Information Processing in Medical Imaging, pp. 27-39, 1995

USING GROUND-TRUTH DATA TO DESIGN PRIORS IN BAYESIAN SPECT RECONSTRUCTION

S.J. LEE AND G.R. GINDI
Departments of Electrical Engineering and Radiology
SUNY at Stony Brook, Stony Brook NY 11794 USA

AND

I.G. ZUBAL AND A. RANGARAJAN
Departments of Diagnostic Radiology and Computer Science
Yale University, New Haven CT 06510 USA

Abstract. We propose a method for estimating hyperparameters for prior models used in SPECT reconstruction. Our method uses a maximum-likelihood technique to fit hyperparameters to a training set that is itself representative of ground-truth objects to be reconstructed. We argue that autoradiography may be a good source for such exemplars. For a simple smoothing prior, we derive a closed-form expression for the hyperparameter, and for a more complex prior that promotes piecewise smoothness, we evaluate the likelihood expression numerically. Numerical experiments demonstrate that parameters derived from training sets are fairly close to parameter values optimal in the sense of various measures of image quality.

1. Introduction

Bayesian methods in SPECT are attractive in that they support accurate models of the imaging system, and allow incorporation of prior information into the solution. Many studies demonstrate that Bayesian reconstruction methods can result in improved, relative to conventional methods, image quality according to various metrics. The transition to practical use is, however, hampered by computational barriers. The optimizations resulting from Bayesian methods are numerically intensive and may be impractical to run if implemented on conventional hardware. Evolution in hardware and the prospects of utilizing special-purpose parallel computers have dampened these objections; but, a second computational obstacle remains.

Algorithms based on Bayesian approaches are usually iterative and contain free parameters, and the performance of these algorithms usually depends crucially on the values of these parameters. Some parameters, such as step size in a descent procedure, or number of iterations, are associated with the iteration procedure itself, and others, termed hyperparameters, are associated with the mathematical model for the prior probability. Empirical studies show a tradeoff in reconstruction bias and variance as these parameters are varied, but, depending on a particular task, there is often an optimal choice for a parameter setting. Many studies have assessed bias/variance tradeoffs as a function of iteration number for ML-EM and other, Bayesian, algorithms, but much less work has been directed towards determining the values of hyperparameters.

The choice of a prior model is tantamount to a belief in the spatial structure of the underlying object whose noisy and corrupted projection image we observe. One class of well-studied "bootstrap" methods [1, 2] seeks to determine the optimal values from the noisy and corrupted projection data. A second strategy [3] involves determining these values from a training set of noiseless objects, chosen judiciously to reflect the kinds of objects we expect to encounter in a clinical setting. The latter method is much less computationally demanding, but its utility ultimately relies on its ability to generalize from training data.

2. Autoradiography

Ideally, an exemplar should be a record of the radionuclide spatial density in a human patient. Since there is no way to record ground-truth functional patterns in humans, one often resorts to physical or software phantoms. These are typically idealized versions of the expected functional pattern [4]. An opportunity to record ground-truth functional patterns, does exist, however, for animal models in the form of autoradiography. Successful reconstruction of ground truth biological data is appealing, but the validity of the extension to humans from animal models, even advanced primates, remains open. Results in our work, however, may be applied to *any* functional phantom, biological or synthetic, that is deemed useful.

While autoradiography has become a standard research tool in much of biological science, it is less familiar in radiological imaging. In autoradiography [5], an animal is injected with a radiopharmaceutical, sacrificed, and thin (20μ) tissue sections are apposed to film in order to record high-resolution images of the radionuclide density. Though the image formation process for autoradiographs is characterized by its own quantities of noise and blur, these are so small, that from the standpoint of radiology, the autoradiograph can be considered as a ground-truth phantom. In our work, we have used digitized (64×64) versions of autoradiographs ("autorads") such as those in Fig. 1 (a)(b)(c). These are scaled geometrically to approximately match adult human dimensions so as to more realistically simulate the physical effects. For the studies reported here, we simply compute deterministic line integrals of the phantom and add an amount of independent Poisson noise appropriate for the count level desired. Attenuation is simulated by assuming a constant attenuation coefficient (0.12 cm^{-1}) equal to that of water. We note that these 2-D phantoms are crude; a better autorad phantom would be 3-D and contain spatial detail only up to a degree of resolution consistent with the imaging system under study.

3. Mechanical Models as Priors

Though the methods here apply to any form of prior, we have been exploring a class of "mechanical" models [6] inspired by observations of the spatial structure of autorads. The *weak membrane* (WM) prior encourages the formation of piecewise flat regions while the *weak plate* (WP) prior encourages piecewise linear regions. In addition, we consider a third prior, the *membrane*(MM), that encourages overall smoothness with no discontinuities.

