Conference Record of the 1994 IEEE Nuclear Science Symposium and Medical Imaging Conference

Uniform Quadratic Penalties Cause Nonuniform Spatial Resolution

Jeffrey A. Fessler and W. Leslie Rogers
3480 Kresge III, Box 0552, University of Michigan, Ann Arbor, MI 48109-0552

Abstract

This paper examines the spatial resolution properties of penalized maximum-likelihood image reconstruction methods by analyzing the *local impulse response*. We show that for emission image reconstruction using the ordinary uniform quadratic regularization penalty, the local impulse response is *spatially variant*. Paradoxically, the *local resolution is poorest in high activity regions*. The analysis leads naturally to a modified quadratic regularization penalty that achieves nearly uniform resolution. The modified penalty also provides a very practical method for *choosing the smoothing parameter to obtain a specified resolution.*

I. Introduction

Penalized maximum-likelihood methods for image reconstruction have two important advantages over alternative methods for reducing noise (such as stopping rules and sieves). First, the penalty function improves the conditioning of the problem, so certain iterative algorithms converge very quickly. Second, one has wide flexibility in choosing penalty functions, so one can control properties of the reconstructed images, such as preserving edges or incorporating anatomical information. In contrast, the smoothness that one obtains through stopping rules is limited by the characteristics of the iterative algorithm. A possible disadvantage of penalized-likelihood methods has been the lack of an easy method for choosing the value of the regularization parameter, even for simple quadratic penalties. One contribution of this paper is a new object-independent method for specifying the regularization parameter in terms of the desired resolution of the reconstructed image.

This paper describes another undesirable property of penalized-likelihood image reconstruction methods, and then proposes a solution to the problem. In brief, we demonstrate that even when one uses the ordinary shift-invariant roughness penalty, the reconstructed images have *object-dependent nonuniform spatial resolution, even if the tomographic system is spatially invariant.* For emission imaging the resolution is poorest in high-count regions, which is directly opposite what one might expect or prefer. We then propose a modified roughness penalty that yields nearly uniform spatial resolution.

This work was supported in part by DOE grant DE-FG02-87ER60561 and NIH grant CA-60711-01.

This paper is somewhat in the spirit of the early studies by Stamos *et al.* [1] and the recent analysis of Barrett, Wilson, and Tsui [2,3], which used the *local impulse response* to quantify spatial resolution properties of the unregularized maximum-likelihood expectation-maximization (ML-EM) algorithm. Liow and Strother also studied ML-EM using an effective local Gaussian resolution [4]. Since the ML-EM algorithm is rarely iterated until convergence, most authors study its spatial resolution properties *as a function of iteration*. In contrast, since there are now fast and *globally convergent* algorithms for maximizing both penalized maximum-likelihood [5–7] and penalized weighted least squares [8,9] objective functions, rather than studying the properties of the *algorithms* as a function of iteration, we study directly the properties of the *estimator* as specified by the *objective function* (Section II). This simplifies the practical use and interpretation of our results since the specifics of the maximization algorithm are unimportant (provided one uses a globally convergent method).

In conventional FBP image reconstruction, one controls the tradeoff between resolution and noise by adjusting the cutoff frequency f_c of a filter. In practice, one determines the (monotonic) relationship between f_c and full-width half maximum (FWHM) through the following empirical approach. First, acquire a sinogram using a point or line source, possibly at several locations within the scanner. Then pick a filter type (e.g. Hanning) and reconstruct images for several different values of f_c. Finally, compute the FHWM of the point spread function (PSF) for each case, and record a table of (f_c, FWHM) value pairs. In subsequent studies, one typically chooses the desired resolution (FWHM) through experience and by visually observing the resolution-noise tradeoff, and then obtains the appropriate f_c from the table. This tabulation process needs only to be done once for a given scanner and window function, since FBP is linear (and hence its resolution properties are object-independent).

