Spatially-Variant Roughness Penalty Design for Uniform Resolution in Penalized-Likelihood Image Reconstruction

J. Webster Stayman and Jeffrey A. Fessler
4415 EECS Dept., University of Michigan, Ann Arbor, MI 48109
email: stayman@eecs.umich.edu

Abstract

Traditional space-invariant regularization schemes in tomographic image reconstruction using penalized-likelihood estimators produce images with nonuniform resolution properties. The local point spread functions that quantify the local smoothing properties of such estimators are not only space-variant and asymmetric, but are also object-dependent even for space-invariant systems. We propose a new regularization scheme for increased spatial uniformity and demonstrate the resolution properties of this new method versus conventional regularization schemes through an investigation of local point spread functions.

1 Introduction

Statistical image reconstruction methods provide improved noise and resolution properties over conventional nonstatistical methods such as filtered backprojection (FBP). However, methods based purely on the maximum-likelihood estimate produce overly noisy images. This noise may be reduced by stopping the iterative procedure used to find the maximum-likelihood estimate before convergence [1], iterating until convergence followed by post-smoothing [2], or including a penalty term in the likelihood objective function [3].

Penalized-likelihood methods have the advantage of allowing arbitrary regularizations including edge-preserving penalties and penalties incorporating anatomical side or boundary information. Regularization can also improve the conditioning of the problem leading to faster convergence.

However, there are also disadvantages with conventional regularization schemes. Space-invariant penalties lead to object-dependent nonuniform resolution properties [4, 5]. Specifically, such estimators tend to smooth the image more in high count regions than in low count regions. The local point spread functions that quantify this space-variant smoothing due to the estimator can also be highly asymmetric, indicating a preferred smoothing direction.

Since conventional regularizations produce images with nonuniform resolution properties, one cannot select the regularization parameter intuitively. For example, with FBP the noise-resolution tradeoff is controlled through the cutoff frequency, f_c, of the filter. There is a direct relationship between f_c and the *global* full-width half-maximum (FWHM) resolution of the reconstructed image. Such a direct relation does not exist with penalized-likelihood reconstructions with conventional regularizations.

One attempt to correct for the resolution nonuniformity has been presented in [4]. This regularization, based on the aggregate certainty of measurement rays intersecting each pixel, provides increased spatial uniformity over conventional space-invariant regularization. However, the local point spread functions are still highly asymmetric.

We present a new parameterized penalty function, whose coefficients are found by a least-squares fitting of the parameterized local impulse response to a shift-invariant response. This new method provides increased spatial uniformity compared to the certainty-based method of [4] and conventional regularization techniques. We demonstrate this increased uniformity through an investigation of the local point spread functions.

2 Background

Let $\lambda = [\lambda_1, \ldots, \lambda_p]'$ represent the nonnegative emission rates for an object discretized into p pixels. Detectors surrounding the object count photons (SPECT) or photon pairs (PET) that are emitted from the object. Measurements are denoted by the random vector $Y = [Y_1, \ldots, Y_N]'$. These measurements are Poisson with a mean given by

$$\bar{Y}_i(\lambda) = \sum_{j=1}^{p} a_{ij}\lambda_j + r_i,$$

where a_{ij} represent nonnegative constants that characterize the tomographic system, and r_i are nonnegative constants that specify the contribution due to background events (background radiation, scatter, etc.). Given measurements y, we would like to reconstruct λ.

We will focus on penalized-likelihood estimators of the form

$$\hat{\lambda}(y) = \arg \max_{\lambda \in \Lambda} L(\lambda, y) - \beta R(\lambda),$$

where Λ is the set of feasible images, $L(\lambda, y)$ is the log-likelihood, β is the regularization parameter which controls the noise-resolution tradeoff, and $R(\lambda)$ is a roughness penalty.

