}$

(A) Original object image
(B) Degraded image
(C) Nonstationary mean image
(D) Restored image by the MAP filter

produced noticeably better results than the restored images of Fig. 5.8. However, the CPU time of the general non-separable filter is approximately a factor of 100 longer than that for a separable filter with a 256x256 image and section sizes indicated. The results shown in Fig. 5.9 were done with a preliminary version of the algorithm to show feasibility only. Processing the entire 256x256 image would take approximately 100 minutes. It is likely that considerable savings in computer time would result from a very carefully written algorithm which would recursively perform Newton-Raphson solutions between windows.

5.7 Conclusions

We have developed an MAP filter for a Poisson noise model with blurring degradations. The implementation and method of solution for the MAP filter are heavily dependent on the form of the blurring degradation matrix H and the covariance matrix of the object image. The overlap-save sectioning method with a Newton-Raphson solution has been shown to be an effective fast approach to the suboptimal MAP estimate. The sampled infinite area superposition model is used for the blurring degradation. Both the one-dimensional blurring and two-dimensional blurring situations with different levels of Poisson noise were simulated.

From these experimental results, we find that the nonstationary mean carries the most of the structured background information and that the covariance matrix leads to a stable Newton-Raphson solution especially for higher $\overline{(SNR)_{rms}}$ image signals. It has been found that there is an optimal weight in the global adaptive MAP filter which produces the best quality restoration. Too much weight on the ML term solution will gives rise to ill-conditioning. From Fig. 5.9, we see that the quality of the restored image with the MAP filter assuming a non-separable PSF gives better results than the separable MAP filter. However, the CPU time for the non-separable filter is much longer than for the separable filter. The overlap-save sectioned MAP filter has been shown to be useful for overcoming problems of large dimensionality and nonlinearity which are inherent in MAP estimation.

CHAPTER 6

THE QUALITY OF THE MAP ESTIMATE

6.1 Introduction

In this chapter, we investigate the quality of the MAP estimate for Poisson noise. The quality of the estimate depends on the performance criteria chosen. There are two types of performance criteria [6-1,6-2]: one depends on the estimator structure and the other depends on the performance itself. The MAP estimate and ML estimate belong to the former one because the MAP estimate is the mode of the posteror probability density and the ML estimate is the mode of the a priori probability density. The Bayes estimate belongs to latter because it minimizes the risk of the estimate. The MMSE (minimum mean square error) estimate is a special case of the Bayes estimate when the cost function is proportional to the mean square error. However, it is customary to choose the conditional or unconditional expected squared error as a universal measure of quality of all estimates. Unfortunately, the expectation operator leading to this measure is, in general, very complicated due to the complexity of variance estimates. However, it is possible to derive an expression

for a lower bound on the variance in terms of only the statistical properties of the observed signal and estimation bias. This quality measure can be formed for any estimator without detailed knowledge of its structure as long as the estimate is unbiased. This lower bound for the estimation error variance is the well known Cramer-Rao lower bound (CRLB) [6-3,6-15].

There are two measures that are used together to determine the quality of an estimate. These are the expectation of the estimate and the variance of the estimation error. The first of these measures the bias inherent in an estimate, and the second is equivalent to the mean-squared error between the estimate and the original data. In general, we try to find unbiased estimates with small estimation error variance.

In section 6.2 and 6.3, we discuss biased and unbiased estimates and show that the $\underline{\hat{f}}_{MAP}$ estimate for the Poisson noise model is an unbiased estimate. In sections 6.4, 6.5 and 6.6, we derive the Cramer-Rao lower bound of the estimate for non-random scalar parameters and random vector parameters with the Poisson noise model. In the final section, we present the conclusions of this chapter.

6.2 Biased and Unbiased Estimates

A conditional unbiased estimate is one whose expected

value is equal to the true value of the quantity being estimated. An unconditional unbiased estimate is one whose expected value is equal to the expected value of the quantity being estimated. We denote the estimate by a random variable $\hat{X}$ which is a function of the observations $\underline{Y}$. If $\hat{X}$ is a conditional unbiased estimate, then

$$E_{\underline{Y}}[\hat{X}] = \int \hat{X}(\underline{Y}) P(\underline{Y}|X) d\underline{Y} = X. \tag{6.1}$$

and if $\hat{X}$ is an unconditional unbiased estimate, then

$$E_{\underline{Y}}[\hat{X}] = \int \hat{X}(\underline{Y}) P(\underline{Y}) d\underline{Y} = E(X) = \overline{X}. \tag{6.2}$$

On the other hand, biased estimates do not possess this desirable feature; their expected values contain an additional function B(X) of the parameter to be estimated. Accordingly, for biased estimates we have

$$E_{\underline{Y}}[\hat{X}] = X+B(X), \tag{6.3}$$

or

$$E_{\underline{Y}}[\hat{X}] = \overline{X+B(X)}. \tag{6.4}$$

for the conditional biased estimate and the unconditional biased estimate, respectively [6-3, 6-4, 6-5].

6.3 $\hat{\underline{f}}_{MAP}$ is an Unconditional Unbiased Estimate Vector

From previous chapters, we have

$$\hat{\underline{f}}_{MAP} = \overline{\underline{f}}+\lambda R_{\underline{f}} H^T (\underline{q}-\underline{1}) \tag{6.5}$$

where $\hat{\underline{f}}_{MAP}$ is the N^2x1 estimate vector

$\overline{\underline{f}}$ is the N^2x1 nonstationary mean vector

R_f is the covariance matrix of the image

and H is the N^2xN^2 discrete blurring matrix.

Taking the expectation on both sides of Eq. (6.5), we have

$$E[\hat{f}_{MAP}] = \overline{\underline{f}}+E[\lambda R_f H^T(\underline{q}-\underline{1})]. \tag{6.6}$$

Since

$$\begin{aligned} q_i &= \frac{d_i}{b_i} \\ \overline{g}_i &= \lambda b_i, \end{aligned} \tag{6.7}$$

then

$$\begin{aligned} E[q_i] &= E_{b_i}\{E_{g_i|b_i}(\frac{\alpha g_i}{b_i})\} \\ &= \alpha E_{b_i}\{E_{g_i|b_i}(\frac{g_i}{b_i})\} \end{aligned} \tag{6.8}$$

where $E_{g_i|b_i}$ denotes conditional expectation over g_i for a given b_i. E_{b_i} denotes expectation over b_i, hence

$$E[q_i] = \alpha E_{b_i}[\frac{\lambda b_i}{b_i}] = \alpha\lambda = 1 \tag{6.9}$$

Thus

$$E[\underline{q}] = \underline{1} \tag{6.10}$$

and substituting Eq. (6.10) into Eq. (6.6), we get

$$E[\hat{\underline{f}}_{MAP}] = \overline{\underline{f}} \tag{6.11}$$

Therefore, $\hat{\underline{f}}_{MAP}$ is an unconditional unbiased estimate vector.

6.4 Cramer-Rao Lower Bound (CRLB)

For notational and mathematical simplicity, we first focus our attention on the CRLB for the case of non-random scalar parameters, an example is sample mean and sample variance of the amplitude, phase, and Doppler frequency of the estimation of the radar signals assuming these parameters unknown but not random variables. Then, for the random vector parameters case, the CRLB can be derived by a straightforward modification of the derivation for the non-random parameter cases. From the derivation of the non-random parameter case, we can easily understand the fundamental concept of the CRLB and the basic relation between the CRLB and the variance of the estimation error.

6.4.1 CRLB for Non-random Variable Case

First, we assume X is an unknown constant parameter to be estimated from a sequence of measurements $y_1, y_2, \ldots, y_K$ as shown in Fig. 6.1 where $\underline{y} = [y_1, y_2, \ldots, y_K]^T$. Assuming $\hat{X}$ is an unbiased parameter estimate we have

$$\int \hat{X} f(\hat{X}|X) d\hat{X} = X \tag{6.12}$$

or from Eq. (6.1), we have

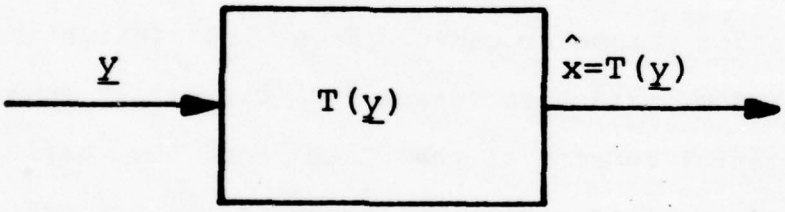


Figure 6.1 Block diagram of nonrandom parameter estimate

$$\underset{\text{K-fold integral}}{\int\int\cdots\int} T(\underline{y})\,f(\underline{y}|X)\,d\underline{y} = X \tag{6.13}$$

where $\hat{X}=T(\underline{y})$. Differentiating both sides of Eq. (6.13) with respect to X we have

$$\underset{\text{K-fold integral}}{\int\int\cdots\int} T(\underline{y})\frac{\partial f(\underline{y}|X)}{\partial X}d\underline{y} = 1. \tag{6.14}$$

