

# PENALIZED MAXIMUM LIKELIHOOD IMAGE RECONSTRUCTION WITH MIN-MAX INCORPORATION OF NOISY SIDE INFORMATION

*Robinson Piramuthu and Alfred O. Hero III*

4229 EECS, University of Michigan, Ann Arbor, MI 48109

## ABSTRACT

A method for incorporating anatomical MRI boundary side information into penalized maximum likelihood (PML) Emission Computed Tomography (ECT) image reconstructions using a set of averaged Gibbs weights was proposed in [6]. A quadratic penalty based on Gibbs weights was used to enforce smoothness constraints everywhere in the image except across the estimated boundary of ROI. In this methodology, a limiting form of the posterior distribution of the MRI boundary parameters was used to average the Gibbs weights obtained as in [10]. There is an improvement in performance over the method proposed in [10], when the variance of boundary estimates from the MRI data becomes significant. Here, we present the empirical performance analysis of the proposed method of averaged Gibbs weights.

## 1. INTRODUCTION

ECT plays an important role in functional imaging, where the functional information of living organs are studied. This is done by estimating the uptake of radio-tracer in the region of interest (ROI). However, ECT images have poor resolution owing to trade-offs between detection sensitivity and collimator resolution, limited photon rates and fluctuations in photon statistics. Better estimates of the ROI can be obtained using high resolution imaging modalities such as magnetic resonance imaging (MRI) or X-ray computed tomography (CT). In order to improve the resolution and to get a reliable estimate of uptake in ROI, many researchers (for example [5]) have proposed to use priors from MRI or CT images.

The method proposed in [10, 11] incorporates anatomical MRI boundary information into penalized likelihood (PL) ECT image reconstructions. The PL uses a quadratic penalty term based on Gibbs weights obtained from the MRI prior. This smoothing penalty is spatially-variant and thus improves the resolution, especially across the boundary. The Gibbs weights thus obtained do not make any corrections for the noise in boundary estimates. In other words, it uses

the MRI prior “blindly”.

In contrast, the method in [6] uses an asymptotic marginalization [8] to average the PL over a normal approximation to the posterior distribution of the side information. This underemphasizes the side information in regions where its variance is high. A theorem from [6] gives the asymptotic posterior distribution of the spline parameters to be a multivariate Gaussian centered at the boundary estimate  $\hat{\theta}$  and the inverse of observed Fisher information as the covariance matrix. The penalty thus obtained is non-polynomial and is very computationally intensive. A simple penalty function with averaging of weight maps is derived using a lower bound on the resultant marginal by applying the Jensen’s inequality.

In this paper, we show that this method of [6] can be interpreted as a min-max optimal principle to average the side information. We provide a quantitative comparison between performance of method [6] using empirical Fisher information matrix versus the expected Fisher information matrix. We illustrate and compare the aforementioned methods for both hot and cold spot uptake estimations.

## 2. PRELIMINARY DETAILS

Let  $\mathbf{Y}_{\text{MRI}}$  represent the noisy NMR spin density image. We model the MRI system as linear, spatially shift invariant, with a symmetric Gaussian 2D impulse response and with additive white Gaussian noise of variance  $\sigma_n^2$  [2]. We parametrize the closed boundary by a periodic B-spline model. The boundary ( $\hat{\theta}$ ) is estimated by non-linear maximum likelihood estimation. This method of boundary extraction was shown [11] to be approximately unbiased over a wide range of SNR and was shown to achieve the CR lower bound on attainable estimator variance for reasonable noise levels ( $\sigma_n$  less than 15% of edge contrast).

Let  $\mathbf{Y}_{\text{E}}$  be the noisy (Poisson) ECT dataset of detected photons and  $\lambda$  be the unknown emission distribution. Then the penalized likelihood is given by

$$J(\lambda) = \ln f(\mathbf{Y}_{\text{E}}|\lambda, \hat{\theta}) - \beta R(\lambda; \hat{\theta}) \quad (1)$$

The penalty  $R(\lambda; \hat{\theta})$  is quadratic and is given by

$$R(\lambda; \hat{\theta}) = \sum_{i,j} w_{ij}(\hat{\theta})(\lambda_i - \lambda_j)^2 \quad (2)$$

---

This work was supported by National Cancer Institute, ROI-CA-54362-02.

where  $w_{ij}$  are the Gibbs weights which are obtained through the binary weight mapping scheme proposed in [10].