In contrast, in penalized maximum-likelihood image reconstruction, the smoothing parameter β controls the tradeoff between resolution and noise, but β is essentially unitless. Therefore it is not obvious how to specify the smoothing parameter, and one finds that for a fixed β, the reconstructed spatial resolution varies between subjects, and even within the same subject (Section III). This paper contributes a practical method for normalizing the weights in the penalty function that nearly eliminates the object-dependent component. This allows one to build an object-independent table relating β to resolution (FWHM) for a given PET system, so that one can select β to achieve a consistent specified resolution within planes, between planes, and even between subjects. The task of choosing the "optimal" resolution is left to the user.

Nonuniform resolution properties are not unique to penalized-likelihood methods. Wilson and Tsui [2] reported resolution variation and asymmetry for the ML-EM algorithm as well. An advantage of the penalized-likelihood approach is that one can modify the penalty to overcome the resolution nonuniformity.

PET and SPECT systems have intrinsically nonuniform resolution [10] (although PET systems are fairly spatially-invariant near the center of the scanner). The simulations we report are for an idealized spatially invariant PET system. Thus the resolution nonuniformity is due solely to the interaction between the log-likelihood and the penalty terms of the objective function, and not due to the system response. Further work is needed to study the effects of penalty functions in spatially-variant systems.

II. Theory

A. Local Impulse Response

Let $Y = [y_1, \ldots, y_N]'$ denote a random measurement vector with density function $f(y;\theta)$, where $\theta = [\theta_1, \ldots, \theta_p]'$ is an unknown parameter in a p-dimensional parameter space Θ, and $'$ denotes vector transpose. Typically $\Theta = \{\theta : \theta_j \geq 0,\ j = 1, \ldots, p\}$. Given a particular realization $Y = y$, an estimator $\hat{\theta} = \hat{\theta}(y)$ has mean:

$$\mu(\theta) = E_\theta\{\hat{\theta}(Y)\} = \int \hat{\theta}(y) f(y;\theta)\ dy. \tag{1}$$

For nonlinear, spatially-variant estimators, one can characterize the spatial resolution properties of μ by analyzing the *local impulse response* (cf [1, 2]). We define the local impulse response at the jth pixel to be:

$$\begin{aligned} l^j(\theta) &= \lim_{\delta \to 0} \frac{1}{\delta}[\mu(\theta + \delta e^j) - \mu(\theta)] \\ &= \frac{\partial}{\partial \theta_j} \mu(\theta), \end{aligned} \tag{2}$$

for $j = 1, \ldots, p$, where e^j is the jth unit vector of length p. This impulse response is local in two different senses. First, it is a function of the index j, reflecting the spatially-variant nature of nonlinear estimation. Second, it depends on the location in the parameter space Θ through the argument θ, reflecting the nonlinear object dependence. The local impulse response also depends on the measurement distribution through (1). Thus, the local impulse response characterizes the object, system, and estimator dependent properties. In words, the local impulse response measures the *change* in the mean reconstructed image due to the perturbation of a particular pixel in the noiseless object.

As a specific example, consider the estimator:

$$\hat{\theta} = \hat{\theta}(y) = \arg\min_\theta\ (y - \mathbf{A}\theta)'\mathbf{W}(y - \mathbf{A}\theta) + \beta\theta'\mathbf{R}\theta,$$

where $\mathbf{W}$ and $\mathbf{R}$ are nonnegative definite matrices and the null space of $\mathbf{R}$ is orthogonal to that of $\mathbf{WA}$. This penalized weighted least-squares estimator is linear:

$$\hat{\theta}(y) = (\mathbf{A}'\mathbf{WA} + \beta\mathbf{R})^{-1}\mathbf{A}'\mathbf{W}y.$$

Since $E_\theta\{Y\} = \mathbf{A}\theta$, one can evaluate (2) to show that

$$l^j = (\mathbf{A}'\mathbf{WA} + \beta\mathbf{R})^{-1}\mathbf{A}'\mathbf{WA}e^j. \tag{3}$$

For such linear estimators, the local impulse response is independent of θ. As we show below, the local impulse response of the nonlinear estimators for image reconstruction have approximately the same form as (3).

