Bayesian Estimator for Angle Recovery: Event Classification and Reconstruction in Positron Emission Tomography

Angela M K Foudray[*,†] and Craig S Levin[*]

[*]*Stanford University*
Molecular Imaging Program at Stanford
Department of Radiology
Palo Alto, CA, USA
[†]*University of California, San Diego*
Department of Physics
La Jolla, CA, USA

Abstract. PET at the highest level is an inverse problem: reconstruct the location of the emission (which localize biological function) from detected photons. Ideally, one would like to directly measure an annihilation photon's incident direction on the detector. In the developed algorithm, Bayesian Estimation for Angle Recovery (BEAR), we utilized the increased information gathered from localizing photon interactions in the detector and developed a Bayesian estimator for a photon's incident direction. Probability distribution functions (PDFs) were filled using an interaction energy weighted mean or center of mass (COM) reference space, which had the following computational advantages: (1) a significant reduction in the size of the data in measurement space, making further manipulation and searches faster (2) the construction of COM space does not depend on measurement location, it takes advantage of measurement symmetries, and data can be added to the training set without knowledge and recalculation of prior training data, (3) calculation of posterior probability map is fully parallelizable, it can scale to any number of processors. These PDFs were used to estimate the point spread function (PSF) in incident angle space for (i) algorithm assessment and (ii) to provide probability selection criteria for classification. The algorithm calculates both the incident θ and ϕ angle, with $\sim$16 degrees RMS in both angles, limiting the incoming direction to a narrow cone. Feature size did not improve using the BEAR algorithm as an angle filter, but the contrast ratio improved 40% on average.

Keywords: 3D detectors, Compton scatter, event filtering, angular resolution, Bayes, PET, medical imaging, measurement estimation, event classification
PACS: 02.30.Zz, 02.50.Tt, 29.40.Gx, 87.58.Fg, 87.58.Pm, 87.58.Ce

INTRODUCTION

Positron emission tomography (PET) imaging has become a workhorse of disease detection and management. To resolve and detect smaller volumes of molecular function, requisite to image early stage cancer, higher resolution cameras are being developed. PET assumes the anti-co-linear (exactly opposite trajectory) production of two 511 keV photons from the annihilation of a positron emitted from an introduced isotope. Detecting both photons from one annihilation then localizes the possible emission positions to a line. All of the lines formed from the photon pairs, called lines of response (LORs), are used to tomographically reconstruct the emission's three-dimensional location(s) within the body. Traditional PET detectors have relatively large detection volumes over which

CP954, *27th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*,
edited by K. H. Knuth, A. Caticha, J. L. Center, Jr., A. Giffin, and C. C. Rodríguez

the multiple interactions involved in high energy annihilation photon transport cannot be distinguished. The next generation PET detectors are beginning to detect these individual interactions, giving rise to more complicated, but potentially higher accuracy, event classification.

This work developed a Maximum Likelihood algorithm that includes a realistic model of the detector system (material interaction physics; detector energy, position and time resolution; device centroiding; and crystal binning and gaps) and incorporates Bayesian methods for extracting a single photon's incident angle with respect to the detector using the information from the detected multiple interactions.

Annihilation Photon Transport

The two major sources of background noise in a coincidence-collimated emission tomography system are present due to the finite energy and time resolution of the detection system, and will be discussed below. The finite position resolution of the detection system contributes to local intensity blurring in the reconstructed image and could also conceivably be addressed with a Bayesian estimation scheme, but wasn't developed here. Photons, at the energies used in emission imaging, are generally not fully absorbed upon interaction with matter. Annihilation photons interact most often via Compton interaction, so we will begin there.

Compton Scattering

Finite energy resolution in the detection system contributes to the inability to distinguish photons that have undergone small angle scatter before being detected. The photon's energy after a Compton Scatter is a function of its pre-interaction energy and the angle at which it scattered:

$$E'(E,\theta) = \frac{E}{1+\alpha(E)(1-cos(\theta))} \tag{1}$$

where E and E' are the energies before and after scattering respectively, θ is the scattering angle, and α(E) is the relative energy of the photon to the rest mass of an electron. For an incoming annihilation (511 keV) photon, α = 1. Detectors used for positron emission tomography have energy resolutions anywhere from 3 to 20%, but typically around 15% at 511 keV for the best clinical systems. Therefore by energy filtering, we are rejecting the photons that have undergone relatively large angle scatter. An event pair that contains one or two scattered photons, but are not filtered out by energy discrimination due to this energy blur, we call "Scatters". They cannot be labeled in a traditional system (if they could, those events would be eliminated!), but can be estimated. They can be directly quantified in a Monte Carlo simulation, and are described below (figure 1).

