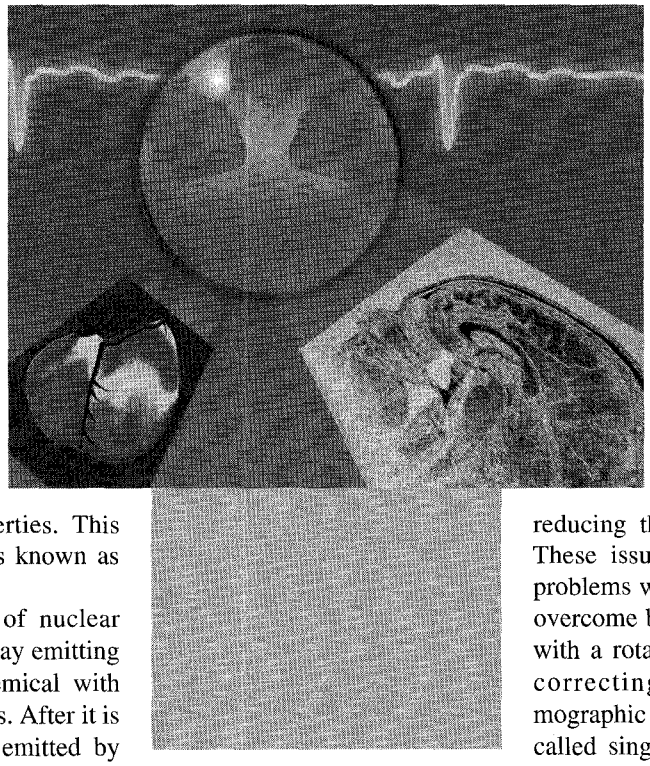
Positron-Emission Tomography

JOHN M. OLLINGER AND JEFFREY A. FESSLER



Medical imaging is often thought of as a way of viewing anatomical structures of the body. Indeed, x-ray computed tomography (CT) and magnetic resonance imaging (MRI) yield exquisitely detailed images of such structures. It is often useful, however, to acquire images of physiologic function rather than of anatomy. Such images can be acquired by imaging the decay of radio-isotopes bound to molecules with known biological properties. This class of imaging techniques is known as nuclear medicine imaging.

The most common form of nuclear medicine scan uses a gamma-ray emitting radio-isotope bound to a chemical with known physiological properties. After it is administered, single photons emitted by the decaying isotope are detected with a gamma camera [1]. These cameras consist of a lead collimator to ensure that all detected photons propagated along parallel paths, a crystal scintillator to convert high-energy photons to visible light, and photo-multiplier tubes and associated electronics to determine the position of each incident photon from the light distribution in the crystal. A two-dimensional (2D) histogram of the detected events forms a projection image of the distribution of the radio-isotope and hence of the chemical compound. An example of such a procedure would be a cardiac study using thallium-201. Image intensity is indicative of cardiac perfusion and can be used to diagnose defects in the blood supply. It is widely used to screen for bypass surgery.

Planar imaging with gamma cameras has three major shortcomings. First, the images are projection images, so the organ of interest can be obscured by activity in front of or behind the organ of interest. Moreover, photons originating in the organ of interest can be attenuated by overlying tissue. This is a problem, for example, in scans of obese women, where *attenuation* in the breast can be misinterpreted as a cardiac defect. Second, the radiopharmaceuticals must incorporate relatively heavy isotopes such as thallium-201 and technetium-99m. Since these elements do not occur naturally in biologically active molecules, the synthesis of physiologically useful tracers incorporating them is a challenging technical problem. This restricts the number of available radiopharmaceuticals. Finally, the lead collimator absorbs many photons, thereby reducing the sensitivity of the camera. These issues are being addressed. The problems with projection imaging can be overcome by acquiring tomographic data with a rotating gamma camera and then correcting for attenuation in a tomographic reconstruction. This method is called single-photon emission computed tomography (SPECT) [1]. Continuing research in radiochemistry has made more radiopharmaceuticals available. Finally, newer SPECT cameras with two or three rotating heads have improved the sensitivity. Nevertheless, single-photon imaging still suffers from problems of poor sensitivity and poor quantitative accuracy.

In this article we review positron-emission tomography (PET), which has inherent advantages that avoid these shortcomings. PET image reconstruction methods with origins in signal and image processing are discussed, including the potential problems of these methods. A summary of statistical image reconstruction methods, which can yield improved image quality, is also presented.

Advantages of PET

One of the advantages of PET that allow it to avoid the above-mentioned shortcomings is that attenuation correction is easily accomplished in PET. Also, positron-emitting isotopes of carbon, nitrogen, oxygen, and fluorine occur natu-

1053-5888/97/$10.00©1997IEEE

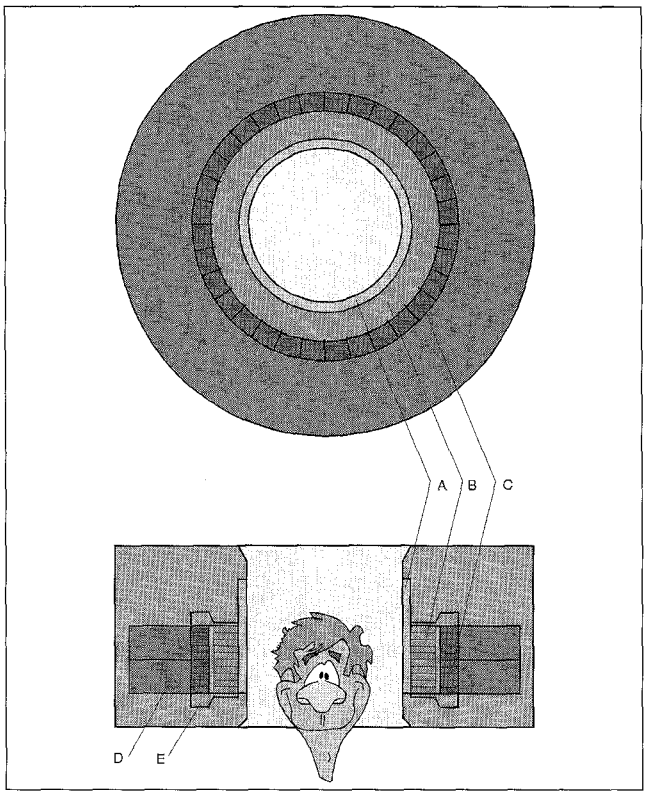
rally in many compounds of biological interest and can therefore be readily incorporated into a wide variety of useful radiopharmaceuticals, and collimation is done electronically, so no collimator is required, leading to relatively high sensitivity. The major problem with PET is its cost. The short half-life of most positron emitting isotopes requires an on-site cyclotron, and the scanners themselves are significantly more expensive than single-photon cameras. Nevertheless, PET is widely used in research studies and is finding growing clinical acceptance, primarily for the diagnosis and staging of cancer.

A PET study begins with the injection or inhalation of a radiopharmaceutical. The scan is begun after a delay ranging from seconds to minutes to allow for transport to and uptake by the organ of interest. When the radio-isotope decays, it emits a positron, which travels a short distance before annihilating with an electron. This annihilation produces two high-energy (511 keV) photons propagating in nearly opposite directions. If two photons are detected within a short (~10 ns) timing window (the *coincidence timing window*), an event (called a *true coincidence* if neither photon is *scattered*) is recorded along the line connecting the two detectors (sometimes referred to as a line of response (LOR)). Summing many such events results in quantities that approximate line integrals through the radio-isotope distribution. The validity of this approximation depends, of course, on the number of counts collected. For 2D imaging, these line integrals form a discrete approximation of the Radon transform [3] of a cross-section of the radio-isotope concentration and can be inverted to form an image of the radioisotope distribution.

If they are suitably calibrated, PET images yield quantitative estimates of the concentration of the radiopharmaceutical at specific locations within the body. The kinetics of the pharmaceutical can be modeled as a linear dynamic system with the arterial concentration of radio-isotope in the blood as the input and the PET measurement as the output. The state variables are the concentrations in different *compartments* of the tissue, where examples of compartments would be blood, the interstitial space between cells, and the interiors of cells. Compartments need not be related to physical spaces and can represent, for example, bound and unbound states of the radiopharmaceutical. The exchange rates between the compartments are parameters of the models. Acquiring a series of images sequentially after injection yields a time-course of the sum of the quantity of tracer in each compartment, i.e., of the output of the model, which can be used to estimate the model's parameters. These parameters can then be used to calculate physiological parameters of interest, such as blood flow, glucose metabolism, receptor binding characteristics, etc. Thus, PET can be used for precise quantitative measurements of specific physiological quantities.

The Physics of PET

A diagram of a PET scanner is shown in Fig. 1. The subject is surrounded by a cylindrical ring of detectors with a diameter of 80-100 cm and an axial extent of 10-20 cm. The



1. A transaxial view of a PET scanner (upper panel) top view (lower panel) showing the rod sources used for attenuation correction (A), the septa used for scatter reduction (B), the detector blocks consisting of crystals (C) and photomultiplier tubes (D), and the end-shields (E).

detectors are shielded from radiation from outside the field of view by relatively thick, lead end-shields. Most scanners can be operated in either a slice-collimated mode, where axial collimation is provided by thin annular rings of tungsten called septa, or in a fully three-dimensional (3D) mode where the septa are retracted and coincidences can be collected between all possible detector pairs. (All commercially available PET scanners simultaneously acquire data for 3D images, either by imaging the entire volume as a unit or by stacking adjacent 2D slices.)

