MAXIMUM LIKELIHOOD IMAGING

Imaging science is a rich and vital branch of engineering in which electromagnetic or acoustic signals are measured, processed, analyzed, and interpreted in the form of multidimensional images. Because these images often contain information about the physical, biological, or operational properties of remote objects, scenes, or materials, imaging science is justly considered to be a fundamental component of that branch of engineering and science known as remote sensing. Many subjects benefit directly from advances in imaging science—these range from astronomy and the study of very large and distance objects to microscopy and the study of very small and nearby objects.

The photographic camera is probably the most widely known imaging system in use today. The familiar imagery recorded by this device usually encodes the spectral reflectance properties of an object or scene onto a two-dimensional plane. The familiarity of this form of imagery has led to a common definition of an image as "an optical reproduction of an object by a mirror or lens." There are, however, many other imaging systems in use and the object or scene properties encoded in their imagery can be very different from

those recorded by a photographic camera. Temperature variations can, for instance, be "imaged" with infrared sensing, velocity variations can be "imaged" with radar, geological formations can be "imaged" with sonar, and the physiological function of the human brain can be "imaged" with positron emission tomography (PET).

A photographic camera forms images in a manner very similar to the human eye, and, because of this, photographic images are easily interpreted by humans. The imagery recorded by an infrared camera might contain many of the features common to visible imagery; however, the phenomena being sensed are different and some practice is required before most people can faithfully interpret raw infrared imagery. For both of these modalities, though, the sensor data is often displayed as an image without the need for significant signal processing. The data acquired by an X-ray tomograph or synthetic aperture radio telescope, however, are not easily interpreted, and substantial signal processing is required to form an "image." In these situations, the processing of raw sensor data to form imagery is often referred to as image reconstruction or image synthesis (1), and the importance of signal processing in these applications is great. To confirm this importance, the 1979 Nobel prize in physiology and medicine was awarded to Alan M. Cormack and Sir Godfrey N. Hounsfield for the development and application of the signal processing methods used for X-ray computed tomography, and the 1974 Nobel prize in physics was awarded to Sir Martin Ryle for the development of aperture synthesis techniques used to form imagery with radio telescope arrays. For both of these modalities the resulting images are usually very different from the visible images formed by photographic cameras, and significant training is required for their interpretation.

Imagery formed by photographic cameras, and similar instruments such as telescopes and microscopes, can also be difficult to interpret in their raw form. Focusing errors, for example, can make imagery appear blurred and distorted, as can significant flaws in the optical instrumentation. In these situations, a type of signal processing known as image restoration (2,3) can be used to remove the distortions and restore fidelity to the imagery. Processing such as this received national attention after the discovery of the Hubble Space Telescope aberrated primary mirror in 1990, and one of the most successful and widely used algorithms for restoring resolution to Hubble imagery was based on the maximum-likelihood estimation method (4). The motivation for and derivation of this image-restoration algorithm will be discussed in great detail later in this article.

When signal processing is required for the formation or improvement of imagery, the imaging problem can usually be posed as one of statistical inference. A large number of estimation-theoretic methods are available for solving statistical-inference problems (5–9), and the method to be used for a particular application depends largely on three factors: (1) the structure imposed on the processing; (2) the quantitative criteria used to define image quality; and (3) the physical and statistical information available about the data collection process.

Structure can be imposed on processing schemes for a variety of reasons, but the most common is the need for fast and inexpensive processing. The most common structure imposed for this reason is linear processing, whereby imagery is formed or improved through linear combinations of the measured data. In some situations structural restrictions such as these are acceptable, but in many others they are not and the advent of faster and more sophisticated computing resources has served to greatly lessen the need for and use of structural constraints in imaging problems.

Many criteria can be used to quantify image quality and induce optimal signal-processing algorithms. One might ask, for example, that the processed imagery produce the "correct" image on average. This leads to an unbiased estimator, but such an estimator may not exist, may not be unique, or may result in imagery whose quality is far from adequate. By requiring that the estimated image also have, in some sense, the smallest deviations from the correct image this criterion could be modified to induce the minimum variance, unbiased estimator (MVUE), whose imagery may have desirable qualities, but whose processing structure can be difficult or impossible to derive and implement. The maximum-likelihood method for estimation leads to an alternative criterion whereby an image is selected to optimize a mathematical cost function that is induced by the physical and statistical model for the acquired data. The relative simplicity of the maximum-likelihood estimation method, along with the fact that maximum-likelihood estimates are often asymptotically unbiased with minimum variance, makes this a popular and widely studied method for statistical inference. It is largely for this reason that the development and utilization of maximum-likelihood estimation methods for imaging are the focus of this article.

One of the most important steps in the utilization of the maximum-likelihood method for imaging is the development of a practical and faithful model that represents the relationship between the object or scene being sensed and the data recorded by the sensor. This modeling step usually requires a solid understanding of the physical and statistical characteristics of electromagnetic- or acoustic-wave propagation, along with an appreciation for the statistical characteristics of the data acquired by real-world sensors. For these reasons, a strong background in the fields of Fourier optics (10,11), statistical optics (12–14), basic probability and random-process theory (15,16), and estimation theory (5–9) is essential for one wishing to apply maximum-likelihood methods to the field of imaging science.

Statistical inference problems such as those encountered in imaging applications are frequently classified as ill-posed problems (17). An image-recovery or -restoration problem is ill posed if it is not well posed, and a problem is well posed in the classical sense of Hadamard if the problem has a unique solution and the solution varies continuously with the data. Abstract formulations of image recovery and restoration problems on infinite-dimensional measurement and parameter spaces are almost always ill posed, and their ill-posed nature is usually due to the discontinuity of the solution. Problems that are formulated on finite-dimensional spaces are frequently well-posed in the classical sense—they have a unique solution and the solution is continuous in the data. These problems, however, are often ill conditioned or badly behaved and are frequently classified as ill posed even though they are technically well posed.

For problems that are ill posed or practically ill posed, the original problem's solution is often replaced by the solution to a well-posed (or well-behaved) problem. This process is referred to as regularization and the basic idea is to change the

problem in a manner such that the solution is still meaningful but no longer badly behaved (18). The consequence for imaging problems is that we do not seek to form a "perfect" image, but instead settle for a more stable—but inherently biased—image. Many methods are available for regularizing maximum-likelihood estimation problems, and these include: penalty methods, whereby the mathematical optimization problem is modified to include a term that penalizes unwanted behavior in the object parameters (19); sieve methods, whereby the allowable class of object parameters is reduced in some manner to exclude those with unwanted characteristics (20); and stopping methods, whereby the numerical algorithms used to solve a particular optimization problem are prematurely terminated before convergence and before the object estimate has obtained the unwanted features that are characteristic of the unconstrained solution obtained at convergence (21). Penalty methods can be mathematically, but not always philosophically, equivalent to the maximum a posteriori (MAP) method, whereby an a priori statistical model for the object is incorporated into the estimation procedure. The MAP method is appealing and sound provided that a physically justified model is available for the object parameters. Each of these regularization methods is effective at times, and the method used for a particular problem is often a matter of personal taste.