The other major contribution to large scale noise is due to the finite time resolution of a PET system. Events are constructed from pairing detected incident single photons. An annihilation event produces two photons at precisely the same time. A blur in timing information leads to possible pairing of detected incident photons from separate annihilations, which incurs contrast loss. Events that incorrectly pair photons from separate annihilation events are call "Randoms". One from a pair of generated photons can go undetected by (1) passing through the detectors without (completely) interacting, (2) traveling along a path that isn't directed toward detectors, (3) being absorbed in the body before getting to the detectors. "Singles" are events in which only one photon was detected. "Multiples" occur when multiple annihilation events have produced three or more detected photons within a chosen time window. Figure 1 shows the fractional rate of the "Trues", "Randoms" and "Scatters" events at typical activities used in laboratory studies, where "Trues" are coincidences formed from photons from the same emission event (neither photon has scattered). The events in figure 1 have survived the energy and time windows, and in a traditional system, would all be used in the reconstruction.

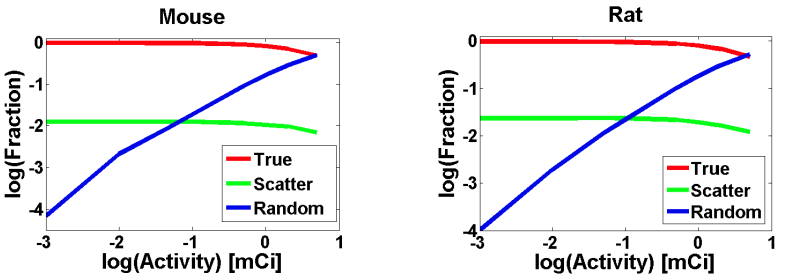

FIGURE 1. Fraction of events that are true, scattered and random coincidences for mouse and rat-sized objects, normalized to total events.

MULTIPLE INTERACTION BASED ELECTRONIC COLLIMATION (MIBEC)

Having information about the characteristics of an incoming photon could allow a higher degree of certainty about how that detected photon (and ultimately event) should be used, if at all. Benefits of having direction and energy information about a photon incident on the detectors include: more precise event typing (Singles, Randoms, Scatters, Multiples, Trues) for 1) filters, 2) reclamation of previously unusable events [1] and 3) for LOR assignment (event positioning). This paper studies the first benefit. The interactions occurring in the imaged object (Compton, Photoelectric, etc.), also occur in the detector. However, in the detector, the deposited energy is recorded, leaving information about the path of the photon. This group of detected interactions we will call clusters. This study undertakes using this information to estimate the incident direction of the photon and how this information affects spatial resolution and contrast.

Instrumentation Considerations

Often, tomographic emission imaging detectors are read-out in a multiplexed fashion, generally to reduce cost and complexity. Device multiplexing will inhibit the ability to distinguish individual interactions that occurred within the multiplexed arrangement during the finite acquisition integration time. In the position-sensitive avalanche photodiodes (PSAPDs) used in this study[2], multiple interactions within a single detector module during the charge integration period cannot be distinguished, but inter-detector interactions can be separately localized. Since the PSAPD photodetector has the highest quantum efficiency when detecting optical photons (~550-800 nm), the scintillation crystals are used to "convert" high energy photon interactions into optical wavelength photons. Optical transport in the scintillation crystal also contributes to spatial blurring, discretization and perceived energy loss from optical photon absorption. Photoelectric interactions and their characteristic x-rays, Bremsstrahlung, and Doppler broadening can shift the magnitude and location of energy deposited. The complex forward model coupled with device and electronic multiplexing make methods that have a less strict (or non-strict) requirement on the details of the forward model, such as Bayesian estimation, ideal candidates[3] for event processing.