Detectors

The most critical components of a PET camera are the detectors [4]. In some cases these are similar to those used in single-photon imaging: large crystals of sodium-iodide coupled to many photo-multiplier tubes (PMTs) [5]. A more commonly used configuration is shown in Fig. 2. In these detectors a rectangular bundle of crystals, a *block*, is optically coupled to several PMTs. When a photon interacts in the crystal, electrons are moved from the valence band to the conduction band. These electrons return to the valence band at impurities in the crystal, emitting light in the process. Since the impurities usually have metastable excited states, the light output decays exponentially at a rate characteristic of the crystal. The ideal crystal has high density so that a large fraction of incident photons scintillate, high light output for

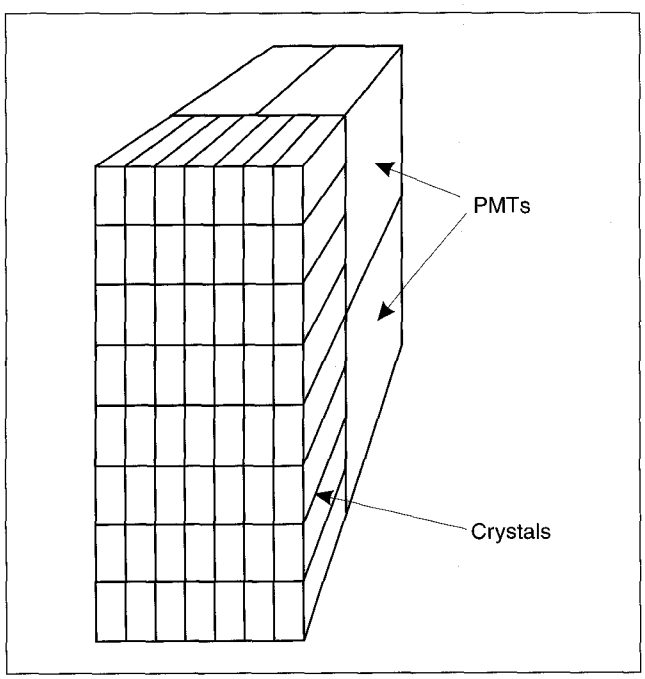
positioning accuracy, fast rise time for accurate timing, and a short decay time so that high counting rates can be handled. Most current scanners use bismuth-germanate (BGO), which generates approximately 2500 light photons per 511 keV photon, and has a decay time (i.e., time-constant) of 300 ns. One such block, for example, couples a 7x8 array of BGO crystals to four PMTs where each crystal is 3.3 mm wide in the transverse plane, 6.25 mm wide in the axial dimension, and 30 mm deep. The block is fabricated in such a way that the amount of light collected by each PMT varies uniquely depending on the crystal in which the scintillation occurred [4]. Hence integrals of the PMT outputs can be decoded to yield the position of each scintillation. The sum of the integrated PMT outputs is proportional to the energy deposited in the crystal.

Resolution

If the data are acquired in the slice-collimated (2D) mode, the LORs connecting crystals can be binned into sets of parallel projections at evenly spaced angles as shown in Fig. 3. Two characteristics are evident. First, samples are unevenly spaced, with finer sampling at the edges of the field of view than at the center. Second, the samples along the heavy solid line at angles one and three are offset by one-half of the detector spacing from samples at angle two. Therefore, adjacent parallel projections can be combined to yield one-half the number of projection angles with a sampling distance of one-half the detector width. (There is a degradation of image quality associated with this approximation, but it is imperceptible for realistic imaging situations.) A typical block might have 3.3 mm thick crystals, so the resulting sampling distance would be 1.65 mm.

The Nyquist criterion is usually stated in medical imaging applications as requiring that the sampling distance be one-half the spatial resolution expressed as the full-width-at-half-maximum (FWHM). (The full-width-at-half-maximum is defined as the distance between the half-value points of the impulse response. It is the minimum separation required for two distinct points to be resolved.) Hence, this block would support a spatial resolution of 3.3 mm. In fact, a scanner with this crystal size has a measured resolution that is somewhat worse, varying from 3.6 mm at the center of the field of view to 5.0 mm at 20 cm from the center. This occurs because scintillations usually consist of one or more Compton interactions followed by photoelectric absorption (assuming the photon is not scattered out of the crystal). Since a 511-keV photon travels on average 7.5 mm in BGO before interacting, the light output is spatially distributed, especially at large radial distances where it is often distributed across two crystals.

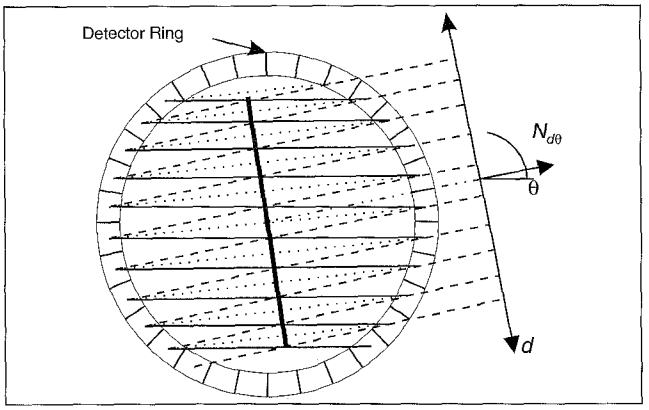
The best obtainable resolution is termed the *intrinsic resolution.* This resolution is rarely achieved in practice because unfiltered images are usually very noisy. Although current scanners have intrinsic resolutions of less than 5 mm, the final resolution of the image is usually greater than 8 mm because the reconstruction algorithms trade off resolution for reduced image variance (as discussed later in this article). This final resolution is called the *reconstructed resolution.* Therefore, the resolution of PET images as they are typically used is not determined by the detectors, but by the degree to which resolution must be degraded to achieve an acceptable image variance. Since the variance is determined by the numbers of counts that can be collected during the scan, the constraints that govern the clinically useful resolution of PET images are the dosage of the radiopharmaceutical, the duration of the scan, the sensitivity of the scanner, and the count-rate capability of the scanner.



2. A block detector consisting of a 7x8 array of crystals coupled to four PMTs.

Positron Range

When the radio-isotope decays, it emits a positron with some nonzero energy. The positron interacts with electrons as it travels through the body, losing energy with each interaction. When its momentum is nearly zero, it annihilates with an electron to produce two annihilation photons, each with an energy of 511 keV. These photons propagate along nearly collinear paths, with the degree of noncollinearity depending on the momentum of the positron and electron when they annihilated. The divergence from collinearity is on the order of one degree or less and is usually ignored. The distance the positron travels before annihilating is termed positron range. The magnitude of this range depends on the positron energy, which varies widely among isotopes. The distribution of positron ranges is very sharply peaked with FWHMs ranging from 0.22-0.31 mm and full-width-at-tenth-maximums (FWTMs) ranging from 0.38-1.6 mm [4, 6] in body tissues. They are much larger in the lungs and other regions containing a significant fraction of air. Since positron range is much smaller than the resolution of most scanners, it is not a serious source of error and is usually ignored.



3. Sampling pattern in the transaxial plane for a PET scanner. Each segment in the detector ring represents one crystal. The solid lines show the parallel projections for the first angle, the dotted lines for the second angle, and the dashed lines for the third angle.

Attenuation

The two possible interactions at 511 keV are photoelectric absorption and Compton scatter. The incidence of photoelectric absorption is negligible for 511-keV photons in body tissues. In a Compton interaction, the photon interacts with an outer shell electron such that its path is deflected and it loses some of its energy. Most scattered photons are scattered out of the field of view and are never detected. The effect of these interactions is termed *attenuation.* The survival probability, i.e., the probability of a photon not interacting as it propagates along the line l at transverse distance d and angle θ, is given by

$$P_{d\theta} = \exp\left(-\int_l \mu(\underline{x})dx\right) \tag{1}$$

where $\mu(\underline{x})$ is the linear attenuation coefficient at position $\underline{x}$. Typical minimum survival probabilities are 0.15 for head scans and 0.003 for body scans. The survival probability given by Eq. (1) is also referred to as the *narrow-beam attenuation.*

The significance of Eq. (1) is that the attenuation experienced by a given pair of annihilation photons is independent of the position of the annihilation along the LOR. This makes possible a simple precorrection of the data. Equation (1) does not hold for SPECT, necessitating approximate methods or computationally expensive iterative methods of attenuation correction. This is the major reason for the relatively poor quantitative accuracy of SPECT relative to PET.

Scattered Events

Those annihilations for which one or both photons are scattered, but both are still detected, are termed *scattered* events, as shown in Fig. 4. These events are incorrectly positioned because the photons' paths are not collinear. A relatively small 30-degree scatter at the center of a typical scanner mispositions the event by 10 cm. The overall effect is to add an error signal to the data at low spatial frequencies.

Since photons lose a fraction of their energy when they undergo a Compton interaction, they can be discriminated from unscattered photons by measuring the energy they deposit in the crystal. This can be estimated by the sum of the integrated PMT outputs. Although this measurement is only accurate to within approximately +/- 10% on most scanners, it can be used with a simple threshold to reject a significant fraction of the scattered events. For scanners using sodium-iodide detectors [5], this accuracy improves to +/- 5%. This not only improves the effectiveness of energy discrimination, but also improves the accuracy of the scatter correction [7, 8].