We use the familiar Gibbs distribution to model prior distributions:

$$\Pr(\mathbf{F}=\mathbf{f}, \mathbf{L}=\mathbf{l}) = \frac{1}{Z}\exp(-E(\mathbf{f},\mathbf{l})),$$

where $\mathbf{f}$ is a 2-D random field realization for the source intensity, with $\mathbf{F}$ the associated random field, and $\mathbf{L}$,$\mathbf{l}$ are the random field and 2-D realization for binary line-process variables [7] used to signal the suspension of continuity. Also, Z is a normalization factor (partition function) and $E(\mathbf{f},\mathbf{l})$ the prior energy: $E_{WP}(\mathbf{f},\mathbf{l})$ for WP, $E_{WM}(\mathbf{f},\mathbf{l})$ for WM, and $E_{MM}(\mathbf{f})$ for simple membrane (MM). Our definitions of the priors are given by the energy functions:

$$E_{WM}(\mathbf{f},\mathbf{l}) = \lambda \sum_{i,j} \left\{ V^v_{i,j}(1-l^h_{i,j}) + V^h_{i,j}(1-l^v_{i,j}) \right\} + \alpha \sum_{i,j} \left(l^h_{i,j} + l^v_{i,j} \right) \quad (1)$$

$$E_{WP}(\mathbf{f},\mathbf{l}) = \lambda \sum_{i,j} \{ V_{i,j}(1-l_{i,j}) \} + \alpha \sum_{i,j} l_{i,j}, \quad (2)$$

$$E_{MM}(\mathbf{f},\mathbf{l}) = \lambda \sum_{i,j} \left\{ V^v_{i,j} + V^h_{i,j} \right\} \quad (3)$$

where $V^v_{i,j} \stackrel{\text{def}}{=} f^2_v(i,j)$ and $V^h_{i,j} \stackrel{\text{def}}{=} f^2_h(i,j)$, and $V_{i,j} \stackrel{\text{def}}{=} f^2_{vv}(i,j) + 2f^2_{hv}(i,j) + f^2_{hh}(i,j)$. The binary variable l is line process (for WM, l^v and l^h are vertical and horizontal line processes, respectively), and λ and α are our positive hyperparameters. In (1), f_v and f_h are the partial first derivatives along the vertical and the horizontal directions, respectively. The term $V_{i,j}$ in (2) is quadratic variation for a thin plate [6]. Here f_{vv} and f_{hh} are partial second derivatives along the vertical and horizontal directions, respectively, and f_{hv} is a partial cross second derivative. The terms involving $V^v_{i,j}$ and $V^h_{i,j}$ in (3) and $V_{i,j}$ in (2) encourage smoothness except where discontinuities occur (i.e. $l = 1$) and include a penalty α for the creation of discontinuities.

It turns out [6] that the above energies can be expressed in terms of **f** only, with the line processes accounted for implicitly by the form of a "broken parabola" function ϕ defined as

$$\phi_{\lambda,\alpha}(\xi) \stackrel{\text{def}}{=} \begin{cases} \xi^2 & \text{if } \xi^2 \le \alpha/\lambda \\ \alpha & \text{if } \xi^2 > \alpha/\lambda \end{cases} \tag{4}$$

In this case, the WM and WP energies may be expressed as:

$$\begin{aligned} E_{WM}(\mathbf{f}) &= \sum_{i,j} \left[\phi_{\lambda,\alpha}(\sqrt{V^v_{i,j}}) + \phi_{\lambda,\alpha}(\sqrt{V^h_{i,j}})\right] \\ E_{WP}(\mathbf{f}) &= \sum_{i,j} \left[\phi_{\lambda,\alpha}(\sqrt{V_{i,j}})\right] \end{aligned}$$

Figure 2 [8] shows regional bias and variance results on reconstructions using these priors with an autorad phantom. These optimistic results were obtained with a laborious manual search for hyperparameters.