B. Brute Force Evaluation

Unlike in the simple penalized weighted least squares estimator described above, there is usually no explicit analytical form for estimators $\hat{\theta}(y)$ that are defined implicitly as the maximum of some objective function:

$$\hat{\theta} = \hat{\theta}(y) = \arg\max_{\theta \in \Theta} \Phi(\theta, y). \tag{4}$$

When there is no explicit form for $\hat{\theta}(y)$, there is usually no explicit form for its mean $\mu(\theta)$ either. Thus it would at first appear that to investigate the local impulse response of a nonlinear estimator of interest, one must resort to a "brute force" numerical approach based on (1) and (2), with the expectation in (2) replaced by the sample mean in a computer simulation. However, often we would accept an *approximation* to the local impulse response if we could avoid performing extensive simulations. The remainder is devoted to approximations suitable for likelihood-based estimators in emission tomography.

C. Implicitly Defined Estimators

The mean of a likelihood-based estimator often approximately equals the estimate obtained from noiseless data:

$$\mu(\theta) = E_\theta\{\hat{\theta}(Y)\} \approx \hat{\theta}(\bar{Y}(\theta)) \stackrel{\triangle}{=} \bar{\theta}, \tag{5}$$

where

$$\bar{Y}(\theta) = E_\theta\{Y\} = \int y f(y;\theta)\ dy$$

denotes the measurement mean and $\bar{\theta}$ denotes the value of the estimator when applied to noiseless data $\bar{Y}(\theta)$. This approximation is equivalent to assuming that the estimator is *locally linear* [2], and is the basis for the remainder of this paper. Substituting (5) into (2) yields the following approximation to the local impulse response:

$$l^j(\theta) \approx \frac{\partial}{\partial \theta_j} \hat{\theta}(\bar{Y}(\theta)). \tag{6}$$

This expression depends on the partial derivatives of the implicitly defined estimator $\hat{\theta}(y)$, which one can determine using the implicit function theorem and the chain rule [11].

Disregarding the nonnegativity constraint the maximizer of Φ satisfies:

$$\nabla^{10}\Phi(\hat{\theta}(y), y) = 0, \quad \forall y, \tag{7}$$

where $\nabla^{10} = [\frac{\partial}{\partial\theta_1} \; \cdots \frac{\partial}{\partial\theta_p}]$ is the row gradient operator (with respect to the first argument of Φ). Now differentiate again with respect to y using the chain rule:

$$\nabla^{20}\Phi(\hat{\theta}(y), y)\nabla_y\hat{\theta}(y) + \nabla^{11}\Phi(\hat{\theta}(y), y) = 0. \tag{8}$$

For simplicity we drop the dependence of $\bar{Y}$ on θ except where explicitly needed. Under the mild assumption that $-\nabla^{20}\Phi(\bar{\theta}, \bar{Y})$ is positive definite, we substitute $y = \bar{Y}$ into (8) and solve:

$$\nabla_y\hat{\theta}(\bar{Y}(\theta)) = [-\nabla^{20}\Phi(\bar{\theta}, \bar{Y})]^{-1}\nabla^{11}\Phi(\bar{\theta}, \bar{Y}).$$

Combining with the chain rule applied to (6):

$$\begin{aligned} l^j(\theta) &\approx \frac{\partial}{\partial\theta_j}\hat{\theta}(\bar{Y}(\theta)) = \nabla_y\hat{\theta}(\bar{Y}(\theta))\frac{\partial}{\partial\theta_j}\bar{Y}(\theta) \\ &= [-\nabla^{20}\Phi(\bar{\theta}, \bar{Y})]^{-1}\nabla^{11}\Phi(\bar{\theta}, \bar{Y})\frac{\partial}{\partial\theta_j}\bar{Y}(\theta). \end{aligned} \tag{9}$$

This approximation specifies the local impulse response solely in terms of the derivatives of the objective function and the measurement mean, i.e., we have eliminated the dependence on the implicitly defined estimator $\hat{\theta}(y)$.