Accidental Coincidences

Given the large number of scattered photons and the relatively small solid angle subtended by the detector ring, it is apparent that for many annihilations only one of the photons will be detected. These events are termed *singles.* If two singles arising from separate annihilations are detected within the same coincidence timing window, they will be recorded as shown in Fig. 4. These events are termed *accidental coincidences* or *randoms.* The rate of accidental coincidences can be related to the singles rate by noting that for each single detected at detector i, on average τR_j singles occur at detector j during the coincidence timing window τ, where R_j is the singles rate at detector j. Since each of these τR_j singles results in a coincidence, there are $\tau R_i R_j$ coincidences per unit time for which the first detected photon is incident on detector i. The total number of accidental coincidences is the sum of those for which the first photon is detected at detector i and those for which the first photon is detected at detector j. Hence, the rate of random coincidences along the LOR connecting detectors i and j is given by

$$R_r = 2\tau R_i R_j. \tag{2}$$

Examination of Eq. (2) shows that reducing the coincidence timing window reduces the counting rate of accidental coincidences. However, timing inaccuracies due to variations in the rise-time of the crystal light output require a timing window of 10-15 ns for BGO. Since the incident singles rates are proportional to the amount of injected isotope, the accidental coincidence rate increases as the square of the amount of isotope in the field of view (for counting rates that do not saturate the detectors). This count-rate limitation, along with detector deadtime, determines the upper limit on the injected dose for many studies.

Detector Efficiency

The efficiency of photon detection varies not only from block detector to block detector, but also varies widely across the elements of a block detector [6]. Referring to Fig. 2, it is apparent that a photon scattering in a central element will

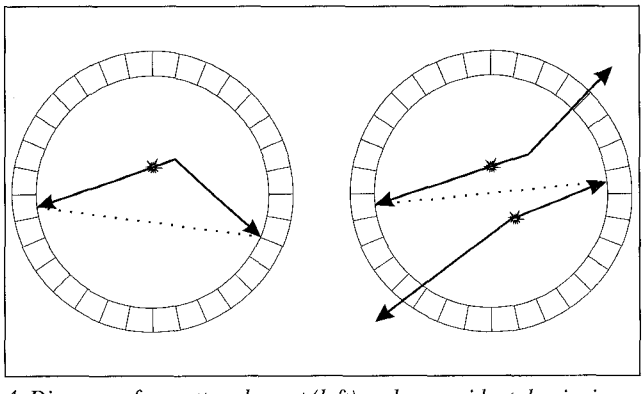
probably deposit the remainder of its energy in adjacent crystals. It might not be positioned as accurately as an event that deposits all of its energy in a single crystal, but it will be detected. A photon scattering in an edge crystal, on the other hand, has a significant probability of scattering out of the entire block and not being detected at all. This results in a decrease of detection efficiency in the edge crystals relative to the center crystals. This efficiency is different for scattered and true events because scattered events have different photon energies and, for a given line of response, the scattered photons arrive from a wide range of angles while unscattered photons detected along a given LOR all arrive at nearly the same angle.

Detector Deadtime

The time required to process a single event limits the counting rate of a PET scanner [9]. Event processing begins with the rising edge of the pulse for the first detector involved. The pulse is integrated for some time interval, then position calculations and energy discrimination are performed. The detector is "dead" to new events during this time. At very low counting rates, randoms are negligible and the number of true events is linearly related to the amount of activity in the field of events. The number of randoms increase as the square of the activity in the field of view until deadtime becomes significant. Then the number of true events begins to saturate. As the counting rate increases further, the numbers of trues and randoms peak and then decline because of detector saturation. Deadtime is the dominant effect that limits the injected dose.

Fully Three-Dimensional PET

In the foregoing discussion we have assumed that data are collected in 2D planes. Current scanners have retractable septa so that coincidences can be acquired between all possible pairs of detectors, a mode called *fully 3D* PET [10]. The effective sensitivity of such a scanner increases by up to a factor of eight, resulting in significant reductions in either the image variance or the injected dose. This improvement comes at the cost of a large increase in the scatter fraction and singles rates (due to less shielding for annihilations inside the field of view and a larger acceptance angle for annihilations outside the field of view of the camera). Until recently, fully 3D data were not used for two reasons: the unavailability of image reconstruction algorithms and the necessity of rejecting scattered events with axial collimation. The advent of appropriate reconstruction algorithms [11] and scatter-correction algorithms [12-14] has changed this. Fully 3D imaging is finding increasing use except in studies where counting rates are very high or shielding from regions just outside the field of view is required.



4. Diagram of a scattered event (left) and an accidental coincidence (right). Photons shown leaving the ring are scattered through an oblique angle such that their paths do not intersect a detector.

A Physical Model

If statistical effects are ignored, these factors can be incorporated into a model for the total number of recorded events to yield

$$Y_{d\theta} = \gamma_{d\theta}\left[\eta^{t}_{d\theta} P_{d\theta} M_{d\theta} + \eta^{r}_{d\theta} r_{d\theta} + \eta^{s}_{d\theta} s_{d\theta}\right] \tag{3}$$

where $M_{d\theta}$ is the number of annihilations with photons emitted along the LOR specified by (d, θ) in Fig. 3, $P_{d\theta}$ is the survival probability as defined in Eq. (1), $r_{d\theta}$ is the number of accidental coincidences, $s_{d\theta}$ is the number of scattered events, $\eta^{t}_{d\theta}$ is the probability of detection for true events, $\eta^{r}_{d\theta}$ is the probability of detection for accidental coincidences, $\eta^{s}_{d\theta}$ is the probability of detection for scattered events, and $\gamma_{d\theta}$ is the probability of an event not being lost due to deadtime.

Of the effects included in Eq. (3), attenuation is not only the most pronounced but also the most straightforward to characterize. Prior to the emission scan, a transmission scan is performed. Here a rotating line source containing a long-lived isotope rotates around the subject to provide a nonzero flux of photons along each line of response. The measured data yield the number of transmitted events, $T_{d\theta}$, along each LOR. Every morning a blank scan, i.e., a transmission scan with nothing in the scanner, is performed to yield a data set, $B_{d\theta}$. The survival probability given by Eq. (1) is approximated by their ratio,

$$\hat{P}_{d\theta} = \frac{T_{d\theta}}{B_{d\theta}}. \tag{4}$$

This estimate of survival probabilities would be exact if the data were noiseless. However, they are not noiseless, so they contribute significantly to the overall image variance unless noise-reduction algorithms are applied. These algorithms utilize smoothing [15], segmentation and reprojection [16, 17], or statistical image reconstruction and reprojection [18-21].

A simple way to estimate the accidental coincidences is to note that the arrival times of the photons due to randoms are uniformly distributed in time while those of true coincidences fall within the coincidence timing window. Collecting data in a second coincidence timing window that is offset in time such that it collects no true coincidences yields data with

nearly the same mean as that of the accidental coincidences falling in the trues timing window. The measured data are given by the product $\gamma_{d\theta}\eta^{r}_{d\theta}r_{d\theta}$, so the detector efficiencies for accidental coincidences, $\eta^{r}_{d\theta}$, do not have to be estimated. Therefore, not only is the method simple to implement, but it can be performed in hardware before the data are stored. The major drawback of this approach is that the variance of the estimate is of the same order of magnitude as the variance of the data if a significant fraction of detected events are accidental coincidences. In this case, the subtraction can lead to a significant increase in the variance of the data unless noise reduction methods are used [22]. This variance increase can be avoided by counting the number of singles at each detector and using Eq. (2). Since there are many more singles than true coincidences, the effect on variance is relatively minor. This approach is not widely used because of the additional requirements placed on the acquisition hardware and because singles rates often vary over the course of an acquisition.

For septa-extended scans, the fraction of scattered events is low (approximately 15% of the total number of collected events). They are usually estimated as an integral transformation of the measured data using an empirically determined kernel [23]. For fully 3D scans, the scatter fraction rises to 30-50% of the total number of events. It can be estimated using a mathematical model of the scanner and scattering process [13, 24] or by utilizing data collected in a second, lower energy window that acquires a higher fraction of scattered events [12, 14].

The detector efficiencies for true and scattered events are estimated from a scan of a calibration source with known characteristics [9]. Deadtime is dependent on many factors related to the architecture and design of a specific machine, so its estimation is tailored to the scanner [25]. It is usually assumed to be constant over the duration of the scan.

These parameters can be used to estimate the number of emitted photons by using the expression

$$\hat{M}_{d\theta} = \frac{1}{\gamma_{d\theta}\eta^{T}_{d\theta}\hat{P}_{d\theta}}\left((Y_{d\theta} - R_{d\theta}) - \eta^{s}_{d\theta}S_{d\theta}\right) \tag{5}$$

where we assume that $R_{d\theta} = \gamma_{d\theta}\eta^{r}_{d\theta}\mathrm{E}[r_{d\theta}], S_{d\theta} = \gamma_{d\theta}\mathrm{E}[s_{d\theta}]$ and E[·] denotes expectation.

The data modeled by Eq. (3) are often stored in 2D arrays with the columns indexed by d and the rows by θ. These data arrays are often called *sinograms*, because, for a point source, d varies sinusoidally with θ.

Image Reconstruction Assuming Deterministic Data

The goal of image reconstruction is to recover the radiotracer concentration from the measurements. This inverse problem is not unlike the classical signal processing problem of *deconvolution* [26]. However, straightforward application of "off-the-shelf" signal processing and image restoration methods yields suboptimal results for PET image reconstruction.

In this section we summarize some of the methods that have been proposed for PET image reconstruction, with a particular emphasis on those with origins in signal and image processing. This review is by no means complete and is primarily intended to describe the potential "pitfalls" of each approach. Most of the discussion also applies to SPECT image reconstruction. We begin by deriving a widely used linear algorithm, then we discuss pre- and post-processing techniques proposed for use with it and end with a survey of statistical image reconstruction methods.