4. Calculating Hyperparameters via Maximum Pseudolikelihood

Given ground truth data in the form of an autoradiograph, and a prior model, we may use a maximum-likelihood (ML) approach to estimate the hyperparameters. With **f** now denoting the (known) ground truth data, we formulate a maximum-likelihood criterion for optimal parameters $\hat{\alpha}$ and $\hat{\lambda}$:

$$(\hat{\lambda}, \hat{\alpha}) = \arg\min_{\alpha,\lambda}(-\log \Pr[\mathbf{F} = \mathbf{f}; \lambda, \alpha])$$

Note that, typically, we are estimating only one or two hyperparameters from a large number (64×64) of data points. For Gibbs priors, we may express the likelihood term as

$$\Pr[\mathbf{F} = \mathbf{f}; \lambda, \alpha] = \frac{1}{Z(\lambda, \alpha)} \exp(-E(\mathbf{f}; \lambda, \alpha)),$$

where the partition function $Z(\lambda, \alpha)$ is given by

$$Z(\lambda, \alpha) = \sum_{\{\mathbf{w}\}} \exp[-E(\mathbf{w}; \lambda, \alpha)].$$

Summation over $\{\mathbf{w}\}$ denotes a sum over all possible object vectors. For example, for a discretization of one pixel to 256 values and a 64x64 image, this sum has $256^{64\times 64}$ terms, clearly intractable! To sidestep this intractability, we utilize a "pseudolikelihood" approximation to ML [9] given, in our context, by

$$\begin{aligned} PL(\mathbf{f};\lambda,\alpha) &= \prod_{i,j} \Pr[F_{i,j}=f_{i,j} \mid F_{k,l}=f_{k,l},(k,l)\epsilon\mathcal{N}_{i,j};\lambda,\alpha] \\ &= \prod_{i,j} \frac{\exp\{-E(f_{i,j} \mid F_{k,l}=f_{k,l},(k,l)\epsilon\mathcal{N}_{i,j};\lambda,\alpha)\}}{\sum_w \exp\{-E(w \mid F_{k,l}=f_{k,l},(k,l)\epsilon\mathcal{N}_{i,j};\lambda,\alpha)\}}, \end{aligned}$$

where indices k,l refer to the pixels in the neighborhood $\mathcal{N}_{i,j}$ at the site i,j and w spans all possible (scalar) values that may occur at a site. Neighborhood $\mathcal{N}_{i,j}$ consists of all pixels that interact with site (i,j) through E. We again convert to a minimization problem by taking a negative log:

$$\begin{aligned} -\log PL(\lambda,\alpha) &= \sum_{i,j} E(f_{i,j} \mid F_{k,l}=f_{k,l},(k,l)\epsilon\mathcal{N}_{i,j};\lambda,\alpha) \\ &\quad + \sum_{i,j}\left[\log\sum_w \exp(-E(w \mid F_{k,l}=f_{k,l},k,l\epsilon\mathcal{N}_{i,j};\lambda,\alpha))\right] \quad (5) \end{aligned}$$

For the WM prior, the above evaluates to:

$$\begin{aligned} -\log PL(\lambda,\alpha) &= \\ & \sum_{i,j}\Big\{\phi_{\lambda,\alpha}(\sqrt{V^h_{i-1,j}})+\phi_{\lambda,\alpha}(\sqrt{V^h_{i,j-1}})+\phi_{\lambda,\alpha}(\sqrt{V^h_{i,j+1}})+\phi_{\lambda,\alpha}(\sqrt{V^h_{i+1,j}})\Big\} \\ & +\sum_{i,j}\Big[\log\sum_w \exp\big\{-\phi_{\lambda,\alpha}(w-f_{i-1,j})-\phi_{\lambda,\alpha}(w-f_{i,j-1}) \\ & \qquad -\phi_{\lambda,\alpha}(w-f_{i,j+1})-\phi_{\lambda,\alpha}(w-f_{i+1,j})\big\}\Big]. \qquad (6) \end{aligned}$$

For the simple membrane prior, we derive a closed-form solution for the single parameter $\hat{\lambda}$. To streamline the derivation, we denote four nearest neighbor pixels to $f_{i,j}$ as N_{ij}, S_{ij}, E_{ij} and W_{ij} (north,south,east,west). The pseudolikelihood solution is then:

$$\begin{aligned} \hat{\lambda} &= \arg\min_{\lambda}\sum_{i,j}\Big\{\lambda\left\{(f_{ij}-N_{ij})^2+(f_{ij}-S_{ij})^2+(f_{ij}-E_{ij})^2+(f_{ij}-W_{ij})^2\right\} \\ & +\sum_{i,j}\log\Big[\sum_w \exp\big\{-\lambda[(w-N_{ij})^2+(w-S_{ij})^2+(w-E_{ij})^2+(w-W_{ij})^2]\big\}\Big]\Big\}. \end{aligned}$$

(7)

We may complete the square in the sum in [small brackets] in (7) and reexpress it as $\left(\frac{w-\mu}{\sigma}\right)^2 + A$, where $\mu(N,S,E,W) = \frac{N+S+E+W}{4}$, $\sigma = \frac{1}{2}$, and $A(N,S,E,W) \stackrel{\text{def}}{=} N^2 + S^2 + E^2 + W^2 - \frac{(N+S+E+W)^2}{4}$. (In the definitions for A and μ we drop subscripts i and j on N, S, E, W for convenience.)