Now we rewrite Eq. (6.14) as

$$\underset{\text{K-fold integral}}{\int\int\cdots\int} T(\underline{y})\,f(\underline{y}|X)\left[\frac{\frac{\partial f(\underline{y}|X)}{\partial X}}{f(\underline{y}|X)}\right]d\underline{y} = 1 \tag{6.15}$$

or

$$\underset{\text{K-fold integral}}{\int\int\cdots\int} \left[T(\underline{y})\frac{\partial \ell n f(\underline{y}|X)}{\partial X}\right]f(\underline{y}|X)\,d\underline{y} = 1 \tag{6.16}$$

Inspection of Eq. (6.16) shows that it is

$$E\left[T(\underline{y})\frac{\partial \ell n f(\underline{y}|X)}{\partial X}\right] = 1. \tag{6.17}$$

Next, we examine the normalized correlation coefficient between $\hat{X}=T(\underline{y})$ and $\frac{\partial \ell n f(\underline{y}|X)}{\partial X}$. From the definition of ρ, we have

$$\rho \triangleq \frac{\mu_{11}}{\sigma_1 \sigma_2} = \frac{E\{[(T(\underline{y})-E(T(y))][\frac{\partial \ell nf(\underline{y}|X)}{\partial X} - E(\frac{\partial \ell nf(\underline{y}|X)}{\partial X})]\}}{\sqrt{\text{var } T(\underline{y})} \cdot \sqrt{\text{var}(\frac{\partial \ell nf(\underline{y}|X)}{\partial X})}} \quad (6.18)$$

where var denotes variance. Now,

$$E\left[\frac{\partial \ell nf(\underline{y}|X)}{\partial X}\right] = E\left[\frac{1}{f(\underline{y}|X)} \frac{\partial f(\underline{y}|X)}{\partial X}\right]$$

$$= \int\int \cdots \int \frac{f(\underline{y}|X)}{\partial X} \cdot \frac{1}{f(\underline{y}|X)} f(\underline{y}|X) d\underline{y} \quad (6.19)$$

$$= \frac{\partial}{\partial X} \int\int \cdots \int f(\underline{y}|X) d\underline{y} = \frac{\partial}{\partial X} [1] = 0,$$

hence

$$E[\frac{\partial}{\partial X} \ell nf(\underline{y}|X)] = 0. \quad (6.20)$$

From Eq. (6.20) and simple manipulations, Eq. (6.18) becomes

$$\rho = \frac{E[T(\underline{y}) \frac{\partial \ell nf(\underline{y}\ X)}{\partial X}]}{\sqrt{\text{var}T(\underline{y})} \cdot \sqrt{\text{var}(\frac{\partial \ell nf(\underline{y}|X)}{\partial\ X})}} \quad (6.21)$$

From Eq. (6.17), then

$$\rho = \frac{1}{\sqrt{\text{var}T(\underline{y})} \cdot \sqrt{\text{var}(\frac{\partial \ell nf(\underline{y}\ X)}{\partial\ X})}} \quad (6.22)$$

From the definition of ρ, we know that ρ is equal to or less than 1. Therefore, we obtain

$$\text{var}\ (T(\underline{y})) = \text{var}\ \hat{X} \geq \frac{1}{\text{var}[\frac{\partial \ell n f(\underline{y}|X)}{\partial X}]} \quad . \tag{6.23}$$

or equivalently

$$\text{var}\ \hat{X} = E[(\hat{X}-X)] \geq \frac{1}{E\{[\frac{\partial \ell n f(\underline{y}|X)}{\partial X}]^2\}} \quad . \tag{6.24}$$

Equation (6.24) is called the Cramer-Rao inequality for the unbiased estimate [6-15]. Note that the CRLB is a bound on the mean-square error. For a biased estimate, the CRLB is

$$\text{var}\ \hat{X} \geq \frac{1+\frac{dB(X)}{dX}}{E\{[\frac{\partial \ell n f(\underline{y}|X)}{\partial X}]^2\}} \quad , \tag{6.25}$$

where B(X) is the bias function of X.

For the non-random variable vector case, we have directly

$$\text{var}\ [\hat{x}_i - x_i] \geq J^{ii} \quad , \tag{6.26a}$$

where J^{ii} is the iith element in the KxK square matrix J^{-1}. The elements of J are

$$J_{ij} \triangleq E\left[\frac{\partial \ell n p(\underline{y}|\underline{x})}{\partial x_i} \cdot \frac{\partial \ell n p(\underline{y}|\underline{x})}{\partial x_j}\right] \quad ,$$

$$\underline{x} = [x_1, x_2, \ldots, x_N]^T, \quad \underline{y} = [y_1, y_2, \ldots, y_N]^T, \tag{6.26b}$$

where $\underline{x}$ is a vector of non-random variables, to be

estimated and $\underline{y}$ is the measurement data vector. Equivalently,

$$J \triangleq E\{(\nabla_{\underline{x}}[\ell np(\underline{y}|\underline{x})])(\nabla_{\underline{x}}[\ell np(\underline{y}|\underline{x})])^T\} \qquad (6.27)$$

where

$$\nabla_{\underline{x}} \triangleq [\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_k}]^T$$

The J matrix is commonly called Fisher's information matrix [6-1, 6-2, 6-3, 6-4, 6-15].

6.4.2 CRLB for Random Variable Vector Case [6-1].

For the random variable vector case, the information matrix J_T now consists of two parts

$$J_T \triangleq J_D + J_P \quad , \qquad (6.28)$$

where

$$J_D \triangleq E(\{\nabla_{\underline{x}}[\ell np(\underline{y}|\underline{x})]\}\{\nabla_{\underline{x}}[\ell np(\underline{y}|\underline{x})]\}^T), \qquad (6.29)$$

$$J_P \triangleq E(\{\nabla_{\underline{x}}\ell np(\underline{x})\}\{\nabla_{\underline{x}}\ell np(\underline{x})\}^T). \qquad (6.30)$$

and

$$J_{D_{ij}} \triangleq E\left[\frac{\partial \ell np(\underline{y}|\underline{x})}{\partial x_i} \cdot \frac{\partial \ell np(\underline{y}|\underline{x})}{\partial x_j}\right] ,$$

$$J_{P_{ij}} \triangleq E\left[\frac{\partial \ell np(\underline{x})}{\partial x_i} \cdot \frac{\partial \ell np(\underline{x})}{\partial x_j}\right] .$$

The matrix J_D represents information obtained from the data or from the probability density $P(\underline{y}|\underline{x})$ of the MAP estimate. The matrix J_P represents information obtained *a priori*.

The correlation matrix of the error is

$$R_{\varepsilon} \triangleq E(\underline{x}_{\varepsilon}\underline{x}_{\varepsilon}^{T}) \tag{6.31}$$

where $\underline{x}_{\varepsilon}=(\hat{\underline{x}}-\underline{x})$. The diagonal elements represent the mean-square errors and the off diagonal elements are the cross correlations. The mean-square error of the estimate as a function of the information matrix is

$$E[x_{\varepsilon_i}^2] \geq (J_T^{-1})^{ii}. \tag{6.32}$$

The diagonal elements in the inverse of the total information matrix J_T are the lower bounds on the corresponding mean-square errors, and this is the situation of interest here.

6.5 Derivation of the CRLB for MAP Estimates with a Poisson Noise Model

The estimate of error covariance is, in general, very complicated to find due to the complexity of the posterior density. However, for the MAP estimate it is possible to derive an expression for a lower bound on the variance because we know the *a priori* density $p(\underline{y}|\underline{x})$ and probability density $p(\underline{x})$ of $\underline{x}$. From Eq. (6.28), we have

$$J_T = J_D + J_P, \tag{6.28}$$

where J_D and J_P are defined in Eqs. (6.29) and (6.30) respectively. From the last chapter, we have

$$p(\underline{d}|\underline{f}) = \prod_i \frac{(\lambda b_i)^{d_i/\alpha} e^{-\lambda b_i}}{\alpha(\frac{d_i}{\alpha})!} \quad , \tag{6.33}$$

and

$$p(\underline{f}) = K_b \exp\{-\frac{1}{2}(\underline{f}-\bar{\underline{f}})^T R_f^{-1}(\underline{f}-\bar{\underline{f}})\} \ . \tag{6.34}$$

From Eqs. (6.30) and (6.34), we can obtain [6-1, 6-10, 6-11]

$$J_P = R_f^{-1} \quad , \tag{6.35}$$

where R_f is the covariance matrix of the image. From Eqs. (6.29) and (6.33), we have

$$\begin{aligned} J_D &= E\{[\nabla_{\underline{f}} \ell n p(\underline{d}|\underline{f})][\nabla_{\underline{f}} \ell n p(\underline{d}|\underline{f})]\} \\ &= E\{[\lambda H^T(\underline{q}-\underline{1})][\lambda H^T(\underline{q}-1)]^T\} \\ &= \lambda^2 H^T E[(\underline{q}-\underline{1})(\underline{q}-\underline{1})^T]H, \end{aligned} \tag{6.36}$$

where E denotes expectation. From Eq. (6.10) we have $E[\underline{q}]=\underline{1}$, thus,

$$J_D = \lambda^2 H^T E[(\underline{q}-\bar{\underline{q}})(\underline{q}-\bar{\underline{q}})^T]H \tag{6.37}$$

$$J_D = \lambda^2 H^T R_{\underline{q}} H \quad , \tag{6.38}$$

where

$$R_{\underline{q}} \triangleq E\{(\underline{q}-\bar{\underline{q}})(\underline{q}-\bar{\underline{q}})^T\} \ .$$

Substituting Eqs. (6.38) and (6.35) into Eq. (6.28), we have

$$J_T = \lambda^2 H^T R_q H + R_f^{-1} \quad . \tag{6.39}$$

This is the total information matrix J_T for the MAP estimate with the Poisson noise model.