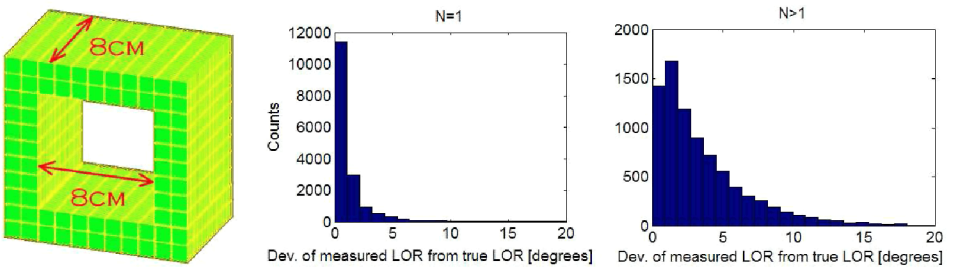

FIGURE 2. Left: an illustration of the high resolution PET detection system[2]. Center: histogram over all LORs of the angular difference between lines formed from the COM of a measured event cluster and 1) the COM of its coincident cluster, and 2) the annihilation photon emission point, for the N=1 case, Right: for the N>1 case. Events used in (center) and (right) were gathered for a source uniformly filling the interior of the detector.

Shown in figure 2b,c is the convergence of all of the particle transport and instrumentation blurring, quantified in angle space, knowing the emission position. This is, in effect, the intrinsic ability of the system (detection + clustering + positioning scheme) to localize the original photon emission (using the COM positioning algorithm), or the histogrammed LOR deviation due to intercrystal scatter and positioning. This deviation is histogrammed in figure 2b,c for N=1 and N>1, where N is the number of detected interactions in the cluster. Since this paper is assessing a MIBEC algorithm (N>1), for this instrumentation[2], figure 2c indicates the distribution of the LOR deviation we can expect inherently in the data to which the algorithm will be applied.

BEAR: A NAIVE BAYESIAN CLASSIFIER

Incident Angle Estimation

We consider a cluster of N interactions from a single photon entering the system, each interaction defined by their relative positions and energies (x_i,y_i,z_i,E_i) abbreviated X_i, and computing the probability of finding these interactions jointly for a given incident photon direction (θ,ϕ) abbreviated Θ (figure 3):

$$P(\Theta|X_1,...,X_N) = \frac{P(X_1,...,X_N|\Theta) \cdot P(\Theta)}{P(X_1,...,X_N)} \tag{2}$$

where the prior P(Θ), using the theoretical assumption that the singles are generated isotropically throughout the volume, was assumed to be one (all equally likely). Using the joint probability definitions, this can then be written as:

$$P(\Theta|X_1,...,X_N) = \prod_{i=1}^{N>1} \frac{P(X_i|X_j,\Theta)}{P(X_i|X_j)} \tag{3}$$

where X_j is (X_{i-1},X_{i-2}, ..., X_1). When i=1, X_j is $\oslash$. The decision rule for angle recovery is simply

$$\Theta_{cluster} = sup\{P(\Theta|X_1,...,X_N)\} \tag{4}$$

Event Space Representation

Separate methods were used to discretize the position and energy values for each interaction in the cluster[4]. An unsupervised static local discretization was used for the position dimensions, using equal width intervals in a space relative to the center of mass of the cluster. This COM reference space had a number of advantages: (1) a significant reduction in the size of the data in measurement space, making further manipulation and searches faster (2) the construction of COM space does not depend on measurement location, it takes advantage of measurement symmetries, and data can be added to the training set without knowledge and recalculation of prior training data, (3) calculation of posterior probability map is fully parallelizable, it can scale to any number of processors. A supervised static global discretization was used for the energy dimension. The energy bins near to large angle scatter and photoelectric interactions were sampled coarsely, whereas the small energy deposition interactions were sampled with more resolution to distinguish processes at low energies, while still keeping the number of bins low.

Probability Density Maps

The probabilities used in eq. 3 were calculated for each interaction cluster in event space for each incident (θ, ϕ). Two sets of interaction locations and energies for each

location and angle studied were generated - one for the training set (likelihood and evidence calculations, right hand side of eq. 3) and one for the test set (posterior probabilities, left hand side of eq. 3).

Training sets were generated over incident angle space using the forward model of the annihilation photon transport physics in the Geant4 Application for Tomographic Emission (GATE) Monte Carlo package; the most validated simulation software for detailed PET physics [5]. A pencil beam with zero angular extent was directed into the detector system from a point source for every θ and ϕ angle value in the range detected by the system, which is nearly π radians in both θ and ϕ (see figure 3 for the trained locations).