Filtered Backprojection

One way to greatly (over)simplify the problem is to ignore the measurement noise altogether, and to assume that the measured data approximate line integrals through the radioisotope distribution. This leads to the classical filtered-backprojection (FBP) method for tomographic image reconstruction [27]. This method is used routinely for x-ray CT, as well as for PET and SPECT. Its widespread popularity stems from historical reasons of computational simplicity, not because of any widely accepted advantage in image quality. Since it is derived without any statistical information, it is unsurprising that use of the unmodified FBP method leads to unacceptable noise amplification in PET.

Filtered backprojection was first applied to PET by Shepp et al. [28]. Introductory treatments of the algorithm can be found in [27] and [29] and more comprehensive treatments in [3] and [30]. The distribution of the radio-isotope is modeled by the function $\lambda(x,y,z) \in L^2$. For a given 2D slice, we assume that the mean of an individual measurement $Y_{d\theta}$ is given by $g_\theta(d) = \int_{l(d,\theta)} \lambda(x,y,z)dxdydz$ where $l(d,\theta)$ is the line connecting the two detectors involved in the coincidence. In practice, it is assumed that the mean $g_\theta(d)$ is equal to the corrected data, $\hat{M}_{d\theta}$, in Eq. (5). In the rotated coordinate system of Fig. 5, $d = x_\theta$, so the line integral can be expressed as

$$g_\theta(x_\theta) = \int_{-\infty}^{\infty} \lambda(x_\theta, y_\theta)dy_\theta, \ \theta \in (0,\pi), x_\theta, y_\theta, \in \Re \tag{6}$$

where x_θ represents transverse distance in the rotated coordinate system shown in Fig. 5. We will refer to the function $g_\theta(x_\theta)$ (and the data it approximates) as a *projection.*

The Fourier transform of each projection is given by

$$\begin{aligned} G_\theta(u_\theta) &= \int_{-\infty}^{\infty} g_\theta(x_\theta)e^{-j2\pi u_\theta x_\theta}dx_\theta \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\lambda(x_\theta,y_\theta)dy_\theta\right]e^{-j2\pi u_\theta x_\theta}dx_\theta \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\lambda(x_\theta,y_\theta)e^{-j2\pi u_\theta x_\theta}dy_\theta dx_\theta \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\lambda(x_\theta,y_\theta)e^{-j2\pi(u_\theta x_\theta + v_\theta y_\theta)}dy_\theta dx_\theta\Big|_{v_\theta=0} \\ &= \Lambda(u_\theta, v_\theta)\Big|_{v_\theta=0} \end{aligned} \tag{7}$$

This result, known as the *projection-slice theorem*, has two implications. First, the Fourier transform of a projection

yields samples of the 2D Fourier transform of the image, and second, these samples lie along a line at the same angle, θ, in the frequency domain as that of the projection in the spatial domain. This result can be written in more standard notation as

$$G_\theta(\rho) = \Lambda(\rho,\gamma)\big|_{\gamma=\theta,\rho\in\Re} \tag{8}$$

where the Fourier transform of the image is now expressed in polar coordinates (ρ, γ). Eq. (7) can be used to reconstruct the image by constructing the Fourier transform in polar coordinates, interpolating to rectangular coordinates, and then taking the inverse transform. A more efficient method can be derived as follows. The image $\lambda(x,y)$ is given by

$$\lambda(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Lambda(u,v) e^{j2\pi(ux+vy)}\,du\,dv. \tag{9}$$

Transforming to polar coordinates as shown in Fig. 5 using the expressions $u = \rho\cos\theta$, $v = \rho\cos\theta$, $x = r\cos\phi$, and $y = r\sin\phi$ yields

$$\lambda(r,\phi) = \int_0^{2\pi}\int_0^{\infty} \rho\Lambda(\rho\cos\theta,\rho\sin\theta) e^{j2\pi\rho r\cos(\phi-\theta)}\,d\rho\,d\theta. \tag{10}$$

Rewriting $\Lambda(\rho\cos\theta, \rho\sin\theta)$ as $\Lambda(\rho,\theta)$ and using the facts that $\cos(\phi-\theta) = -\cos(\phi-\theta+\pi)$ and $G_\theta(\rho) = G_{\theta+\pi}(-\rho)$, this can be rewritten as

$$\lambda(r,\phi) = \int_0^{\pi}\int_{-\infty}^{\infty} |\rho|\Lambda(\rho,\theta) e^{j2\pi\rho r\cos(\phi-\theta)}\,d\rho\,d\theta. \tag{11}$$

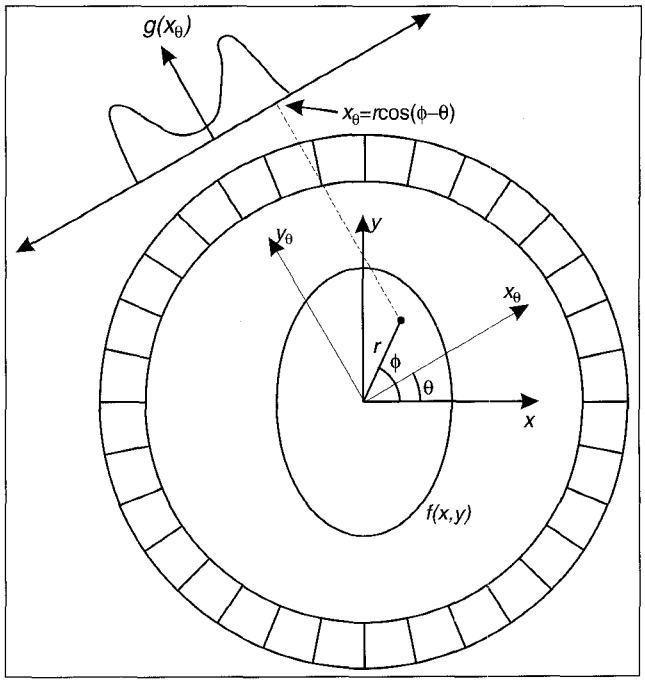
Applying the projection-slice theorem leads to

$$\begin{aligned}\lambda(r,\phi) &= \int_0^{\pi}\int_{-\infty}^{\infty} |\rho| G_\theta(\rho) e^{j2\pi\rho r\cos(\phi-\theta)}\,d\rho\,d\theta \\ &= \int_0^{\pi} \tilde{g}_\theta(r\cos(\phi-\theta))\,d\theta\end{aligned} \tag{12}$$

where $\tilde{g}_\theta(x) = \Im^{-1}\{|\rho| G_\theta(\rho)\}$ and $\Im^{-1}\{\cdot\}$ denote the inverse Fourier transform. Discretizing leads to the expression

$$\lambda(r,\phi) = \frac{\pi}{N_\theta}\sum_{i=0}^{N_\theta-1} \tilde{g}_i(r\cos(\phi-\theta_i)). \tag{13}$$

Equation 13 shows that the value of the image at a point ($r\cos\phi$, $r\sin\phi$) in Fig. 5 can be found by first filtering the projections with a ramp filter, then summing the filtered values at the coordinate $x_{\theta_i} = r\cos(\theta_i - \phi)$ over all projection angles θ_i. Note that the value at x_{θ_i} will contribute to all pixels along the LORs that contributed to the measurement at this point. The algorithm can be efficiently implemented by filtering each estimated projection, $g_\theta(d) \approx \hat{M}_{\theta d}$, with a ramp filter to yield $\tilde{g}_\theta(d)$ and then adding each filtered value into all voxels along the corresponding LOR as shown by the dashed line in Fig. 5. (Note that the discretization and the finite support of the image and projections necessitate modifications to the filter [27].) The latter operation is called *backprojection*, so the algorithm is unsurprisingly called



5. Projection geometry.

filtered-backprojection. This algorithm and its extension to three dimensions [11, 31] is used almost exclusively for image reconstruction in PET. It is identical to the algorithm used in x-ray CT except for modifications to the filter necessitated by the noise properties of PET data.

There are several problems with this algorithm. First, although the intensity is known to be non-negative, the algorithm yields negative values, particularly if the data are noisy. Second, models for the detector response must be space invariant and can only be incorporated into the algorithm as a deconvolution with the attendant noise amplification. Finally, and most importantly, the ramp filter accentuates high frequency noise. This effect can be seen by examining the magnitude spectrum of the typical and low-noise projections of the same image shown in Fig. 6. (The low-noise projection was found by reconstructing the image and reprojecting it to form an estimated projection. The variance of the noise in this estimated projection will be reduced by a factor approximately equal to the number of projections. In this case 192 projections were used.) It is apparent that reconstructing with an unwindowed ramp filter is unwise.

For frequencies above 0.8 cm^{-1} the data are dominated by noise so the resulting images would be too noisy, as shown at the top left in Fig. 7. Moreover, in many systems, frequencies near the foldover frequency are significantly aliased and should be rejected. Therefore, the ramp is often truncated at one-half the foldover frequency as shown at the bottom in Fig. 6. The effect on the image is shown at the top right in Fig. 7. Although this image is still too noisy for visualization, it would be useful for quantitative measurements that involve averaging over a region. This window degrades the resolution from an intrinsic resolution of 5.2 mm to a reconstructed resolution of 5.6 mm FWHM. For visualization purposes, the ramp filter is often apodized with a Hanning, Parzen, or

Butterworth window. An image reconstructed with a fifth-order Butterworth window with a cutoff frequency of 1 cm^{-1} is shown at the bottom left in Fig. 7. Examination of the image shows what appears to be a small defect in the thalamus, as shown by the arrow. This particular subject was scanned again in the fully 3D mode three minutes after the first scan, yielding the image shown at the bottom right in Fig. 7. (The fully 3D image differs from the 2D image in that more counts were collected and the image was sampled by many more LORs.) There is no evidence of the defect in this image. The apparent defect is probably due to noise at spatial frequencies near 1 cm^{-1}, which are not attenuated by the Butterworth filter. In this case, filtering gives the impression of a noise-free image by reducing high-frequency noise but does not eliminate low-frequency artifacts. Concern over such issues leads naturally to the development of more sophisticated algorithms.