D. Penalized Likelihood Estimators

In the remainder, we focus on objective functions Φ that are of the penalized likelihood form. If we define $L(\theta, y) = \log f(\theta; y)$ to be the log-likelihood under the assumed distribution, the penalized likelihood objective function is

$$\Phi(\theta, y) = L(\theta, y) - \beta V(\theta), \tag{10}$$

where V is a roughness penalty function.

Define $\mathbf{R}(\theta) = \nabla^2 V(\theta)$, and note that $\nabla^{11}V = 0$. Then we have from (9) the following approximation to the local impulse response of penalized likelihood estimators:

$$l^j(\theta) \approx [-\nabla^{20}L(\bar{\theta}, \bar{Y}) + \beta\mathbf{R}(\bar{\theta})]^{-1}\nabla^{11}L(\bar{\theta}, \bar{Y})\frac{\partial}{\partial\theta_j}\bar{Y}(\theta). \tag{11}$$

E. Poisson Statistics

Here we focus on the emission case, see [12] for the transmission case. In emission tomography θ_j denotes the radioisotope concentration in the jth voxel, and the measurements have independent Poisson distributions with means:

$$\bar{Y}_i(\theta) = \sum_{j=1}^{p} a_{ij}\theta_j + r_i, \tag{12}$$

where $\{a_{ij}\}$ are nonnegative constants that characterize the system, and $\{r_i\}$ are nonnegative constants that represent the contribution of background events (random coincidences, scatter, etc.). The Poisson log-likelihood is: $L(\theta, y) = \sum_i y_i \log \bar{Y}_i(\theta) - \bar{Y}_i(\theta)$, and one can show [12]:

$$\begin{aligned} -\nabla^{20}L(\theta, y) &= \mathbf{A}'D\left\{y_i/\bar{Y}_i^2(\theta)\right\}\mathbf{A} \\ \nabla^{11}L(\theta, y) &= \mathbf{A}'D\left\{1/\bar{Y}_i(\theta)\right\}, \end{aligned}$$

where $D\{u_i\}$ denotes a $N \times N$ diagonal matrix with elements $u_1, \ldots, u_N$ along the diagonal. Note that $\frac{\partial}{\partial\theta_j}\bar{Y}_i(\theta) = a_{ij}$, so substituting into (11) yields an approximation for the local impulse response:

$$l^j(\theta) \approx [\mathbf{A}'D\left\{\frac{\bar{Y}_i(\theta)}{\bar{Y}_i^2(\bar{\theta})}\right\}\mathbf{A} + \beta\mathbf{R}(\bar{\theta})]^{-1}\mathbf{A}'D\left\{\frac{1}{\bar{Y}_i(\bar{\theta})}\right\}\mathbf{A}e^j.$$

For moderate or small values of β, $\bar{\theta}$ is a slightly blurred version of θ (see (5)). Since a projection induces smoothing, $\bar{Y}(\theta) \approx \bar{Y}(\bar{\theta})$, and if the model mismatch is small, then $\bar{Y}(\theta) \approx \bar{Y}(\theta)$. Therefore, we simplify the above expression to the approximation

$$l^j(\theta) \approx [\mathbf{A}'D\{u_i(\theta)\}\mathbf{A} + \beta\mathbf{R}(\bar{\theta})]^{-1}\mathbf{A}'D\{u_i(\theta)\}\mathbf{A}e^j, \tag{13}$$

where (cf (3))

$$u_i(\theta) = 1/\bar{Y}_i(\theta) \tag{14}$$

is the inverse of the variance of Y_i under the assumed Poisson model. This is our final approximation to the local impulse response for penalized likelihood estimators in emission tomography[2].