We replace the sum over w with an integral: $\sum_w \to \int_{-\infty}^{\infty} dw$, and the integration over w becomes:

$$\begin{aligned} \exp\{-\lambda A(N,S,E,W)\} \int_{-\infty}^{\infty} \exp\left\{-\lambda \left(\frac{w-\mu}{\sigma}\right)^2\right\} dw \\ = \frac{1}{2}\sqrt{\frac{\pi}{\lambda}} \exp\{-\lambda A(N,S,E,W)\}. \end{aligned} \tag{8}$$

At this point we have evaluated the term in $\Big[$ large brackets $\Big]$ in (7). Application of $\sum_{i,j}\log(\cdot)$ to the right side of (8) results, after dropping terms independent of λ, in the expression

$$-\frac{N}{2}\log\lambda - \lambda \sum_{i,j} A(N,S,E,W). \tag{9}$$

Here, N, the number of pixels in the autoradiograph, does not include pixels outside the region of support of the image as these lead to an estimate of λ based partly and incorrectly on the "flat" parts of the image field.

The insertion of (9) into (7) results in the following restatement of the minimization problem:

$$\begin{aligned} \hat{\lambda} &= \arg\min_{\lambda}\{\lambda \sum_{i,j} \left[(f_{ij}-N_{ij})^2 + (f_{ij}-S_{ij})^2 + (f_{ij}-E_{ij})^2 + (f_{ij}-W_{ij})^2\right] \\ &\qquad -\frac{N}{2}\log\lambda - \lambda \sum_{i,j} A(N_{ij}, S_{ij}, E_{ij}, W_{ij})\} \\ &= \arg\min_{\lambda}\{-\frac{N}{2}\log\lambda + \lambda \sum_{i,j} G_{ij}\}, \end{aligned} \tag{10}$$

where the term G_{ij} may, after some algebra, be identified as

$$G_{ij} = \frac{1}{4}\left[(2f_{ij} - N_{ij} - S_{ij}) + (2f_{ij} - E_{ij} - W_{ij})\right]^2. \tag{11}$$

Note that in (11), the two terms in (parentheses) may be identified as discrete second derivatives in the horizontal and vertical directions, respectively. We identify the latter term in (10) as the discrete version of the number:

$$G \equiv \sum_{i,j} G_{ij} \to \frac{1}{4} \iint \left(\nabla^2 f(x,y)\right)^2 dxdy. \tag{12}$$

The minimization problem has been further reduced to

$$\hat{\lambda} = \arg\min_{\lambda}\{-\frac{N}{2}\log\lambda + \lambda G\} = \arg\min_{\lambda}\{PL(\lambda)\}$$

which may be solved via $\frac{dPL(\lambda)}{d\lambda} = 0$ to get the solution

$$\hat{\lambda} = N/2G. \tag{13}$$

Equation (13) together with the definition in (12) is our solution. Note that since

$$\frac{d^2PL}{d\lambda^2} = \frac{N}{2\lambda^2} > 0$$

our solution for $\hat{\lambda}$ is indeed a minimum, and that since N and G are clearly nonnegative, $\hat{\lambda}$ is also nonnegative as desired.

Intuitively, $\hat{\lambda}$ should be dependent upon the amount of noise (total bin counts) in the projection data. Yet the value of $\hat{\lambda}$ was derived from noise*less* training data. The noise level does enter, however, through the scaling of the autoradiograph values $\mathbf{f}$, which are the means of Poisson densities for γ-emission at each location. With proper scaling of $\mathbf{f}$ to account for the typical number of bin counts, λ depends on noise after all. It is easy to show that $\mathbf{f}$ should be scaled according to

$$f_{i,j} = \frac{K}{\sum_{t,\theta} H_{t,\theta,i,j}} f_{i,j}^{\mathrm{norm}}, \tag{14}$$

where K equals the desired total bin counts, $H_{t,\theta,i,j}$ is the probability of a count at (t,θ) from pixel (i,j), and f^{norm} is a version of the autorad normalized so that $\sum_{i,j} f_{i,j}^{\mathrm{norm}} = 1$. In (13), $\hat{\lambda}$ scales with the inverse square of the count level, i.e. more counts means less smoothing as expected. The value of $\hat{\lambda}$ also depends on geometrical scaling; if $\mathbf{f}$ is scaled by factor Δ (in 2-D), then $\hat{\lambda} \rightarrow \hat{\lambda}\Delta^4$. In summary, one first inserts noise level and geometrical information via (14) to obtain a scaled autoradiograph $f_{i,j}$ and inserts this into (13) to obtain $\hat{\lambda}$.