Sinogram Preprocessing

The apodization window applied to the reconstruction filter is equivalent to smoothing the projections prior to reconstruction. Although this smoothing does reduce the noise variance, it is suboptimal since PET measurement statistics are nonstationary because they follow a Poisson distribution. There have been several attempts to improve the sinogram smoothing using both iterative [32-34] and noniterative [35] nonstationary methods. While requiring less computation than the iterative methods described below, these preprocessing methods are still suboptimal since object constraints such as non-negativity and piecewise smoothness are not naturally expressed in the sinogram domain.

Image Post-processing

The radiotracer distribution estimate computed by any reconstruction method is typically represented by a discrete image. This certainly invites the application of many an image processing method, both those classical (such as Wiener filtering) as well as those trendy (such as wavelets, neural nets, etc.). Unfortunately, most image processing methods are based on the (often implicit) assumption that the noise is Gaussian, or at least independent from pixel to pixel. The noise in tomographic images is generally highly correlated between neighboring pixels (since each measurement "ray" transects many pixels). For the (linear) FBP method the correlation function can be calculated [36-38]. In principle the noise correlation function can also be determined for some statistical image reconstruction methods [39-41], although the correlation functions may be expensive to compute. In our experience, classical image processing methods perform poorly for images with such correlated noise. Furthermore, the correlation structure is often nonstationary, so noise prewhitening is usually impractical. On the other hand, post-processing methods that specifically account for the correlation structure have shown some promise, e.g., [42].

Statistical Image Reconstruction

A More Complete Model of the Data

This summary of statistical reconstruction methods is condensed from [43]. The measurement statistics are quite complex, so any treatment (including ours) must make simplifying assumptions. However, many papers in the signal processing and statistics literature oversimplify the problem, e.g., [44], so we attempt to be somewhat more complete here. We modify the notation used earlier to emphasize functional dependencies.

Since PET measurements are based on a counting process, a reasonable statistical model is that the measurements have independent Poisson distributions (If a deterministic finite number of nuclei are injected into the patient, then, strictly speaking, a multinomial distribution would be more precise than the Poisson assumption. However, in practice the exact number of nuclei is unknown and may well be considered a random variable with a Poisson distribution. In this case the radioactive decay will be a Poisson process; furthermore, a Poisson process "thinned" by Bernoulli trials remains Poisson [45], all of which leads to the Poisson model.):

$$Y_i \sim \text{Poisson}\{\bar{Y}_i(\lambda)\}, \quad i = 1, \ldots, n, \tag{14}$$

where n is the number of coincident detector pairs, λ is the spatial distribution of radio-tracer (typical units are counts/s/cm^3), and $\bar{Y}_i(\lambda)$ is the mean of the ith measurement. (Note that each i corresponds to a unique $d\theta$ pair in the notation used above.) The measurement means depend on the radio-tracer distribution $\lambda(\underline{x})$ through the physical model described above; for low to moderate counting rates, the dependence is nearly linear in λ:

$$\bar{Y}_i(\lambda) = T\left(\int p_i(\underline{x})\lambda(\underline{x})d\underline{x} + s_i(\lambda) + r_i(\lambda)\right) \tag{15}$$

where T is the scan time, $p_i(\underline{x})$ is the (unitless, scatter-free) point-response function of the ith detector pair ($p_i(\underline{x})$ is probability that a positron emitted from a nuclei at position $\underline{x}$ will produce a pair of annihilation photons that are detected by the ith detector pair without scattering (including geometric effects, attenuation, and detector efficiencies)), $s_i(\lambda)$ is mean rate of detected scattered events for the ith detector, $r_i(\lambda)$ is the mean rate of detected random coincidences for the ith detector pair, and the integral is over the scanner field of view. (For detector i indexed by $d\theta$ in our previous notation, $r_i(\lambda) = \gamma_{d\theta}\eta^r_{d\theta}r_{d\theta}$ and $s_i(\lambda) = \gamma_{d\theta}\eta^s_{d\theta}s_{d\theta}$, and $[x,y]$ is replaced by the vector $\underline{x}$.) Although the scatter contribution $s_i(\lambda)$ is linear in λ, the random coincidences $r_i(\lambda)$ depend nonlinearly on λ (if the detectors are not saturated, the singles rates increase monotonically with λ and the randoms increase as the square of the singles rates as described by Eq. (2)). For most scanners, the singles rates required to model this dependence directly are not available, so the estimates obtained with the

delayed coincidence window [46] are used to obtain information about $r_i(\lambda)$.

For moderate counting rates, the linearity in λ implied by the first term in Eq. (15) is reasonable. However, for high count rates, the measurement means are highly nonlinear functions (they are, in fact, nonmonotonic functions) of the activity in the patient due to scanner deadtime [25]. In practice, the effect of this nonlinearity is often assumed to be reducible to a single "deadtime correction factor" for each plane, or, more accurately, by different correction factors for different detector pairs or detector blocks. This type of correction implicitly separates the nonlinear deadtime loss from the ideal linear relationship between λ and Y_i. We are unaware of any attempts to include the deadtime nonlinearity directly in the forward model. We also take the "separable" approach here.

Classical Estimation Methods

Since a PET scanner collects only a finite number of measurements, one must, in general, also represent the radiotracer distribution $\lambda(\underline{x})$ by a finite parameterization, e.g., in terms of a set of basis functions:

$$\lambda(\underline{x}) = \sum_{j=1}^{p} \theta_j b_j(\underline{x}) \tag{16}$$

where $\theta = [\theta_1 \cdots \theta_p]'$ is the vector of unknown coefficients that must be computed from the Y_i's. (Typically $b_j(x)$ is just the indicator function for the jth voxel, so we will refer to θ_j as the jth pixel value hereafter.) With such a discretization, the reconstruction problem is equivalent to a parameter estimation problem. If one assumes the scatter and random contributions are predetermined values s_i and r_i, respectively (i.e., if they are determined separately), and if the deadtime nonlinearity is approximated by a single "known" loss factor d_i, then the measurement mean is linear in θ:

$$\bar{Y}_i(\theta) = \sum_{j=1}^{p} a_{ij}\theta_j + s_i + r_i \tag{17}$$

where

$$a_{ij} = d_i T \int p_i(\underline{x}) b_j(\underline{x}) d\underline{x}. \tag{18}$$

Dozens of papers have been published based on this model, most of which not only ignored the d_i, r_i, and s_i terms, but also used very simple approximations for $p_i(\underline{x})$. The linear form above invites application of the two most common tools from statistical signal processing: maximum likelihood estimation and linear least-squares estimation. The linear least-squares estimate is easily written as

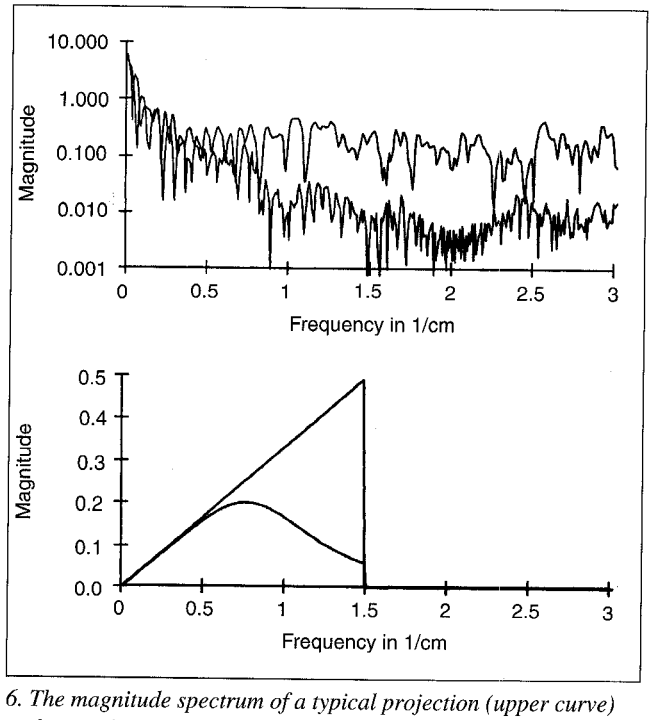
$$\hat{\theta}_{LS} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'(\underline{Y} - \underline{s} - \underline{r}), \tag{19}$$

but this expression is impractical for computation due to the large size of the matrix $\boldsymbol{A} = \{a_{ij}\}$. (The dimension of matrix $\boldsymbol{A}$ is on the order of 16384 x 60000 for a single plane of a typical scanner.) Furthermore, the conventional linear least-squares estimate produces negative pixel values, which are physically impossible. This can be a significant problem in low activity regions of the image. Both the size of $\boldsymbol{A}$ and incorporation of the non-negativity constraint necessitate iterative algorithms.