When the J_T^{-1} exists, from Eq. (6.32) we have

$$E[(\hat{f}_i - f_i)^2] \geq \{J_T^{-1}\}_{ii} \quad , \tag{6.40}$$

where $\{J_T^{-1}\}$ are the diagonal elements of J_T^{-1} and $\hat{f}_i$ is an estimate of the ith component of the restored image vector $\underline{f}$. Inspection of Eq. (6.40) indicates that the error bound depends on four quantities: the Rate function constant λ, the discrete blurring matrix H, the covariance matrix R_q, and the covariance matrix of the image R_f. To obtain some physical meaning from this expression, we assume that

$$R_q = \sigma_q^2 I \quad \text{and} \quad R_f = \sigma_f^2 I \quad ,$$

to obtain

$$\begin{aligned} J_T^{-1} &= [(\lambda^2 H^T H \sigma_q^2 I + (\sigma_f^2)^{-1} I)]^{-1} \\ &= [(\sigma^2 ||H||^2 \sigma_q^2 I) + (\sigma_f^2)^{-1} I]^{-1} \quad , \end{aligned} \tag{6.41}$$

Thus

$$E[(\hat{f}_i - f_i)^2] = \frac{\sigma_f^2}{\lambda^2 ||H||^2 \sigma_q^2 \sigma_f^2 + 1} \quad . \tag{6.42}$$

From Eq. (6.42), we divide out σ_f^2 on both the numerator and the denominator of the right hand side of Eq. (6.42) and

get

$$E[(\hat{f}_i - f_i)^2] = \frac{1}{\lambda^2 \|H\|^2 \sigma_q^2 + \frac{1}{\sigma_f^2}} \quad . \tag{6.43}$$

From Eq. (6.43), we observe that

(1) When σ_f^2 is large, we can drop the term $1/\sigma_f^2$ from the denominator. Multiplying both the numerator and denominator of Eq. (6.43) by the mean intensity $\bar{b}_i^2$, gives

$$E[(\hat{f}_i - f_i)^2] \cong \frac{\bar{b}_i^2}{(\lambda \bar{b}_i)^2 \|H\|^2 \sigma_q^2} \quad . \tag{6.44}$$

In this situation, $\lambda \bar{b}_i$, H, and σ_q^2 play more important roles in the error bound of the Poisson noise model than the variance of the object σ_f^2;

(2) The error bound decreases with the square of the ensemble mean rate function ($\lambda \bar{b}_i$). From the experimental results of Fig. 6.2, we see that increasing the ensemble mean rate function of the Poisson process $\lambda \bar{b}_i$, reduces the Poisson noise degradation. Consequently, when the $\overline{(SNR)}_{rms}$ is greater than or equal to 10 db ($\lambda \bar{b}_i \geq 100$), then

$$E[(\hat{f}_i - f_i)^2] \geq 10^{-4} \quad , \tag{6.45}$$

and the Poisson noise degradation effects are small in a practial sense.

(3) The error bound is also inversely proportional to the squared norm of the point spread function H and the error

Figure 6.2 Images with Poisson noise at different $\overline{(SNR)}_{rms}$

(A) Original object image
(B) The Poisson noisy image with $\overline{(SNR)}_{rms}=\sqrt{5}$
(C) The Poisson noisy image with $\overline{(SNR)}_{rms}=\sqrt{10}$
(D) The Poisson noisy image with $\overline{(SNR)}_{rms}=\sqrt{20}$

due to noise is "amplified" by the point spread function. This results from the ill-conditioning of the restoration process.

6.6 Conclusions

We conclude that the MAP estimate for the Poisson noise model is an unbiased estimate and have found the Cramer-Rao lower bound (CRLB) for the variance of the estimation error. The CRLB is the tightest lower bound for an efficient estimate. When an efficient estimate does not exist, the lower bound can be improved compared to the Cramer-Rao inequality. Better lower bounds may be the Bhattacharyya bound and Barankin bound, but these bounds are very difficult and tedious computationally. The Barankin bound does not require the probablity density to be differentiable and it gives the greatest lower bound. It requires a maximization over a function to obtain the bound and the procedure for finding this maximum is usually not straightforward.

Use of the statistical estimation method is particularly desirable from the viewpoint of error analysis because known techniques can be applied to compute the error bound. We have developed the CRLB for MAP estimation with the Poisson noise model and also shown the behavior of the CRLB approximation. From these facts, we are able to find the algoritm whose mean-squared errors is closest to

the CRLB. It must be remembered that the CRLB is a lower bound and that the actual restoration error will be greater. It is possible that a better suboptimal sectioned MAP algorithm can be found to reduce the actual restoration error closer to the Cramer-Rao lower bound.

CHAPTER 7

COMPARISON BETWEEN THE LMMSE RESTORATION FILTER AND THE MAP RESTORATION FILTER

7.1 Introduction

The most common goal for restoring a degraded image is to reduce the measurement error and to make the information more visible. A true comparison between filters should follow some objective criteria. However, the mechanism of human information extraction is not well understood and there are no universally agreed upon criteria by which to judge the quality of a proposed image restoration filter. There are two simple criteria which are commonly accomodated. One is the numerical closeness of the restored image to the undegraded original object image in terms of mean square error, and the other is the visual subjective appearance of the restored image compared to the original. These two criteria are often in conflict because of the complex, nonlinear, and adaptive properties in the psycho-physical processes of human vision [7-1]. Pearlman has proposed a new compromise criterion [7-2] which is a weighted mean square error. The amplitude weights are not constants but are dependent upon the contrast ratio of the

image and also upon an exponential function of average intensity. This new weighted mean square error criterion reflects some degree of the nonlinearity and complex properties of the human visual system, although it is complicated to evaluate. The numerical closeness criterion is most often employed because it is well defined and mathematically tractable. There is a fundamental difference in the estimation criterion between the MAP and the LMMSE image restoration filters. The LMMSE filter is based upon minimizing linear minimum mean-square error under perfect a priori knowledge of the object and image, while the MAP filter is based upon the maximization of the a posteriori density of the image. Although it is theoreticaly hard to make any comparison between them because of this fundamental criterion difference, we will compare them based upon the first two quality criteria and computation time.

The structure of the LMMSE filter is discussed and analyzed in the next section. In section 7.3 we discuss the implementation and illustrate additional experimental results. In section 7.4 we discuss image quality measures and compute the normalized mean square error of the restorations. In the final section we make comparisons and state the conclusions.