III. Resolution Properties

A. Nonuniform Resolution

Consider the conventional quadratic roughness penalty:

$$V(\theta) = \sum_{j=1}^{p}\frac{1}{2}\sum_k w_{jk}\frac{1}{2}(\theta_j - \theta_k)^2,$$

where typically one chooses $w_{jk} = 1$ for horizontal or vertical neighbors, $w_{jk} = 1/\sqrt{2}$ for diagonal neighbors, and $w_{jk} = 0$ otherwise. The elements of the θ-independent Hessian $\mathbf{R} = \{r_{jk}\}$ for this penalty are given by:

$$r_{jk} = \begin{cases} \sum_l w_{jl}, & k = j \\ -w_{jk}, & k \neq j \end{cases}.$$

This penalty is *shift invariant* in the sense that if the entire image is shifted, the value of the penalty $V(\theta)$ would remain unchanged. However, upon evaluating (13) one finds that the local impulse response is very nonuniform. To explain this nonuniformity, suppose for argument's sake that

[2] Preconditioned conjugate gradient and Gauss-Siedel [8] work well for evaluating (13).

$D\{u_i\}$ were simply a scaled identity matrix: $\mathbf{D} = \nu^{-1}\mathbf{I}$. Then the local impulse response (13) would be

$$\begin{aligned} l^j(\theta) &\approx [\nu^{-1}\mathbf{A}'\mathbf{A} + \beta\mathbf{R}(\bar{\theta})]^{-1}\nu^{-1}\mathbf{A}'\mathbf{A}e^j \\ &= [\mathbf{A}'\mathbf{A} + \nu\beta\mathbf{R}(\bar{\theta})]^{-1}\mathbf{A}'\mathbf{A}e^j. \end{aligned}$$

Thus, changes in the variance ν lead to an impulse response with an "effective" smoothing parameter $\nu\beta$. In other words, the influence of the penalty is *not invariant to changes in the noise variance*, which partly explains the difficulty in choosing β. The situation is further complicated since the values of $D\{u_i\}$ in fact vary along the diagonal. Since a given pixel is primarily affected by the detectors whose rays intersect it, each pixel sees a different "effective variance" and hence a different effective smoothing parameter.

B. Nonuniformity Correction

To develop a modified penalty function that yields uniform resolution, we study the problem of resolution nonuniformity more closely for quadratic penalties. First, note that in PET, the elements a_{ij} can be factored as: $a_{ij} = c_i^{-1}g_{ij}$, where $\{g_{ij}\}$ are fixed geometrical factors describing the tomographic system, and $\{c_i\}$ are correction terms that change between studies, including attenuation factors, dead time, radioisotope decay, and detector normalization factors. We assume $c_i > 0$. In matrix notation $\mathbf{A} = D\{c_i\}^{-1}\mathbf{G}$, so substituting into (13):

$$l^j(\theta) \approx (\mathbf{G}'D\{q_i(\theta)\}\mathbf{G} + \beta\mathbf{R})^{-1}\mathbf{G}'D\{q_i(\theta)\}\mathbf{G}e^j,$$

where

$$q_i(\theta) = 1/(\bar{Y}_i(\theta) \cdot c_i^2). \tag{15}$$

Next, we note that *the local impulse response l^j depends predominantly on those elements i of $D\{q_i\}$ that correspond to rays that intersect pixel j.* Define

$$d_j(\theta) = \sum_i g_{ij}q_i(\theta) \Big/ \sum_i g_{ij}, \tag{16}$$

then if we content ourselves with the *average* local resolution, i.e., assuming the local impulse response is approximately circularly symmetric, then we can replace the diagonal matrix $D\{q_i(\theta)\}$ with an identity matrix multiplied by $d_j(\theta)$. Thus

$$l^j(\theta) \approx (\mathbf{G}'\mathbf{G} + (\beta/d_j(\theta))\mathbf{R})^{-1}\mathbf{G}'\mathbf{G}e^j. \tag{17}$$

This approximation illuminates the paradoxical oversmoothing of high-count regions with the uniform penalty. If pixel j is transected by rays with high counts, then from (15) we see that $q_i(\theta)$ and hence $d_j(\theta)$ will be small, so the effective smoothing parameter $\beta/d_j(\theta)$ in (17) will be large, causing lower resolution. As θ_j increases, the rays that intersect it will also increase, so the local resolution decreases. This is certainly undesirable, and may explain in part why quadratic penalties have had such bad press.