For the Poisson model, the log-likelihood is

$$L(\lambda, y) = \sum_{i=1}^{N} y_i \log \bar{Y}_i(\lambda) - \bar{Y}_i(\lambda).$$

Pairwise roughness penalties have the following form

$$R(\lambda) = \sum_{j=1}^{p} \frac{1}{2} \sum_{k \in \mathcal{N}_j} w_{jk} \psi(\lambda_j - \lambda_k),$$

where $\mathcal{N}_j$ is a neighborhood of pixels around pixel j, ψ is a symmetric convex function, and $w_{jk} = w_{kj}$.

In the case of a quadratic penalty, $\psi(x) = x^2/2$ and the roughness penalty may be written in matrix form, $R(\lambda) = \frac{1}{2} \lambda' R \lambda$, where the matrix R has elements defined by

$$R_{jk} = \begin{cases} \sum_{l \in \mathcal{N}_j} w_{lj}, & k = j \\ -w_{jk}, & k \neq j. \end{cases}$$

A conventional choice for a space-invariant penalty using a first-order neighborhood is to choose $w_{jk} = 1$ for the horizontal and vertical neighbors. For a second-order penalty, one often includes $w_{jk} = 1/\sqrt{2}$ for the diagonal neighbors in addition to the first-order neighbors.

The mean reconstruction of an estimator is designated by

$$\mu(\lambda) = E_\lambda[\hat{\lambda}(Y)] = \int \hat{\lambda}(y) f(y; \lambda) dy.$$

The local impulse response [4] is defined as

$$l^j = \lim_{\delta \to 0} \frac{\mu(\lambda + \delta e^j) - \mu(\lambda)}{\delta} = \frac{\partial}{\partial \lambda_j} \mu(\lambda),$$

where e^j represents the jth unit vector. This definition of the local impulse response is dependent on the estimator, the object, and the pixel position j. From [4], for penalized-likelihood with quadratic penalties, the local impulse response may be well approximated by

$$l^j \approx [A' D_\lambda A + \beta R]^{-1} A' D_\lambda A e^j, \qquad (1)$$

where A is a matrix of the $\{a_{ij}\}$ elements, and D_λ is a diagonal matrix with elements $1/\bar{Y}_i$.

Typically, $A'A$ is a shift-variant operator even without D_λ. However, for PET systems it is often possible to

factor A such that $A = D[c_i]G$, where $G'G$ is approximately shift-invariant and represents the geometric system response. The diagonal matrix $D[c_i]$, contains ray-dependent effects such as detector efficiency and attenuation. With this factorization, equation (1) becomes

$$l^j \approx [G'WG + \beta R]^{-1} G'WG e^j, \qquad (2)$$

where diagonal matrix W has elements $c_i^2/\bar{Y}_i$.

These approximations of the local impulse response form a very important set of tools for the design and evaluation of different regularization methods.

3 Methods

Ideally, we would like to be able to find a penalty that yields an arbitrary desired space-invariant response. For example, we may desire penalty functions that produce a global impulse response with a Gaussian shape and some specified FWHM resolution. If we restrict ourselves to the quadratic penalty, we can formulate this problem in terms of the design of the penalty matrix R.

Specifically, we would like to find R as

$$R = \arg \min_{R \geq 0} \sum_{j=1}^{p} d(l^j(R), l^\star), \qquad (3)$$

where $d(l^j, l^\star)$ is some distance measure between the actual response, l^j and a desired space-invariant response, $l^\star$. Theoretically, we could solve (3) by plugging in (2), however this procedure is not computationally tractable.

In practice, conventional penalties use only a small neighborhood of pixels for the penalty support (i.e. first- and second-order neighborhoods). Therefore we reformulate the problem in terms of these small support neighborhoods.