7.2 The Structure of the LMMSE Filter and MAP Filter

7.2.1 Structure of the LMMSE Restoration Filter

From chapter 3, the LMMSE filter transfer function $W(u,v)$ is derived by minimizing $E\{\int\int_{-\infty}^{\infty} e^2(x,y)dxdy\}$. By Parseval's theorem, it is equivalent to minimizing $E[\int\int_{-\infty}^{\infty} |\varepsilon(u,v)|^2 dudv]$, where ε is the Fourier transform of $e(x,y)$. Use of this minimizing mean-square error criterion and the orthogonality principle [7-10, 7-11, 7-12] yields the transfer function of the LMMSE filter as

$$W_p(u,v) = \frac{\Phi_{fg}(u,v)}{\Phi_{gg}(u,v)} \quad , \tag{7.1}$$

where Φ_{fg} is the cross-spectral density of the detected image $g(x,y)$ and object $f(x,y)$ while Φ_{gg} is the spectral density of the detected image $g(x,y)$. From a straight forward substitution into Eq. (7.1) [7-3,7-4], we have

$$W_p(u,v) = \frac{\bar{N}\mathcal{H}^*(u,v)\Phi_f(u,v)}{1+\bar{N}||\mathcal{H}(u,v)||^2\Phi_f(u,v)} \quad , \tag{7.2}$$

where $\bar{N}$ is mean number of photon counts in the detected image $\mathcal{H}(u,v)$ is the Fourier transform of the PSF $h(x,y)$ and $\Phi_f(u,v)$ is the spectral density of the object. The detailed derivation of Eq. (7.2) is described in [7-3]. For a linear Gaussian additive noise model, the Wiener filter is [7-9, 7-13, 7-14, 7-15, 7-16]

$$W_W(u,v) = \frac{\mathcal{H}^*(u,v)\Phi_f(u,v)}{|\mathcal{H}(u,v)|^2\Phi_f(u,v)+\Phi_n(u,v)} \qquad (7.3)$$

where $\Phi_n(u,v)$ is the spectral density of noise which is statistically independent of the signal, $\Phi_f(u,v)$ is the spectral density of the object and $\mathcal{H}(u,v)$ is the Fourier transform of $h(x,y)$. Rewriting Eq. (7.3), the Wiener filter takes the most familiar form

$$W_W(u,v) = \frac{\mathcal{H}^*(u,v)}{|\mathcal{H}(u,v)|^2+\frac{1}{\beta}} \quad , \qquad (7.4)$$

where

$$\beta \triangleq \frac{\Phi_f(u,v)}{\Phi_N(u,v)}$$

β is called the signal-to-noise (SNR) for the linear additive Gaussian noise model.

Since the linear additive Gaussian model assumes signal-independent noise, it is reasonable to define β as the signal-to-noise ratio. Similarly, from Eq. (7.2) for the Poisson noise model, the LMMSE filter is

$$W_p(u,v) = \frac{\mathcal{H}^*(u,v)}{|\mathcal{H}(u,v)|^2+\frac{1}{\alpha}} \qquad (7.5)$$

where $\alpha=\bar{N}\Phi_f(u,v)$. The function $W_p(u,v)$ is the same as $W_W(u,v)$ except that α is defined differently from β. Since Poisson noise is signal-dependent, the SNR is not well defined. However α can be called the equivalent

signal-to-noise ratio $(SNR)_{eq}$. From Eq. (7.5), when the value of α is very large, which is the case when the rate function is high, then Eq. (7.5) becomes

$$W_p(u,v) \cong \mathcal{H}^{-1}(u,v) \quad . \tag{7.6}$$

Thus, the LMMSE filter approaches the inverse filter in the absence of Poisson noise. Indeed, the larger the rate function, the lesser is the degradation due to Poisson noise. In this case, the LMMSE filter only needs to remove the blurring degradation effects. As discussed in chapter 5, Poisson noise effects are much more pronounced at low light levels when the value of α is smaller. In this case, the LMMSE filter W_p is dominated by Poisson noise and image signals will be seriously distorted. Thus the performance of the LMMSE filter will expected to be worse at lower equivalent SNR's. Although the LMMSE filter is based upon the minimum mean-square error criterion, this estimation error is a minimum under the condition that the a priori knowledge is perfect. Functions such as the spectral density of the object Φ_f and the mean number of photon counters $\bar{N}$ must be perfectly known. In reality, Φ_f and $\bar{N}$ are never perfectly known and must be estimated from the observation [7-7, 7-18]. Hence, the actual LMMSE error does not reach the minimum. In short, the LMMSE filter tries to force the solution toward the inverse solution with some sort of smoothness controlled by the equivalent

SNR α.

7.2.2 Structure of the MAP Restoration Filter

It has been noted in preceding chapters that the fundamental MAP estimate contains maximum likelihood (ML) and *a priori* terms in its solution. From previous chapters, the MAP estimate equation is

$$\lambda H^T(\underline{g}-\underline{l})-R_f^{-1}(\underline{f}-\bar{\underline{f}}) = \underline{0} \quad , \tag{7.7}$$

or equivalenty

$$\hat{\underline{f}}_{MAP} = \bar{\underline{f}}+\lambda R_f H^T(\underline{g}-\underline{l}) \quad . \tag{7.8}$$

The physical interpretation of the MAP estimate is that maximizing the probability $p(\underline{d}|\underline{f})$ forces the solution toward the inverse solution which is the maximum likelihood solution, while maximizing the probability $p(\underline{f})$ is equivalent to enforcing a smoothness criterion. Thus, the MAP filter tries to balance the inverse solution with a smoothness constraint [7-5]. Another physical interpretation from Eq. (7.8) is that the MAP estimate tries to move the solution of the estimate $\hat{\underline{f}}_{MAP}$ from the *a priori* nonstationary mean $\hat{\underline{f}}$ to a maximum likelihood solution $\hat{\underline{f}}_{ML}$.

7.2.3 Conclusions

The LMMSE filter and MAP filter are performing similar

functions of balancing the inverse solution with a smoothness constraint. The LMMSE filter uses the equivalent SNR to control the balancing, while the MAP filter uses the covariance matrix R of the object as a measure of the confidence in the nonstationary mean $\hat{\underline{f}}$ and the maximum likelihood solution $\hat{\underline{f}}_{ML}$ as a solution to the restoration filter. The LMMSE filter is based on the assumption that the detected image intensity can be approximately modeled by a stationary random field. For a typical image, each part of an image generally differs sufficiently from other parts so that the stationarity is not generally valid. Moreover, the LMMSE filtering process is insensitive to abrupt changes of image signals. This results in edge smoothing and contrast reduction. The MAP filter does consider the nonstationary properties of random image fields because it contains an ML term and an a priori term. Also it is an adaptive processor which depends on the local nonstationary properties of image signals. The MAP filter theoretically needs more a priori knowledge, but in actual implementation, the MAP filter uses less a priori knowledge than the LMMSE filter. Both filters require knowledge of the blurring matrix H and mean number of photon counts $\bar{N}$ but the LMMSE filter also requires the spectral density of the object Φ_f. It can be intuitively concluded that the MAP filter should perform better than the LMMSE filter for the Poisson noise model.

7.3 Implementation and Experimental Results of the LMMSE Filter and the MAP Filter

Our experimental implementation of the LMMSE filter is based on Eq. (7.5) using a fast Fourier transform algorithm. The ensemble mean photon counts $\bar{N}$ and the object spectral density Φ_f are estimated from observed photon count data. The estimate of Φ_f can be made by substituting a similar "prototype" spectral density suggested by Cannon in the blind deconvolution process [7-6], or estimated by an iteration method suggested by Limb in a method of image restoration called spectral subtraction (SSIR) [7-7]. Unfortunately, they assume the *noise is linear signal-independent additive.* Because the Poisson noise is signal-dependent, the estimation of Φ_f is different.

The spectral density of the object Φ_f is related to the spectral density of the detected image by

$$\Phi_d(u,v) = \bar{N} + (\bar{N})^2 ||\mathcal{H}(u,v)||^2 \Phi_f(u,v) \qquad (7.9)$$

where $\bar{N}$ is ensemble mean number of photon counts and $\mathcal{H}(u,v)$ is the Fourier transform of the PSF $h(x,y)$. Thus, the spectral density of the object Φ_f can be estimated by an iterative method, although this involves an inverse filter with $\mathcal{H}$. We do not investigate the estimation of the spectral density Φ_f. Instead we assume that Φ_f is a white

spectral density in implementating the LMMSE filter. The white object spectral density assumption extracts more higher spatial frequency information such as edge and fine detail of the image because the spatial frequency content of many images falls rapidly at high spatial frequencies.

For generality, a two-dimensional moving average blurring point spread function (PSF) is chosen for the simulation rather than a Gaussian blurring PSF because it has singularities and phase reversals in the frequency response. Restored images with the LMMSE filter for diferent equivalent SNR's are illustrated in Fig. 7.1. The amount of restoration is controlled by the $\hat{\alpha}$. The serious effects of ill-conditioning can be seen in Fig. 7.1. The restored image gradually blows up as $\hat{\alpha}$ goes to higher values, while it becomes more noisy as $\hat{\alpha}$ goes to lower values. Because the Poisson noise degradation is very pronounced for small values of $\hat{\alpha}$ and the ill-conditioning of the inverse filter is worse at larger values of $\hat{\alpha}$, implementation of the LMMSE filter is very sensitive to the equivalent SNR $\hat{\alpha}$.