SCALAR FIELDS AND COHERENCE

Because most imaging problems involve the processing of electromagnetic or acoustic fields that have been measured after propagation from a remote object or scene, a good place to begin our technical discussion is with a review of scalar waves and the concept of coherence. The scalar-wave theory is widely used for two reasons: (1) acoustic wave propagation is well-modeled as a scalar phenomenon; and (2) although electromagnetic wave propagation is a vector phenomenon, the scalar theory is often appropriate, particularly when the dimensions of interest in a particular problem are large in comparison to the electromagnetic field wavelength.

A scalar field is in general described by a function in four dimensions $s(x, y, z; t)$, where x, y, and z are coordinates in three-dimensional space, and t is a coordinate in time. In many situations, the field fluctuations in time are concentrated about some center frequency f_0, so that the field can be conveniently expressed as

$$s(x, y, z; t) = a(x, y, z; t) \cos [2\pi f_0 t + \theta(x, y, z; t)] \tag{1}$$

or, in complex notation, as

$$s(x, y, z; t) = \mathrm{Re}\{u(x, y, z; t) e^{j2\pi f_0 t}\} \tag{2}$$

where

$$u(x, y, z; t) = a(x, y, z; t) e^{j\theta(x, y, z; t)} \tag{3}$$

is the complex envelope for the field. Properties of the field amplitude a, phase θ, or both are often linked to physical or operational characteristics of a remote object or scene, and the processing of remotely sensed data to determine these properties is the main goal in most imaging applications.

Coherence is an important concept in imaging that is used to describe properties of waveforms, sensors, and processing algorithms. Roughly speaking, coherence of a waveform refers to the degree to which a deterministic relationship exists between the complex envelope phase $\theta(x, y, z; t)$ at different time instances or spatial locations. Temporal coherence at time delay τ quantifies the relationship between $\theta(x, y, z; t)$ and $\theta(x, y, z; t + \tau)$, whereas the spatial coherence at spatial shift $(\Delta_x, \Delta_y, \Delta_z)$ quantifies the relationship between $\theta(x, y, z; t)$ and $\theta(x + \Delta_x, y + \Delta_y, z + \Delta_z; t)$. A coherent sensor is one that records information about the complex-envelope phase of a waveform, and a coherent signal-processing algorithm is one that processes this information. Waveforms that are coherent only over vanishingly small time delays are called temporally incoherent; waveforms that are coherent only over vanishingly small spatial shifts are called spatially incoherent. Sensors and algorithms that neither record nor process phase information are called incoherent.

Many phenomena in nature are difficult, if not impossible within our current understanding, to model in a deterministic manner, and the statistical properties of acoustic and electromagnetic fields play a fundamental role in modeling the outcome of most remote sensing and imaging experiments. For most applications an adequate description of the fields involved is captured through second-order averages known as coherence functions. The most general of these is the mutual coherence function, which is defined mathematically in terms of the complex envelope for a field as

$$\Gamma_{12}(\tau) = E[u(x_1, y_1, z_1, t + \tau) u^*(x_2, y_2, z_2, t)] \tag{4}$$

The proper interpretation for the expectation in this definition depends largely on the application, and much care must be taken in forming this interpretation. For some applications a definition involving time averages will be adequate, whereas other applications will call for a definition involving ensemble averages.

The mutual coherence function is often normalized to form the complex degree of coherence as

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{[\Gamma_{11}(0)\Gamma_{22}(0)]^{1/2}} \tag{5}$$

and it is tempting to define a coherent field as one for which $|\gamma_{12}(\tau)| = 1$ for all pairs of spatial locations, (x_1, y_1, z_1) and (x_2, y_2, z_2), and for all time delays, τ. Such a definition is overly restrictive and a less restrictive condition, as discussed by Mandel and Wolf (22), is that

$$\max_{\tau} |\gamma_{12}(\tau)| = 1 \tag{6}$$

for all pairs of spatial locations, (x_1, y_1, z_1) and (x_2, y_2, z_2). Although partial degrees of coherence are possible, fields that are not coherent are usually called incoherent. In some situations a field is referred to as being fully incoherent over a particular region and its mutual coherence function is modeled over this region as

$$\Gamma_{12}(\tau) \simeq \kappa I(x_1, y_1, z_1) \delta_3(x_1 - x_2, y_1 - y_2, z_1 - z_2) \delta_1(t - \tau) \tag{7}$$

where $I(\cdot)$ is the incoherent intensity for the field, $\delta_3(\cdot, \cdot, \cdot)$ is the three-dimensional Dirac impulse, $\delta_1(\cdot)$ is the one-di-

mensional Dirac impulse, and κ is a constant with appropriate units. Most visible light used by the human eye to form images is fully incoherent and fits this model. Goodman (13) and Mandel and Wolf (22) provide detailed discussions of the coherence properties of electromagnetic fields.

INCOHERENT IMAGING

Astronomical telescopes, computer assisted tomography (CAT) scanners, PET scanners, and many forms of light microscopes are all examples of incoherent imaging systems; the waveforms, sensors, and algorithms used in these situations are all incoherent. The desired image for these systems is typically related to the intensity distribution of a field that is transmitted through, reflected by, or emitted from an object or scene of interest. For many of these modalities it is common to acquire data over a variety of observing scenarios, and the mathematical model for the signal acquired by these systems is of the form

$$I_k(y) = \int h_k(y,x) I(x)\, dx, \quad k = 1, 2, \ldots, K \tag{8}$$

where $I(\cdot)$ is the object incoherent intensity function—usually related directly to the emissive, reflective, or transmissive properties of the object, $h_k(\cdot\,,\,\cdot)$ is the measurement kernel or system point-spread function for the kth observation, $I_k(\cdot)$ is the incoherent measurement signal for the kth observation, x is a spatial variable in two- or three-dimensions, and y is usually a spatial variable in one-, two-, or three-dimensions. The mathematical forms for the system point-spread functions $\{h_k(\cdot\,,\,\cdot)\}$ are induced by the physical properties of the measurement system, and much care should be taken in their determination. In telescope and microscope imaging, for example, the instrument point-spread functions model the effects of diffraction, optical aberrations, and inhomogeneities in the propagation medium; whereas for transmission or emission tomographs, geometrical optics approximations are often used and the point-spread functions model the system geometry and detector uncertainties.

For situations such as astronomical imaging with ground-based telescopes, each measurement is in the form of a two-dimensional image, whereas for tomographic systems each measurement may be in the form of a one-dimensional projection of a two-dimensional transmittance or emittance function. In either situation, the imaging task is to reconstruct the intensity function $I(\cdot)$ from noisy measurements of $I_k(\cdot)$, $k = 1, 2, \ldots, K$.