For the shift-invariant quadratic penalty, the matrix R can be thought of as a space-invariant filtering operator. Therefore the operation of R on the image λ can be represented as the convolution of the image with a kernel,

$$R\lambda = k_R ** \lambda.$$

For example, the conventional first-order penalty described beforehand has the kernel,

$$k_{R_0} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \qquad (4)$$

The design of R can be thought of as a filter design problem. However, selection of a kernel, k_R, cannot be completely arbitrary. Since R must be symmetric, k_R must be symmetric through the origin. Additionally, R should yield a zero penalty for uniform regions. This is equivalent to specifying a zero DC gain on the filter represented

686

by k_R. (The coefficients must sum to zero.) Lastly, we require that R be nonnegative definite to guarantee finite solutions.

The following equation forms basis for valid kernels of space-invariant R,

$$b_{(k,l)}(\underline{x}) = 2\delta(x_1, x_2) - \delta(x_1 - k, x_2 - l) - \delta(x_1 + k, x_2 + l),$$

where $\delta(\cdot)$ represents a Kronecker delta function and, x_1 and x_2 represent spatial coordinates. For example, for a first-order neighborhood,

$$b_{(1,0)} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad b_{(0,1)} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

is a valid basis for the kernels of R. In general, any valid penalty kernel may be specified by

$$k_R(x_1, x_2) = \sum_{k,l \in \mathcal{N}} r_{kl} b_{(k,l)}(x_1, x_2) = B\underline{r},$$

where r_{kl} represent the basis coefficients. Letting $\underline{r}$ represent the set of all r_{kl} for a given neighborhood of support, $\mathcal{N}$, and B represent a matrix of basis functions, we write this linear combination more compactly. For nonnegative definite R, the elements of $\underline{r}$ must be nonnegative.

In general the penalty specified by R does not have to be space-invariant. In fact, we *require* a space-variant regularization for uniform resolution properties. Since this is the case, we extend idea of a kernel representation and let R be represented by the space-variant set of kernels $k_R^j = B\underline{r}^j$.

We may now rewrite (3) as

$$\{\underline{r}^k\}_{k=1}^p = \arg \min_{\{\underline{r}^k\}_{k=1}^p \geq 0} \sum_{j=1}^p d(l^j(\{\underline{r}^k\}_{k=1}^p), l^\star). \quad (5)$$

Although this requires less computations than (3), this optimization is still impractical since all $\underline{r}^j$ must be found simultaneously.

3.1 Proposed Penalty Design

Since $G'G$ is approximately shift-invariant, we may approximate $G'G$ by $Q'\Omega Q$, where Q is a 2D discrete Fourier matrix operator and Ω is a diagonal matrix representing a frequency domain filtering operator. (This is the well-known $1/\rho$ response of the backprojected projection operator.) Although $G'WG$ is not shift-invariant, it is approximately locally shift-invariant and we make the following approximation to (2)

$$\begin{aligned} l^j(R) &\approx [Q'\Omega_j Q + \beta Q'\Psi_j Q]^{-1} Q'\Omega_j Q e^j \\ &= Q' \left[\frac{\Omega_j}{\Omega_j + \beta\Psi_j} \right] Q e^j, \end{aligned}$$

where the division is an element-by-element division, $\Omega_j = F\{G'WGe^j\}$, and $\Psi_j = F\{Re^j\}$. ($F\{\cdot\}$ represents the discrete 2D Fourier operator.) Therefore, since $Re^j = k_R^j$,

$$l^j(\underline{r}^j) = F^{-1} \left\{ \frac{F\{G'WGe^j\}}{F\{G'WGe^j\} + \beta F\{B\underline{r}^j\}} \right\}. \quad (6)$$

Combining (5) and (6) yields a separable minimization and we may find $\underline{r}^j$ individually by

$$\underline{r}^j = \arg \min_{\underline{r}^j \geq 0} d(l^j(\underline{r}^j), l^\star). \quad (7)$$

If $d(x, y) = \|x - y\|^2$, then this is a constrained nonlinear least-squares (CNLLS) problem which must be solved for each pixel j. Since this is fairly time consuming to calculate, we would like to simplify this optimization.