Although necessary because of existing instrumentation, the real-time correction for random coincidences using the delayed-window method renders the data non-Poisson. For such measurements, estimates based on (weighted) least-squares may be suitable [47]. (Also see [48] for more accurate approaches.) For scans that are not precorrected for randoms, the least-squares methods are suboptimal since they do not fully accommodate the Poisson distribution. (Often the number of counts per ray is sufficiently low that the Gaussian approximation to the Poisson distribution is inapplicable.) Furthermore, data-based weighted least-squares methods lead to systematic biases for low-count Poisson measurements [18, 41]. This problem can be avoided by using the measurement log-likelihood $L(\theta)$ rather than the weighted least-squares criterion, where

$$L(\theta) = \sum_{i=1}^{n} \left(Y_i \log \bar{Y}_i(\theta) - \bar{Y}_i(\theta) - \log Y_i! \right). \tag{20}$$

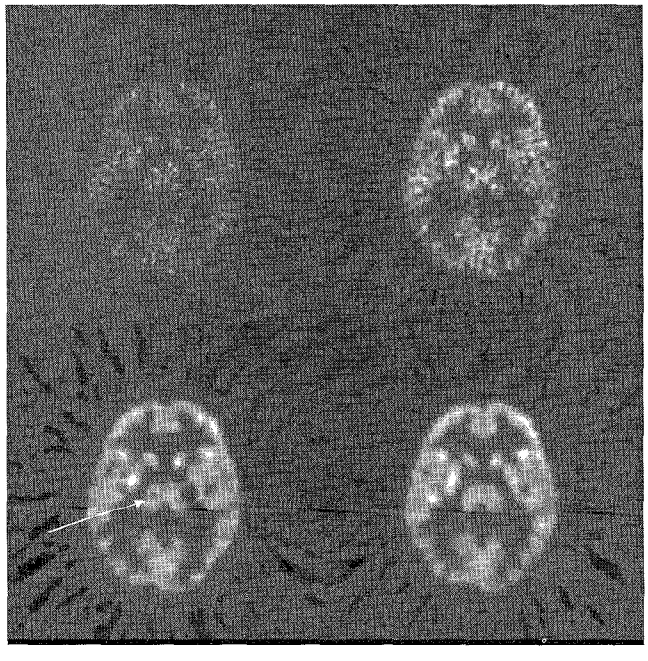
Unfortunately, there is no closed-form expression for the estimate $\hat{\theta}_{ML}$ that maximizes the likelihood, which again necessitates iterative algorithms. Unfortunately, *each itera-*



6. The magnitude spectrum of a typical projection (upper curve) and a nearly noiseless projection (bottom curve) are shown at the top. Two practical filters are shown at the bottom: a ramp filter cut off at 50% of the Nyquist rate and the same filter windowed with a Butterworth filter.

tion of these algorithms requires computation time roughly comparable to that required by the FBP method. This has hampered their clinical acceptance. The oldest of these algorithms (for PET) is an expectation-maximization (EM) algorithm [49], which converges very slowly to $\hat{\theta}_{ML}$. This slow convergence has not greatly diminished the popularity of the EM algorithm, however, because the intermediate images generated during the iterations toward $\hat{\theta}_{ML}$ are usually more appealing then $\hat{\theta}_{ML}$ itself. (Determining which of the many iterates is the best one is nontrivial however.) The problem of determining $\lambda(\underline{x})$ from $\{Y_i\}$ is inherently ill-posed [50], so, after parameterization, the problem of estimating θ is generally very ill-conditioned. Thus $\hat{\theta}_{ML}$ is usually extremely noisy [50].

Naturally, one simple way to reduce this noise is to postsmooth $\hat{\theta}_{ML}$. Such postsmoothing is a special case of the more general *method of sieves* [50] and is in fact by far the most popular version of the sieve method. Postsmoothing has two disadvantages. First, in its usual form of space-invariant filtering, the nonstationarity of the measurement statistics cannot be modeled. And second, although postsmoothing reduces noise, the problem of slow convergence of the EM algorithm remains, and hundreds to thousands of EM iterations may be required for the postsmoothed images to converge [51]. This problem has spawned a variety of methods for accelerating the EM algorithm, which vary in the extent to which convergence is guaranteed (see [52, 53]).



7. A glucose metabolism image reconstructed with a ramp filter (top left), a ramp filter cut off at one-half the Nyquist frequency (top right), a ramp filter cut off at one-half the Nyquist frequency windowed with a fifth order Butterworth filter with a cutoff frequency of 1 cm^{-1} (bottom left), and data of the same subject acquired in the fully 3D mode (bottom right). The fully 3D image was reconstructed from more finely sampled data containing a higher number of counts.

Classical Regularization Methods

Another way to overcome the problems of slow convergence and to reduce the image noise is to replace the log-likelihood criterion by a penalized-likelihood objective function:

$$\hat{\theta}_{PL} = \arg\max_{\theta} \Phi(\theta) \text{ where } \Phi(\theta) = L(\theta) - \beta R(\theta),$$

where $R(\theta)$ is a measure of image roughness. Larger values of β encourage smoother images with less noise. When first investigated for PET, the penalty function posed a computational challenge since the M-step of the EM algorithm has no closed form [54-56]. However, now there are a variety of fast algorithms (compared to EM) available for maximizing such objective functions, e.g., [47, 52, 53, 57, 58]. These algorithms converge rapidly in part because the penalty function greatly improves the conditioning of the reconstruction problem.

In the context of least-squares problems, such regularization methods date at least to the early '70s [59], so many may well be considered "classical." The most classical penalty function simply measures the norm of the image:

$$R(\theta) = \|\theta\|^2 = \sum_{j=1}^{p} \theta_j^2,$$

which has its origins in ridge-regression. This simple penalty leads to images that are "squashed down" since even the DC component is penalized. For reducing noise, a more suitable penalty is to discourage neighboring pixels from having disparate values:

$$R(\theta) = \sum_{j=1}^{p} \sum_{k \in N_j} \psi(\theta_j - \theta_k),$$

where N_j is the set of pixel indices in the neighborhood of pixel j, and $\psi(t)$ is a symmetric function typically chosen to be nondecreasing for $t \leq 0$. Such penalty functions (or "priors" in the Bayesian terminology) have yielded good results in image restoration and image segmentation problems. However, in PET the nonstationary noise statistics again complicate the problem. Although $R(\theta)$ above is a shift-invariant function, recent analysis shows that images reconstructed by maximizing $\Phi(\theta)$ have nonuniform spatial resolution, due to interactions between the log-likelihood and penalty terms [60, 61]. (Such effects are absent in image restoration problems with white Gaussian noise.) Although modified penalty functions have been proposed that reduce the resolution nonuniformity, these modifications cause more nonuniform noise variance [61, 62].

Another challenge in penalized-likelihood methods is choosing β. This problem is comparable to that of choosing the width of the apodizing window in FBP or the resolution of the filter used when postsmoothing ML images. However,

in the latter two problems the parameter that one varies to tradeoff resolution and noise is one that is naturally related to spatial resolution, whereas β has essentially arbitrary units. Automatic or data-based methods for choosing β, e.g., [63, 64]. have shown some potential, but may also be unstable in imaging problems [65].

There is also no consensus on the best choice for $\psi(t)$. Quadratic penalties lead to oversmoothing, and nonquadratic penalties require additional parameters that must be chosen. Nonconvex penalties cause additional problems with algorithm convergence, but have led to impressive results in image restoration problems in images with sharply defined regions [66]. However, in medical images one must take care to avoid turning smooth transitions into stair steps [67].

Model Errors

Nearly all papers on model-based methods for PET image reconstruction assume that the measurement model is known, particularly the "system matrix" $\boldsymbol{A}$. In practice this matrix is occasionally measured, or more commonly simply computed based on an approximate geometry. In either case $\boldsymbol{A}$ contains errors, and the effect of this model mismatch on $\hat{\theta}$ is poorly understood. The errors in $\boldsymbol{A}$ might invite the application of the *total least-squares* (TLS) estimation method, e.g., [68]. However, TLS assumes that the errors in $\boldsymbol{A}$ are normally distributed, which is questionable in PET. Furthermore, $\boldsymbol{A}$ usually includes attenuation factors that are determined from separate noisy transmission scans. Understanding the effects of both deterministic and random errors in the model remains an important problem.

Attenuation Correction

As described above, the conventional attenuation correction method in PET uses the ratio of the measurements in the blank and transmission scans. The transmission measurements can be very noisy, and, with randoms subtraction, can even take on negative or zero values. This noise is usually reduced either by smoothing with a space-invariant filter [15] or by reconstructing an image of the attenuation coefficient from the line integrals in Eq. (1), segmenting it, and reprojecting it [16, 17]. However, these methods introduce bias and ignore the nonstationary statistics. More accurate attenuation correction factors can be computed by first using statistical methods to reconstruct an attenuation image while incorporating nonlinear constraints such as non-negativity and piecewise smoothness, and then reprojecting this image along all of the LORs [18-21].

SPECT

Most of the above discussion also applies to SPECT imaging. Statistical methods are perhaps even more useful in SPECT than in PET for two major reasons. First, attenuation is depth dependent and cannot be precorrected [69], and second, the resolution of collimators degrades with distance [70]. Both of these effects can be incorporated directly into statistical models [71]. In fact, for SPECT cardiac studies, statistical methods are now in routine use at some centers, e.g., [72], and the EM algorithm is available commercially.