Quantum Noise in Incoherent Imagery

Light and other forms of electromagnetic radiation interact with matter in a fundamentally random manner, and, because of this, statistical models are often used to describe the detection of optical waves. Quantum electrodynamics (QED) is the most sophisticated theory available for describing this phenomenon; however, a semiclassical theory for the detection of electromagnetic radiation is often sufficient for the development of sound and practical models for imaging applications. When using the semiclassical theory, electromagnetic energy is transported according to the classical theory of wave propagation—it is only during the detection process that the field energy is quantized.

When optical fields interact with a photodetector, the absorption of a quantum of energy—a photon—results in the release of an excited electron. This interaction is referred to as a photoevent, and the number of photoevents occurring over a particular spatial region and time interval are referred to as photocounts. Most detectors of light record photocounts, and although the recorded data depend directly on the image intensity, the actual number of photocounts recorded is a fundamentally random quantity. The images shown in Fig. 1 help to illustrate this effect. Here, an image of Simeon Poisson (for whom the Poisson random variable is named) is shown as it might be acquired by a detector when 1 million, 10 million, and 100 million total photocounts are recorded.

Statistical Model

For many applications involving charge coupled devices (CCD) and other detectors of optical radiation, the semiclassical theory leads to models for which the photocounts recorded by each detector element are modeled as Poisson random variables whose means are determined by the measurement intensity $I_k(\cdot)$. That is, the expected number of photocounts acquired by the nth photodetector during the kth observation interval is

$$I_k[n] = \gamma \int_{\mathscr{Y}_n} I_k(y)\, dy \tag{9}$$

where n is a two-dimensional discrete index to the elements of the detector array, $\mathscr{Y}_n$ is the spatial region over which the nth detector element integrates the image intensity, and γ is a nonnegative scale factor that accounts for overall detector efficiency and integration time. Furthermore, the number of photocounts acquired by different detector elements are usually statistically independent, and the detector regions are often small in size relative to the fluctuations in the image intensity so that the integrating operation can be well-modeled by the sampling operation

$$I_k[n] \simeq \gamma |\mathscr{Y}_n| I_k(y_n) \tag{10}$$

where y_n is the location of the nth detector element and $|\mathscr{Y}_n|$ is its integration area.

Other Detector Effects

In addition to the quantum noise, imaging detectors introduce other nonideal effects into the imagery that they record. The efficiency with which detectors convert electromagnetic energy into photoevents can vary across elements within a detector array, and this nonuniform efficiency can be captured by attaching a gain function to the photocount mean

$$I_k[n] = a[n]\gamma |\mathscr{Y}_n| I_k(y_n) \tag{11}$$

Seriously flawed detector elements that fail to record data are also accommodated with this model by simply setting the gain to zero at the appropriate location. If different detectors are used for each observation the gain function may need to vary with each frame and, therefore, be indexed by k.

Because of internal shot noise, many detectors record photoevents even when the external light intensity is zero. The resulting photocounts are usually modeled as independent

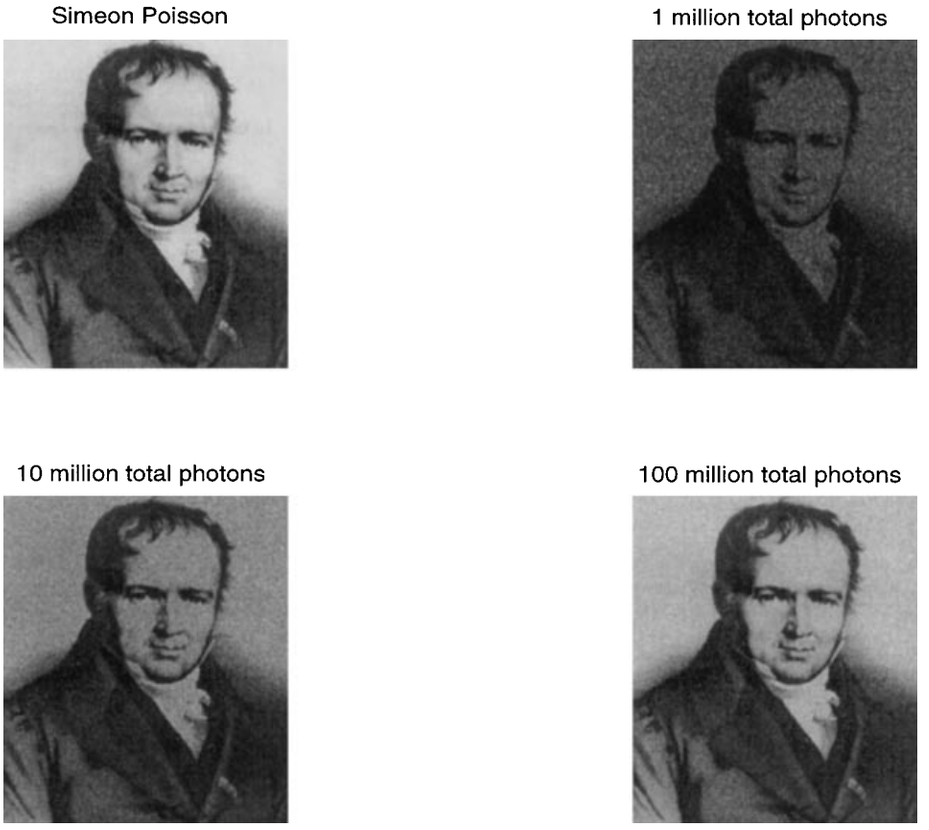

Figure 1. Image of Simeon Poisson as it might be acquired by a detector when 1 million, 10 million, and 100 million total photocounts are recorded.

Poisson random variables, and this phenomenon is accommodated by inserting a background term into the imaging equation

$$I_k[n] \simeq a[n]\gamma|\mathscr{Y}_n|I_k(y_n) + I_b[n] \tag{12}$$

As with the gain function, if different detectors are used for each observation this background term may need to vary with each frame and, therefore, be indexed by k. With the inclusion of these background counts, the number of photocounts acquired by detector element n is a Poisson random variable with mean $I_k[n]$ and is denoted by $N_k[n]$.

The data recorded by many detectors are also corrupted by another form of noise that is induced by the electronics used for the data acquisition. For CCD detectors, this is *read-out* noise and is often approximated as additive, zero-mean Gaussian random variables so that the recorded data are modeled as

$$d_k[n] = N_k[n] + g_k[n] \tag{13}$$

where $g_k[n]$ models the read-out noise at the nth detector for the kth observation. The variance of the read-out noise $\sigma^2[\cdot]$ may vary with each detector element, and the read-out noise for different detectors is usually modeled as statistically independent.

The appropriate values for the gain function $a[\cdot]$, background function $I_b[\cdot]$, and read noise variance $\sigma^2[\cdot]$ are usually selected through a controlled study of the data acquisition system. A detailed discussion of these and other camera effects for optical imaging is given in Ref. 23.

