Scanning linear estimation: improvements over region of interest (ROI) methods

To cite this article: Meredith K Kupinski et al 2013 Phys. Med. Biol. 58 1283

View the article online for updates and enhancements.

Related content
- Task-based measures of image quality and their relation to radiation dose and patient risk
  Harrison H Barrett, Kyle J Myers, Christoph Hoeschen et al.
- Comparison of the scanning linear estimator (SLE) and ROI methods for quantitative SPECT imaging
  Arda Könik, Meredith Kupinski, P Hendrik Pretorius et al.
- Noise and signal properties in PSF-based fully 3D PET image reconstruction
  S Tong, A M Alessio and P E Kinahan

Recent citations
- Optimal channels for channelized quadratic estimators
  Meredith K. Kupinski and Eric Clarkson
- Task-based measures of image quality and their relation to radiation dose and patient risk
  Harrison H Barrett et al
- Singular value decomposition for photon-processing nuclear imaging systems and applications for reconstruction and computing null functions
  Abhinav K Jha et al
Scanning linear estimation: improvements over region of interest (ROI) methods

Meredith K Kupinski¹, Eric W Clarkson¹,² and Harrison H Barrett¹,²

¹ College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA
² Department of Medical Imaging, University of Arizona, Tucson, AZ 85721, USA

E-mail: meredith@optics.arizona.edu

Received 21 August 2012, in final form 19 December 2012
Published 6 February 2013
Online at stacks.iop.org/PMB/58/1283

Abstract

In tomographic medical imaging, a signal activity is typically estimated by summing voxels from a reconstructed image. We introduce an alternative estimation scheme that operates on the raw projection data and offers a substantial improvement, as measured by the ensemble mean-square error (EMSE), when compared to using voxel values from a maximum-likelihood expectation–maximization (MLEM) reconstruction. The scanning-linear (SL) estimator operates on the raw projection data and is derived as a special case of maximum-likelihood estimation with a series of approximations to make the calculation tractable. The approximated likelihood accounts for background randomness, measurement noise and variability in the parameters to be estimated. When signal size and location are known, the SL estimate of signal activity is unbiased, i.e. the average estimate equals the true value. By contrast, unpredictable bias arising from the null functions of the imaging system affect standard algorithms that operate on reconstructed data. The SL method is demonstrated for two different tasks: (1) simultaneously estimating a signal’s size, location and activity; (2) for a fixed signal size and location, estimating activity. Noisy projection data are realistically simulated using measured calibration data from the multi-module multi-resolution small-animal SPECT imaging system. For both tasks, the same set of images is reconstructed using the MLEM algorithm (80 iterations), and the average and maximum values within the region of interest (ROI) are calculated for comparison. This comparison shows dramatic improvements in EMSE for the SL estimates. To show that the bias in ROI estimates affects not only absolute values but also relative differences, such as those used to monitor the response to therapy, the activity estimation task is repeated for three different signal sizes.
1. Introduction

Reconstructions from single-photon emission computed tomography (SPECT) data are often utilized as point-wise estimates of the object’s activity distribution. Estimates of signal activity are calculated by operating on the reconstructed values within a boundary or region of interest (ROI) that specifies the signal shape and location. This method appeals to one’s optimistic interpretation of reconstructed data as an estimate of the object, but the estimation rule is not derived according to a cost minimization, and it is unclear if the statistical properties of ROI estimators are the best that can be achieved (Keyes 1995, Thie 2004, Jaskowiak et al 2005, Boellaard et al 2004, Huang 2000). Furthermore, both phantom and patient imaging studies have revealed significant differences when comparing the ROI selection methods and choice of pixel statistic therein (mean, median, or maximum) (Vanderhoek et al 2012, Tylski et al 2010). From first principles, it is known that null functions of the imaging system prevent an unbiased estimation from reconstructed data (Barrett and Myers 2004). An example of a null function is the fine detail beyond the resolution capabilities of the imaging system. If the template defining the ROI contains null functions with respect to the imaging system, then any algebraic manipulation of the reconstructed values in this ROI will be a biased estimate of the signal’s true activity (Barrett and Myers 2007, Müller et al 1990). Any binary ROI that is designed to abruptly demarcate the signal from background will have high-frequency content that is beyond the resolving power of current tomographic imaging systems.

Unfortunately, this bias cannot be corrected because it depends upon the null component of the object which is, by definition, undetected by the imaging system. When reporting the change in ROI statistics between a pair or sequence of images, this bias is not a systematic error that can be corrected by subtracting two equally biased values (Kessler et al 1984). The ROI shape and location within the field of view (FOV), as well as the given distribution of activity, will have different null functions and hence different biases as reported in Boellaard et al (2004) and Feuardent et al (2005). The variability of quantities based upon ROI voxel values, such as the standardized uptake value (SUV) in positron emission tomography, is well documented (Fahey et al 2010, Doot et al 2007). Community-adopted protocols which standardize associated imaging study variables are also widely discussed (Shankar et al 2006, Boellaard 2012, Adams et al 2010, Hicks 2008, Buckler and Boellaard 2011, Boellaard et al 2008). When study variables are highly controlled, SPECT/CT imaging studies have reported that the error of activity estimates calculated from ROI statistics is in the neighborhood of 10% and sometimes lower (Shcherbinin et al 2008, Buvat 2007, Ritt et al 2011, McQuaid et al 2011).