### 3. MIN-MAX THEORY

Let us assume that  $\lambda$  is a Gibbs random field. Then the probability distribution function (pdf) of  $\lambda$  is of the form

$$f_{\theta}(\lambda) = \alpha \cdot \exp(-\beta R(\lambda, \theta)) \quad (3)$$

where  $\alpha$  is a normalization constant. When  $\theta$  is known exactly, the posterior mode (MAP) estimate of  $\lambda$  is a function of  $\mathbf{Y}_E$  and  $\mathbf{Y}_{\text{MRI}}$  which achieves the minimum

$$\min_{\hat{\lambda}} P_{\theta} \left( \left\| \hat{\lambda} - \lambda \right\| > \epsilon \right), \quad \text{for small } \epsilon.$$

When  $\theta$  is unknown, we can either substitute the extracted boundary  $\hat{\theta}$ , which is the maximum likelihood estimate of  $\theta$ , or use the following min-max criterion: choose  $\hat{\lambda}$  to minimize the maximum probability of error (*i.e.*)  $\hat{\lambda}$  achieves

$$\min_{\hat{\lambda}} \max_{\theta} P_{\theta} \left( \left\| \hat{\lambda} - \lambda \right\| > \epsilon \right), \quad \text{for small } \epsilon.$$

This min-max problem can be shown [1] to be equivalent to

$$\min_{\hat{\lambda}} \int_{\Theta} P_{\theta} \left( \left\| \hat{\lambda} - \lambda \right\| > \epsilon \right) \tilde{f}(\theta) d\theta$$

where  $\tilde{f}(\theta)$  is a function of  $\theta$ , called the *equalizer density* and can be derived as the solution to a related integral equation. Note that this objective function is equivalent to the marginal probability of error under a stochastic model for  $\theta$  having a pdf  $\tilde{f}(\theta)$ .

Assume that  $\mathbf{Y}_{\text{MRI}}$  and  $\mathbf{Y}_E$  are conditionally independent, given  $\lambda$ . We can then show using a factorization property that the min-max  $\hat{\lambda}$  maximizes a marginalized posterior

$$\begin{aligned} \ln \tilde{f}(\lambda | \mathbf{Y}_E, \mathbf{Y}_{\text{MRI}}) &= \ln f(\mathbf{Y}_E | \lambda) + (\text{a constant } c) + \\ &\ln \int_{\Theta} \left\{ \exp(-\beta R(\lambda; \theta)) \tilde{f}(\theta | \mathbf{Y}_{\text{MRI}}) d\theta \right\} \quad (4) \end{aligned}$$

where

$$\tilde{f}(\theta | \mathbf{Y}_{\text{MRI}}) = \frac{f_{\theta}(\mathbf{Y}_{\text{MRI}})}{\int f_{\theta}(\mathbf{Y}_{\text{MRI}}) \tilde{f}(\theta) d\theta}$$

We recognize (4) as a modified PML function for estimating  $\lambda$ . The penalty term in (4) remains convex in  $\lambda$ , but is non-linear in weights  $w_{ij}(\theta)$  and it requires a high dimensional integration for every candidate value  $\hat{\lambda}$  of  $\lambda$ . We can simplify this using the Jensen's inequality which provides a lower bound. By Jensen's inequality,

$$\begin{aligned} E_{\theta | \mathbf{Y}_{\text{MRI}}} [\exp \{-\beta R(\lambda; \theta)\}] &\geq \exp \{-\beta E_{\theta | \mathbf{Y}_{\text{MRI}}} [R(\lambda; \theta)]\} \\ \ln E_{\theta | \mathbf{Y}_{\text{MRI}}} [\exp \{-\beta R(\lambda; \theta)\}] &\geq -\beta \sum_{ij} \tilde{w}_{ij}(\hat{\theta}) (\lambda_i - \lambda_j)^2 \end{aligned}$$

where

$$\tilde{w}_{ij}(\hat{\theta}) = \int w_{ij}(\hat{\theta}) \tilde{f}(\theta | \mathbf{Y}_{\text{MRI}}) d\theta \quad (5)$$

are the smoothed weights. The smoothed weights are thus obtained from the unsmoothed weights  $w_{ij}(\theta)$  through an integral operator that is independent of  $\lambda$ . The modified PML is now

$$\hat{\lambda} = \arg \max_{\lambda} \left\{ \ln f(\mathbf{Y}_E | \lambda) - \beta \sum_{ij} \tilde{w}_{ij}(\hat{\theta}) (\lambda_i - \lambda_j)^2 \right\}$$

The next step is to compute  $\tilde{f}(\theta | \mathbf{Y}_{\text{MRI}})$ . The following theorem gives us an asymptotic form for the posterior distribution  $\tilde{f}(\theta | \mathbf{Y}_{\text{MRI}})$ .