5. Numerical Studies

In order to test the performance of our optimal $\hat{\lambda}$ for the membrane prior, we created 40 noisy projection data sets of the phantom in Fig. 1(a). The count level, approx. 730,000 counts, was equivalent to that of a typical SPECT brain scan on a per-pixel basis. For each projection data set, reconstructions using the membrane prior were performed for nine values $\lambda = \lambda_0 \times 2^n, n = 0, \pm 1, \pm 2 \pm 3 \pm 4$, thus yielding 360 reconstructions ($\lambda_0 = 0.02$).

Figure 3(a) shows a plot of root mean-squared error (RMSE) vs. λ for one projection data set, and indicates the position of the value for $\hat{\lambda}$ obtained by training on the same phantom. As seen, the derived value is not far from the one that minimizes RMSE. Figures 3 (b) and (c) show the reconstructions using the λ that minimizes RMSE and the $\hat{\lambda}$ obtained from maximum pseudolikelihood (MPL), respectively.

Figure 4 shows bias-variance variation with λ. The definitions for regional bias and standard deviation are: $b_R = \sqrt{\sum_{i,j \in \mathrm{ROI}} b_{i,j}^2}$, where $b_{i,j}$ is the pointwise bias at (i, j), and $s_R = \sqrt{\sum_{i,j \in \mathrm{ROI}} \sigma_{i,j}^2}$, where $\sigma_{i,j}^2$ is the pointwise variance at (i, j). Figure 5 shows bias variance results for membrane reconstructions. Here, bias and variance were calculated for a random variable equal to the number of counts in an ROI.

Table 1. summarizes a few trials used to anecdotally explore issues of generalization. Each row displays $\hat{\lambda}$ and also λ_{RMSE}, the value that minimizes, for the phantoms in Fig. 1. Phantom A is a sagittal autorad primate phantom. Since phantoms B and C are transverse slice autorads, from a single primate, spaced at 5 mm, we expect the parameter values to be similar for each, and they are. (Training and reconstruction were done on only central regions of these phantoms.) By contrast, we expect the Hoffman brain phantom D, which consists of flat regions with sharp edges, to have fairly different values. This is indeed the case. Phantom E consists of pixels sampled independently from a uniform $[0, 255]$ distribution. It has correspondingly different λ values. Note that there are general trends: “Rougher” (more edge content) phantoms have lower values of $\hat{\lambda}$ and λ_{RMSE}, as expected, and $\hat{\lambda}$ and λ_{RMSE} are correlated for a given phantom.

TABLE 1.

Phantom	$\hat{\lambda}$ obtained from MPL	λ for minimum RMSE
A	0.033	0.020
B	0.039	0.022
C	0.031	0.025
D	0.00167	0.00175
E	0.00228	0.00228

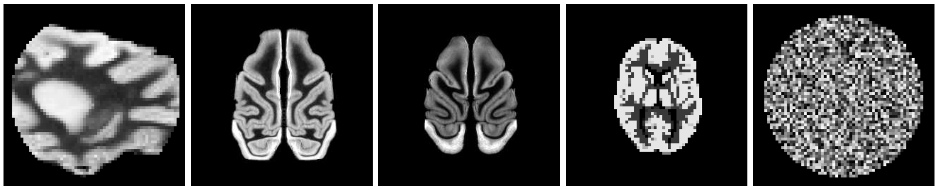
We evaluated RMSE's for WM and WP reconstructions of phantom A as a function of α and λ. These are shown in Fig. 6(a)(b). For these plots, iterations were terminated at minimum RMSE relative to the phantom. We also evaluated (6) numerically to explore the variation of WM pseudolikelihood with λ and α. The contour plot in Fig. 6 (c) shows a smooth unimodal surface over the range of α and λ used. This result implies that a descent procedure used to calculate $\hat{\lambda}$ and $\hat{\alpha}$ may work well.