In this paper, we use a measured small-animal SPECT imaging matrix to illustrate the dramatic errors induced by relying on ROI voxel values for quantitative studies. These conclusions have a direct implication for the clinical interpretation of SUV and other imaging metrics calculated from ROI voxels. We will demonstrate that operating on the raw projection data and accurately modeling the statistics of the object, as opposed to ad hoc characterization of the reconstructed voxel values, offer a solution toward the unbiased estimation of diagnostically relevant signal features. See Rossi and Willsky (1984), Huesman (1984), Wang and Qi (2009) and Kamasak et al (2005) for other examples in medical imaging of estimation tasks performed directly on projection data. The statistical properties of the estimation rules under consideration are presented in section 2. Section 3 describes the choice of object and image ensembles which are related to the ground truth and difficulty level of the estimation task. The numerical methods used to invert a large covariance matrix and perform a high-dimensional optimization are briefly summarized in section 4, with references cited for further details about the computational methodologies. Results are presented in section 5.
followed by conclusions in section 6. The prior knowledge required for a clinically relevant application of scanning-linear (SL) estimation is outlined in section 7.

2. Limits on estimability

2.1. Scanning linear estimation

A reconstructed image provides a visual interpretation of the data, but post-processing cannot increase the mutual information between the data and the object parameters to be estimated (Shao 1999). Therefore, we seek an estimation rule that operates directly on the raw projection data. As in Barrett and Myers (2004), a linear imaging equation can be written as

\[
(\mathbf{g}(\theta))_m = \int_S d^3 r f(r, \theta) h_m(r),
\]

where \(\theta\) is the vector of parameters to be estimated, \(m\) is an index for the detector elements, the overline on \(\mathbf{g}\) indicates that the data are averaged over measurement noise, \(f(r, \theta)\) is the three-dimensional continuous object being imaged, \(S\) is the FOV and the scalar-valued function \(h_m(r)\) specifies the \(m\)th detector’s sensitivity to the object activity at the location \(r\) within the system’s FOV. An alternative short hand form represents the imaging system as a deterministic operator, which maps a function of a continuous variable to a finite-dimensional vector, plus a noise term. This can be compactly expressed as

\[
g(\theta) = \mathcal{H} f(\theta) + \mathbf{n},
\]

where \(\mathcal{H}\) is a continuous-to-discrete (CD) linear operator, \(f\) is an infinite-dimensional vector in a Hilbert space that represents the continuous object and the vector \(\mathbf{n}\) is a random perturbation to each data element with statistics that may in general be object dependent. In this study, we will consider sources of randomness from both measurement noise and a stochastic object ensemble.

Maximum-likelihood (ML) estimation yields optimal performance for many estimation tasks. By definition, an efficient estimator of \(\theta\) is unbiased with variance that equals the Cramér–Rao lower bound (Shao 1999). Efficient estimators do not exist for every problem, but from theoretical principles, we know that if an efficient estimator does exist, the ML estimator is efficient (Van Trees 2001). Fortunately, even when no efficient estimator exists, but the parameter is estimable, the ML estimator is efficient in the asymptotic limit of more counts collected during a given acquisition or of a greater quantity of noisy data sets. For a treatment of ML estimation performance at high noise levels, see Müller et al (2005). It has been shown in Clarkson (2007a, 2007b) that for a joint detection/estimation task, the area under the estimation receiver operating characteristic curve is maximized by the ML estimate when the prior \(\operatorname{pr}(\theta)\) is flat and the tolerance for correct answers is chosen to be very narrow. Obtaining these desirable statistical properties of ML estimation requires, for many cases, an impractical amount of computation. A search for the maximum of the data likelihood requires a method for point-wise evaluation at candidate solutions of the parameter. Furthermore, a single evaluation requires averaging over sources of image variability which are present for a fixed value of \(\theta\); these are referred to as nuisance parameters. Consider an object decomposition into a signal which depends on the parameters of interest and a background which does not. When background variability is a consideration, the likelihood requires marginalizing over the background ensemble (Kupinski et al 2003)

\[
\operatorname{pr}(g|\theta) = \int d^M b \operatorname{pr}(g|\theta ; b) \operatorname{pr}(b),
\]
where \( b = Hf_{\text{bkgnd}} \) is the noise-free image of a background. The measurement noise is described by \( pr(g|\theta; b) \), the probability of an image when the object is fixed. The ML estimation requires numerous evaluations of the integral expression in equation (3) to find the value of \( \theta \) that maximizes the likelihood. The optimization and associated integral evaluations require repeated evaluation for each image \( g \).

Here we present an SL estimator that is an approximate ML estimator formed by a series of assumptions designed to make the calculation tractable. In the previous work, we showed that parameters that are non-affine transforms of the object, such as location and signal shape, could not be estimated by linear (or affine) image operations (Whitaker et al 2008). For such cases, SL estimation offers superior performance by using a merit function that is an affine transform of the image. Performing the optimization of this merit function requires an inner product between a template and the raw projection data at each candidate solution. Because the template depends upon the candidate solution, but not the image data, a majority of the computation can be performed prior to image acquisition, and the same templates can be used on an ensemble of images.