Computing Speed and the Future

Since computers are continually increasing in speed and memory, it might seem at first that it is only a matter of time before iterative reconstruction methods become routinely used. However, the same advances in technology that lead to faster computers also lead to bigger and harder problems! For example, although computing speed certainly has reached the point where iterative methods are clinically feasible for 2D problems, the focus is now on 3D PET where the size of $\boldsymbol{A}$ is 11-15 times larger than in 2D (after exploiting symmetries). Similar considerations apply to cone-beam SPECT, or even to parallel collimator SPECT with 3D compensation for detector response. Thus, there is continuing need for new ideas in image reconstruction algorithm development. Although some of those ideas will undoubtedly be borrowed from signal and image processing work, the algorithms must be based on accurate models of the physics and statistics of PET if they are to be fully effective. Convincingly demonstrating that new methods are truly more effective than previous methods requires careful matching of the resolution or noise properties of the methods compared. The medical imaging community is generally unconvinced by the type of anecdotal, single-image comparisons often found in image processing papers. There is increasing emphasis on formal statistical evaluations of different image reconstruction methods [73-75], which are also being applied to image processing [76].

Conclusion

The image formation process in PET lends itself well to relatively simple algorithms that yield accurate results when there are good counting statistics. Statistical methods can yield improved image quality but have not been widely adopted, largely because of their computational complexity. They play a more significant role in SPECT because they accurately incorporate models of attenuation and collimator resolution.

Acknowledgment

This work was supported by grants CA-60711, CA-54362, and CA-54959 from the National Cancer Institute, and by grant number 1380 from the National Center for Research Resources.

John M. Ollinger is an Assistant Professor of Biomedical Computing and Radiology at Washington University, St. Louis, Missouri 63110, (jmo@ibc.wustl.edu). *Jeffrey A. Fessler* is an Assistant Professor of Electrical Engineering at

the University of Michigan, 4240 EECS Bldg., Ann Arbor, Michigan 48109-2122 (fessler@umich.edu).

References

1. J.A. Sorenson and M.E. Phelps, *Physics in Nuclear Medicine*, 2nd ed. Orlando: Grune & Stratton Inc, 1987.

2. M.M. Ter-Pogossian, M.E. Raichle, and B.E. Sobel, "Positron-Emission Tomography," *Scientific American*, vol. 243, pp. 170-181, 1980.

3. S.R. Deans, *The Radon Transform and Some of its Applications*. New York: Wiley, 1983.

4. M. Dahlbom and E.J. Hoffman, "An Evaluation of a Two-Dimensional Array Detector for High Resolution PET," *IEEE Trans. Med. Imag.*, vol. 7, pp. 264-272, 1988.

5. J.S. Karp, G. Muehllehner, D.A. Mankoff, C.E. Ordonez, J.M. Ollinger, et al., "Continuous-slice PENN-PET: A positron tomograph with volume imaging capability," *J. Nucl. Med.*, vol. 31, pp. 617-627, 1990.

6. S.E. Derenzo, "Mathematical Removal of Positron Range Blurring in High Resolution Tomography," *IEEE Trans. Nucl. Sci.*, vol. 33, pp. 565-569, 1986.

7. R.L. Harrison, D.R. Haynor, and T.K. Lewellen, "Dual Energy Window Scatter Corrections for Positron Emission Tomography," *Conf. Rec. of the 1991 IEEE Med. Imag. Conf.*, pp. 1684-1688, 1992.

8. J.M. Ollinger, "The Effect of Energy Threshold on Image Variance in Fully 3D PET," *Proc. 1995 Intl. Meeting on Fully 3D Image Reconstruction in Radiology and Nucl. Med.*, vol. 1, pp. 221-224, 1995.

9. E.J. Hoffman, T.M. Guerrero, G. Germano, W.M. Digby, and M. Dahlbom, "PET System Calibration Corrections for Quantitative Spatially Accurate Images," *IEEE Trans. Nucl. Sci.*, vol. 36, pp. 1108-1112, 1989.

10. T.J. Spinks, T. Jones, D.L. Bailey, D.W. Townsend, S. Grootoonk, et al., "Physical Performance of a Positron Tomograph for Brain Imaging with Retractable Septa," *Phys. Med. Biol.*, vol. 37, pp. 1637-1655, 1993.

11. P.E. Kinahan and J.G. Rogers, "Analytic Three-Dimensional Image Reconstruction Using all Detected Events," *IEEE. Trans. Nucl. Sci.*, vol. 36, pp. 964-968, 1989.

12. S. Grootoonk, T.J. Spinks, T. Jones, C. Michel, and A. Bol, "Correction for Scatter Using a Dual Energy Window Technique with a Tomograph Operating without Septa," *IEEE 1991 Medical Imaging Conference Record*, pp. 1569-1573, 1992.

13. J.M. Ollinger, "Model-Based Scatter Correction for Fully 3D PET," *Phys Med Biol*, vol. 41, pp. 153-176, 1996.

14. L. Shao, R. Freifelder, and J. S. Karp, "Triple Energy Window Scatter Correction Method for PET," *IEEE Trans. Med. Imag.*, vol. 13, pp. 641-648, 1994.

15. M.R. Palmer, J.G. Rogers, M. Bergstrom, M.P. Beddoes, and B.D. Pate, "Transmission Profile Filtering for Positron Emission Tomography," *IEEE Trans. Nucl. Sci*, vol. 33, pp. 478-481, 1986.

16. S.R. Meikle, M. Dahlbom, and S.R. Cherry, "Attenuation Correction Using Count-Limited Transmission Data in Positron Emission Tomography," *J. Nucl. Med*, vol. 34, pp. 143-144, 1993.

17. M. Xu, W.K. Luk, P.D. Cutler, and W.M. Digby, "Local Threshold for Segmented Attenuation Correction of PET Imaging of the Thorax," *IEEE Trans. Nucl. Sci.*, vol. 41, pp. 1532-1537, 1994.

18. J.A. Fessler, "Hybrid Poisson/polynomial Objective Functions for Tomographic Image Reconstruction from Transmission Scans," *IEEE Trans. Imag. Proc.*, vol. 4, pp. 1439-50, 1995.

19. J.M. Ollinger, "The Use Of Maximum A Posteriori And Maximum Likelihood Transmission Images For Attenuation Correction In PET," *Proc. 1992 IEEE Medical Imaging Conf.*, pp. 1185-1187, Orlando, 1992.

20. C.A. Bouman and K. Sauer, "A Unified Approach to Statistical Tomography Using Coordinate Descent Optimization," *IEEE Trans. Imag. Proc.*, vol. 5, pp. 480-92, 1996.

21. J.A. Fessler, E.P. Ficaro, N.H. Clinthorne, and K. Lange, "Grouped-coordinate Ascent Algorithms for Penalized-likelihood Transmission Image Reconstruction," *IEEE Trans. Med. Imag*, in press, 1996.

22. M.E. Casey and E.J. Hoffman, "Quantitation in Positron Emission Tomography: 7. A Technique to Reduce Noise in Accidental Coincidence Measurements and Coincidence Efficiency Calibration," *J. Computr. Asst. Tomogr.*, vol. 10, pp. 845-850, 1986.

23. M. Bergstrom, L. Eriksson, C. Bohm, G. Blomqvist, and J. Litton, "Corrections for Scattered Radiation in a Ring Detector Positron Camera by Integral Transformation of the Projections," *J. Computr. Assist. Tomogr.*, vol. 7, pp. 42-50, 1983.

24. C.C. Watson, D. Newport, and M.E. Casey, "A Single Scatter Simulation Technique for Scatter Correction in 3D PET," *Proc. 1995 Intl. Meeting on Fully 3D Image Reconstruction in Radiology and Nucl. Med.*, vol. 1, pp. 215-220, 1995.

25. M.E. Daube-Witherspoon and R.E. Carson, "Unified Deadtime Correction Model for PET," *IEEE Trans. Med. Imag.*, vol. 10, pp. 267-275, 1991.

26. G. Demoment, "Image Reconstruction and Restoration: Overview of Common Estimation Structures and Problems," *IEEE Trans. Acoust. Sp. Sig. Proc.*, vol. 37, pp. 2024-2036, 1989.

27. A.C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. New York, 1988.

28. L.A. Shepp and B.F. Logan, "The Fourier Reconstruction of a Head Section," *IEEE Trans Nucl Sci*, vol. 21, pp. 21-43, 1974.

29. G.T. Herman, *Image Reconstruction from Projections*. New York: Academic Press, 1980.

30. F. Natterer, *The Mathematics of Computerized Tomography*. New York: Wiley, 1986.

31. M. Defrise, D.W. Townsend, and R. Clack, "Three-Dimensional Image Reconstruction from Complete Projections," *Phys. Med. Biol.*, vol. 34, pp. 573-587, 1989.

32. T.J. Hebert and S.S. Gopal, "An Improved Filtered Back-projection Algorithm Using Pre-processing," *Proc. IEEE Nuc. Sci. Symp. Med. Imag.*, pp. 2068-2072, Santa Fe, 1991.

33. T.J. Hebert, "A Union of Deterministic and Stochastic Methods for Image Reconstruction," *Proc. IEEE Nuc. Sci. Symp. Med. Imag. Conf*, pp. 1117-1119, Orlando, 1992.

34. Z. Liang and R.E. Coleman, "Restoration for Detector Response in High Resolution PET Image Reconstruction," *J. Nucl. Med.*, vol. 33, pp. 872, 1992.

35. J.A. Fessler, "Tomographic Reconstruction Using Information Weighted Smoothing Splines," in *Lecture Notes in Computer Science*, vol. 687, H.H. Barrett and A.F. Gmitro, Eds. Berlin: Springer Verlag, 1993, pp. 372-386.

