Bayesian Tomographic 재구성에 있어서 Gibbs Smoothing Priors의 효과에 대한 비교연구

이수진
배재대학교 공학연구소 및 컴퓨터전자정보공학부

A Comparative Study of the Effects of Gibbs Smoothing Priors in Bayesian Tomographic Reconstruction

S. J. Lee
Department of Electronic Engineering, Paichai University

Abstract

Bayesian reconstruction methods for emission computed tomography have been a topic of interest in recent years, partly because they allow for the introduction of prior information into the reconstruction problem. Early formulations incorporated priors that imposed simple spatial smoothness constraints on the underlying object using Gibbs priors in the form of four-nearest or eight-nearest neighbors. While these types of priors, known as "membrane" priors, are useful as stabilizers in otherwise unstable ML-EM reconstructions, more sophisticated prior models are needed to model underlying source distributions more accurately. In this work, we investigate whether the "thin plate" model has advantages over the simple Gibbs smoothing priors mentioned above. To test and compare quantitative performance of the reconstruction algorithms, we use Monte Carlo noise trials and calculate bias and variance images of reconstruction estimates. The conclusion is that the thin plate prior outperforms the membrane prior in terms of bias and variance.

I. INTRODUCTION

The objective of emission computed tomography (ECT), such as single-photon emission computed tomography (SPECT) and positron emission tomography (PET), is to determine the 3-D distribution of radionuclide concentrations within the body using 2-D projectional views acquired at many different angles about the patient. Since ECT systems in practice, however, inherently involve noise and several physical factors that degrade the projection data, the quality of image reconstructions is extremely poor when the projection is modeled as a simple line integral, as in the case of deterministic methods. On the other hand, probabilistic reconstruction methods are attractive in that the statistical nature of emission processes can naturally be expressed in reconstruction algorithms by modeling the statistical character of the observed data.

Although well-known maximum likelihood (ML) approaches using the expectation maximization (EM) algorithm is attractive in that it can naturally express accurate system models of physical effects, and can accurately model the statistical character of the data, it is known to be unstable for the noise levels and numbers of measurements that characterize ECT. In contrast, maximum *a posteriori* (MAP) approaches in the context of a Bayesian framework overcome this instability by incorporating prior information while retaining the above advantages of ML-EM approaches. The prior information may also be regarded as reflecting assumptions about the spatial properties of the underlying source distribution. Many priors have been proposed; some of these implicitly model the underlying radionuclide density as globally smooth [1, 2], and others extend the smoothness model by allowing for spatial discontinuities [3, 4]. Discontinuity preservation is associated with a smoothing penalty that is a *non*quadratic function [3] of nearby pixel differences, whereas conventional (e.g. membrane) smoothing priors use quadratic penalties. The nonquadratic priors can exhibit good performance, but suffer difficulties in optimization and hyperparameter estimation. On the other hand, quadratic smoothing priors may not perform as well in edge regions as nonquadratic priors, but the simpler quadratic versions are more amenable to useful theoretical analyses even when couched in nonlinear Bayesian algorithms, and present an easier hyperparameter estimation problem [5]. In this paper we compare, by using Monte Carlo noise trials, the quantitative performance of the quadratic smoothing priors – the membrane (MM) and the thin plate (TP) [6].

II. BAYESIAN RECONSTRUCTION MODELS

The MAP approach in the context of a Bayesian framework is to estimate the underlying source field **f**

by maximizing the posterior probability, given as

$$\Pr(\mathbf{F}=\mathbf{f} \mid \mathbf{G}=\mathbf{g}) = \frac{\Pr(\mathbf{G}=\mathbf{g}|\mathbf{F}=\mathbf{f})\Pr(\mathbf{F}=\mathbf{f})}{\Pr(\mathbf{G}=\mathbf{g})},$$

where $\mathbf{f}$ and $\mathbf{g}$ are 2-D vector fields for the source intensities and projection data, respectively. Given the posterior distribution in (??), maximizing the posterior distribution is equivalent to minimizing $-\log$ of the posterior probability and the MAP estimation reduces to

$$\hat{\mathbf{f}} = \arg\min_{\mathbf{f}}[-\log\Pr(\mathbf{G}=\mathbf{g}|\mathbf{F}=\mathbf{f}) - \log\Pr(\mathbf{F}=\mathbf{f})],$$

where the two terms on the right side are the likelihood and the prior, respectively. For the likelihood, Poisson statistics are applied in a conventional way.