The SL estimation rule assumes that the marginalization over nuisance parameters (as in equation (3)) produces a Gaussian likelihood:

\[
pr(g|\theta) = \frac{\exp\left[-\frac{1}{2}(g - \tilde{g}(\theta))^tK_{g\theta}^{-1}(g - \tilde{g}(\theta))\right]}{\sqrt{(2\pi)^M|K_{g\theta}|}},
\]

where the \( M \times 1 \) vector \( g \) is the noisy projection data, \( \theta \) is the vector of randomly varying signal parameters to be estimated and \( K_{g\theta} \) is the covariance of the data conditioned on a particular value of the signal’s parameters. We use the imaging equation

\[
g = s(\theta) + b(\theta) + n,
\]

where \( s \) is the noise-free projection of the signal, \( b \) is the noise-free projection of the random background and \( n \) is a zero-mean vector representing measurement noise. Here both the signal and the background are dependent upon \( \theta \) because the presence of a signal displaces the background. This displacement in the object model is an assumption of ROI estimation that is preserved to idealize the conditions for ROI estimation to succeed. The mean data vector, conditioned on a given value of the signal activity, is

\[
\tilde{g}(\theta) = s(\theta) + \tilde{b}(\theta),
\]

where a noise-free projection of the average background is denoted \( \tilde{b} \). Finding the maximum of the approximated likelihood in equation (4) would, in general, require an evaluation of the determinant and inverse of the conditional covariance. Instead, we substitute the conditional covariance with its evaluation at the prior mean, i.e. \( K_{g\theta} \approx K_{g\theta}|_{\theta} \), thereby erasing the dependence of this matrix on the parameters we seek to estimate. An alternative means of removing the dependence of the covariance matrix on the parameters of interest is to normalize over the ensemble, as in \( K_{g\theta} \approx \langle K_{g\theta(\theta)} \rangle_{\theta} \). The choice between these two approximations will depend upon the available prior knowledge. In this paper, we will not assume knowledge of the distribution \( pr(\theta) \), only the mean value \( \theta \).

Now only the quadratic exponent of the Gaussian needs to be maximized, and after dropping terms that do not depend on \( \theta \), we are left with an estimate of the form

\[
\hat{\theta}_{SL}(g) = \arg\max_{\theta} \left\{ \tilde{g}(\theta)^tK_{g\theta}^{-1}\left[g - \frac{1}{2}\tilde{g}(\theta)\right]\right\},
\]

where the SL estimate \( \hat{\theta}_{SL}(g) \) earns its name because the quantity to be maximized is linear with respect to the noisy image \( g \), and the search for the maximum of this linear functional can be accomplished by scanning the space of possible solutions to \( \theta \). The SL estimator is a...
nonlinear optimization of a merit function which is an affine transform of the image data. This form is simply expressed by recasting equation (7) into

$$\hat{\theta}_{SL}(g) = \arg \max_{\theta} \{ W(\theta)^t g + \epsilon(\theta) \}.$$  

(8)

Here the inner product of the $M \times 1$ vector $W(\theta)$ with a given image is added to the scalar $\epsilon(\theta)$ to produce the SL estimator merit function at the candidate solution $\theta$. This highlights a feature very useful for computation. The templates $W(\theta)$ are independent of the image data and can be computed once for the entire image ensemble. If used in the course of an imaging study, the SL estimator templates would be computed prior to imaging, thus not delaying post-processing. Figures 1 and 2 display example SL estimator templates used in this simulation study.

In the special case of the parameters of interests being linear, with respect to the object, and a linear imaging system, scanning is not necessary. For the task of estimating the activity of a signal of known location and size, the relations above become $\theta = \alpha$ and the average over noise and nuisance parameters becomes

$$\tilde{g}(\alpha) = s(\alpha) + \bar{b} = \alpha s' + \bar{b},$$  

(9)

where $s'$ is a normalized image of the signal such that $\alpha$ is the signal’s true activity. For a signal of known location and size, the background is zero within the support of the signal and independent of the signal’s activity. Now the partial derivative of the objective function in equation (7) with respect to $\alpha$ will vanish at the arg max solution; thus, the estimate can be calculated by

$$\hat{\alpha}_{SL}(g) = \frac{s'^t K_{g \alpha}^{-1} (g - \bar{b})}{s'^t K_{g \alpha}^{-1} s'},$$  

(10)
which is an unbiased estimate of signal activity

$$\langle \hat{\alpha}_{SL}(g) \rangle_{g|\alpha} = \frac{s' K_{g\alpha}^{-1} (\alpha s' + \bar{b} - \bar{b})}{s' K_{g\alpha}^{-1} s'} = \alpha. \quad (11)$$

This unbiased property of the SL estimate of activity is not dependent upon the likelihood approximation of equation (4). Here, the SL estimate is unbiased since the parameter to be estimated, $\alpha$, is linear with respect to the image data and the average image equals equation (9).