The form of (17) immediately suggests that a natural method for encouraging the local resolution to be uniform is to use modified weights:

$$w_{jk}^{\star} = w_{jk}\sqrt{d_j d_k}. \tag{18}$$

With this modified penalty function, the local impulse response is approximately:

$$l^j(\theta) \approx (\mathbf{G}'\mathbf{G} + \beta\mathbf{R})^{-1}\mathbf{G}'\mathbf{G}e^j. \tag{19}$$

This expression is independent of the object θ, and of all the patient-dependent c_i factors. Thus, to with the accuracy of our (possibly fairly crude) approximations, *the modified quadratic penalty function yields an object-independent impulse response.* If the geometric response $\mathbf{G}$ of the imaging system is approximately spatially-invariant, then (19) implies that the impulse response will also be approximately spatially invariant, i.e. the resolution will be uniform. In [12] we provide empirical results that confirm that, despite the many approximations in our development, the modified penalty function greatly improves the resolution uniformity of penalized maximum likelihood image reconstruction.

C. Projection Dependence

The definition (15) depends on θ, which is unknown for real data. Fortunately, a remarkable property of the local impulse response approximation given by (13),(14) is that if $V(\theta)$ is a quadratic form so that $\mathbf{R}$ is independent of θ, then *$l^j(\theta)$ depends on the object θ only through its projections $\bar{Y}(\theta)$.* Even if the object is unknown, its projections are approximately known through the noisy measurements y. Thus, even for real noisy measurements, we can predict the local impulse response simply by replacing $\bar{Y}(\theta)$ with y in (13). This simplistic approach is effective primarily because the diagonal term in (13) is sandwiched between the backprojection and projection operators $\mathbf{A}'$ and $\mathbf{A}$, so the noise in y is significantly smoothed out. Therefore, in implementation with measured data, we use the following approximation when computing $\{d_j\}$:

$$q_i(\theta) \approx 1/(\max\{y_i, 10\} \cdot c_i^2).$$

The "10" factor ensures that the denominator is not too close to zero.

D. Choosing β

Since the local impulse response (19) for the modified quadratic penalty is independent of the object θ, the process of choosing the smoothing parameter β is significantly simplified by the following approach. Let j be a pixel in the center of the image, for example. For a given system geometric response $\mathbf{G}$, precompute

$$l^j \approx (\mathbf{G}'\mathbf{G} + \beta\mathbf{R})^{-1}\mathbf{G}'\mathbf{G}e^j.$$

for a range of values of β. For each β, tabulate some measure of resolution, such as the full-width half-maximum (FWHM) of l^j. Then, when presented with a new data set to be reconstructed at a *specified resolution*, simply interpolate the table to determine the appropriate value for β. Finally, reconstruct the object using the modified quadratic penalty based on (18).

IV. Discussion

By analyzing the local impulse response, we have shown that for emission tomographic image reconstruction with the ordinary uniform quadratic penalty, the reconstructed image resolution is nonuniform. Fig. 1 gives an empirical example of the nonuniform resolution, and shows how the proposed modified penalty function improves the resolution uniformity. We chose the value of β for this example to match the resolution of the Hanning windowed FBP image using the proposed method described in Section III.D.

We believe the spatially-variant spatial resolution properties are not restricted to the quadratic penalty. From (13), one can see that for non-quadratic penalties the local resolution will depend both on the "effective smoothing parameter" $\beta/d_j(\theta)$, as well as on the object-dependent curvature of the penalty for mean object, i.e. $\mathbf{R}(\bar{\theta})$. The latter effect is presumably desired, such in edge-preserving methods. The former will manifest itself as a spatial variance in "the propensity to preserve edges," which is probably undesirable. Future work will explore whether for non-quadratic edge-preserving penalties, a modified penalty can achieve a more uniform *local step response*.