To incorporate the MM and TP priors in a MAP approach, we model the prior probability as a Gibbs distribution

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

where $\mathbf{f}$ is the 2-D source distribution comprising pixel components f_{ij}, E the associated Gibbs prior energy function, Z a normalization of no concern here, and λ the positive hyperparameter that weights the prior relative to the likelihood term. The energy $E(\mathbf{f})$, is given, for MM, as:

$$E_{MM}(\mathbf{f}) = \sum_{i,j}\left[f_h^2(i,j) + f_v^2(i,j)\right]$$

with discretizations of first partial derivatives given by $f_h(i,j) = f_{i,j+1} - f_{i,j}$ and $f_v(i,j) = f_{i+1,j} - f_{i,j}$. For the thin plate, the energy is

$$E_{TP}(\mathbf{f}) = \sum_{i,j}\left[f_{hh}^2(i,j) + 2f_{hv}^2(i,j) + f_{vv}^2(i,j)\right].$$

Here, $f_{vv}(i,j)$ and $f_{hh}(i,j)$ denote the discrete second partial derivatives of the source distribution in the vertical and horizontal directions, respectively, and $f_{hv}(i,j)$ is the second partial cross derivative. Our choices for discretization of the derivatives are:

$$\begin{aligned} f_{hh}(i,j) &= f_{i,j+1} - 2f_{i,j} + f_{i,j-1} \\ f_{vv}(i,j) &= f_{i+1,j} - 2f_{i,j} + f_{i-1,j} \\ f_{hv}(i,j) &= f_{i+1,j+1} - f_{i+1,j} - f_{i,j+1} + f_{i,j}. \end{aligned}$$

Since the noise properties of the MAP reconstruction estimate depend on the algorithm, we tried two algorithms. One was the iterative MAP-EM OSL (one-step-late) algorithm derived by Green [7]. The OSL algorithm is not derivable directly from a MAP principle, but can be shown to converge to the MAP solution if it converges at all. It has the advantage of simplicity (it is a simple modification of ML-EM) but does not always converge. (The OSL simulations reported here were results of convergent iterations.) The OSL algorithm is given by:

$$\hat{f}_{ij}^{k+1} = \frac{\hat{f}_{ij}^k}{\sum_{t\theta}\mathcal{H}_{t\theta,ij} + \lambda\left.\frac{\partial E(\mathbf{f})}{\partial f_{ij}}\right|_{f_{ij}=\hat{f}_{ij}^k}} \sum_{t\theta}\frac{\mathcal{H}_{t\theta,ij}g_{t\theta}}{\sum_{kl}\mathcal{H}_{t\theta,kl}\hat{f}_{kl}^k},$$

where $\hat{f}_{ij}^k$ is the object estimate at location (i,j) and iteration k, $g_{t\theta}$ the number of detected counts in the detector bin indexed by t at angle θ, $\mathcal{H}_{t\theta,ij}$ the probability that a photon emitted from source location (i,j) hits detector bin t at angle θ, and E stands for the energy of either prior.

We also used a version of ICM (iterated conditional modes), that avoids the convergence problems of OSL, but takes longer to stabilize. A derivation that follows the one in [8] yields the following update equations:

$$\hat{f}_{ij}^{k+1} = \frac{-(\sum_{t\theta}\mathcal{H}_{t\theta,ij} - 2\lambda X_3) + \sqrt{(\sum_{t\theta}\mathcal{H}_{t\theta,ij} - 2\lambda X_3)^2 + 8\lambda X_2 X_1}}{4\lambda X_2},$$

where

$$X_1 \stackrel{\text{def}}{=} \sum_{t\theta} g_{t\theta}\frac{\mathcal{H}_{t\theta,ij}\hat{f}_{ij}^k}{\sum_{kl}\mathcal{H}_{t\theta,kl}\hat{f}_{kl}^k},$$

and, for the MM prior,

$$\begin{aligned} X_2 &= 4 \\ X_3 &= \hat{f}_{i-1,j}^{k+1} + \hat{f}_{i,j-1}^{k+1} + \hat{f}_{i,j+1}^{k} + \hat{f}_{i+1,j}^{k}, \end{aligned}$$

while for the TP prior,

$$\begin{aligned} X_2 &= 20 \\ X_3 &= 8(\hat{f}_{i-1,j}^{k+1} + \hat{f}_{i,j-1}^{k+1} + \hat{f}_{i,j+1}^{k} + \hat{f}_{i+1,j}^{k}) \\ &\quad -2(\hat{f}_{i-1,j-1}^{k+1} + \hat{f}_{i-1,j+1}^{k} + \hat{f}_{i+1,j-1}^{k} + \hat{f}_{i+1,j+1}^{k}) \\ &\quad -(\hat{f}_{i-2,j}^{k+1} + \hat{f}_{i,j-2}^{k+1} + \hat{f}_{i,j+2}^{k} + \hat{f}_{i+2,j}^{k}). \end{aligned}$$

Note that a ICM uses a raster-scan update in which each pixel is replaced as soon as it is updated. The superscripting in the expressions for X_2 and X_3 reflects this.