When the parameters to be estimated are random variables, the ensemble mean-square error (EMSE), for a given estimation rule $\hat{\theta}$, is defined as

$$\text{EMSE}(\hat{\theta}) = \langle \langle (\hat{\theta}(g) - \theta)^2 \rangle_{g|\theta} \rangle_{\theta} \quad (12)$$

and the ensemble bias as

$$\text{B}(\hat{\theta}) = \langle (\langle \hat{\theta}(g) - \theta \rangle_{g|\theta} - \theta) \rangle_{\theta} \quad (13)$$

From these definitions, we see that the ensemble bias of the SL estimate of activity (i.e. $\theta = \alpha$) is zero and the EMSE is given by

$$\text{EMSE}(\hat{\alpha}_{SL}) = \frac{1}{s' K_{\alpha\alpha}^{-1} s'}. \quad (14)$$

Analytic results for the EMSE and ensemble bias of the SL estimator are not available for more complicated estimation tasks, such as when the task (and therefore the vector $\theta$) includes signal size and location. For these cases, sample statistics are reported in section 5.

For a comparison between the unbiased linear estimation rule which minimizes EMSE (i.e. the Wiener estimator) and SL activity estimates, see Whitaker et al (2008). An important distinction is that the Wiener estimator requires second-order statistics on the parameter to be estimated (i.e. the variance of the signal activity) and SL requires only first-order statistics.
2.2. ROI estimation from reconstructions

Null functions of the imaging system prevent the unbiased estimation of integrated activity from ROIs of reconstructed data (Barrett and Myers 2004). For practically any real imaging system, one could imagine two distinct objects producing identical noise-free images. The difference between such objects is referred to as a null function of the imaging system (Clarkson and Barrett 1998). Consider two sine wave patterns that, due to aliasing, produce the same image data; the difference between the sine waves is a null function of the imaging system. Regardless of the size of the signal, any binary ROI that is designed to abruptly demarcate the signal from background will have high-frequency content that is beyond the resolving power of current SPECT imaging systems. If the template defining the ROI contains null functions with respect to the imaging system, then any algebraic manipulation of the reconstructed values in this ROI will be a biased estimate of the signal’s true activity. One way that the singularity of an imaging operator can be described is using singular-value decomposition (SVD) techniques (Palit et al. 2009). The SVD analysis of a CD operator $H$ on the Hilbert space $L^2(S)$ leads to the eigenvalue equation:

$$H^\dagger H u_n = \mu_n u_n.$$  \hfill (15)

Here $H^\dagger$ denotes the operator’s adjoint and $\mu_n$ and $u_n$ are the $n$th singular value and singular vector in object space, respectively. Since $H^\dagger H$ is a Hermitian operator, its eigenvectors can serve as an orthonormal basis for $L^2(S)$ and an arbitrary object $f$ can be expressed as a weighted sum of each $u_n$:

$$f = \sum_{n=1}^{\infty} a_n u_n,$$  \hfill (16)

where $a_n = (u_n, f)$ is a scalar product. The rank of the operator $H$ is equal to the number of nonzero eigenvalues. If $\mu_n = 0$, then $u_n$ is a null vector of the operator $H$. The eigenvalues can be rank-ordered largest to smallest, such that $\mu_n \geq \mu_{n+1}$. We can use this result to write an arbitrary object $f$ in terms of measurement and null components:

$$f = \sum_{n=1}^{R} a_n u_n + \sum_{n=R+1}^{\infty} a_n u_n$$

$$= f_{\text{meas}} + f_{\text{null}},$$  \hfill (17)

where $R$ is the rank of the $H$ operator, $f_{\text{meas}}$ is the measurement component of the object, $f_{\text{null}}$ is the null component of the object, $R \leq M$ and $\mu_n = 0$ for $n > R$.

To formulate the bias due to null functions, consider a single scalar which is defined linearly with respect to the object

$$\beta(f) = \chi^\dagger f = (\chi, f).$$  \hfill (19)

Here $f$ is the object, $\chi$ is a binary ROI template separating signal from background and $\beta$ is the total ROI output formed by the inner product of these two vectors. For a given imaging system, the object and ROI can be decomposed into their respective measurement-space and null-space components, given as

$$\beta(f) = [\chi_{\text{meas}} + \chi_{\text{null}}]^\dagger [f_{\text{meas}} + f_{\text{null}}]$$

$$= (\chi_{\text{meas}}, f_{\text{meas}}) + (\chi_{\text{meas}}, f_{\text{null}}) + (\chi_{\text{null}}, f_{\text{null}}) + (\chi_{\text{null}}, f_{\text{meas}})$$

$$= (\chi_{\text{meas}}, f_{\text{meas}}) + (\chi_{\text{null}}, f_{\text{null}}),$$  \hfill (20)

where the inner product between any two vectors’ measurement-space and null-space components is identically zero since the eigenvectors of a Hermitian operator are orthogonal,
i.e. $(u_1, u_2) = \delta_{m,n}$. The inner product between the two null vectors is a fundamental source of bias for any estimator of the quantity $\beta$ because a change in the term $(x_{null}, f_{null})$ changes the value of $\beta$ but not the image data $g$ and thus not the estimate $\hat{\beta}(g)$. The condition for estimability, that an unbiased estimator exists, is only satisfied if either $x_{null} = 0$ or $f_{null} = 0$.