### 4. ASYMPTOTIC MARGINALIZATION

The following limit theorem can be shown using techniques of [7] and [8].

**Theorem 1** *Assume that  $f(\mathbf{Y}_{\text{MRI}} | \theta)$  is a smooth function of  $\theta$  in the sense of satisfying the regularity conditions [7] (p. 131) and that  $f(\theta)$  is a smooth function in the neighborhood of  $\hat{\theta}$ . Then the maximum likelihood estimator  $\hat{\theta}$  is an asymptotically consistent estimator of  $\theta \in \mathfrak{R}^p$*

$$E_{\theta} \left[ \|\hat{\theta} - \theta\|^2 \right] = \delta \rightarrow 0$$

and

$$\begin{aligned} f(\theta | \mathbf{Y}_{\text{MRI}}) &= \frac{|\hat{\mathbf{F}}_{\hat{\theta}}|^{1/2}}{(\sqrt{2\pi})^p} \cdot \exp \left\{ -\frac{1}{2} (\theta - \hat{\theta})^T \hat{\mathbf{F}}_{\hat{\theta}} (\theta - \hat{\theta}) \right\} \\ &\cdot (1 + O(\delta)) \quad (6) \end{aligned}$$

where  $\hat{\mathbf{F}}_{\hat{\theta}}$  is the observed [3] Fisher information matrix

$$\hat{\mathbf{F}}_{\hat{\theta}} = -\nabla_{\hat{\theta}}^2 \ln f(\theta | \mathbf{Y}_{\text{MRI}}).$$

We observe that the mean of this asymptotic distribution is the maximum likelihood estimate of  $\theta$  and that the asymptotic posterior distribution  $f(\theta | \mathbf{Y}_{\text{MRI}})$  is independent of the explicit form of the prior density  $f(\theta)$ . When the factorization identity is valid, (6) can be shown to be identical to the profile posterior approximation proposed in [9].

### 5. RESULTS

We compare the reconstructed images using the ‘‘unsmoothed weights’’  $w_{ij}$ , the averaged weights  $\tilde{w}_{ij}$  and the ‘‘ideal’’ weights extracted from noiseless MRI image. The Gibbs weights  $w_{ij}$  were obtained through binary weight mapping described in [10], which assigns unity weights to neighbor pixel pairs and zero weights otherwise. Second order neighborhood

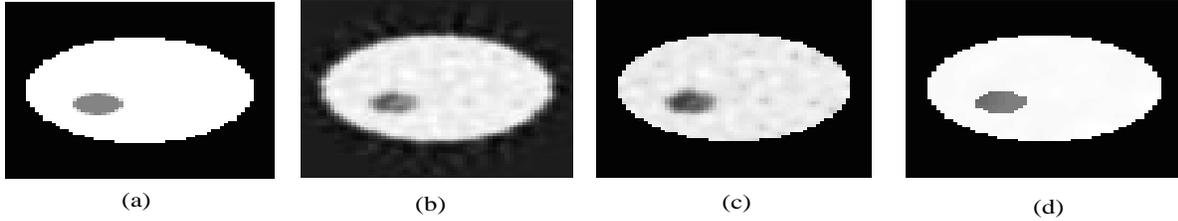


Figure 1: (a) True Emission phantom. (b) FBP. (c) No MRI side information . (d) Unsmoothed weights ( $\beta = 1, \sigma_n = 0.36$ ).

relations were considered. A modified version of PML-SAGE3 [4] algorithm was used to maximize the PL objective.

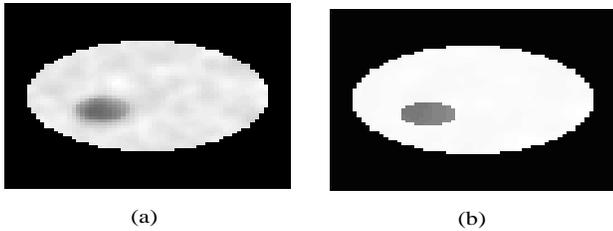


Figure 2: (a) Smoothed weights without leakage prevention boundary . (b) Smoothed weights with leakage prevention boundary ( $\beta = 1, \sigma_n = 0.36$ ).