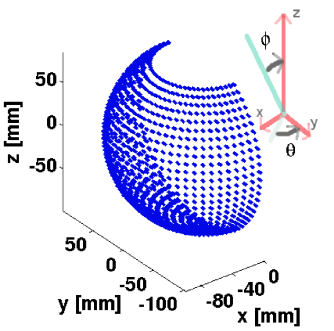

FIGURE 3. Illustration of coordinate system and the beam locations of the simulation annihilation photon source locations in (x,y,z). Each blue dot represents a training set of ~25,000 events. In the reference space used in the figure, the x-direction is the direction normal to the face of the detector.

Posterior probability maps were determined for each event by calculating the right hand side of eq. 3 at each (θ, ϕ), using the probabilities of the cluster interactions determined by the training set data. Figure 4 shows an example of these PDFs during the calculation of equation 4 for a typical cluster (one event) interacting in a central module of one sides of the system.

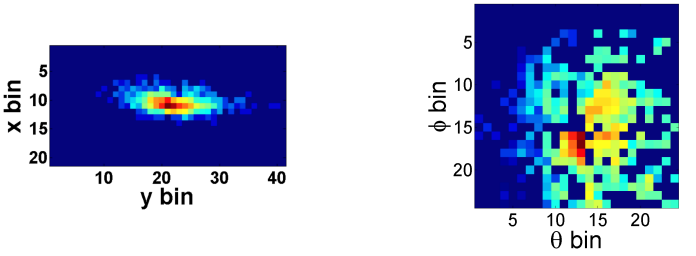

FIGURE 4. Log of the probability density maps for (a) likelihood in X_i space (an x-y plane slice), for θ=17 and ϕ=13, and (b) posterior probability in Θ space, for a particular interaction cluster.

Prediction Capabilities

The algorithm's ability to predict photon incident angle, only knowing the interaction locations and energies, was measured using test sets of known incident angle data, and calculating the point spread functions in (θ,ϕ) space (see figure 5). ϕ and θ are the angles that are made with the axial and radial direction (z and x in figure 3), respectively. The root mean squared (RMS) deviation from the true angle bin was used to calculate the angular resolution of the algorithm shown in figure 6a and b.

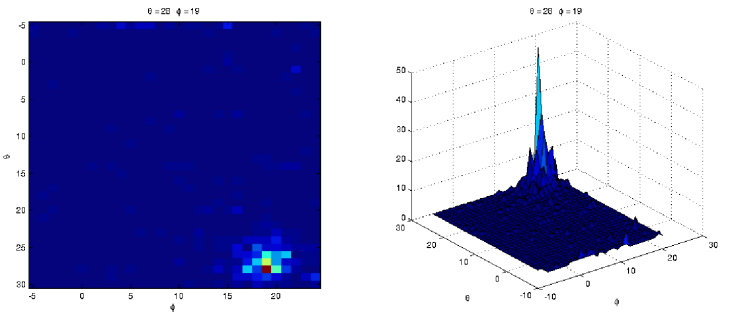

FIGURE 5. Point spread function for a single incident θ and ϕ value (θ=17 and ϕ=13).

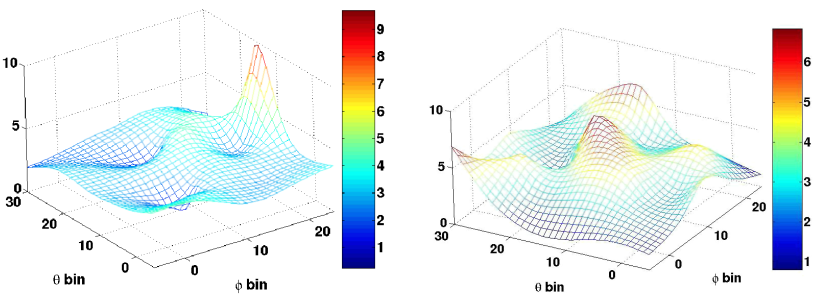

FIGURE 6. Shown is the RMS deviation of the BEAR-estimated angle from the angle of the line that goes through the annihilation in (a) ϕ and (b) θ angles in angle bins (vertical axis) where 1 bin = 4 degrees. The distribution of sources for the test data set was the same as that for the training data set.

The mean and variance of the RMS values were $15.2^o \pm 5.1^o$ and $15.7^o \pm 6.9^o$ for the theta and phi angle estimation respectively. The effective angular resolution from this algorithm provides a substantial degree of photon collimation information. The variation in the RMS values across all angular space in θ and ϕ would most likely go down as a result of including more event interaction data in the training set. With higher training set counts, the RMS values themselves will likely asymptotically approach a constant but non-zero value due to the finite energy resolution and position blurring effects (binning, scintillation light transport) in the detector modules.