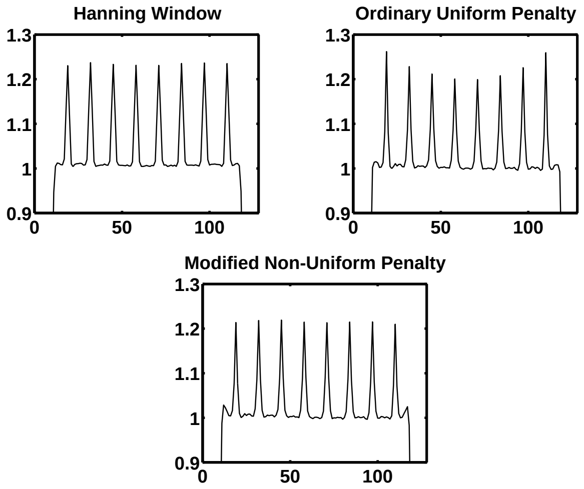

Figure 1: Horizontal profile through reconstructed images (from noiseless PET projections) of a phantom consisting of a uniform ellipse (intensity 1) and several single pixel point sources (intensity 2). Conventional FBP with a Hanning window demonstrates spatially-invariant resolution, as expected. Penalized maximum likelihood reconstruction with the ordinary uniform quadratic smoothness penalty induces *nonuniform* resolution (upper right). By appropriately modifying the roughness penalty using (18), we achieve uniform resolution (lower). We do *not need to know the object* to perform this modification; the necessary information is contained in the projections.

References

[1] J A Stamos, W L Rogers, N H Clinthorne, and K F Koral. Object-dependent performance comparison of two iterative reconstruction algorithms. *IEEE Trans. Nuc. Sci.*, 35(1):611–614, February 1988.

[2] D W Wilson and B M W Tsui. Spatial resolution properties of FB and ML-EM reconstruction methods. In *Conf. Rec. of the IEEE Nuc. Sci. Symp. Med. Im. Conf.*, volume 2, pages 1189–1193, 1993.

[3] H H Barrett, D W Wilson, and B M W Tsui. Noise properties of the EM algorithm: I. Theory. *Phys. Med. Biol.*, 39:833–846, 1994.

[4] J S Liow and S C Strother. The convergence of object dependent resolution in maximum likelihood based tomographic image reconstruction. *Phys. Med. Biol.*, 38(1):55–70, January 1993.

[5] J A Fessler and A O Hero. Space-alternating generalized expectation-maximization algorithm. *IEEE Trans. Sig. Proc.*, 1994. In press - October issue.

[6] J A Fessler and A O Hero. Space-alternating generalized EM algorithms for penalized maximum-likelihood image reconstruction. Technical Report 286, Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. of Michigan, Ann Arbor, MI, 48109-2122, February 1994.

[7] C Bouman and K Sauer. Fast numerical methods for emission and transmission tomographic reconstruction. In *Proc. 27th Conf. Info. Sci. Sys., Johns Hopkins*, pages 611–616, 1993.

[8] K Sauer and C Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Trans. Sig. Proc.*, 41(2):534–548, February 1993.

[9] J A Fessler. Penalized weighted least-squares image reconstruction for positron emission tomography. *IEEE Trans. Med. Im.*, 13(2):290–300, June 1994.

[10] E J Hoffman, S C Huang, D Plummer, and M E Phelps. Quantitation in positron emission tomography: 6 effect of nonuniform resolution. *J. Comp. Assisted Tomo.*, 6(5):987–999, October 1982.

[11] J A Fessler. Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): applications to tomography, 1994. Submitted to IEEE Trans. Image Proc.

[12] J A Fessler et al. Spatial resolution properties of penalized maximum-likelihood image reconstruction methods, 1994. In preparation.