Figures 7.2 and 7.3 show results with the nonlinear sectioned MAP filter implemented with Newton-Raphson iterative techniques as before. Figure 7.2 has results with no blur, and Figure 7.3 shows results with linear blur. The experiments for both cases of degradation are

Figure 7.1 The restored images with the LMMSE filter for different estimated $(SNR)_{eq}$ $\hat{\alpha}$

(A) Restored image with the LMMSE filter for $\hat{\alpha}=10$
(B) Restored image with the LMMSE filter for $\hat{\alpha}=20$
(C) Restored image with the LMMSe filter for $\hat{\alpha}=40$
(D) Restored image with the LMMSE filter for $\hat{\alpha}=100$

Figure 7.2a Images restored by the LMMSE filter and the MAP filter at $\overline{(SNR)}_{rms}=\sqrt{2.5}$

(A) Original object image
(B) Poisson noisy image
(C) Image restored by the LMMSE filter
(D) Image restored by the MAP filter

Figure 7.2b Images restored by the LMMSE filter and the MAP filter at $\overline{(SNR)}_{rms}=\sqrt{5}$

(A) Original object image
(B) Poisson noisy image
(C) Restored image by the LMMSE filter
(D) Restored image by the MAP filter

Figure 7.2c Images restored by the LMMSE filter and the MAP filter at $\overline{(SNR)}_{rms}=\sqrt{10}$

(A) Original object image
(B) Poisson noisy image
(C) Restored image by the LMMSE filter
(D) Restored image by the MAP filter

Figure 7.2d Images restored by the LMMSE filter and the MAP filter at $\overline{(SNR)}_{rms}=\sqrt{20}$

(A) Original object image
(B) Poisson noisy image
(C) Restored image by the LMMSE filter
(D) Restored image by the MAP filter

Figure 7.3a Images restored by the LMMSE filter and the MAP filter with two-dimensional linear blurring degradation at $\overline{(SNR)}_{rms}=\sqrt{2.5}$

(A) Original object image
(B) Degraded image
(C) Restored image by the LMMSE filter
(D) Restored image by the MAP filter

Figure 7.3b Images restored by the LMMSE filter and the MAP filter with two-dimensional linear blurring degradation at $\overline{(SNR)}_{rms}=\sqrt{5}$

(A) Original object image
(B) Degraded image
(C) Restored image by the LMMSE filter
(D) Restored image by the MAP filter

Figure 7.3c Images restored by the LMMSE filter and the MAP filter with two-dimensional linear blurring degradation at $\overline{(SNR)}_{rms}=\sqrt{10}$

(A) Original object image
(B) Degraded image
(C) Restored image by the LMMSE filter
(D) Restored image by the MAP filter

Figure 7.3d Images restored by the LMMSE filter and the MAP filter with two-dimensional linear blurring degradation at $\overline{(SNR)}_{rms}=\sqrt{20}$

(A) Original object image
(B) Degraded image
(C) Restored image by the LMMSE filter
(D) Restored image by the MAP filter

performed with different $\overline{(SNR)}_{rms}$.

7.4 Numerical Values of Restored Image Quality Measures

One of the most commonly used numerical quality measures is the mean square error (MSE). It can be defined as

$$MSE = \sum_{m=1}^{N} \sum_{n=1}^{N} \overline{[f(m,n)-\hat{f}(m,n)]^2} \quad . \tag{7.10}$$

Although this measure is attractive because it is tractable, it does not match human evaluation on many types of images. It is also possible to define a measure based on the MSE and energy normalization [7-9]. This normalized mean square error (NMSE) is defined as

$$NMSE = \frac{\sum_{m=1}^{N} \sum_{n=1}^{N} [f(m,n)-\hat{f}(m,n)]^2}{\sum_{m=1}^{N} \sum_{n=1}^{N} [f(m,n)]^2} \quad . \tag{7.11}$$

The NMSE performs somewhat better than MSE and retains the analytic tractability. For these reasons, we use it in this chapter. Other numerical measures such as normalized error (NE), and Laplacian mean square error (LMSE) [7-8,7-9] etc. will not be used here. Based on Eq. (7.11), the NMSE of the restored image between the LMMSE filter and the MAP filter will be computed. The results are shown in Table 7.1 and Table 7.2 for non-blurring and blurring cases respectively.

7.5 Comparisons and Conclusions

The image quality of the pictures restored by the MAP filter shown in Figs. 7.2 and 7.3 seems superior to those processed by the LMMSE filter. Images restored by the LMMSE filter have an excessive enhancement of Poisson noise, especially for higher values of $\hat{\alpha}$. In addition, the NMSE of the MAP filter is lower than that of the LMMSE filter. The LMMSE filter strongly depends on perfect *a priori* knowledge of the object and it is very sensitive to parameters in the filter implementation. Table 7.1 and Table 7.2 show that the MAP filter has advantages over the LMMSE filter, particularly at lower $\overline{(SNR)_{rms}}$. Furthermore, the NMSE of the LMMSE filter does not improve greatly at increasing $\overline{(SNR)_{rms}}$ as it does with the MAP filter in the blurring case. A possible explanation for this is that the LMMSE filter performs very well for the higher (SNR), but the MAP filter also works better for the lower (SNR) at which Poisson noise dominates. The disadvantage is that the MAP filter is a nonlinear spatial estimate which needs iterative methods for solution. Thus, the computing time of the MAP filter is longer than that of the LMMSE filter. The better performance is achieved at the expense of additional computing time.

Table 7.1

The NMSE for nonblurring cases

$\overline{(SNR)}_{rms}$	LMMSE filter	MAP filter
$\sqrt{2.5}$	0.3601287	0.15927
$\sqrt{5}$	0.18131	0.11352
$\sqrt{10}$	0.090158	0.072803
$\sqrt{20}$	0.0449264	0.04180208

Table 7.2

The NMSE for blurring cases

$\overline{(SNR)}_{rms}$	LMMSE filter	MAP filter
$\sqrt{2.5}$	0.269397	0.190431
$\sqrt{5}$	0.1502184	0.0994055
$\sqrt{10}$	0.1466819	0.054126
$\sqrt{20}$	0.1433148	0.036138

CHAPTER 8

SUMMARY AND CONCLUSIONS

In this work we have modeled photon resolved image signals as a Poisson point random process and developed an optimal spatial restoration filter for the Poisson noise model.

Poisson noise is an inherent part of any detected image and is particularly evident in low level image signals. Because it results from the discrete random nature of quantum limitations, it is signal-dependent. The optimal spatial filter was based on a criterion of maximizing the a posteriori probability density. The formulation and solution of the MAP estimation problem have been presented. It has been found that the overlap-save sectioning method with a Newton-Raphson iterative solution is the most efficient way of coping with the nonlinearity and large dimensionality of the MAP estimation equations. The implementation of the MAP filter with the Poisson noise model was made for both blurring and non-blurring degradation cases. It has been demonstrated that the MAP filter with the Poisson noise model has improved performance because the MAP filter can be generalized to

linear or nonlinear image models and to noise models different from additive Gaussian noise. In addition, the MAP filter can be a local adaptive processing filter and can be extended to space-variant blurring. It also has been shown that modeling images with a nonstationary mean and stationary variance gives useful a priori information for the MAP filter.

The Cramer-Rao lower bound (CRLB) on the mean-square estimation error of the MAP unbiased estimate was derived for the Poisson noise model. It is likely that the CRLB may be useful for finding the best suboptimal sectioned MAP filter. A comparison was made between the LMMSE filter and the MAP filter with the Poisson noise model. It has been shown that the quality of the restored image of the MAP filter is superior to that of the LMMSE filter by simple subjective evaluation and by numerical closeness criteria.

Boulter [8-1] has shown that even with large amounts of blurring present in a detected image, a low noise level permits almost complete restoration. However, photon resolved image signals have a very low signal-to-noise ratio, making it impossible to obtain perfect restoration for image signals suffering from both blurring and Poisson noise.

The research pursued in this dissertation may be extended in several areas. A more detailed study of an

optimal local adaptive MAP filter using the local properties of the first and second moment of image signals would be of considerable interest. Use of these local statistics is interesting because image fields are inherently nonstationary. Another area of practical importance is to develop a fast algorithm for the MAP filter or a recursive MAP filter for saving computing time and memory space. A recursive MAP filter would not only offer computational advantages over a non-recursive filter but also could be applied to space-variant and nonstationary models. However it is expected that the recursive spatial restoration filter would be very sensitive to errors in the knowledge of the PSF. Another intereting area is to apply these results to other types of photon resolved image signals such as nuclear medicine, medical images, astronomical images and the projection reconstruction image signals [8-3]. In these cases, the image formation system model and data acquisition system must be carefully modeled to identify the parameters of the photon counting system. Another interesting possibility for the future is to use Lebedev's "composite" image model to develop a "multicategory" spatial MAP filter [8-4, 8-5]. This model involves multiple categories of random fields in the image with each category distinguished by its covariance. A Gaussian probability density is associated with the occurence of each category in the data. With this

model, the optimal estimate is obtained by filtering the measurement data for each category, and forming the estimate as the weighted sum of all the filter outputs. The weights are the *a posteriori* probabilites that the point is a member of the respective category. The composite image model locally decomposes the image signals, and it should be able to model the local nonhomogeneous information in images.