Maximum-Likelihood Image Restoration

Consistent with the noise models developed in the previous sections, the data recorded by each detector element in a photon-counting camera are a mixture of Poisson and Gaussian random variables. Accordingly, the probability of receiving N photocounts in the nth detector element is

$$\Pr\{N_k[n] = N; I\} = \exp(-I_k[n])(I_k[n])^N/N! \tag{14}$$

where

$$\begin{aligned} I_k[n] &= a[n]\gamma|\mathscr{Y}_n|I_k(y_n) + I_b[n] \\ &= a[n]\gamma|\mathscr{Y}_n|\int h_k(y_n, x)I(x)\,dx + I_b[n] \end{aligned} \tag{15}$$

contains the dependence on the unknown intensity function $I(\cdot)$. Furthermore, the probability density for the read-out noise is

$$p_{g_k[n]}(g) = (2\pi\sigma^2[n])^{-1/2}\exp[-g^2/(2\sigma[n])] \tag{16}$$

so that the density for the measured data is

$$\begin{aligned} p_{d_k[n]}(d; I) &= \sum_{N=0}^{\infty} p_{g_k[n]}(d - N)\Pr\{N_k[n] = N; I\} \\ &= \frac{(2\pi\sigma^2[n])^{-1/2}}{N}\sum_{N=0}^{\infty}\exp[-(d-N)^2/(2\sigma[n])] \\ &\quad \exp(-I_k[n])(I_k[n])^N \end{aligned} \tag{17}$$

For a given data set $\{d_k[\cdot]\}$, the maximum-likelihood estimate of $I(\cdot)$ is the intensity function that maximizes the likelihood

$$l(I) = \prod_{k=1}^{K} \prod_{n} p_{d_k[n]}(d_k[n]; I) \tag{18}$$

or, as is commonly done, its logarithm (the log-likelihood)

$$\begin{aligned} \mathscr{L}(I) &= \ln l(I) \\ &= \sum_{k=1}^{K} \sum_{n} \ln p_{d_k[n]}(d_k[n]; I) \end{aligned} \tag{19}$$

The complicated form for the measurement density $p_{d_k[n]}(\cdot\,; I)$ makes this an overly complicated optimization. When the read-out noise variance is large (greater than 50 or so), however, $\sigma^2[n]$ can be added to the measured data to form the modified data

$$\begin{aligned} \tilde{d}_k[n] &= d_k[n] + \sigma^2[n] \\ &= N_k[n] + g_k[n] + \sigma^2[n] \\ &\simeq N_k[n] + M_k[n] \end{aligned} \tag{20}$$

where $M_k[n]$ is a Poisson-distributed random variable whose mean value is $\sigma^2[n]$. The modified data at each detector element are then similar (in distribution) to the sum of two Poisson-distributed random variables $N_k[n]$ and $M_k[n]$ and, as such, are also Poisson-distributed with the mean value $I_k[n] + \sigma^2[n]$. This approximation is discussed by Snyder et al. in Refs. 23 and 24. The probability mass function for the modified data is then modeled as

$$\Pr[\tilde{d}_k[n] = D; I] = \exp\{-(I_k[n] + \sigma^2[n])\}(I_k[n] + \sigma^2[n])^D / D! \tag{21}$$

so that the log-likelihood is

$$\begin{aligned} \mathscr{L}(I) = \sum_{k=1}^{K} \sum_{n} \{ &-(I_k[n] + \sigma^2[n]) \\ &+ \tilde{d}_k[n] \ln(I_k[n] + \sigma^2[n]) - \ln d_k[n]! \} \end{aligned} \tag{22}$$

Two difficulties are encountered when attempting to find the intensity function $I(\cdot)$ that maximizes the log-likelihood $\mathscr{L}(I)$: (1) the recovery of an infinite-dimensional function $I(\cdot)$ from finite data is a terribly ill-conditioned problem; and (2) the functional form of the log-likelihood does not admit a closed form, analytic solution for the maximizer even after the dimension of the parameter function has been reduced.

To address the dimensionality problem, it is common to approximate the parameter function in terms of a finite-dimensional basis set

$$I(x) \simeq \sum_{m} I[m] \psi_m(x) \tag{23}$$

where the basis functions $\{\psi_m(\cdot)\}$ are chosen in an appropriate manner. When expressing the object function with a predetermined grid of image pixels, for example, $\psi_m(\cdot)$ might be an indicator function that denotes the location of the mth pixel. For the same situation, the basis functions might alternatively be chosen as two-dimensional impulses co-located with the center of each pixel. Many other basis sets are possible and a clever choice here can greatly affect estimator performance, but the grid of two-dimensional impulses is probably the most common. Using this basis, the data mean is expressed as

$$\begin{aligned} I_k[n] &= a[n]\gamma|\mathscr{Y}_n| I_k(y_n) + I_b[n] \\ &= a[n]\gamma|\mathscr{Y}_n| \int h_k(y_n, x) \sum_{m} I[m] \delta_2(x - x_m)\, dx + I_b[n] \\ &= a[n]\gamma|\mathscr{Y}_n| \sum_{m} h_k(y_n, x_m) I[m] + I_b[n] \end{aligned} \tag{24}$$

where y_n denotes the location of the nth measurement, x_m denotes the location of the mth object pixel, and $\delta_2(\cdot)$ is the two-dimensional Dirac impulse. The estimation problem, then, is one of estimating the discrete samples $I[\cdot]$ of the intensity function from the noisy data $\{d_k[\cdot]\}$. Because $I[\cdot]$ represents samples of an intensity function, this function is physically constrained to be nonnegative.

Ignoring terms in the log-likelihood that do not depend upon the unknown object intensity, the optimization problem required to solve for the maximum-likelihood object estimate is

$$\begin{aligned} \hat{I}[n] = \arg\max_{I \geq 0} \Bigg\{ &-\sum_{k=1}^{K} \sum_{n} (I_k[n] + \sigma^2[n]) \\ &+ \sum_{k=1}^{K} \sum_{n} \tilde{d}_k[n] \ln(I_k[n] + \sigma^2[n]) \Bigg\} \end{aligned} \tag{25}$$

where $\tilde{d}_k[n] = d_k[n] + \sigma^2[n]$ is the modified data and

$$I_k[n] = a[n]\gamma|\mathscr{Y}_n| \sum_{m} h_k(y_n, x_m) I[m] + I_b[n]$$

is the photocount mean. The solution to this problem generally requires the use of a numerical method, and a great number of techniques are available for this purpose. General-purpose techniques such as those described in popular texts on optimization theory (25,26) can be applied. In addition, specialized numerical methods devised specifically for the solution of maximum-likelihood and related problems can be applied (27,28)—a specific example is discussed in the following section.

The Expectation-Maximization Method. The expectation-maximization (EM) method is a numerical technique devised specifically for maximum-likelihood estimation problems. As described in Ref. 27, the classical formulation of the EM procedure requires one to augment the measured data—commonly referred to as the *incomplete data*—with a set of *complete data* which, if measured, would facilitate direct estimation of the unknown parameters. The application of this procedure then requires one to alternately apply an *E-step,* wherein the conditional expectation of the complete-data log-likelihood is determined, and an *M-step,* wherein all parameters are simultaneously updated by maximizing the expectation of the complete-data log-likelihood with respect to all of the unknown parameters. In general, the application of the EM procedure results in an iterative algorithm that produces a sequence of parameter estimates that monotonically increases the measured data likelihood.