Image Reconstruction in the Presence of Biologically Relevant Background

The BEAR algorithm was used in this study as a simple filter to reject lines whose estimated angles deviated more than a selected degree from the angle calculated from coincidence pairing. This deviation was chosen to be a 24 degree window (±12 degrees), to enable a high number of accepted events for good statistics. Further, the PDF could be used as a weighted projector in reconstruction for a myriad of functions: multiples selection, single and scatter reconstruction, leading to new advances in tomographic reconstruction.

Detection in an emission tomography system is dependent on the system's ability to resolve sources of activity, in this paper quantified by the perceived spread of activity (FWHM) and relative intensity to background (contrast ratio (CR)). In this study, a plane of sphere sources; each quadrant with an array of a single diameter (1.25, 1.5, 2.5 and 3.5 mm in quadrant I, II, III, IV respectively), centers separated by 4 times their diameter; were simulated in a 6cm diameter, 8cm long, cylinder of scattering medium (water). The activity concentration in the spheres was 9 times that of the uniform background in the cylinder.

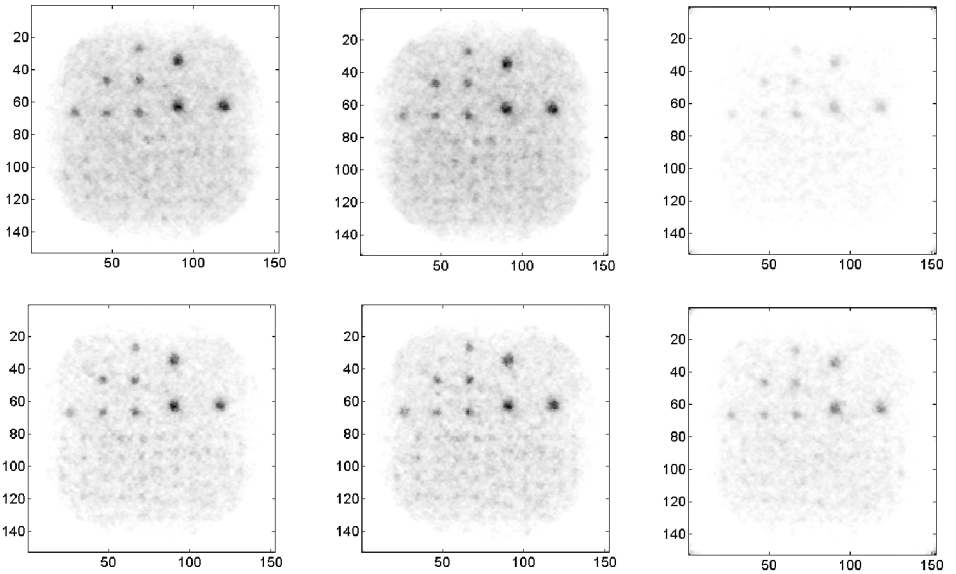

FIGURE 7. Shown are the unfiltered images (top) and filtering with the BEAR algorithm (bottom); for the 1% (left), 18% (middle), and 50% (right) acquisitions; each reconstructed by the iterative image reconstruction method Ordered Subset Expectation Maximization [6]: 12 subsets, 1 iteration, keeping the number of true coincidences the same in each study. Inverse grey scale: Darker represents higher intensity.

As the total activity imaged in the detector effects the fraction of random counts accepted by the system (figure 1), three studies, with different total activities were acquired: 100μCi, 1mCi, and 5mCi, leading to acquisitions with about 1%, 18% and

50% of the counts being random events (with volumetrically uniform background). The energy and arrival time were distributed according to Gaussian distributions with 12% FWHM, and 2 ns FWHM, respectively, and the window settings were twice those resolution values. The UF (unfiltered) data was filtered before reconstruction via BEAR to produce the BEAR data. The number of counts used in reconstruction were: 144 (UF) and 94 (BEAR) for the 1% case, 174 (UF) and 106 (BEAR) for the 18% case, and 348 (UF) and 197 (BEAR) for the 50% case, in millions of counts, respectively.

The features that are apparent in all images are the 2.5 and 3.5 mm spheres. In the 1% and 18% randoms fraction images (left and center in figure 7), some of the smaller spheres would potentially gain further contrast with more counts (increased SNR). A fit was made to each 2D image using a 2D Gaussian plus a contrast background (5 fit parameters) using MATLAB's fminsearch. The FWHM of the 2.5 and 3.5 mm features were extracted from the fit (table 1), and didn't improve significantly by using the BEAR algorithm.