36. N.M. Alpert, D.A. Chesler, J.A. Correia, R.H. Ackerman, J.Y. Chang, et al., "Estimation of the Local Statistical Noise in Emission Computed Tomography," *IEEE Trans. Med. Imag.*, vol. 1, pp. 142-146, 1982.

37. M.F. Kijewski et al., "The Noise Spectrum of CT Images," *Phys. Med. Biol.*, vol. 32, pp. 565-575, 1987.

38. S.J. Riederer, N.J. Pelc, and D.A. Chelser, "The Noise Power Spectrum in Computed X-ray Tomography," *Phys. Med. Biol.*, vol. 23, pp. 446-454, 1978.

39. H.H. Barrett, D.W. Wilson, and B.M.W. Tsui, "Noise Properties of the EM Algorithm: I. Theory," *Phys. Med. Biol.*, vol. 39, pp. 833-846, 1994.

40. D.W. Wilson, B.M.W. Tsui, and H.H. Barrett, "Noise properties of the EM algorithm: II. Monte Carlo simulations," *Phys. Med. Biol.*, vol. 39, pp. 847-872, 1994.

41. J.A. Fessler, "Mean and Variance of Implicitly Defined Biased Estimators Such as Penalized Maximum Likelihood: Applications to Tomography," *IEEE Trans. Imag. Proc.*, vol. 5, pp. 493-506, 1996.

42. F. O'Sullivan, Y. Pawitan, and D. Haynor, "Reducing Negativity Artifacts in Emission Tomography: Post-processing Filtered Backprojection Solutions," *IEEE Trans. Med. Imag.*, vol. 12, pp. 653-663, 1993.

43. J.A. Fessler and J.M. Ollinger, *Signal Processing Pitfalls in Positron Emission Tomography*. Univ. of Michigan, Ann Arbor, MI, 48109-2122 Available on WWW from http://www.eecs.umich.edu/~fessler, 1996.

44. Y. Vardi, L.A. Shepp, and L. Kaufman, "A Statistical Model for Positron Emission Tomography," *J. Am. Stat. Assoc.*, vol. 80, pp. 8-37, 1985.

45. A. Macovski, *Medical Imaging Systems*. New Jersey: Prentice-Hall, 1983.

46. E.J. Hoffman, S.C. Huang, M.E. Phelps, and D.E. Kuhl, "Quantitation in Positron Emission Computed Tomography: 4 Effect of accidental coincidences," *J. Comp. Assisted Tomogr.*, vol. 5, pp. 391-400, 1981.

47. J.A. Fessler, "Penalized Weighted Least-squares Image Reconstruction for Positron Emission Tomography," *IEEE Trans. Med. Imag.*, vol. 13, pp. 290-300, 1994.

48. M. Yavuz and J.A. Fessler, "Objective Functions for Tomographic Reconstruction from Randoms - Precorrected PET Scans," presented at Proc. IEEE Nuc. Sci. Symp. Med. Imag. Conf., Anaheim, 1996.

49. K. Lange and R. Carson, "EM Reconstruction Algorithms for Emission and Transmission Tomography," *J. Comp. Assist. Tomogr.*, vol. 8, pp. 306-316, 1984.

50. D.L. Snyder, M.I. Miller, L.J. Thomas, and D.G. Politte, "Noise and Edge Artifacts in Maximum-likelihood Reconstructions for Emission Tomography," *IEEE Trans. Med. Imag.*, vol. 6, pp. 228-238, 1987.

51. T.R. Miller and J.W. Wallis, "Clinically Important Characteristics of Maximum-likelihood Reconstruction," *J. Nucl. Med.*, vol. 33, pp. 1678-1684, 1992.

52. J.A. Fessler and A.O. Hero, "Space-alternating Generalized Expectation-maximization Algorithm," *IEEE Trans. Sig. Proc.*, vol. 42, pp. 2664-2677, 1994.

53. J.A. Fessler and A.O. Hero, "Penalized Maximum-likelihood Image Reconstruction Using Space-alternating Generalized EM Algorithms," *IEEE Trans. Image Proc.*, vol. 4, pp. 1417-29, 1995.

54. P.J. Green, "Bayesian Reconstructions from Emission Tomography Data Using a Modified EM Algorithm," *IEEE Trans. Med. Imag.*, vol. 9, pp. 84-93, 1990.

55. T. Hebert and R. Leahy, "A Generalized EM Algorithm for 3-D Bayesian Reconstruction from Poisson Data using Gibbs Priors," *IEEE Trans. Med. Imag.*, vol. 8, pp. 194-202, 1989.

56. K. Lange, "Convergence of EM Image Reconstruction Algorithms with Gibbs Smoothing," *IEEE Trans. Med. Imag.*, vol. 9, pp. 439-436, Corrections June 1991, 1990.

57. K. Sauer and C. Bouman, "A Local Update Strategy for Iterative Reconstruction from Projections," *IEEE Trans. Sig. Proc.*, vol. 41, pp. 534-548, 1993.

58. E.U. Mumcuoglu, R. Leahy, S.R. Cherry, and Z. Zhou, "Fast Gradient-based Methods for Bayesian Reconstruction of Transmission and Emission PET Images," *IEEE Trans. Med. Imag.*, vol. 13, pp. 687-701, 1994.

59. A. Tikhonov and V. Arsenin, *Solution of Ill-posed Problems*. New York: Wiley, 1977.

60. J.A. Fessler and W.L. Rogers, "Uniform Quadratic Penalties Cause Nonuniform Image Resolution and Sometimes Vice versa," *Proc. IEEE Nuc. Sci. Symp. Med. Imag. Conf*, pp. 1915-1919, Norfolk, 1994.

61. J.A. Fessler and W.L. Rogers, "Spatial Resolution properties of Penalized-likelihood Image Reconstruction Methods: Space-invariant Tomographs," *IEEE Trans. Imag. Proc.*, vol. 5, 1996.

62. J.A. Fessler, "Resolution Properties of Regularized Image Reconstruction Methods," in *Technical Report 297, Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. of Michigan*. Ann Arbor, MI, 48109-2122, 1995.

63. Z. Zhou, R.M. Leahy, and E.U. Mumcuoglu, "Maximum Likelihood Hyperparameter Estimation for Gibbs Priors with Applications to PET," in *Information Processing in Medical Imag.*, pp. 39-51, C.B.Y. Bizais, and R.D. Paola, editors, Ed. Netherlands: Kluwer, 1995.

64. Y. Pawitan and F. O'Sullivan, "Data-dependent Bandwidth Selection for Emission Computed Tomography Reconstruction," *IEEE Trans. Med. Imag.*, vol. 12, pp. 167-172, 1993.

65. J.W. Hilgers and W.R. Reynolds, "Instabilities in the Optimal Regularization Parameter Relating to Image Recovery Problems," *J. Opt. Soc. Amer.* Ser. A, vol. 9, pp. 1273-1279, 1992.

66. S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and Bayesian Restoration of Images," *IEEE Trans. Patt. Anal. Mach. Int.*, vol. 6, pp. 721-741, 1984.

67. S.J. Lee, A. Rangarajan, and G. Gindi, "Bayesian Image Reconstruction in SPECT Using Higher Order Mechanical Models as Priors," *IEEE Trans. Med. Imag.*, vol. 14, pp. 669-80, 1995.

68. V.Z. Mesarovic, N.P. Galatsanos, and A.K. Katsaggelos, "Regularized Constrained Total Least Squares Image Restoration," *IEEE Trans. Imag. Proc.*, vol. 4, pp. 1096-1109, 1995.

69. G.T. Gullberg, *The Attenuated Radon Transform: Theory Application in Medicine and Biology*. Berkeley, CA: University of California, 1979.

70. S.H. Manglos, F.D. Thomas, and R.B. Capone, "Attenuation Compensation of Cone beam SPECT Images Using Maximum Likelihood Reconstruction," *IEEE Trans. Med. Imag.*, vol. 10, pp. 66-73, 1991.

71. M.I. Miller, D.L. Snyder, and T.R. Miller, "Maximum Likelihood Reconstruction for Single-Photon Emission Computed-Tomography," *IEEE Trans. Nucl. Sci*, vol. 32, pp. 769-778, 1985.

72. E.P. Ficaro, J.A. Fessler, P.D. Shreve, J.N. Kritzman, P.A. Rose, et al., "Simultaneous Transmission/emission Myocardial Perfusion Tomography: Diagnostic-accuracy of Attenuation Corrected Tc-99m sestamibi SPECT," *Circulation*, vol. 93, pp. 463-73, 1996.

73. C.E. Metz, "ROC Methodology in Radiological Imaging," *Invest. Radiol.*, vol. 21, pp. 720-733, 1986.

74. H.H. Barrett, J.L. Denny, R.F. Wagner, and K.J. Myers, "Objective assessment of Image Quality. II. Fisher Information, Fourier Crosstalk, and Figures of Merit for Task Performance," *J. Opt. Soc. Amer.* Ser. A, vol. 12, pp. 834-52, 1995.

75. A.O. Hero, J.A. Fessler, and M. Usman, "Exploring Estimator Bias-variance Tradeoffs Using the Uniform CR Bound," *IEEE Trans. Sig. Proc.*, vol. 44, pp. 2026-41, 1996.

76. T. Kanungo, M.Y. Jaisimha, J. Palmer, and R.M. Haralick, "A Methodology for Quantitative Performance Evaluation of Detection Algorithms," *IEEE Trans. Imag. Proc.*, vol. 4, pp. 1667-1674, 1995.