The application of the EM procedure to the incoherent imaging problems has been proposed and described for numerous applications (29–32). The general application of this method is outlined as follows. First, recall that the measured (or incomplete) data $\tilde{d}_k[n]$ for each observation k and detector element n are independent Poisson variables with the expected value

$$E\{\tilde{d}_k[n]\} = a[n]\gamma|\mathscr{Y}_n|\sum_m h_k(y_n, x_m)I[m] + I_b[n] + \sigma^2[n] \quad (26)$$

Because the sum of Poisson random variables is still a Poisson random variable (and the expected value is the sum of the individual expected values), the incomplete data can be statistically modeled as

$$\tilde{d}_k[n] = \sum_m N_k^c[n, m] + M_k^c[n] \quad (27)$$

where for all frames k, detector locations n, and object pixels m, the data $N_k^c[n, m]$ are Poisson random variables, each with the expected value

$$E\{N_k^c[n, m]\} = a[n]\gamma|\mathscr{Y}_n|h_k(y_n, x_m)I[m] \quad (28)$$

and for all frames k and detector locations n, the data $M_k^c[n]$ are Poisson random variables, each with the expected value

$$E\{M_k^c[n]\} = I_b[n] + \sigma^2[n] \quad (29)$$

In the terminology of the EM method, these data $\{N_k^c[\cdot, \cdot], M_k^c[\cdot]\}$ are the complete data, and although they cannot be observed directly, their measurement, if possible, would greatly facilitate direct estimation of the underlying object intensity.

Because the complete data are independent, Poisson random variables, the complete-data log-likelihood is

$$\begin{aligned}\mathscr{L}^c(I) = &-\sum_k\sum_n\sum_m a[n]\gamma|\mathscr{Y}_n|h_k(y_n, x_m)I[m] \\ &+\sum_k\sum_n\sum_m N_k^c[n, m]\ln(a[n]\gamma|\mathscr{Y}_n|h_k(y_n, x_m)I[m])\end{aligned} \quad (30)$$

where terms not dependent upon the unknown object intensity $I[\cdot]$ have been omitted. Given an estimate for the object intensity $I^{\text{old}}[\cdot]$, the EM procedure makes use of the complete data and their corresponding log-likelihood to update the object intensity estimate in such a way that $I^{\text{new}}[\cdot]$ increases the measured data log-likelihood. The E-step of the EM procedure requires the expectation of the complete-data log-likelihood, conditional on the measured (or incomplete) data and using the old object intensity estimate $I^{\text{old}}[\cdot]$

$$\begin{aligned}Q(I; I^{\text{old}}) &= E[\mathscr{L}^c(I)|\{\tilde{d}_k[n]\}; I^{\text{old}}] \\ &= -\sum_k\sum_n\sum_m a[n]\gamma|\mathscr{Y}_n|h_k(y_n, x_m)I[m] \\ &\quad +\sum_k\sum_n\sum_m E[N_k^c[n, m]|\{\tilde{d}_k[n]\}; I^{\text{old}}] \\ &\quad \ln(a[n]\gamma|\mathscr{Y}_n|h_k(y_n, x_m)I[m])\end{aligned} \quad (31)$$

The intensity estimate is then updated in the M-step by maximizing this conditional expectation over I

$$I^{\text{new}} = \arg\max_{I\geq 0} Q(I; I^{\text{old}}) \quad (32)$$

It is straightforward to show that the object estimate is then updated according to

$$I^{\text{new}}[m] = \frac{\sum_k\sum_n E[N_k^c[n, m]|\{\tilde{d}_k[n]\}; I^{\text{old}}]}{\sum_k\sum_n a[n]\gamma|\mathscr{Y}_n|h_k(y_n, x_m)} \quad (33)$$

As described in Ref. 29, the conditional expectation is evaluated as

$$\begin{aligned}&E[N_k^c[n, m]|\{\tilde{d}_k[n]\}; I^{\text{old}}] \\ &\quad = \frac{a[n]\gamma|\mathscr{Y}_n h_k(y_n, x_m)I^{\text{old}}[m]}{\sum_{m'} a[n]\gamma|\mathscr{Y}_n h_k(y_n, x_{m'})I^{\text{old}}[m'] + I_b[n] + \sigma^2[n]}\tilde{d}_k[n]\end{aligned} \quad (34)$$

so that the iterative formula for updating the object estimate is

$$\begin{aligned}I^{\text{new}}[m] = I^{\text{old}}[m] &\\ \frac{\sum_k\sum_n h_k(y_n, x_m)\left[\frac{a[n]\gamma|\mathscr{Y}_n|\tilde{d}_k[n]}{\sum_{m'} a[n]\gamma|\mathscr{Y}_n h_k(y_n, x_{m'})I^{\text{old}}[m'] + I_b[n] + \sigma^2[n]}\right]}{\sum_k\sum_n a[n]\gamma|\mathscr{Y}_n|h_k(y_n, x_m)} &\end{aligned} \quad (35)$$

For the special case of uniform gain with no background or detector noise, the iterative algorithm proposed by Richardson (33) and Lucy (34) has the same form as these iterations. An excellent historical perspective of the application of the EM method to imaging problems is presented in Ref. 35, and detailed discussions of the convergence properties of this algorithm along with the pioneering derivations for applications in emission tomography can be found in Ref. 36.

Figures 2 and 3 illustrate the use of this technique on imagery acquired by the Hubble Space Telescope (HST). Shortly after the launch of the HST with its aberrated primary mirror in 1990, the imagery acquired by this satellite became a focus of national attention. Whereas microscopic flaws in the telescope's mirror resulted in the severely distorted imagery, image restoration methods were successful in restoring much of the lost resolution (4). Figure 2, for example, shows imagery of the star cluster R136 in a star formation called 30 Doradus as acquired by the telescope and as restored using the methods described in this article. Also shown in this figure are imagery acquired by the telescope after its aberrated mirror was corrected, along with a processed image showing the potential advantage of applying image restoration methods to imagery acquired after the correction. Figure 3 contains an image of Saturn along with restorations formed by simple inverse filtering, Wiener filtering, and by the maximum-likelihood method. According to scientific staff at the Space Telescope Science Institute, the maximum-likelihood restoration