The true cold-spot phantom contains a background ellipse with major and minor axes 27,19 pixels respectively and centered at (-10,-5). The ROI is represented by a cold spot ellipse with major and minor axes 6,4 pixels respectively. Identical contrast of 6 units (3 for hot spot) was used for both MRI and ECT images. The ROI was chosen to have an intensity of -6, 6 units respectively for MRI and ECT images (6, 9 for hot spot). The parameter  $\sigma_s$  of the symmetric Gaussian spatial blurring function used to model the MRI scanner was 0.75 (15% of average radius of ROI). A 16-knot B-spline model was used to parametrize the boundary. The ECT data with a total of a million counts was sampled by a parallel beam tomograph corresponding to PET projections over 64 radial bins, and 60 equispaced projection angles over  $180^\circ$ . Poisson noise was added and 15% random coincidences were added.

Figure (1) shows the true cold-spot phantom and the reconstructed image using FBP, no side information and unsmoothed weights for  $\sigma_n = 0.36$  (6% of contrast). Figure (2) shows the corresponding reconstruction for smoothed weights. Since the smoothed weights introduce neighbors to be connected across the boundary (hence leakage, as in (a)), we hard limited the weights to zero at the boundary which surrounds the pixels that are likely to be atleast 50% in the interior. This preserves the overall intensity level within

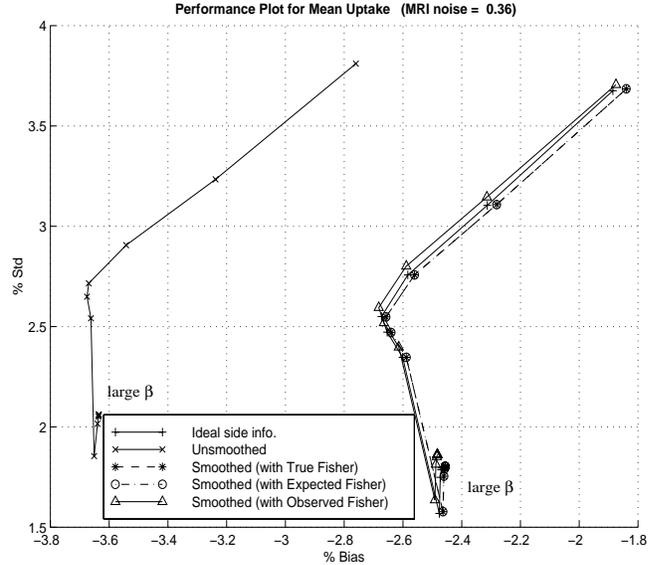


Figure 3: % Std. Vs % Bias of mean uptake for various  $\beta \in (2^{-7}, 2^{15})$  for the cold spot phantom.

the ROI and lowers the bias of the uptake estimate. Figure (3) shows the bias-variance trade-off curve of the estimated mean uptake for  $\sigma_n = 0.36$  for the cold-spot phantom. This was obtained from 400 realizations of both simulated noisy MRI and PET data. We observe that the smoothed weights outperform the unsmoothed weights and the trade-off curve for smoothed weights is close to the corresponding curve for ideal side weights (no MRI noise). The plots are shown for 3 kinds of Fisher information used, namely (a) true (b) expected and (c) observed; (a) was obtained using the true boundary and (b),(c) were obtained using the extracted boundary. We do not see much difference in performance using these choices. Choice (a) is not practical since we will never know the true boundary. Choice (b) can be shown to have fewer computations than choice (c). So, in practice, we can use the expected Fisher information and get reasonable results. Plots (5), (4) show % bias and %std for various MRI noise ( $\sigma_n$ ). We also present the bias-variance trade-off plot for the hot-spot phantom as in figure (6). We get similar results.

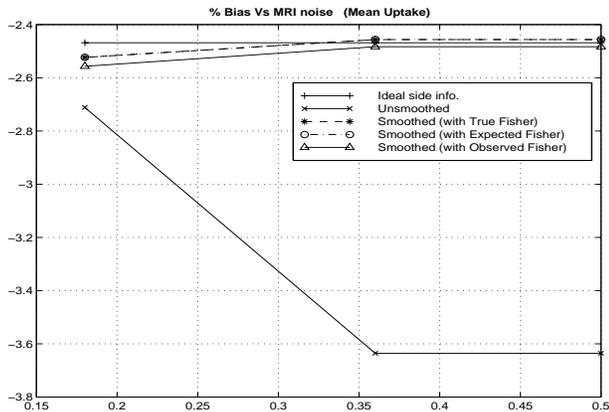


Figure 4: % Bias of mean uptake Vs MRI noise  $\sigma_n$  (large  $\beta$ ) for the cold spot phantom.

## 6. CONCLUDING REMARKS

We have shown that the variance corrected penalty obtained using the averaged Gibbs weights by asymptotic marginalization performs well for a reasonable range of MRI noise for cold spot phantom. Similar results were achieved for a hot spot phantom. We would like to extend this technique to 3-D PET reconstructions by parametrizing the boundary using a suitable model for the boundary.