TABLE 1. Fitted FWHM of the 3.5mm and 2.5mm spheres in all reconstructed images.

% Randoms	Filter	3.5mm FWHM	3.5mm StDev	2.5mm FWHM	2.5mm StDev
1	None	2.83	0.37	2.29	0.14
1	BEAR	2.80	0.22	2.11	0.26
18	None	2.80	0.47	2.11	0.27
18	BEAR	2.88	0.30	2.11	0.32
50	None	3.17	0.78	2.46	0.49
50	BEAR	3.06	0.24	2.35	0.12

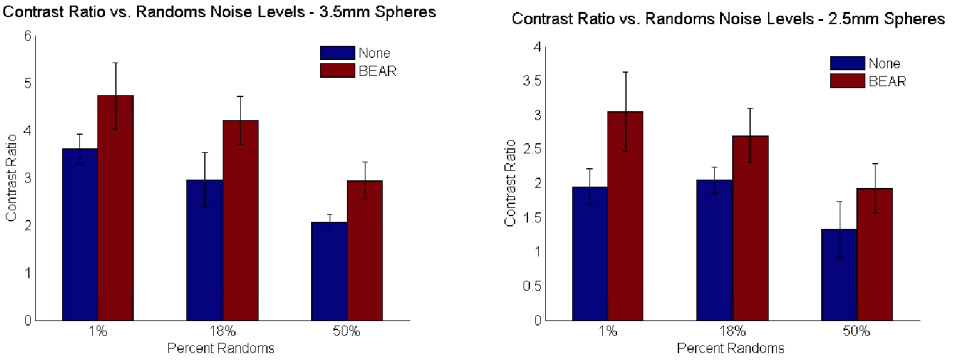

FIGURE 8. Shown are the contrast ratios for the 3.5 and 2.5 mm spheres in the three noise regimes for both unfiltered and BEAR filtered images.

Contrast ratio, on the other hand, improved markedly. The contrast ratio was defined in this paper by the peak height of the Gaussian fit divided by the constant background fit. Since a source concentration ratio of 9 used in the study, the reconstructed contrast ratio would ideally also be 9. However blurring effects in the detection and filtering schemes attenuate and spread the activity away from the point of origin. Using the BEAR algorithm, the 3.5mm spheres improved in CR by 38.4%±6.41%, and the 2.5mm spheres by 44.1%±12.4% across the images.

Removing object scatter and random counts alone does not account for the number of counts filtered or the contrast ratio improvement using this algorithm in the low activity data set. Mispositioning of the LOR due to intercrystal scatter is not activity dependent and would contribute to the background near the correct LOR (feature location) as seen from figure 2. These intercrystal scatter events, which have angular deviation from the true LOR, are also filtered by this algorithm. Further study will be performed using a contrast ratio phantom to investigate this issue.

ACKNOWLEDGMENTS

The authors would like to thank the Molecular Imaging Program at Stanford, and particularly the Sam Gambhir lab, for the use of the 32-node reconstruction cluster. We would also like to thank Garry Chinn for the use of the list-mode Ordered Subset Expectation Maximization reconstruction code used in this work.

Angela M K Foudray would like to thank PEO International for the Scholar Award, which partially funded this work. Angela M K Foudray would also like to thank Kevin Knuth and the MaxEnt organization for full sponsorship to attend and present this work at the 27th International MaxEnt Workshop.

This work was also supported in part by the following grants: NIH-National Cancer Institute R21 CA098691, NIH-NIBIB R21/R33 EB003283, and NIH-NCI CA119056.

REFERENCES

1. G. Chinn, A. M. K. Foudray, and C. S. Levin, *IEEE Nuclear Science Symposium and Medical Imaging Conference Record* pp. 1746–1751 (2006).
2. A. M. K. Foudray, G. Chinn, and C. S. Levin, *IEEE NSS-MIC Conference Record* **4**, 2108–2111 (2005).
3. D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2005, version 7.2, 4th edn.
4. J. Dougherty, R. Kohavi, and M. Sahami, *Machine Learning: Proceedings* (1995).
5. S. Jan, and et. al., *Phys. Med. Bio.* **49**, 4543–4561 (2004).
6. H. M. Hudson, and R. S. Larking, *IEEE Transactions on Medical Imaging* **13**, 601–609 (1994).