APPENDIX A

POISSON RANDOM NOISE GENERATOR [2-25]

For the simulation of photon count observation data, we need an algorithm for generating a sequence of random numbers from a population conforming to the Poisson distribution with mean. The solution is to make use of a random number generator which returns a random variable z having the uniform distribution $h(z)=1$ for $0 \leq z \leq 1$ and $h(z)=0$ for other values of z. If we form $x_k=z_0,z_1,z_2,\ldots,z_k$ as the product of a sequence of k+1 such random variables, then the lowest value of k which first cause x_k to be less than or equal to $e^{-\lambda}$, will be a random variable which has a Poisson distribution with mean λ. A proof of this property follows. If y and z are independent random variables with pdf g(y) and h(z) respectively, it can be shown that [2-17] x=yz has pdf

$$f(x) = \int g(y)h(\frac{x}{y})\frac{dy}{y} \quad . \tag{A.1}$$

Let $f_k(x_k)$ denote the pdf of x_k. Since $f(x_0)=h(x_0)=1$ we have, for $x_1=x_0z$

$$f_1(x_1) = \int_{x_1}^{1} g(x_0)\, h\left(\frac{x}{x_0}\right) \frac{dx_0}{x_0} \quad , \tag{A.2}$$

$$f_1(x_1) = \int_{x_1}^{1} 1 \cdot 1 \cdot \frac{dx_0}{x_0} = -\ell n\ x_1 \quad . \tag{A.3}$$

By induction, it can be shown that $x_k = x_{k-1} z_k$ has pdf

$$f_k(x_k) = \frac{(-1)^k}{k!}\ \ell n^k x_k \quad . \tag{A.4}$$

By use of the well known recursive relation

$$I_k = x\ \ell n^k x - k\ I_{k-1},$$

where (A.5)

$$I_k = \int \ell n^k x\ dx,$$

the probability that x_k is less than $e^{-\lambda}$ is

$$\begin{aligned} P_r(x_k \leq e^{-\lambda}) &= \int_0^{e^{-\lambda}} f_k(x_k)\, dx_k \\ &= \frac{(-1)^k}{k!}\ [x \ell n^k x - k\ I_{k-1}]^{e^{-\lambda}} \quad . \end{aligned} \tag{A.6}$$

Thus

$$P_r(x_k \leq e^{-\lambda}) = \frac{e^{-\lambda}\lambda}{k!} + P_r(x_{k-1} \leq e^{-\lambda}) \quad . \tag{A.7}$$

Using the same method to calculate $P_r(x_{k-1} \leq e^{-\lambda})$, we then have

$$P_r(x_k \leq e^{-\lambda}) = e^{-\lambda}\left[\frac{\lambda^k}{k!} + \frac{\lambda^{k-1}}{(k-1)!} + \ldots + \frac{\lambda^2}{2!} + \lambda + 1\right] \tag{A.8}$$

and we conclude that if $x_k \leq e^{-\lambda}$ but $x_{k-1} > e^{-\lambda}$, then k obeys the Poisson distribution with mean λ. This algorithm turns out to be a very accurate and fast way to generate the photon counts with given mean image intensities. This algorithm is also available in the IMSL subroutine package (GGPOSH) [2-28].

APPENDIX B

NEWTON-RAPHSON ITERATIVE SOLUTION METHOD FOR NONLINEAR MAP ESTIMATION EQUATIONS

A detailed derivation and procedure for the Newton-Raphson iterative method is described in [4-4]. Here, we summarize the Newton-Raphson solution of the MAP estimate equations. From Eq. (4.14), we get N MAP estimate equations, that is

$$g_1(\underline{f}) = \left(\frac{d_1}{f_1}-1\right)-\frac{r}{(1+\rho^2)}(f_1-\bar{f}_1)+\frac{\beta r}{\lambda}(f_2-\bar{f}_2) = 0, \quad \text{(B.1a)}$$

$$g_i(\underline{f}) = \left(\frac{d_i}{f_i}-1\right)+\frac{\beta r}{\lambda}(f_{i-1}-\bar{f}_{i-1})-\frac{r}{\lambda}(f_i-\bar{f}_i)+\frac{\beta r}{\lambda}(f_{i+1}-\bar{f}_{i+1}) = 0, \quad \text{(B.1b)}$$

$$i = 2,3,4,\ldots,N-1$$

$$g_N(\underline{f}) = \left(\frac{d_N}{f_N}-1\right)+\frac{\beta r}{\lambda}(f_{N-1}-\bar{f}_{N-1})-\frac{r}{(1+\rho^2)}(f_N-\bar{f}_N) = 0, \quad \text{(B.1c)}$$

where $\beta = \frac{\rho}{1+\rho^2}$, $r = \frac{1+\rho^2}{1-\rho^2}\cdot\frac{1}{\sigma_f^2}$ and

ρ is the correlation coefficients between pixels

$\underline{f}$ is the object vector to be estimated

$\bar{\underline{f}}$ is the nonstationary mean vector

Once the problem has been set up as in Eqs.(B.1a) to

(B.1c), the solution procedure is:

Step 1: Choose initial guess of object vector

$$\underline{f}^{(k)} = \underline{f}^{(0)} = [\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_N]^T \tag{B.2a}$$

or

$$\underline{f}^{(k)} = \underline{f}^{(0)} = [d_1, d_2, \ldots, d_N]^T \tag{B.2b}$$

where $\bar{f}$ is the nonstationary mean and $\underline{d}$ is the noisy observation data vector. The choice of initial values affects only the convergence rate because the equations finally converge to a unique set of roots. Here, the superscript k denotes the kth iterative step

Step 2: Solve the linear system Eq. (B.3) to find the solution vector $\delta(\underline{f}^{(k)})$

$$\underline{\phi}(\underline{f}^{(k)})\underline{\delta}^{(k)} = -\underline{g}(\underline{f}^{(k)}) \tag{B.3}$$

where

$$\phi_{ij}(\underline{f}^{(k)}) = \frac{\partial g_i}{\partial f_j}(\underline{f}^{(k)}),$$

$$\underline{g}(\underline{f}^{(k)}) = [g_1(\underline{f}^{(k)}), g_2(\underline{f}^{(k)}), \ldots, g_N(\underline{f}^{(k)})]^T$$

and incremental vector

$$\underline{\delta}(\underline{f}^{(k)}) = [\delta_1(\underline{f}^{(k)}, \delta_2(\underline{f}^{(k)}), \ldots, \delta_N(\underline{f}^{(k)})]^T$$

From Eq. (B.1) and $\Phi_{ij}(\underline{f}^{(k)}) = \frac{\partial g_i}{\partial f_j}(\underline{f}^{(k)})$, we obtain

$$\frac{\partial g_1}{\partial f_1} = -\frac{d_1}{f_1^2} - \frac{r}{(1+\rho^2)} \tag{B.4a}$$

$$\frac{\partial g_1}{\partial f_2} = \frac{\beta r}{\lambda} \tag{B.4b}$$

$$\frac{\partial g_i}{\partial f_{i-1}} = \frac{\beta r}{\lambda} \tag{B.4c}$$

$$\frac{\partial g_i}{\partial f_i} = -\frac{d_i}{f_i^2} - \frac{r}{\lambda} \tag{B.4d}$$

$$\frac{\partial g_N}{\partial f_{N-1}} = \frac{\beta r}{\lambda} \tag{B.4e}$$

$$\frac{\partial g_N}{\partial f_N} = -\frac{d_N}{f_N^2} - \frac{1}{(1+\rho^2)} \tag{B.4f}$$

These equations can be arranged as a matrix

$$\underline{\phi}(\underline{f}^{(k)}) = \begin{bmatrix} \frac{\partial g_1}{\partial f_1} & \frac{\partial g_1}{\partial f_2} & & & 0 \\ \frac{\partial g_2}{\partial f_1} & \frac{\partial g_2}{\partial f_2} & \frac{\partial g_2}{\partial f_3} & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{\partial g_i}{\partial f_{i-1}} & \frac{\partial g_i}{\partial f_i} & \frac{\partial g_i}{\partial f_{i+1}} \\ & & & \ddots & \ddots \\ 0 & & & \frac{\partial g_N}{\partial f_{N-1}} & \frac{\partial g_N}{\partial f_N} \end{bmatrix} \tag{B.5}$$

From Eq. (B.4), we know that the absolute values of

diagonal entries are larger than that of off diagonal entries. Thus, Eq. (B.3) equations will be converged.

Step 2 for solving the linear Eq. (B.3) can be summarized as follows:

(a) Substitute $\underline{f}^{(k)}$ into Eq. (B.1) and (B.4) to find the values of $\underline{g}(\underline{f}^{(k)})$ and $\underline{\phi}(\underline{f}^{(k)})$

(b) Solve Eq. (B.3) using a linear equation subroutine to find the solution $\underline{\delta}(\underline{f}^{(k)})$

Step 3: Update the approximation to the root for the next iteration

$$\underline{f}^{(k+1)} = \underline{f}^{(k)} + \underline{\delta}(\underline{f}^{(k)}) \tag{B.6}$$

Step 4: Check for possible convergence to a real root $\underline{f}$ by applying the test

$$\delta_i(\underline{f}^{(k)}) < \varepsilon \quad \text{for all } i. \tag{B.7}$$

Step 5: If Eq. (B.7) is true for all i, then $\underline{f}^{(k+1)}$ is taken to be the root. If Eq. (B.7) fails for any i, then the process is repeated starting from step 2.