Consider an estimate formed by the inner product between an ROI and a reconstruction

$$\hat{\beta}_{ROI}(g) = (\chi, R(g)),$$

(21)

where $R$ is a reconstruction operator. In section 5, we report estimates from this rule using a nonlinear MLEM reconstruction operator which is a well-established method for ROI estimation in nuclear medicine (Carson 1986). For linear least-squares reconstructions, we can analytically solve for the ensemble bias:

$$B(\hat{\beta}_{ROI}) = \langle \langle (\chi, f) - (\chi, H^+ H f) \rangle f \rangle_f$$

$$= \langle \langle (\chi, f) - (\chi, H^+ H f) \rangle f \rangle_f$$

$$= \langle (\chi, f) - (\chi_{meas}, H^+ H f) \rangle_f$$

$$= \langle (\chi_{null}, f_{null}) \rangle_f,$$

(22)

where $H^+$ is the pseudo-inverse of the imaging operator, the estimation rule is $\hat{\beta}_{ROI} = (\chi, H^+ g)$ and we are aware of the fact that linear least-squares reconstructions are in the measurement space of the imaging operator. Note that in an idealized case of a full-rank imaging operator, the reconstruction, and the associated ROI estimate, would be unbiased but in fact no CD operator can be of full rank. We also see that ROI estimators from linear least-squares reconstructions will be biased, for a given object, when the imaging operator is under-determined. An ensemble of objects could be selected such that this ensemble-averaged bias is zero since null components are not subject to a positivity constraint. However, the null component of the ROI, and therefore the bias in the estimate, depends upon factors that cannot be manipulated during an imaging study, including the signal size and the location of signal in the system FOV. The null components of various ROIs used in this simulation study are displayed in figure 3 and calculated according to methods described in Wilson and Barrett (1998). The changes in the null component with the signal size are visible in this figure. We will see in section 5 that the associated ROI estimation bias, from MLEM reconstructions, is also dependent upon the signal size.

3. Stochastic object models and image formation

The SL estimation rule requires a functional form that relates the parameters of interest to the object being imaged. The estimation rule does not affect an imaging system’s null space. Changes in the value of $\theta$ that do not in turn affect the image data $g$ and thus not the estimate $\hat{\theta}(g)$ still exist in this model-based approach but are not subject to the same systematic bias as ROI estimators. This model can be tailored to the imaging application where, for example, the tracer distribution in a targeted tissue is constrained in size or within a certain anatomical location. Understanding the specificity of the imaging agent can allow anatomical boundaries measured by CT images to be incorporated into the functional relationship between the parameters of interest $\theta$ and the object.

To quantify the performance of SL estimation and ROI estimation, this study will invoke a highly constrained simulation study. This result establishes an upper bound on the capability of these estimation techniques and serves to inform their utility in clinical applications. The SL estimation rule can be extended to higher-dimensional signal models that describe more realistic data sets.
3.1. Object model

The quantities we wish to estimate are the elements of a vector called $\theta$. This vector quantifies features or attributes of the object that we would like to estimate. Consider an object equation given by

$$f(r; \theta) = f^{\text{sig}}(r; \theta) + f^{\text{bkgnd}}(r; \theta).$$  \hspace{1cm} (23)

Here, an object $f(r; \theta)$ is constructed by adding together a signal $f^{\text{sig}}(r; \theta)$ and a background $f^{\text{bkgnd}}(r; \theta)$ term. The object is a doubly stochastic random function because both the statistics on the signal and the statistics on the background determine the overall object statistics. The signal is a deterministic function of the random vector $\theta$. The background is also a function of $\theta$. Even though the background structure is not among the features we seek to estimate, the presence of a signal displaces the background distribution. The image will be treated as a triply stochastic variable: object fluctuations due to signal and background variability, and measurement noise.

For the simulations that follow, we confine our treatment to signals that are related to $\theta$ by the parameterized signal model:

$$f^{\text{sig}}(r; \theta) = \alpha \text{sph} \left( \frac{r - c}{R} \right),$$  \hspace{1cm} (24)

where

$$\text{sph} \left( \frac{r - c}{R} \right) = \begin{cases} 0, & \text{for } |r - c| > R, \\ 1, & \text{otherwise}. \end{cases}$$  \hspace{1cm} (25)
Here, the signal’s size is defined by a sphere of radius $R$, $\alpha$ is a scalar quantity representing the activity of the spherical signal and $c = [c_x, c_y, c_z]$ is the location of the signal’s center. In this study, we will consider two tasks:

1. Estimating specific activity only: $\theta$ reduces to a single scalar $\alpha$ representing a specific activity. Using visual inspection of reconstructed data to select an ROI is tantamount to estimating a location and size of the signal. We remove this potential source of error by estimating the specific activity of a spherical signal whose radius and centroid are known. The known signal size and location fully specify a binary ROI template, $\chi$ in object space as in equation (19). The SL estimator has the form of equation (10) (in this equation, a noise-free image of $\chi$ is equivalent to $s'$) and the performance is compared to ROI estimates.

2. Simultaneously estimating size, location and specific activity: $\theta$ is a $5 \times 1$ vector given by $\theta = (\alpha, R, c_x, c_y, c_z)'$ and the SL estimate has the form of equation (7). After size and location estimates are formed, equation (10) is used to estimate a specific activity.