Consider the unweighted response given by

$$l_0^j = [G'G + \beta R_0]^{-1} G'Ge^j. \quad (8)$$

If R_0 is chosen to be space-invariant, the response, l_0^j, is approximately space-invariant since $G'G$ is nearly shift-invariant. We choose (8) as the desired response, $l^\star$. Defining $L^j = F\{l(\underline{r}^j)\}$, we have

$$\begin{aligned} L^j(\underline{r}^j) &\approx \frac{F\{G'WGe^j\}}{F\{G'WGe^j\} + \beta F\{B\underline{r}^j\}} \\ L_0^j &\approx \frac{F\{G'Ge^j\}}{F\{G'Ge^j\} + \beta F\{k_{R_0}\}}. \end{aligned}$$

We want to choose $\underline{r}^j$ so that $L^j(\underline{r}^j) \approx L_0^j$. Cross-multiplying and simplifying yields

$$F\{G'WGe^j\} \odot F\{k_{R_0}\} \approx F\{G'G\} \odot F\{B\underline{r}^j\},$$

where $\odot$ represents an element-by-element multiplication. Note that this step has eliminated the dependence on β. We now choose $\underline{r}^j$ such that

$$\underline{r}^j = \arg \min_{\underline{r}^j \geq \underline{0}} \left\| \Phi\underline{r}^j - \underline{d}^j \right\|^2 \quad (9)$$

with

$$\begin{aligned} \Phi &= F\{G'Ge^j\} \odot F\{B\} \\ \underline{d}^j &= F\{G'WGe^j\} \odot F\{k_{R_0}\}. \end{aligned}$$

This optimization may be performed by applying a nonnegative least-squares for each pixel position, j, for example, the NNLS algorithm in [6].

687

3.2 Practical Implementation

While the penalty design given by equation (9) gives a simple form for the calculation of $\underline{r}^j$, it still requires too much computation for practical use. Since $G'G$ is nearly shift-invariant, we only need to calculate the 2D-FFT (fast Fourier transform) of $G'Ge^j$ for a single j. Similarly, $F\{B\}$ requires the one-time calculation of the 2D-FFT of each of the 2D basis functions. Therefore Φ may be precalculated. For $\underline{d}^j$, $F\{k_{R_0}\}$ may be precalculated but we must compute the 2D-FFT of $G'WGe^j$ for every pixel j. This step makes direct computation of (9) a slow process.

One can show that in an idealized continuous system, if the continuous equivalent of W is a *radially-constant* sinogram scaling operator, then the continuous equivalent of $G'WG$ can be expressed as a position-independent blurring operation [5]. Therefore, this should be approximately true in the discrete case. If W were radially-constant, we would only need one computation of $F\{G'WGe^j\}$.

In practice we cannot expect the radially-constant assumption to hold. However, since the projection of a single pixel forms a relatively narrow trace in sinogram space (only a few radial bins in width), W can be well approximated by a position-dependent radially-constant operator W^j. Therefore, $G'WGe^j$ approximately equals a shifted $G'W^jGe^{i_0}$ for an arbitrary fixed pixel i_0 (*i.e.*: the center pixel in the image).

Note that for the unconstrained case, equation (9) has the closed form linear solution $\underline{r}^j = [\Phi'\Phi]^{-1}\Phi'd_j$. Also note that the unconstrained solution is a linear function of the elements of the weighting matrix W. Since, the unconstrained solution is linear with respect to W, we can form a linear operator (matrix) mapping $\{w_{ii}\}$ to $\{\underline{r}^j\}$. However, this requires a different mapping for every pixel j.

For the radially-constant W^j, we require only a single linear mapping from $\{w_{ii}^j\}$ to $\{\underline{r}^j\}$. Call this matrix operator T. We may find T by superposition. Define W^{δ_n} to be weighting matrix with radially-constant values having unit values at angle n, and zero otherwise. For each angle let

$$m_n = F\{G'W^{\delta_n}G\} \odot F\{k_{R_0}\}, \quad n = 1, \ldots, n_a$$

and $M = [m_1, \ldots, m_{n_a}]$, where n_a is the number of angular samples. Therefore,

$$T = [\Phi'\Phi]^{-1}\Phi M$$

is a linear operator that solves the unconstrained optimization on (9). It is straightforward to modify the NNLS algorithm of [6] using these ideas to provide the constrained solution.