In short, this iterative algorithm converges very fast, usually in about two to three steps in our simulation.

REFERENCES

[1-1] W.K. Pratt, Digital Image Processing, Wiley-Interscience, New York, 1978.

[1-2] H.C. Andrews and B.R. Hunt, Digital Image Restoration, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[1-3] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Wiley, New York, 1968.

[2-1] W.K. Pratt, Digital Image Processing, Wiley-Interscience, New York, 1978.

[2-2] H.C. Andrews and B.R. Hunt, Digital Image Restoration, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[2-3] B.R. Hunt, "Bayesian Methods in Nonlinear Digital Image Restoration," IEEE Transactions on Computers, vol. C-26, pp. 219-229, March 1977.

[2-4] H.J. Trussell and B.R. Hunt, "Sectioned Methods for Image Restoration," IEEE Transaction on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 157-164, April 1978.

[2-5] H.J. Trussell and B.R. Hunt, "Notes on Linear Image Restoration by Maximizing the A Posteriori Probability," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 174-176, April 1978.

[2-6] H. Kato and J.W. Goodman, "Nonlinear Filtering in Coherent Optical Systems Through Halftone Screen Processes," Appl. Opt., vol. 14, pp. 1813-1834, August 1975.

[2-7] J.F. Walkup and R.C. Choens, "Image Processing In Signal-Dependent Noise," Optical Engineering, vol. 13, pp. 258-266, May/June, 1974.

[2-8] K. Kondo, Y. Ichioka, and T. Suzuki, "Image Restoration by Wiener Filtering in the presence of Signal-Dependent Noise," Applied Optics, vol 16, pp. 2554-2558, September 1977.

[2-9] T. Yatagai, "Optimum Spatial Filter for Image Restoration Degraded by Multiplicative Noise," Optics Communications, vol. 19, pp. 236-239, November 1976.

[2-10] F. Naderi and A.A. Sawchuk, "Estimation of Images Degraded by Film-grain Noise," Applied Optics, vol. 17, pp. 1228-1237, April 1978.

[2-11] G.K. Froehlich, J.F. Walkup and R.B. Asher, "Optimal Estimation in Signal-Dependent Noise," J. Opt. Soc. Am., vol. 68, pp. 1665-1672, December 1978.

[2-12] B.R. Hunt and T.M. Cannon, "Nonstationary Assumptions for Gaussian Models of Images," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-6, pp. 876-882, December 1976.

[2-13] H.J. Trussell and R.P. Kruger, "Comment on Nonstationary Assumptions for Gaussian Models in Images," submitted to IEEE Transactions on Systems, Man and Cybernetics, 1978.

[2-14] S.K. Mitra and M.P. Ekstrom, editor, Two Dimensional Digital Signal Processing, Dowden, Hutchinson & Ross, Inc., Stroudsburg, Pennsylvania, 1978.

[2-15] D.S. Lebedev and L.I. Mirkin, "Digital Nonlinear Smoothing of Images," Institute for Information Transmission Problems, Academy of Sciences, USSR, 1975, pp. 150-157.

[2-16] P.A. Wintz, "Transform Picture Coding," Proc. IEEE, vol. 60, pp. 809-820, July 1972.

[2-17] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1965.

[2-18] R.O. Duda, P.E. Hart, Pattern Classification and Scene Analysis, Wiley-Interscience, New York, 1973.

[2-19] J.W. Goodman, Statistical Optics Notes, to be published by Wiley, New York.

[2-20] J.F. Walkup, "Limitation in Interferometric Measurements and Image Restoration at Low Light Levels," Ph.D. Dissertation, Stanford University, July 1971.

[2-21] J.F. Walkup, J.W. Goodman, "Limitations of Fringe-Parameter Estimation at Low Light Levels," J. of Opt. Soc. of Am., vol. 63, pp. 399-407, April 1973.

[2-22] S. Karp and J.R. Clark, "Photon Counting: A Problem in Classical Noise Theory," IEEE Transactions on Information Theory, vol. IT-16, pp. 672-680, November 1970.

[2-23] L.L. Wang, "Detection and Estimation of An Optical Image by Photon-Counting Techniques," Ph.D. Dissertation, University of California, San Diego, Ca. 1973.

[2-24] A. Rose, Vision: Human And Electronic, Plenum Press, New York, 1974.

[2-25] B. Carnahan, H.A. Luther and J.O. Wilkes, Applied Numerical Methods, Wiley-Interscience, New York, 1969.

[2-26] L. Mandel, "Fluctuations of Photon Beams: The Distribution of the Photon-Electrons," Proc. Phys. Soc., vol. 74, pp. 233-243, September 1959.

[2-27] W.T. Mayo Jr., "Photon Counting Processor for Laser Velocimetry," Applied Optics, vol. 16, pp. 1157-1162, May 1977.

[2-28] International Mathematical and Statistical Libraries, 7500 Bellaire, Houston, Texas 77036, 1977.

[3-1] W.K. Pratt, Digital Image Processing, Wiley-Interscience, New York, 1978.

[3-2] H.C. Andrews and B.R. Hunt, Digital Image Restoration, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[3-3] B.R. Hunt, "Bayesian Methods in Nonlinear Digital Image Restoration," IEEE Transactions on Computers, vol. C-26, pp. 219-229, March 1977.

[3-4] H.J. Trussell and B.R. Hunt, "Sectioned Methods for Image Restoration," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 157-164, April 1978.

[3-5] H.J. Trussell and B.R. Hunt, "Notes on Linear Image Restoration by Maximizing the A Posteriori Probability," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 174-176, April 1978.

[3-6] H. Kato and J.W. Goodman, "Nonlinear Filtering in Coherent Optical Systems through Halftone Screen Processes," Appl. Opt., vol. 14, pp. 1813-1824, August 1975.

[3-7] J.F. Walkup and R.C. Choens, "Image Processing In Signal-Dependent Noise," Optical Engineering, vol. 13, pp. 258-266, May/June, 1974.

[3-8] K. Kondo, Y. Ichioka, and T. Suzuki, "Image Restoration by Wiener Filtering in the presence of Signal-Dependent Noise," Applied Optics, vol. 16, pp. 2554-2558, September 1977.

[3-9] T. Yatagai, "Optimum Spatial Filter for Image Restoration Degraded by Multiplicative Noise," Optics Communications, vol. 19, pp. 236-239, November 1976.

[3-10] F. Naderi and A.A. Sawchuk, "Estimation of Images Degraded by Film-grain Noise," Applied Optics, vol. 17, pp. 1228-1237, April 1978.

[3-11] G.K. Froehlich, J.F. Walkup and R.B. Asher, "Optimal Estimation in Signal-Dependent Noise," J. Opt. Soc. Am., vol. 68, pp. 1665-1672, December 1978.

[3-12] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1965.

[3-13] J.W. Goodman and J.F. Belsher, "Fundamental Limitations in Linear Invariant Restoration of Atmospherically Degraded Images," Proc. SPIE 75, pp. 141-154, 1976.

[3-14] J.W. Goodman, Statistical Optics Notes, 1976.

[3-15] D.L. Snyder, Random Point Processes, Wiley-Interscience, New York, 1975.

[3-16] T. Inouye, "Computer Processing of Scintillation Camera Images," Nuclear Instruments and Method, vol. 124, pp. 215-219, 1975.

[3-17] D.L. Fried, "Signal Processing for a Signal with Poisson Noise," Applied Optics, vol. 13, pp. 1282-1283, June 1974.

[3-18] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Wiley, New York, 1968.

[3-19] J.J. Burke, "Estimating Objects from their Blurred and Grainy Images," Proc. IEEE International Optical Computing Conference, pp. 48-51, Washington, D.C., April 1975.

[3-20] J.F. Boulter, "Digital Restoration of Blurred Photographs," DREV REPORT 4107, Department of National Defense, Canada, February 1978.

[3-21] L. Mandel, "Fluctuation of Photon Beams: The Distribution of the Photo-Electrons," Proc. Phys. Soc., vol. 74, pp. 233-243, September 1959.

[4-1] B.R. Hunt and T.M. Cannon, "Nonstationary Assumptions for Gaussian Models of Images," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-6, pp. 876-882, December 1976.

[4-2] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Wiley, New York, 1968.

[4-3] H.J. Trussell, "Improved Methods of Maximum A Posteriori Image Restoration," Ph.D. Dissertation, University of New Mexico, 1976.