III. SIMULATION PROCEDURE

Figure 1 shows the 2-D (64 × 64) phantoms, A, B and C. The mathematical phantom A is designed to illustrate the bias advantage of the thin plate, particularly in cold regions delimited by soft edges. Phantom B comprises a constant background region with a blobby hot spot and blobby cold spot. The both cold and hot regions have 10 pixel diameter, and their contrast relative to the base circle (53 pixel diameter) is 70%. Figure 1(c) shows a realistic [5] rCBF phantom C obtained from primate autoradiograph [9] of the SPECT agent Tc-99m ECD. In particular, note the variety of edge structures in phantom C.

For a given phantom and noise level, we generated 50 Monte Carlo noise trials by adding independent realizations of Poisson noise to the noiseless projection data. Reconstructions were performed as described

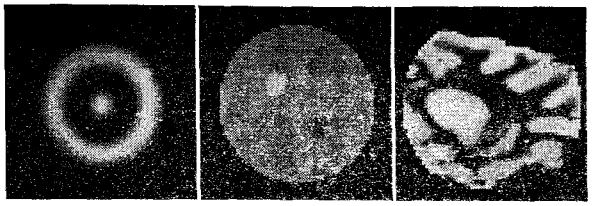

Fig. 1 64 × 64 phantoms used in the simulations. (a) phantom A (b) phantom B (c) phantom C, primate autoradiograph phantom obtained with the blood flow agent (Tc-99m ECD).

below, and the sample bias and variance of the 50 reconstructions computed at each point to derive bias and variance images.

Since ML-EM diverges in rmse, and we need to fix stopping criteria with which to conduct comparisons. (This is qualitatively equivalent to defining an optimal λ for MM and TP.) For a given phantom and count level, we choose two stopping criteria and designate the resulting reconstructions as EM-1 and EM-2, respectively. EM-1 is chosen by observing the iteration number at which reconstructions minimize rmse (this number is quite stable); hence the strategy for EM-1 is also a minimal rmse criterion as for TP and MM. Unlike the stable (after transients) MM and TP reconstructions, ML-EM bias and variance continually trade off as iterations proceed, and EM-1 yields only one choice along the bias-variance curve. We thus included a second (EM-2) criterion based on the simple heuristic of choosing, for given phantom/count level, that iteration number that optimizes qualitative resemblance of reconstruction and phantom, based on our own subjective impression. (This number was surprisingly stable.)

Since the ICM algorithm and (stable) OSL algorithm are iterative, we needed to choose a sufficient number of iterations after which the change in reconstruction was negligible. Thus, iteration number is removed as a parameter in comparisons involving MM and TP.

IV. RESULTS

To get increased dynamic range in the displays, we compute the standard deviation (STD) image ($= \sqrt{\text{variance}}$) at each point; intensity (lighter means greater) codes the positive STD value. Bias images are bipolar, with a value of zero displayed as an intermediate grey, and with darker/lighter regions corresponding to negative/positive bias. A given figure comprises images displayed with same grey scale to allow visual comparisons.

Anecdotal reconstructions for each of the four estimators (EM-1, EM-2, MAP-MM, and MAP-TP) are shown for phantom A (Fig. 2), phantom B (Fig. 3) and phantom C (Fig. 4).

For the bias/variance results for Phantom A, shown in Fig. 5, we used the OSL algorithm for MAP

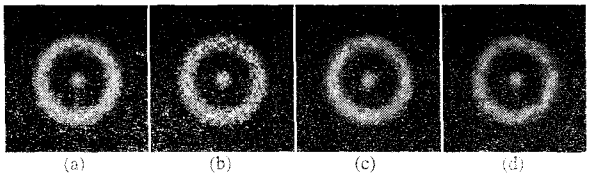

Fig. 2 Anecdotal reconstructions for phantom A. (a) EM-1 (b) EM-2 (c) OSL-MM (d) OSL-TP