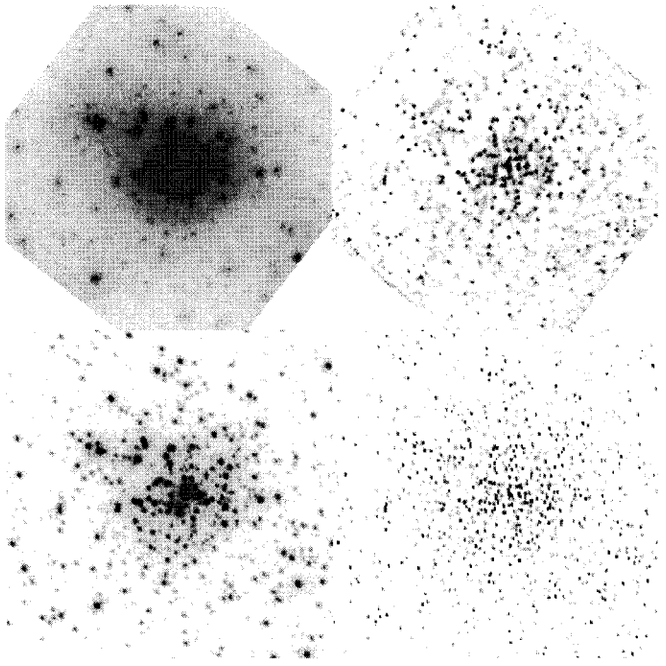

Figure 2. Imagery of the star cluster R136 in the star formation 30 Doradus as acquired by the Hubble Space Telescope both before and after its aberrated primary mirror was corrected. Upper left: raw data acquired with the aberrated primary mirror; upper right: restored image obtained from imagery acquired with the aberrated primary mirror; lower left: raw data acquired after correction; lower right: restored image obtained from imagery acquired after the correction. (Courtesy of R. J. Hanisch and R. L. White, Space Telescope Science Institute and NASA.)

provides the best trade-off between resolution and noise amplification.

Regularization. Under reasonably unrestrictive conditions, the EM method described in the previous section produces a sequences of images that converges to a maximum-likelihood solution (36). Imaging problems for which this method is applicable are often ill-conditioned or practically ill-posed, however, and because of this the maximum-likelihood image estimates frequently exhibit severe noise artifacts. Common methods for addressing this problem are discussed briefly in this section.

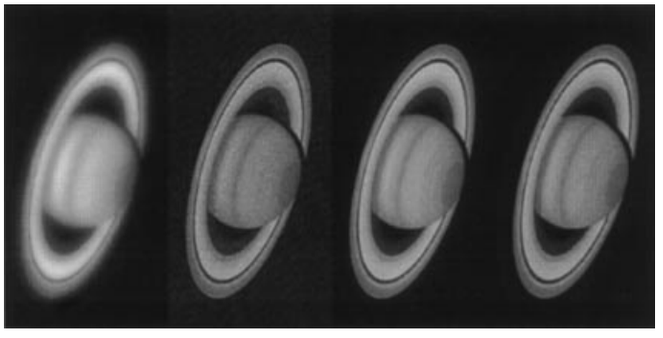

Figure 3. Raw imagery and restorations of Saturn as acquired by the Hubble Space Telescope. From left to right: telescope imagery; restoration produced by simple inverse filtering; restoration produced by Wiener filtering; restoration produced by the maximum-likelihood method. (Courtesy of R. J. Hanisch and R. L. White, Space Telescope Science Institute and NASA.

Stopping Rules. Probably the simplest method to implement for overcoming the noise artifacts seen in maximum-likelihood image estimates obtained by numerical procedures is to terminate the iterative process before convergence. Implementation of such a procedure is straightforward; however, the construction of optimal "stopping rules" can be challenging. Criteria for developing these rules for problems in coherent imaging are discussed in Refs. 21, 37, 38.

Sieve Methods. The basic idea behind the method of sieves is to constrain the set of allowable image estimates to be in a smooth subset called a sieve. The sieve is selected in a manner that depends upon the degree to which the problem is ill-conditioned and upon the noise level. Badly ill-conditioned problems and noisy data require a "small" sieve set containing only very smooth functions. Problems that are better conditioned with little noise can accommodate "large" sieve sets, and the sieve is ideally selected so that its "size" grows with decreasing noise levels in such a manner that the constrained image estimate converges to the true image as the noise level shrinks to zero. Establishing this consistency property for a sieve can, however, be a difficult task.

The general method of sieves as a statistical inference tool was introduced by Grenander (20). The application of this method to problems in incoherent imaging was proposed and investigated by Snyder et al. (39,40). The method is based on a kernel sieve defined according to

$$\mathcal{S} = \left\{ I : I[m] = \sum_p s[m, p]\alpha[p] \right\} \tag{36}$$

where intensity functions within the sieve set $\mathcal{S}$ are determined by the nonnegative parameters $\{\alpha[p]\}$. The sieve-constrained optimization problem then becomes one of maximizing the likelihood subject to the additional constraint $I \in \mathcal{S}$. The smoothness properties of the sieve are induced by the sieve kernel $s[\cdot, \cdot]$. With a Gaussian kernel, for instance, the smoothness of the sieve set is determined by the variance parameter σ

$$s[m, p] = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(m-p)^2}{2\sigma^2}\right) \tag{37}$$

This Gaussian kernel was investigated in Refs. 39, 40, but kernels with other mathematical forms can be used. The EM method can, with straightforward modifications, be applied to problems in which kernel sieves are used for regularization.

Penalty and MAP Methods. Another method for regularizing maximum-likelihood estimation problems is to augment the likelihood with a penalty function

$$\mathcal{C}(I) = \mathcal{L}(I) - \gamma\Phi(I) \tag{38}$$

where Φ is a function that penalizes undesirable qualities (or rewards desirable ones) of the image estimate, and γ is a nonnegative scale factor that determines the relative contribution of the penalty to the optimization problem. The penalized image estimate is then selected to maximize the cost function $\mathcal{C}$, which involves a trade between maximizing the likelihood $\mathcal{L}$

and minimizing the penalty Φ. The choice of the penalty can greatly influence the resulting image estimate, as can the selection of the scale factor γ. A commonly used penalty is the quadratic smoothness penalty

$$\Phi(I) = \sum_n \sum_{m \in \mathcal{N}_n} w_{nm}(I[n] - I[m])^2 \tag{39}$$

where $\mathcal{N}_n$ denotes a neighborhood of pixel locations about the nth object pixel, and the coefficients w_{nm} control the link between pixel n and m. This penalty can also be induced by using a MAP formulation with Gaussian Markov random field (GMRF) prior model for the object. However, because the use of this penalty often results in excessive smoothing of the object edges, alternative penalties have been developed and investigated (41–43). A particularly interesting penalty is induced by using a MAP formulation with the generalized Gaussian Markov random field (GGMRF) model (43). The use of this prior results in a penalty function of the form

$$\Phi(I) = \gamma^q \sum_n \sum_{m \in \mathcal{N}_n} w_{nm}|I[n] - I[m]|^q \tag{40}$$

where $q \in [1, 2]$ is a parameter that controls the smoothness of the reconstruction. For $q = 2$ this is the common quadratic smoothing penalty, whereas smaller values of q will, in general, allow for sharper edges in the object estimates.

Although the EM method is directly applicable to problems in which stopping rules or kernel sieves are used, the EM approach is less simple to use when penalty or MAP methods are employed. The major difficulty arises because the maximization step usually has no closed-form solution; however, approximations and modifications can be used (41,44) to address this problem.