[4-4] B. Carnahan, H.A. Luther and J.O. Wilkes, Applied Numerical Methods, Wiley-Interscience, New York, 1969.

[4-5] E.K. Blum, Numerical Analysis and Computation: Theory and Practice, Addison-Wesley, Reading, Massachusetts, 1972.

[4-6] C.F. Hall and E.L. Hall, "A Nonlinear Model for the Spatial Characteristics of the Human Visual Systems," IEEE Transactions on Systems, Man and Cybernetics, vol. SMC-7, pp. 161-170, March 1977.

[4-7] T.G. Stockham, Jr., "Image Processing in the Context of a Visual Model," Proc. IEEE, vol. 60, pp. 828-842, 1972.

[4-8] H.J. Trussell and B.R. Hunt, "Sectioned Methods for Image Restoration," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 157-164, April 1978.

[4-9] H.J. Trussell and B.R. Hunt, "Image Restoration of Space-Variant Blurs by Sectioned Methods," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 608-609, December 1978.

[4-10] C.M. Lo, A.A. Sawchuk, "Estimation of Image Signals with Poisson Noise Model," USC, IPI Report 840, pp. 135-156, September 1978.

[4-11] W.K. Pratt, Digital Image Processing, Wiley-Interscience, New York, 1978.

[4-12] A.K. Jain, "Some New Techniques in Image Processing," Image Science Mathematics, pp. 201-223, 1976.

[4-13] B.R. Hunt and T.M. Cannon, "Nonstationary Assumptions for Gaussian Models of Images," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-6, pp. 876-882, December 1976.

[4-14] A.S. Willsky, "Relationships Between Digital Signal Processing and Control and Estimation Theory," Proc. IEEE, vol. 66, pp. 996-1017, September 1978.

[5-1] J.B. Thomas, An Introduction to Statistical Communication Theory, Wiley, New York, 1968.

[5-2] W.K. Pratt, Digital Image Processing, Wiley-Interscience, New York, 1978.

[5-3] H.C. Andrews and B.R. Hunt, Digital Image Restoration, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[5-4] H.J. Trussell and B.R. Hunt, "Sectioned Methods for Image Restoration," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 157-164, April 1978.

[5-5] T.G. Stockham, Jr., T.M. Cannon, and R.B. Ingebretsen, "Blind Deconvolution Through Digital Signal Processing," Proc. IEEE, vol. 63, pp. 678-692, April 1975.

[5-6] M. Cannon, "Blind Deconvolution of Spatially Invariant Image Blurs with Phase," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-24, pp. 58-63, February 1976.

[5-7] E.R. Cole, "The Removal of Unknown Image Blurs by Homomorphic Filter," Ph.D. Dissertation, University of Utah, 1973.

[5-8] T.M. Cannon, "Digital Image Deblurring by Nonlinear Homomorphic Filteing," Ph.D. Dissertation, University of Utah, 1974.

[5-9] J.B. Morton, "An Investigation Into An A Posteriori Method of Image Restoration," Ph.D. Dissertation, University of Southern California, 1978.

[5-10] L.R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing, Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[5-11] J.W. Goodman, Statistical Optics Notes, to be published by Wiley, New York.

[5-12] J.W. Goodman, Introduction to Fourier Optics, McGraw-Hill, New York, 1968.

[5-13] H.J. Trussell and B.R. Hunt, "Image Restoration of Space-Variant Blurs by Sectioned Methods," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 608-609, December 1978.

[5-14] R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.

[5-15] J.B. Morton, H.C. Andrews, "A Posteriori Method of Image Restoration," J. Opt. Soc. Amer. , vol.69, pp. 280-290, February 1979.

[6-1] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Wiley, New York, 1968.

[6-2] J.B. Thomas, An Introduction to Statistical Communication Theory, Wiley, New York, 1969.

[6-3] N.E. Nahi, Estimation Theory and Application, Wiley, New York, 1976.

[6-4] R. Deutsch, Estimation Theory, Prentice-Hall, Englewood Cliffs, New Jersey, 1965.

[6-5] D. Middleton, An Introduction to Statistical Communication Theory, McGraw-Hill, New York, 1960.

[6-6] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.

[6-7] W.K. Pratt, Digital Image Processing, Wiley-Interscience, New York, 1978.

[6-8] H.C. Andrews and B.R. Hunt, Digital Image Restoration, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[6-9] H.C. Andrews, Tutorial and Selected Papers in Digital Image Processing, IEEE Computer Society, New York, 1978.

[6-10] B.R. Hunt, "Bayesian Methods in Nonlinear Digital Image Restoration," IEEE Transactions on Computers, vol. C-26, pp. 219-229, March 1977.

[6-11] B.R. Hunt, "Approximate Error Bounds for Bayesian Methods of Nonlinear Digital Image Restoration," Technical Report, University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico.

[6-12] H.J. Trussell and B.R. Hunt, "Sectioned Methods for Image Restoration," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 157-164, April 1978.

[6-13] H.J. Trussell and B.R. Hunt, "Note on Linear Image Restoration by Maximizing the A Posteriori Probability," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 174-176, April 1978.

[6-14] J.W. Goodman and J.F. Belsher, "Fundamental Limitations in Linear Invariant Restoration of Atmospherically Degraded Images," Proc. SPIE 75, pp. 141-154, 1976.

[6-15] H. Cramer, Mathematical Methods of Statistics, Princeton University Press, Princeton, N.J., 1946.

[7-1] T.M. Cannon, H.J. Trussell and B.R. Hunt, "A Comparison of Image Restoration Methods," Submitted to Applied Optics.

[7-2] W.A. Pearlman, "A Visual System Model and A New Distortion Measure in the Context of Image Processing," J. Opt. Soc. Am., vol. 68, pp. 374-386, March 1978.

[7-3] J.W. Goodman and J.F. Belsher, "Fundamental Limitations in Linear Invariant Restoration of Atmospherically Degraded Images," Proc. of SPIE 75, pp. 141-154, 1976.

[7-4] J.W. Goodman, "Precompensation and Postcompensation of Photon Limited Degraded Images," RADC-TR-76-382 ARPA Report, Stanford University, December 1976.

[7-5] H.J. Trussell, "Notes on Linear Image Restoration by Maximizing the A Posteriori Probability," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-26, pp. 174-176, April 1978.

[7-6] T.M. Cannon, "Digital Image Deblurring by Nonlinear Homomorphic Filtering," Ph.D. Dissertation, University of Utah, 1974.

[7-7] J.S. Lim, "Image Restoration by Short Space Spectral Subtraction," Report JA 4943, MIT Lincoln Laboratory, Lexington, Massachusetts, February 1979.

[7-8] C.F. Hall, "Digital Color Image Compression in A Perceptual Space," Ph.D. Dissertation, IPI, University of Southern California, February 1978.

[7-9] W.K. Pratt, Digital Image Processing, Wiley-Interscience, New York, 1978.

[7-10] N.E. Nahi, Estimation Theory and Application, Wiley, New York, 1976.

[7-11] R. Deutsch, Estimation Theory, Prentice-Hall, Englewood Cliffs, New Jersey, 1965.

[7-12] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.

[7-13] C.W. Helstrom, "Image Restoration by the Method of Least Square," J. Opt. Soc. of Ame., vol. 57, pp. 297-303, March 1967.

[7-14] H.C. Andrews and B.R. Hunt, Digital Image Restoration, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[7-15] T.S. Huang, W.F. Schreiber, O.J. Tretiak, "Image Processing," Proc. IEEE, vol. 59, pp. 1586-1609, November 1971.

[7-16] J.F. Boulter, "Use of Two-Dimensional Digital Fourier Transform for Image Processing and Analysis," DREV Report R-4025/75, Department of National Defense, Canada, July 1975.

[7-17] A.V. Oppenheim, J.S. Lim, Gary Kopec, and S.C. Pohlig, "Phase in Speech and Pictures," Technical Report, MIT Lincoln Laboratory, Lexington, Massachusetts, 1978.

[8-1] J.F. Boulter, "Digital Restoration of Blurred Photographs," DREV Report 4107, Department of National Defense, Canada, February 1978.

[8-2] A.S. Willsky, "Relationships Between Digital Signal Processing and Control and Estimation Theory," Proc. IEEE, vol. 66, pp. 96-1017, September 1978.

[8-3] L.A. Shepp and B.F. Logan, "The Fourier Reconstruction of A Head Section," IEEE Transactions on Nuclear Science, vol. NS-21, pp. 21-34, June 1974.

[8-4] S.K. Mitra and M.P. Ekstrom, editors, Two Dimensional Digital Signal Processing, Dowden, Hutchinson and Ross, Inc., Stroudsburg, Pennsylvania, 1978.

[8-5] D.S. Lebedev and L.I. Mirkin, "Digital Nonlinear Smoothing of Images," Institute for Information Transmission Problems, Academy of Sciences, USSR, 1975, pp. 150-157.