For the second task, we will assume that the parameters to be estimated are statistically independent, so the multivariate joint probability density function (pdf) is separable leading to $\text{pr}(\theta) = \text{pr}_\alpha(\alpha) \text{pr}_R(R) \text{pr}_c(c)$. The signal’s radius and specific activity are values naturally constrained to be positive; thus, a shifted gamma distribution is chosen to ensure positivity. The gamma distribution has been shifted such that all signal radii have a lower bound equal to twice the spacing of measured voxel points; $R_0 = 1.0$ mm. Without this constraint, all radii below 1.0 mm would produce nearly the same object and the parameter would be unidentifiable from even noise-free images. For all studies that follow, $R = 2.3$ mm and the average specific activity is selected such that the ratio between average values of $f_{\text{sig}}$ and $f_{\text{bkgnd}}$ is $5.3$. Over the activity ensemble, this ratio varies from around 4.8 to 6.0; this is a physical realistic range given the specificity of common imaging agents in nuclear medicine. The prior distribution on the location of the signal is a multivariate uncorrelated Gaussian centered at the origin. In our simulation study, the linear dimension of the object space FOV is 31 mm, and the width parameter of the Gaussian distribution is set to $\sigma_c = 4$ mm. This limits the probability of signal-absent images while maintaining significant variation in the location of the signal within the FOV.

Samples from $\text{pr}(\theta)$ are drawn to generate an image ensemble; noisy image examples are shown in figure 4. Note that knowledge of $\text{pr}(\theta)$ is not used again in this study after the images have been generated. The SL estimation rule does not use this prior to calculate an estimate from a given image. Also, the reconstruction algorithm and ROI estimation rule do not require this type of prior knowledge.

Object variability due to an unknown background is modeled by sampling $f_{\text{bkgnd}}(r)$ from a Gaussian lumpy background model described in Rolland and Barrett (1992). This models the pdf $\text{pr}(f_{\text{bkgnd}}(r))$ as a Poisson point process filtered by a Gaussian, as opposed to the more simplified model that the pdf itself is Gaussian.

3.2. Forward model

A measured point spread function from our multi-module multi-resolution ($M^3R$) small-animal SPECT system (Hesterman et al 2007) was used to emulate realistic projection data from discrete objects. The FOV of the imaging system is a cylinder 38.5 mm high and 31.0 mm in diameter. Within this FOV, the system response at 211 673 locations is sampled at 0.5 mm intervals. Sixteen cameras, each containing $79 \times 79$ detectors, surround the object. Each aperture is a single 0.5 mm pinhole, and the lateral magnification is 3.33. Sixteen
angular projections are collected for a total of $M = 99,856$ measurements. On average, all 16 cameras combined collect approximately 9.4 million counts. The Poisson noise is added to the projections. Figure 4 shows example noisy projection data.

The reconstructions are calculated from an MLEM iterative algorithm derived from Poisson measurement noise statistics. The results of ROI estimation are from 80 MLEM iterations with no post-filtering. Example digital objects are displayed in figure 5 along with corresponding reconstructions.

4. Simulation methodology

4.1. Covariance approximations

The SL estimator requires the inverse of the data covariance matrix. The data covariance $K_{g|g}$ is an $M \times M$ matrix that would require about 40 Gb to store as floating-point numbers and even more memory to calculate the inverse. Our covariance is influenced by two terms: measurement noise and background variability. We are able to circumvent the storage dilemma by using $S$ sample background images to estimate the background variability term and an analytic Poisson noise model. This technique requires a statistical process from which to draw sample backgrounds. The resulting full-rank estimate of the matrix can be inverted by only calculating an $S \times S$ inverse using the Woodbury matrix inversion lemma (Barrett et al 2001). In the current simulation, $S = 6000$. An analysis of the effect this sample size has on the resulting estimation performance can be found in Whitaker (2008).
Figure 5. Cross-sectional slices of the 3D object space FOV for (a) 1.0 mm, (b) 2.0 mm and (c) 3.0 mm signal radius of known location and size. The top row shows four samples from the object ensemble where only the specific activity is varying and in the bottom row are MLEM reconstructions of the object after 80 iterations (no post-smoothing). Here the displaced signal model is visible.
4.2. Optimization

To implement the optimization required by the SL estimation rule in equation (7), the candidate solutions for location are parsed into 64,000 points equally spaced within the FOV. Based on the inspection of the SL merit function, a sequential optimization routine is employed where the radius is fixed, at a relatively small size, and the activity at a relatively high value. This first sequence of evaluations scans candidate solutions for the location. A cross-section of the merit function, for a given image, is shown in figure 6. The location grid’s extremum serves as the starting point of a simulated annealing routine that searches all five components of $\theta$ simultaneously (Press et al, 2007). The stopping criterion is a threshold on the number of merit-function evaluations. Reported SL estimates meet the criteria of $O(\hat{\theta}_{SL}) \geq O(\theta)$, where $O$ is the SL estimation merit function and $\theta$ is the true value of the parameters. This ensures that the reported estimation error does not also contain error due to an imperfect optimization implementation. A potentially useful hardware implementation for likelihood optimization, which uses an initial grid evaluation and subsequent contractions around each grid’s extremum, is described in Hesterman et al (2012).