Alternative Numerical Approaches

A major difficulty encountered when using the EM method for incoherent-imaging problems is its slow convergence (45). Many methods have been proposed to overcome this problem, and a few of these are summarized briefly here. Because of the similarities of the EM method to gradient ascent, line-search methods can be used to accelerate convergence (45), as can other gradient-based optimization methods (46,47). Substantial improvements in convergence can also be obtained by using a generalization of the EM method—the space-alternating generalized expectation-maximization (SAGE) method (28,48)—whereby convergence is accelerated through a novel choice for the complete data at each iteration. In addition, a coordinate descent (or ascent) optimization method has been shown to provide for greatly reduced computational time (49).

COHERENT IMAGING

For synthetic aperture radar (SAR), ultrasound, and other forms of coherent array imaging, an object or scene is illuminated by a highly coherent source (such as a radar transmitter, laser, or acoustic transducer), and heterodyne, homodyne, or holographic methods are used to record amplitude and phase information about the reflected field. The basic signal model for these problems is of the form:

$$u_p = \int h_p(x)u(x)\,dx + w_p, \quad p = 1, 2, \ldots, P \tag{41}$$

where p is an index to sensor locations (either real or synthetic), u_p is the complex-amplitude measured by the pth sensor, $u(x)$ is the complex-amplitude of the field that is reflected from an object or scene of interest, $h_p(x)$ is a sensor response function for the pth sensor measurement, and w_p accounts for additive sensor noise. The response function accounts for both the sensor characteristics and for wave propagation from the object or scene to the sensor; in the Fraunhofer approximation for wave propagation, these functions take on the form of a Fourier-transform kernel (10).

When the object or scene gives rise to diffuse reflections, the Gaussian speckle model (50) is often used as a statistical model for the reflected field $u(\cdot)$. That is, $u(\cdot)$ is modeled as a complex Gaussian random process (13,51,52) with zero-mean and the covariance

$$E[u(x)u^*(x')] \simeq s(x)\delta_2(x - x') \tag{42}$$

where $s(\cdot)$ is the object incoherent scattering function. The sensor noise is often modeled as zero-mean, independent complex Gaussian variables with variance σ^2 so that the recorded data are complex Gaussian random variables with zero-mean and the covariance

$$E[u_p u^*_{p'}] = \int h_p(x)h^*_{p'}(x)s(x)\,dx + \sigma^2\delta[p - p'] \tag{43}$$

where $\delta[\cdot]$ is the Kronecker delta function. The maximum-likelihood estimation of the object scattering function $s(\cdot)$ then becomes a problem of covariance estimation subject to the linear structure constraint of Eq. (43).

Using vector-matrix notation the data covariance is, as a function of the unknown object scattering function

$$\begin{aligned} \boldsymbol{R}(s) &= E[\boldsymbol{u}\boldsymbol{u}^\dagger] \\ &= \int \boldsymbol{h}(x)\boldsymbol{h}^\dagger(x)s(x)\,dx + \sigma^2\boldsymbol{I} \end{aligned} \tag{44}$$

where $\boldsymbol{u} = [u_1 u_2 \cdots u_P]^T$ is the data vector, $\boldsymbol{h}(x) = [h_1(x)h_2(x) \cdots h_P(x)]^T$ is the system response vector, $[\cdot]^T$ denotes matrix transposition, $[\cdot]\dagger$ denotes Hermitian matrix transposition, and $\boldsymbol{I}$ is the $P \times P$ identity matrix. Accordingly, the data log-likelihood is

$$L(s) = -\ln\det[\boldsymbol{R}(s)] - \mathrm{tr}[\boldsymbol{R}^{-1}(s)\boldsymbol{S}] \tag{45}$$

where $\boldsymbol{S} = \boldsymbol{u}\boldsymbol{u}\dagger$ is the data sample-covariance. Parameterization of the parameter function as in Eq. (23) is a natural step before attempting to solve this problem, but direct maximization of the likelihood is still a difficult problem. Because of this, the EM method has been proposed and discussed in Refs. 53–55 for addressing this problem, and the resulting algorithm has been shown to produce parameter estimates with lower bias and variance than alternative methods (56). A major problem with this method, though, is the high computational cost; however, the application of the SAGE method (28) to this problem has shown great promise for reducing the computational burden (57). The development and application of regularization methods for problems in coherent imaging is an area of active research.

SUMMARY

Imaging science is a rich and vital area of science and technology in which information-theoretic methods can be and

have been applied with great benefit. Maximum-likelihood methods can be applied to a variety of problems in image restoration and synthesis, and their application to the restoration problem for incoherent imaging has been discussed in great detail in this article. To conclude, the future of this field is best summarized by the following quote from Bracewell (58):

> The study of imaging now embraces many major areas of modern technology, especially the several disciplines within electrical engineering, and will be both the stimulus for, and recipient of, new advances in information science, computer science, environmental science, device and materials science, and just plain high-speed computing. It can be confidently recommended as a fertile subject area for students entering upon a career in engineering.