In this way, the majority of calculations may be precomputed for a given system geometry and the proposed regularization specified $\underline{r}^j$ may be calculated very quickly for specific measurement realizations. (Recall the elements of W are given by $c_i^2/\bar{Y}_i$, or estimated by c_i^2/y_i.)
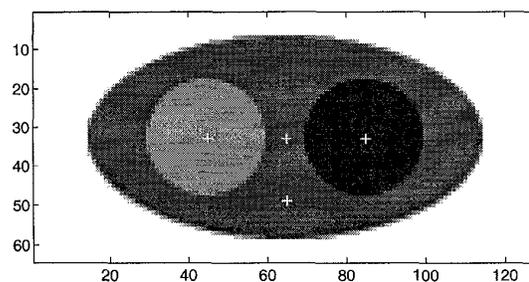

Figure 1: Digital phantom used for investigation of resolution properties of different regularizations.

4 Results

This section provides simulation results comparing the relative resolution uniformity of different regularization schemes. Figure 1 shows the 128×64 emission image (λ) used for the investigation. The image has a warm background ellipse, a cold left disc, and hot right disc with relative emission intensities of 2, 1, and 3, and attenuation coefficients of 0.003, 0.0096, and 0.013/mm, respectively. The PET system model included projection data with 128 radial bins over 110 angles uniformly spread over 180° with 3 mm pixels, 6 mm wide strip integrals (3 mm center-to-center spacing), and detector efficiencies with a standard deviation of 0.3.

We investigate the resolution properties of four different quadratic regularizations: (I) The conventional space-invariant first-order penalty given by the kernel in (4), (II) the certainty-based penalty developed in [4], (III) the constrained nonlinear least-squares (CNLLS) penalty given by equation (7), and (IV) the reduced computationally-efficient penalty we have proposed in this paper. For the CNLLS and proposed penalty design, we choose a second-order basis and select R_0 in equation (8) to be the conventional space-invariant first-order penalty, as above. We specify the regularization parameter (β) to correspond to a target resolution 4.0 pixels FWHM resolution for each of these penalties.

To demonstrate the relative spatial uniformity of these regularization methods we use equation (1) to calculate local point spread functions (PSFs). Since we expect these responses to be space-variant, we choose four different locations in the object for our investigation. These points are represented by the white + marks in Figure 1.

Results of this impulse response survey are presented in Figures 2-5. For each penalty PSF contours at 25, 50, 75, and 99% of peak value are shown. Above each set of contours are estimates of the mean and standard deviation of the FWHM resolution in pixels, which quantify the mean resolution and radial variation at that location.

688

m = 3.566, s=0.511 m = 4.171, s=0.466 m = 4.349, s=0.490 m = 3.902, s=0.527
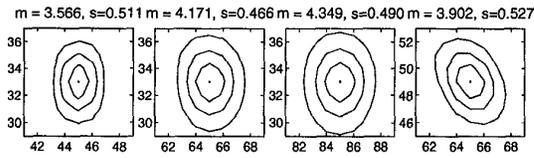
Figure 2: Local PSFs for space-invariant penalty.

m = 4.222, s=0.667 m = 4.160, s=0.465 m = 4.077, s=0.429 m = 4.078, s=0.564
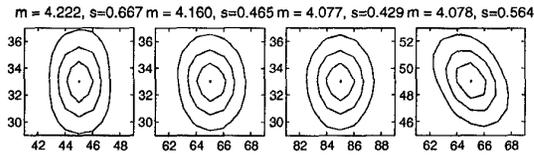
Figure 3: Local PSFs for certainty-based penalty.