5. Simulation results

A quantitative comparison between ROI and SL estimations is performed using the EMSE definition in equation (12).

5.1. Task 1: estimating activity only

For 200 images, the results of both the maximum and average values within the MLEM-reconstructed ROI, for three different signal sizes, are shown in figure 7. The bias in the average ROI estimates can be seen in figure 7(a); simply shifting each estimate by a constant would improve the performance. It is important to note that the bias is a strong function of signal size. As evident from figure 7(a), the estimates of signal activity from larger signals have a lower bias. The performance of the estimator, at all signal sizes, is very poor. The EMSE is 1092, 270 and 143 for signal radius 1.0, 2.0 and 3.0 mm, respectively. The performance of this estimator does increase with the signal size, although it is low overall.

Figure 7(b) is a scatter plot of the maximum ROI estimates. The maximum in the ROI is biased high for the two larger signal sizes and biased low for the smallest. Maximum ROI estimates report the value of a single reconstructed voxel from a noisy ROI; therefore, the estimate’s variance is higher compared to the variance of the average within the ROI.
Figure 7. (a) The mean value within the reconstructed ROI as dependent upon the actual value. (b) The maximum value within the reconstructed ROI as dependent upon the actual value. The blue triangles are the estimates from signals of 1.0 mm in radius (two voxels), the red circles are 2.0 mm (four voxels) and the green squares are 3.0 mm (six voxels). The black line indicates ideal estimation performance; the estimated value equals the true value.

Figure 8. Two hundred SL estimates of activity as dependent upon the actual value for (a) 1.0 mm, (b) 2.0 mm and (c) 3.0 mm signal radius of known location and size.

Averaging numerous reconstructed values smooth out noise variability. The EMSEs are 269, 664 and 898 for the signal radius 1.0, 2.0 and 3.0 mm, respectively. For this task, the ROI maximum performance is better than the ROI average for the smallest signal but still very poor overall.

The SL estimator is calculated on the same 200 images, and the results are given in figure 8. The SL EMSEs are 5.13, 0.22 and 0.04 for the signal radius 1.0, 2.0 and 3.0 mm, respectively. These errors are extremely small compared to those of the ROI estimators.

More angular projections would improve the appearance of the reconstructed data (shown in figure 5) and likely increase the performance of ROI estimation. These results indicate that collecting more projection images is not necessary for this estimation task. When the task is to estimate the activity of a signal whose size and location are known, operating on the raw projection data using the SL estimator yields significantly lower EMSE than using the statistics of a reconstructed ROI.

5.2. Task 2: estimating activity, size and location

For 200 images, the results of simultaneously estimating a signal’s activity, radius and three-dimensional location are shown in separate scatter plots in figures 9 and 10. These results show
Figure 9. (a) Two hundred SL estimates of activity as dependent upon the actual value for a signal of unknown radius and location which are simultaneously estimated and reported in figure 10. (b) Two hundred ROI estimates of activity based on prior knowledge of signal location and radius.

Figure 10. Two hundred SL estimates of (a) radius, (b) x-coordinate, (c) y-coordinate and (d) z-coordinate locations as dependent upon the actual value for a signal of unknown activity which is simultaneously estimated and reported in figure 9.
Figure 11. (a) SL estimates of integrated activity versus the true value and (b) relating error in SL activity estimates to an error in SL radius estimates.

the high SL estimation performance that is possible under these simulation conditions. The sum of EMSE values for SL estimates of signal radius and location is below a voxel spacing. SL estimates of signal amplitude are calculated once the signal’s radius and location have been estimated; the EMSE of SL estimates of activity is 4.67. Figure 9 includes a comparison of the SL activity estimates to the ROI estimates. Note that the ROI estimates have prior knowledge of the radius and location yet the SL estimates still outperform them, as measured by the EMSE, without this prior knowledge.

Figure 11(a) shows that for 92% of the images, the SL errors in activity estimation are under 5% and that the performance of SL activity is directly related to the error in size estimates. The integrated activity is the product of the signal’s uniform activity and the signal volume. This estimate is shown in figure 11(b) and the coupling between signal size and signal activity estimates is evident.

6. Conclusions

The SLE offers a dramatic increase in estimation performance when compared to using the statistics of a reconstructed ROI. The inherent bias in ROI estimators can be understood from considering that any sharp boundary demarcating signal from background will have a null component with respect to tomographic medical imaging systems. If the parameter to be estimated is defined by an inner product between the object and this binary template, then the parameter is not estimable, i.e. no unbiased estimator exists. Furthermore, the correlations between adjacent voxels of an MLEM reconstruction are dependent upon many factors (Barrett et al 1994, Wilson et al 1994). By treating reconstructed data as an estimate of the continuous object and performing an inner product with a binary template, these measurement limitations of the imaging system are ignored at the cost of estimation accuracy. A quantitative image analysis demands greater accuracy than the ROI estimation performance demonstrated in this SPECT simulation study. A solution is offered by instead projecting a binary ROI through the imaging system and minimizing a distance metric between candidate solutions from a parametric model and the measured projection images. In this simulation study, the distance metric is based on an approximate ML estimate. This scanning-linear (SL) estimator earns its name by performing linear operations on the raw projection data and then scanning through
candidate solutions from the parametric model to optimize the data agreement. This offers a
computation that is simultaneously accurate, within the constraints of this study, and tractable to
calculate. In practice, the candidate solutions would be calculated prior to imaging acquisition
and thus SL estimation would not necessarily require significant delays in post-processing.
In this simulation study, we have neglected disagreement between the parametric object
models used in the estimation rule and the objects that generated the image data; this is an
inevitable source of error in clinical imaging. This simulation test has decoupled estimation
robustness from the value of prior information in improving estimation accuracy. By doing
so, the demonstrated accuracy achievable by SL estimation motivates further investigation.
To move toward a clinical application extension to more realistic and higher dimensional
parameterizations of the object is needed along with an analysis of robustness to model
disagreement.