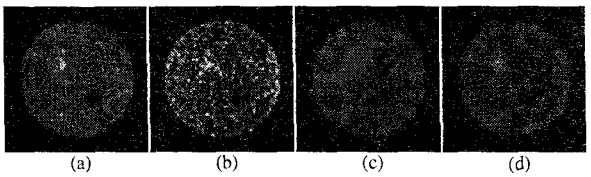

Fig. 3 Anecdotal reconstructions for phantom B. (a) EM-1 (b) EM-2 (c) ICM-MM (d) ICM-TP

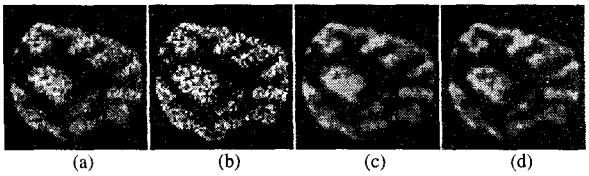

Fig. 4 Anecdotal reconstructions for phantom C. (a) EM-1 (b) EM-2 (c) OSL-MM (d) OSL-TP

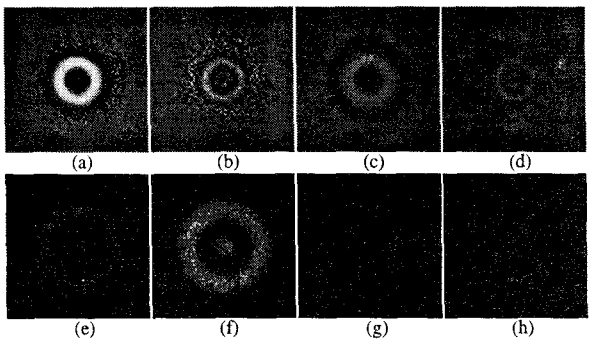

Fig. 5 Pointwise bias-STD images for phantom A. (a) bias EM-1 (b) bias EM-2 (c) bias OSL-MM (d) bias OSL-TP (e) STD EM-1 (f) STD EM-2 (g) STD OSL-MM (h) STD OSL-TP

estimates. We used a noise level corresponding to 500K counts, 18 and 80 iterations for EM-1 and EM-2, respectively, and 200 iterations for both the OSL-MM and OSL-TP algorithms. Values of $\lambda = 0.37$ and 0.12 were computed for OSL-MM and OSL-TP, respectively. Comparison of Figs. 5(a)(e) with Figs. 5(b)(f) shows the usual bias/variance tradeoff inherent in ML-EM. Fewer iterations (Figs. 5(a)(e)) lead to lower variance but high bias, and the opposite is true for the larger number of iterations used in Figs. 5(b)(f). Note that the STD images (Figs. 5(e)(f)) resemble the phantom itself, a well known behavior of ML-EM, and that the bias is negative in high-count regions, and positive in low-count regions. The MAP

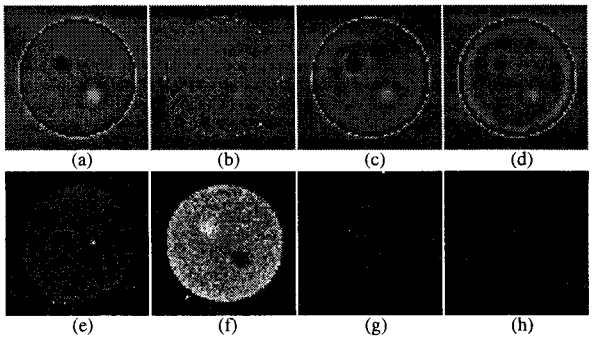

Fig. 6 Pointwise bias-STD images for phantom B. (a) bias EM-1 (b) bias EM-2 (c) bias ICM-MM (d) bias ICM-TP (e) STD EM-1 (f) STD EM-2 (g) STD ICM-MM (h) STD ICM-TP

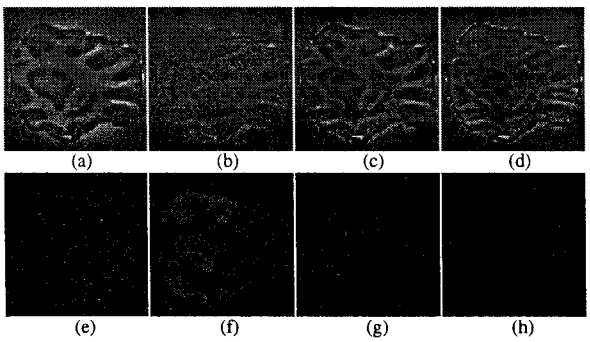

Fig. 7 Pointwise bias-STD images for phantom C. (a) bias EM-1 (b) bias EM-2 (c) bias OSL-MM (d) bias OSL-TP (e) STD EM-1 (f) STD EM-2 (g) STD OSL-MM (h) STD OSL-TP

results, Figs. 5(c)(g) and (d)(h) again show the TP advantage. The bias, especially in the cold region, is considerably reduced relative to MM, and the variance is increases slightly.