BIBLIOGRAPHY

1. H. Stark (ed.), *Image Recovery: Theory and Application,* Orlando, FL: Academic Press, 1987.
2. A. K. Katsaggelos (ed.), *Digital Image Restoration,* Heidelberg: Springer-Verlag, 1991.
3. R. L. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images,* Boston: Kluwer, 1991.
4. R. J. Hanisch and R. L. White (eds.), *The Restoration of HST Images and Spectra—II,* Baltimore, MD: Space Telescope Science Institute, 1993.
5. H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I,* New York: Wiley, 1968.
6. H. V. Poor, *An Introduction to Signal Detection and Estimation,* 2nd ed., New York: Springer-Verlag, 1994.
7. B. Porat, *Digital Processing of Random Signals: Theory and Methods,* Englewood Cliffs, NJ: Prentice-Hall, 1993.
8. L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis,* Reading, MA: Addison-Wesley, 1991.
9. S. M. Kay, *Modern Spectral Estimation: Theory and Applications,* Englewood Cliffs, NJ: Prentice-Hall, 1988.
10. J. W. Goodman, Introduction to Fourier Optics, 2nd edition, New York: McGraw-Hill, 1996.
11. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics,* New York: Wiley, 1978.
12. M. Born and E. Wolf, *Principles of Optics,* 6th edition, Elmsford, NY: Pergamon, 1980.
13. J. W. Goodman, *Statistical Optics,* New York: Wiley, 1985.
14. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics,* New York: Wiley, 1991.
15. A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* 3rd edition, New York: McGraw-Hill, 1991.
16. R. M. Gray and L. D. Davisson, *Random Processes: An Introduction for Engineers,* Englewood Cliffs: Prentice-Hall, 1986.
17. A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems,* Washington, DC: Winston, 1977.
18. W. L. Root, Ill-posedness and precision in object-field reconstruction problems, *J. Opt. Soc. Am., A,* **4** (1): 171–179, 1987.
19. J. R. Thompson and R. A. Tapia, *Nonparametric Function Estimation, Modeling, and Simulation,* Philadelphia: SIAM, 1990.
20. U. Grenander, *Abstract Inference,* New York: Wiley, 1981.
21. E. Veklerov and J. Llacer, Stopping rule for the MLE algorithm based on statistical hypothesis testing, *IEEE Trans. Med. Imaging,* **6**: 313–319, 1987.
22. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics,* New York: Cambridge University Press, 1995.
23. D. L. Snyder, A. M. Hammoud, and R. L. White, Image recovery from data acquired with a charge-coupled-device camera, *J. Opt. Soc. Am., A,* **10** (5): 1014–1023, 1993.
24. D. L. Snyder et al., Compensation for readout noise in CCD images, *J. Opt. Soc. Am., A,* **12** (2): 272–283, 1995.
25. D. G. Luenberger, *Linear and Nonlinear Programming,* Reading, MA: Addison-Wesley, 1984.
26. R. Fletcher, *Practical Methods of Optimization,* New York: Wiley, 1987.
27. A. P. Dempster, N. M. Laird, and D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc., B,* **39**: 1–37, 1977.
28. J. A. Fessler and A. O. Hero, Space-alternating generalized expectation-maximization algorithm, *IEEE Trans. Signal Process.* **42**: 2664–2677, 1994.
29. L. A. Shepp and Y. Vardi, Maximum-likelihood reconstruction for emission tomography, *IEEE Trans. Med. Imaging,* **MI-1**: 113–121, 1982.
30. D. L. Snyder and D. G. Politte, Image reconstruction from list-mode data in an emission tomography system having time-of-flight measurements, *IEEE Trans. Nucl. Sci.,* **NS-30**: 1843–1849, 1983.
31. K. Lange and R. Carson, EM reconstruction algorithms for emission and transmission tomography, *J. Comput. Assisted Tomography,* **8**: 306–316, 1984.
32. T. J. Holmes, Maximum-likelihood image restoration adapted for noncoherent optical imaging, *J. Opt. Soc. Am., A,* **6**: 666–673, 1989.
33. W. H. Richardson, Bayesian-based iterative method of image restoration, *J. Opt. Soc. Am.,* **62** (1): 55–59, 1972.
34. L. B. Lucy, An iterative technique for the rectification of observed distributions, *Astronom. J.,* **79** (6): 745–754, 1974.
35. Y. Vardi and D. Lee, From image deblurring to optimal investments: Maximum likelihood solutions for positive linear inverse problems, *J. R. Stat. Soc. B,* **55** (3): 569–612, 1993.
36. Y. Vardi, L. A. Shepp, and L. Kaufman, A statistical model for positron emission tomography, *J. Amer. Stat. Assoc.,* **80**: 8–37, 1985.
37. T. Hebert, R. Leahy, and M. Singh, Fast MLE for SPECT using an intermediate polar representation and a stopping criterion, *IEEE Trans. Nucl. Sci.,* **NS-34**: 615–619, 1988.
38. J. Llacer and E. Veklerov, Feasible images and practicle stopping rules for iterative algorithms in emission tomography, *IEEE Trans. Med. Imaging,* **MI-8**: 186–193, 1989.
39. D. L. Snyder and M. I. Miller, The use of sieves to stabilize images produced with the EM algorithm for emission tomography, *IEEE Trans. Nucl. Sci.,* **NS-32**: 3864–3871, 1985.
40. D. L. Snyder et al., Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography, *IEEE Trans. Med. Imaging,* **MI-6**: 228–238, 1987.
41. P. J. Green, Bayesian reconstructions from emission tomography data using a modified EM algorithm, *IEEE Trans. Med. Imaging,* **9**: 84–93, 1990.
42. K. Lange, Convergence of EM image reconstruction algorithms with Gibbs priors, *IEEE Trans. Med. Imaging,* **MI-9**: 439–446, 1990.
43. C. A. Bouman and K. Sauer, A generalized Gaussian image model for edge-preserving MAP estimation, *IEEE Trans. Image Process.,* **2**: 296–310, 1993.
44. T. Hebert and R. Leahy, A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors, *IEEE Trans. Med. Imaging,* **MI-8**: 194–202, 1989.

45. L. Kaufman, Implementing and accelerating the EM algorithm for positron emission tomography, *IEEE Trans. Med. Imaging,* **MI-6**: 37–51, 1987.
46. E. U. Mumcuoglu et al., Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images, *IEEE Trans. Med. Imag.,* **MI-13**: 687–701, 1994.
47. K. Lange and J. A. Fessler, Globally convergent algorithm for maximum a posteriori transmission tomography, *IEEE Trans. Image Process.,* **4**: 1430–1438, 1995.
48. J. A. Fessler and A. O. Hero, Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms, *IEEE Trans. Image Process.,* **4**: 1417–1429, 1995.
49. C. A. Bouman and K. Sauer, A unified approach to statistical tomography using coordinate descent optimization, *IEEE Trans. Image Process.,* **5**: 480–492, 1996.
50. J. C. Dainty (ed.), Laser speckle and related phenomena. *Topics in Applied Physics,* vol. 9, 2nd ed., Berlin: Springer-Verlag, 1984.
51. F. D. Neeser and J. L. Massey, Proper complex random processes with applications to information theory, *IEEE Trans. Inf. Theory,* **39**: 1293–1302, 1993.
52. K. S. Miller, *Complex Stochastic Processes: An Introduction to Theory and Applications,* Reading, MA: Addison-Wesley, 1974.
53. D. B. Rubin and T. H. Szatrowski, Finding maximum likelihood estimates of patterned covariance matrices by the EM algorithm, *Biometrika,* **69** (3): 657–660, 1982.
54. M. I. Miller and D. L. Snyder, The role of likelihood and entropy in incomplete-data problems: Applications to estimating point-process intensities and Toeplitz constrained covariances, *Proc. IEEE,* **75**: 892–907, 1987.
55. D. L. Snyder, J. A. O'Sullivan, and M. I. Miller, The use of maximum-likelihood estimation for forming images of diffuse radar-targets from delay-doppler data, *IEEE Trans. Inf. Theory,* **35**: 536–548, 1989.
56. M. J. Turmon and M. I. Miller, Maximum-likelihood estimation of complex sinusoids and Toeplitz covariances, *IEEE Trans. Signal Process.,* **42**: 1074–1086, 1994.
57. T. J. Schulz, Penalized maximum-likelihood estimation of structured covariance matrices with linear structure, *IEEE Trans. Signal Process.,* **45**: 3027–3038, 1997.
58. R. N. Bracewell, *Two-Dimensional Imaging,* Upper Saddle River, NJ: Prentice-Hall, 1995.

TIMOTHY J. SCHULZ
Michigan Technological University

MEASUREMENT. See ACCELERATION MEASUREMENT; DENSITY MEASUREMENT; DISPLACEMENT MEASUREMENT; MAGNETIC FIELD MEASUREMENT; MILLIMETER WAVE MEASUREMENT; Q-FACTOR MEASUREMENT.

MEASUREMENT, ATTENUATION. See ATTENUATION MEASUREMENT.

MEASUREMENT, C-V. See C-V PROFILES.