6. Summary and Discussion

The essence of our approach has been to consider "ground-truth" exemplar phantoms as samples from an underlying Gibbs prior, and to fit hyperparameters of this prior model to the exemplars. Two unproven presumptions are that the optimal values thus derived will lead to reconstructions with an image quality not far from optimal, and that the training will generalize to objects that are similar to the training set. Our initial experiments reported here give optimistic anecdotal evidence for these presumptions. We note that the prospects for generalization is made somewhat more practical in that we show how to scale optimal parameters for count level and geometrical scaling.

References

1. V. E. Johnson, W. H. Wong, X. Hu, and C. T. Chen, "Image Restoration Using Gibbs Priors: Boundary Modeling, Treatment of Blurring, and Selection of Hyperparameter", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-13(5), pp. 413–425, May 1991.
2. Z. Zhou, R. M. Leahy, and E. U. Mumcuoglu, "A Comparative Study of the Effects of Using Anatomical Priors in PET Reconstruction", In *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference*, volume 3, pp. 1749–1753, November 1993.
3. S. Geman and C. Graffigne, "Markov random fields image models and their application to computer vision", In A. M. Gleason, editor, *Proc. of the Intl. Congress of Mathematicians 1986*, pp. 1496–1517, American Mathematical Society, Providence, 1987.
4. E. J. Hoffman, P. D. Cutler, W. M. Digby, and J. C. Mazziotta, "3-D Phantom To Simulate Cerebral Blood Flow and Metabolic Images for PET", *IEEE Trans. Nuclear Sci*, 37, pp. 616–620, 1990.
5. J. L. Lear, "Principles of Single and Multiple Radionuclide Autoradiography", In M. E. Phelps, J. C. Mazziotta, and H. R. Schelbert, editors, *Positron Emission Tomography and Autoradiography*, chapter 5, Raven Press, New York, NY, 1986.
6. A. Blake and A. Zisserman, *"Visual Reconstruction"*, Artificial Intelligence, MIT Press, Cambridge, MA, 1987.
7. S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-6(6), pp. 721–741, November 1984.
8. S. J. Lee, A. Rangarajan, and G. Gindi, "A Comparative Study of the Effects of Using Higher Order Mechanical Priors in SPECT Reconstruction", In *IEEE Nuclear Science Symposium and Medical Imaging Conference (in press)*, November 1994.
9. J. Besag, "Efficiency of Pseudo-likelihood estimation for simple Gaussian Fields", *Biometrika*, 64, pp. 616–618, 1977.



(a) phantom A (b) phantom B (c) phantom C (d) phantom D (e) phantom E

Figure 1. Phantoms used in the simulations. (a) Primate autoradiograph phantom obtained with the blood flow agent (^{99m}Tc-ECD). (b)(c) Primate autoradiograph phantoms obtained with the benzodiazepine neuroreceptor agent Iomazenil (^{123}I). (d) Hoffman brain phantom. (e) Circle phantom filled with uniform noise.

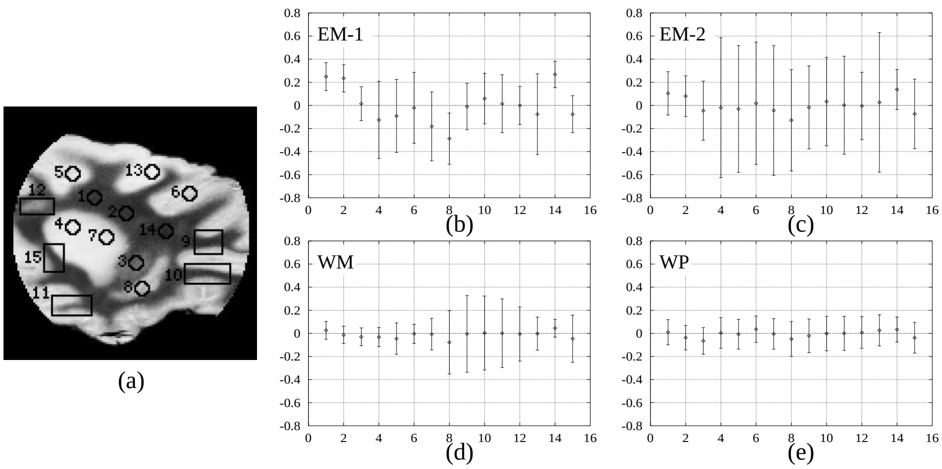


Figure 2. Regional bias-standard deviation plots for phantom A. (a) The phantom with region borders superposed. In (b)-(e), the abscissa indexes the region number, and plotted points with error bars indicate bias ± one standard deviation.