7. Discussions

In addition to the increased computational requirements (relative to ROI estimation), a more
formidable and nuanced challenge to implementing SL estimation in clinical environments is
the prior knowledge required. An SLE ingredient list includes the following.
(I) A well-characterized forward model which enables noise-free images to be calculated
from a given candidate solution to the object distribution (as described in section 3.2).
(II) An analytic model that relates the object to the parameters to be estimated (as in equation
(24)).
(III) The average value of the parameters to be estimated, used to form the conditional
covariance \( K_{g|\theta} \).
(IV) An analytic model from which to estimate the first- and second-order statistics (i.e. \( \bar{g} \) and
\( K_b \)) of the background activity (this simulation relied on a Gaussian lumpy background
model).

Items I and II are paramount to any quantitative imaging study and not unique to medical
tomography or SL estimation. A forward model is needed for reconstruction algorithms;
typically measurements, models, or both are used to compute an image from a discrete object
representation. This calculation is necessary to compute a distance metric between a measured
image and a candidate solution to the unknown object. Estimation rules are predicated on a
parametric object representation that relates the quantities of interest, \( \theta \), to the object \( f(\theta) \).
Usually, these quantities do not uniquely specify an object; in medical imaging, the values
describe a signal of interest, \( f_{\text{sig}}(\theta) \), that is embedded in a background. The form of \( f_{\text{sig}}(\theta) \) can
be specified from knowledge of the disease and mechanism of tracer uptake. Multi-modality
imaging allows the span of candidate object solutions to be narrowed down, by demarcating
anatomical boundaries from CT scans. Similarly, a value of \( \bar{\theta} \), mentioned in item III, could
be informed by supplementary measurements, \textit{a priori} knowledge, or from a rough estimate
using current clinical metrics. Given that the exact value of \( \bar{\theta} \) is not required, rather this value
determines features of the conditional covariance matrix, SL estimation performance might
prove to be more robust to this condition as compared with the other three ingredients in the
above list.

Item IV, the statistical characterization of the background activity (nuisance parameters),
constitutes a significant hurdle to implementing SL estimation in clinical environments.
However, this ambition for more information is timely given that new instrument capabilities
are changing measurement strategies in medical imaging. In Barrett \textit{et al} (2008), an adaptive
SPECT system is described that collects an initial scout image to obtain preliminary
information about the radiotracer distribution and then adjusts the configuration or sizes
of the pinholes, the magnifications, or the projection angles in order to improve performance. In addition to hardware adaptations, such a scout image could also inform an estimate of the background activity, or the correlation length of the covariance, or it could limit the scope of candidate background ensembles. Another key advantage of the SL estimate over ROI estimation is the framework to make use of this type of prior information. For example, an assumed parametric form of the background covariance allows training sets to be used in a preliminary estimation step that yields an estimate of $K_b$. Under constraints of stationarity and isotropy, even a single image could be used to estimate a parameterized power spectral density. Objects composed of several tissue types might obey stationarity within these anatomical boundaries; the statistics within each boundary would be different. Modifying the parameters of a stationary covariance to include spatial dependence yields a quasi-stationary model which is a candidate solution for modeling the covariance of a population of patients. To work toward clinical solutions for SL implementation, these methods of estimating the background covariance will be the subject of upcoming publications.

Acknowledgment

This work has been supported by NIH grants R01 EB000803 and P41 EB002035.

References

Boellaard R 2012 Mutatis mutandis: harmonize the standard! J. Nucl. Med. 53 1–3
Buckler A J and Boellaard R 2011 Standardization of quantitative imaging: the time is right, and 18F-FDG PET/CT is a good place to start J. Nucl. Med. 52 171–2
Clarkson E 2007b Estimation ROC curves and their corresponding ideal observers Proc. SPIE 6515 651504
Hicks R J 2008 PET for therapeutic response monitoring in oncology PET Clin. 3 89–99
Shankar L K et al 2006 Consensus recommendations for the use of 18F-FDG PET as an indicator of therapeutic response in patients in national cancer institute trials J. Nucl. Med. 47 1059–66
Shao J 1999 Mathematical Statistics (Berlin: Springer)
Thie J A 2004 Understanding the standardized uptake value, its methods, and implications for usage J. Nucl. Med. 45 1431–4
Vanderhoek M, Perlman S B and Jeraj R 2012 Impact of the definition of peak standardized uptake value on quantification of treatment response J. Nucl. Med. 53 4–11
Wilson D W and Barrett H H 1998 Decomposition of images and objects into measurement and null components Opt. Express 2 254–60