## 7. REFERENCES

- [1] B. Baygun and A. O. Hero, "Optimal simultaneous detection and estimation under a false alarm constraint," *IEEE Trans. on Information Theory*, vol. 41, no. 3, pp. 688–703, May 1995.
- [2] W. A. Edelstein, "Intrinsic signal to noise ratio in NMR imaging," *Magnetic Resonance in Medicine*, 1986.
- [3] B. Efron and D. V. Hinkley, "Assessing the accuracy of the maximum likelihood estimator: observed versus expected Fisher information," *Biometrika*, vol. 65, no. 3, pp. 457–487, 1978.
- [4] J. A. Fessler and A. O. Hero, "Penalized maximum likelihood image reconstruction using space alternating generalized em algorithms," *IEEE Trans. on Image Processing*, vol. 4, no. 10, pp. 1417–1429, Oct. 1995.
- [5] G. Gindi, M. Lee, A. Rangarajan, and G. Zubal, "Bayesian reconstruction of functional images using anatomical information as priors," *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, , Dec. 1993.
- [6] A. O. Hero and R. Piramuthu, "A method for ECT image reconstruction with uncertain MRI side information using asymptotic marginalization," *Proc. of IEEE/EURASIP Workshop on Nonlinear Signal and Image Processing*, Sept. 1997.
- [7] P. J. Huber, *Robust Statistics*, Wiley, New York, 1981.

- [8] M. A. Tanner, *Tools for statistical inference; Methods for the exploration of posterior distributions and likelihood functions*, Springer-Verlag, New York, 1993.
- [9] L. Tierney and J. B. Kadane, "Accurate approximations for posterior moments and marginal densities," *JASA*, vol. 81, pp. 82–86, 1986.
- [10] S. R. Titus, A. O. Hero, and J. A. Fessler, "Improved penalized likelihood reconstruction of anatomically correlated emission data," *IEEE Int. Conf. on Image Processing*, vol. volume 2, , Laussane, Sep. 1996.
- [11] S. R. Titus, A. O. Hero, and J. A. Fessler, "Penalized likelihood emission image reconstruction with uncertain boundary information," *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, vol. volume 3, , Munich, April 1997.

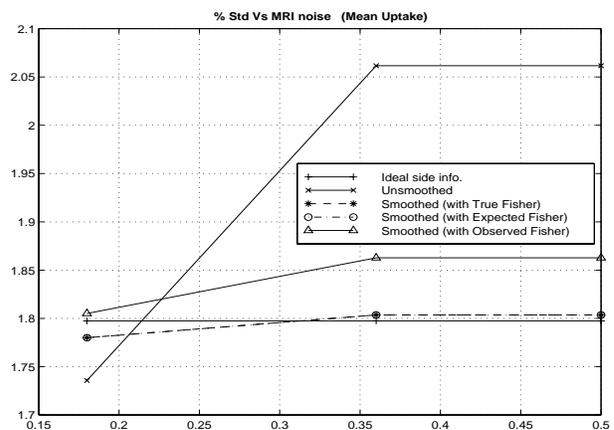


Figure 5: % Std. of mean uptake Vs MRI noise  $\sigma_n$  (large  $\beta$ ) for the cold spot phantom.

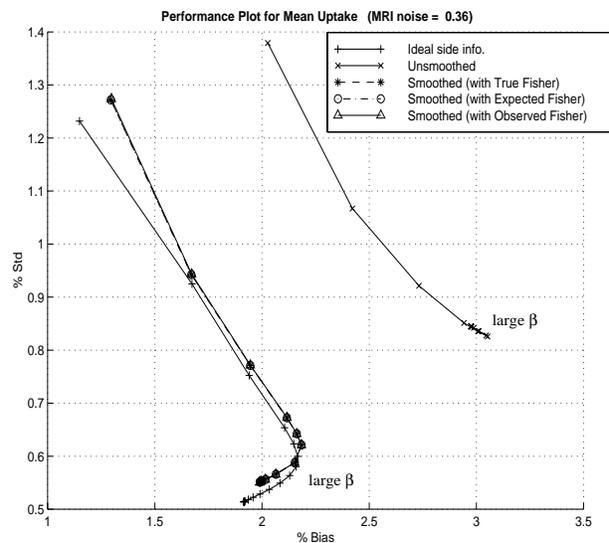


Figure 6: % Std. Vs % Bias of mean uptake for various  $\beta \in (2^{-7}, 2^{15})$  for the hot spot phantom.