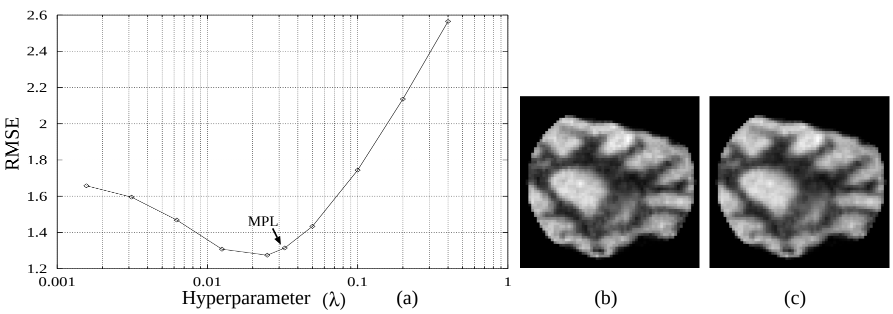


Figure 3. (a) Plot of RMSE vs. hyperparameter (λ) for the membrane reconstruction of phantom A. The arrow denotes λ obtained from MPL. (b) Reconstruction using λ for minimum RMSE. (c) Reconstruction using $\hat{\lambda}$ obtained from MPL.

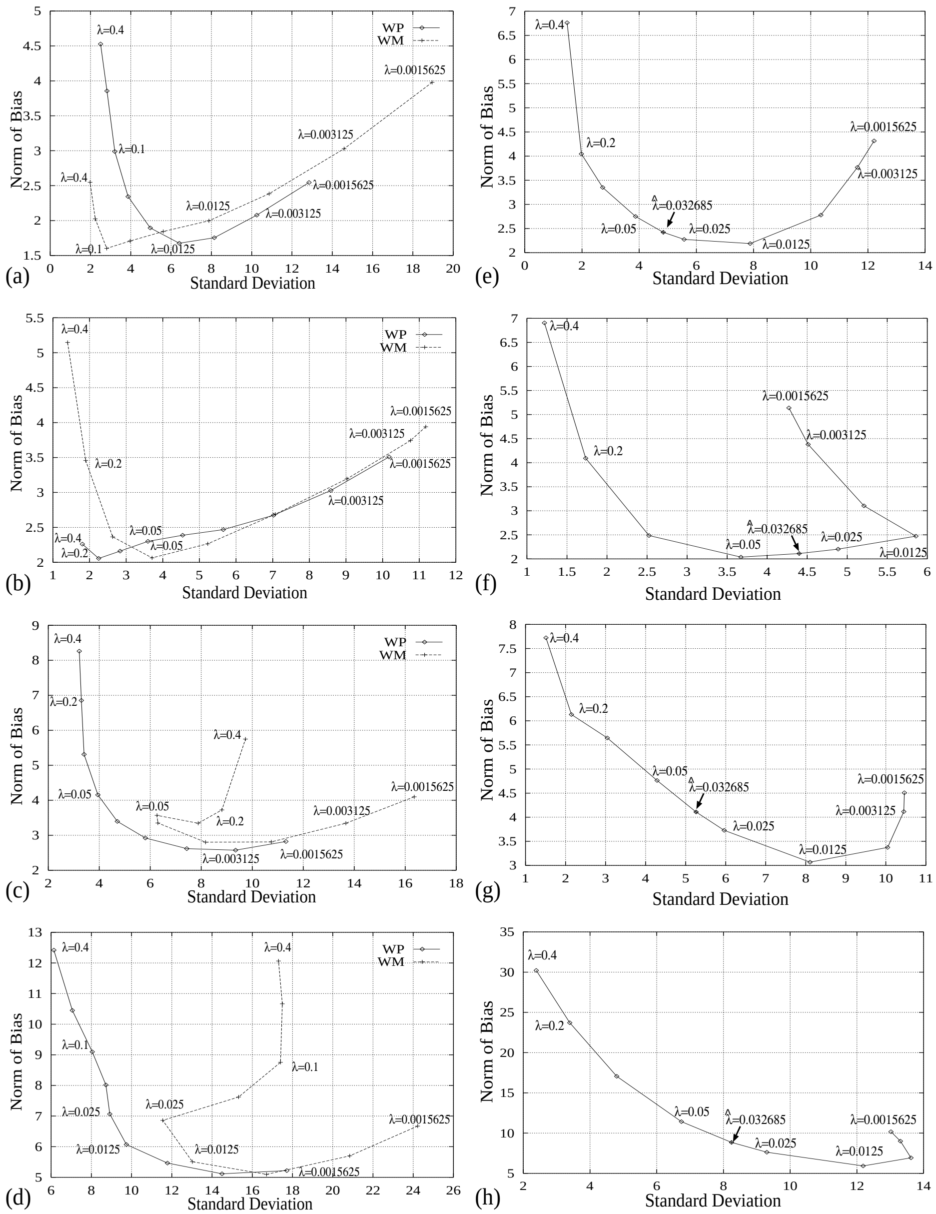


Figure 4. The effects of several different values for λ in WM, WP and membrane reconstructions. (a)-(d): Regional bias (b_R) standard deviation (s_R) plots for WM and WP reconstructions. (e)-(h): b_R and s_R plots for membrane reconstructions. The four representative regions in phantom A – (a)(e) high activity region (grey matter), (b)(f) low activity region (white matter), (c)(g) low contrast (ramplike) region, and (d)(h) high contrast region (anatomical boundary) – were measured. The arrows in (e)-(f) indicate the $\hat{\lambda}$ obtained from MPL.