m = 3.975, s=0.123 m = 3.965, s=0.175 m = 3.918, s=0.089 m = 3.935, s=0.178
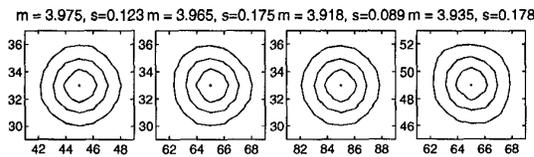
Figure 4: Local PSFs for CNLLS penalty.

m = 3.857, s=0.169 m = 3.994, s=0.215 m = 3.861, s=0.068 m = 3.971, s=0.156
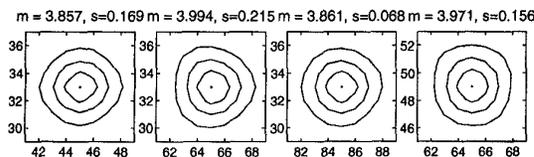
Figure 5: Local PSFs for proposed penalty.

For the space-invariant penalty in Figure 2, the local PSFs are highly asymmetric and space-variant, blurring more in high count regions (85,33) than in low count regions. The certainty-based penalty shown in Figure 3 provides some improvement making the mean FWHM close to 4.0 pixels. However, the responses are still fairly asymmetric. PSFs for the CNLLS penalty are shown in Figure 4. Note these contours are nearly radially-symmetric and near the 4.0 pixel FWHM target resolution. The PSFs of the proposed regularization method are presented in Figure 5. These contours are also highly symmetric and the average FWHM resolution is very close to the target resolution of 4.0 pixels. In this case, there is little to be gained from the computationally expensive CNLLS penalty.

5 Discussion

Conventional space-invariant regularization methods for penalized-likelihood image reconstruction produce images with space-variant resolution properties. At present, the only method available that attempts to provide more uniform resolution is the certainty-based method of [4]. However, as we have seen in our investigations, this method does not provide truly uniform resolution properties.

We have presented a new computationally-efficient regularization scheme for increased spatial uniformity. The proposed method is based on fitting an unweighted response with an arbitrary space-invariant penalty (equation (8)). This method offers nearly space-invariant and nearly symmetric local point spread functions at FWHM resolutions very close to specified target resolutions.

Providing a regularization scheme that yields uniform resolution properties makes the selection of the regularization parameter (β) more intuitive. One may simply specify the desired global resolution of the reconstructed image. While one may arguably desire space-variant resolution properties, one would most likely want to be able to control regional resolution properties, while maintaining radially-symmetric responses. These methods can be modified to provide such control, allowing for predictable and intuitive specification of resolution properties in image reconstruction.

A more complete presentation of the proposed regularization technique including an investigation of the noise properties can be found in [7].

References

[1] E. Veklerov and J. Llacer, "Stopping rule for the MLE algorithm based on statistical hypothesis testing," *IEEE Tr. Med. Im.*, vol. 6, pp. 313–319, Dec. 1987.

[2] D. L. Snyder and M. I. Miller, "The use of sieves to stabilize images produced with the EM algorithm for emission tomography," *IEEE Tr. Nuc. Sci.*, vol. 32, pp. 3864–3871, Oct. 1985.

[3] K. Lange, "Convergence of EM image reconstruction algorithms with Gibbs smoothing," *IEEE Tr. Med. Im.*, vol. 9, pp. 439–446, Dec. 1990. Corrections, June 1991.

[4] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs," *IEEE Tr. Im. Proc.*, vol. 5, pp. 1346–58, Sept. 1996.

[5] J. A. Fessler, "Resolution properties of regularized image reconstruction methods," Tech. Rep. 297, Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. of Michigan, Ann Arbor, MI, 48109-2122, Aug. 1995.

[6] C. L. Lawson and R. J. Hanson, *Solving least squares problems*. Prentice-Hall, 1974.

[7] J. W. Stayman and J. A. Fessler, "Regularization for uniform resolution properties in penalized-likelihood reconstruction," 1998. In preparation.

689