For the bias/variance results for Phantom B, shown in Fig. 6, we used the ICM algorithm for MAP estimates. We used a noise level corresponding to 300K counts, 7 and 40 iterations for EM-1 and EM-2, respectively, and 400 iterations for both the ICM-MM and ICM-TP algorithms. Values of $\lambda = 3.1$ and 3.7 were computed for ICM-MM and ICM-TP, respectively. Comparison of Figs. 6(a)(e) with Figs. 6(b)(f) again shows the usual bias/variance tradeoff inherent in ML-EM. The MAP results, Figs. 6(c)(g) and (d)(h) again show the TP advantage. For each lesion, the typical undershoot/overshoot bias effects seen clearly in Fig. 6(c) are lessened by the TP prior. The variance for both ICM-MM and ICM-TP estimates is uniform and low.

Bias/variance results, shown for the complex phantom C in Fig. 7, illustrate the same general behaviors. Here MAP estimates were computed using the OSL algorithm. We used a noise level corresponding to 300K counts, 18 and 60 iterations for EM-1 and EM-2, respectively, and 200 iterations for both the OSL-MM and OSL-TP algorithms. Values of $\lambda = 0.50$ and 0.27 were computed for OSL-MM and OSL-TP, respectively. The same bias-STD tradeoffs are shown for EM-1 (Figs. 7(a)(e)) and EM-2 (Figs. 7(b)(f)). The MAP results, Figs. 7(c)(g) and (d)(h) again show the TP advantage.

V. CONCLUSION

It is generally difficult to draw firm conclusions regarding smoothing functionals since there are so many variables one might optimize. For example, viewed as an Markov random field, the TP prior corresponds simply to a slightly larger (quadratic) neighborhood with different weights relative to a conventional smoothing MM prior. One could, for example, attempt to optimize over all possible weights in some large neighborhood, but such an exhaustive study would lead to conclusions that are highly object dependent. Here, we have obtained a perhaps more robust result: weights corresponding to second differences in smoothing functionals do have a bias advantage for a variety of meaningful objects and reconstruction procedures.

VI. REFERENCES

[1] L. Kaufman, "Maximum likelihood, least squares, and penalized least squares for PET", *IEEE Trans. Med. Imaging*, MI-12(2), pp. 200–214, June 1993.

[2] S.J. Lee, I.T. Hsiao, and G.R. Gindi, "The Thin Plate as a Regularizer in Bayesian SPECT Reconstruction", *IEEE Trans. Nuclear Science* (to appear), June 1997.

[3] S. J. Lee, A. Rangarajan, and G. Gindi, "Bayesian Image Reconstruction in SPECT Using Higher Order Mechanical Models as Priors", *IEEE Trans. on Medical Imaging*, MI-14(4), pp. 669–680, Dec. 1995.

[4] T. Hebert and R. Leahy, "A Generalized EM Algorithm for 3-D Bayesian Reconstruction for Poisson Data Using Gibbs Priors", *IEEE Trans. on Medical Imaging*, MI-8(2), pp. 194–202, June 1989.

[5] S. J. Lee, G. R. Gindi, I. G. Zubal, and A. Rangarajan, "Using Ground-Truth Data to Design Priors in Bayesian SPECT Reconstruction", In Y. Bizais, C. Barillot, and R. D. Paola, editors, *Information Processing in Medical Imaging*, pp. 27–38, Kluwer Academic Publishers, 1995.

[6] A. Blake and A. Zisserman, *Visual Reconstruction*, Artificial Intelligence, MIT Press, Cambridge, MA, 1987.

[7] P. J. Green, "Bayesian Reconstructions from Emission Tomography Data Using a Modified EM Algorithm", *IEEE Trans. on Medical Imaging*, MI-9(1), pp. 84–93, March 1990.

[8] G. Gindi, A. Rangarajan, M. Lee, P. J. Hong, and G. Zubal, "Bayesian Reconstruction for Emission Tomography via Deterministic Annealing", In H. Barrett and A. Gmitro, editors, *Information Processing in Medical Imaging*, pp. 322–338, Springer–Verlag, 1993.

[9] 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.