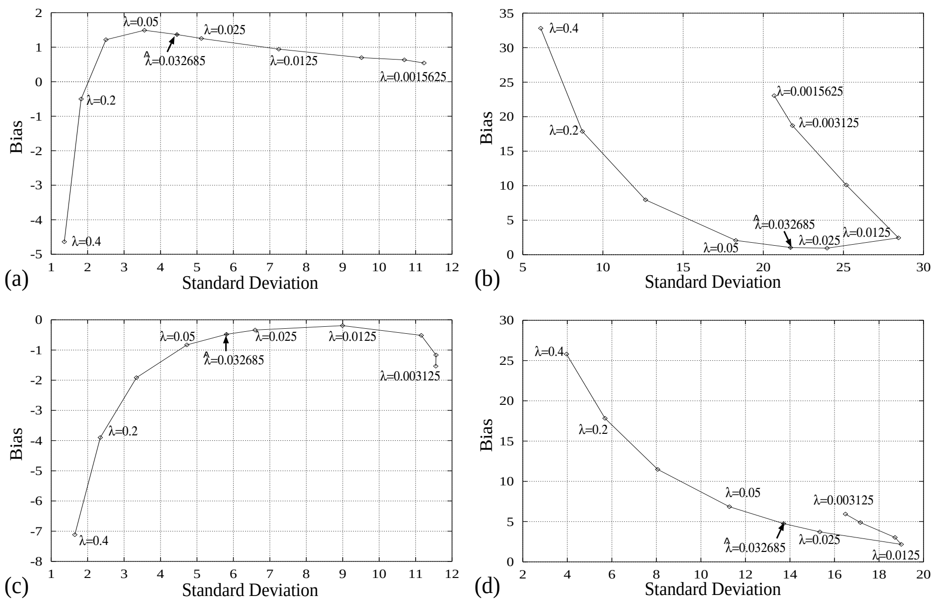


Figure 5. The effects of several different values for λ in membrane reconstructions. The same regions as in Fig. 4 were measured by using the standard definitions for regional bias and standard deviation. The arrows indicate the $\hat{\lambda}$ obtained from MPL.

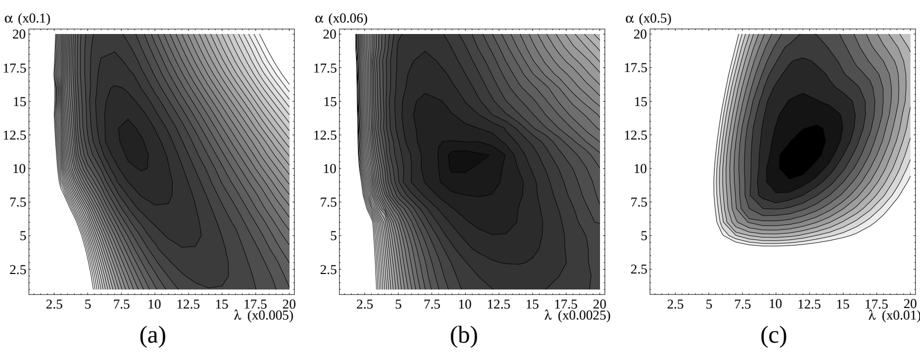


Figure 6. Contour plots for sensitivity of the hyperparameters λ (abscissa) and α (ordinate) in terms of RMSE and maximum pseudolikelihood. (a) Plot for WM. (b) Plot for WP. Each contour level in (a) and (b) indicates the equi-error within a range and the darkest areas correspond to the hyperparameters for the least RMSE's. (c) Plot for WM. Each contour level in (c) indicates the equi-pseudolikelihood within a range and the darkest area corresponds to the hyperparameters for the maximum pseudolikelihood.