A METHOD FOR ECT IMAGE RECONSTRUCTION WITH UNCERTAIN MRI SIDE INFORMATION USING ASYMPTOTIC MARGINALIZATION

Alfred O. Hero III and Robinson Piramuthum

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

ABSTRACT

In [1] a methodology for incorporating extracted MRI anatomical boundary information into penalized likelihood (PL) ECT image reconstructions and tracer uptake estimation was proposed. This methodology used quadratic penalty based on Gibbs weights which enforced smoothness constraints everywhere in the image except across the MRI-extracted boundary of the ROI. When high quality estimates of the anatomical boundary are available and MRI and ECT images are perfectly registered, the performance of this method was shown to be very close to that attainable using ideal side information, i.e. noiseless anatomical boundary estimates. However when the variance of the MRI-extracted boundary estimates becomes significant this penalty function method performs poorly. We give a modified Gibbs penalty function implemented with a set of averaged Gibbs weights, where the averaging is performed with respect to a limiting form of the posterior distribution of the MRI boundary parameters.

1. INTRODUCTION

Radio-tracer uptake estimation is an essential tool in medicine and biological sciences for evaluating metabolic function of living systems. Emission computed tomography (ECT) is very useful in this regard due to its ability to image in three dimensions. Critical to uptake estimation accuracy is a reliable estimate of the anatomical region of interest (ROI), e.g. a target organ within the body. While it is possible to estimate the ROI directly from the acquired radio-isotope image, better estimates can be obtained from higher resolution imaging systems, such as X-ray computed tomgraphy (CT) or magnetic resonance images (MRI), which are specifically adapted to imaging anatomy. Many researchers have suggested ways to incorporate such MRI or CT side information into ECT image reconstructions and tracer uptake estimates [2, 3, 4, 5, 6, 7]. Here we present a method for using noisy MRI side information in ECT which is based on asymptotic marginalization of the penalty likelihood method proposed in [8, 1].

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

The method of [9, 8, 1] performs ECT image reconstrion using a penalized (Poisson) likelihood function approach. A quadratic penalty based on Gibbs weights is introduced which enforces smoothness constraints everywhere in the image except across the MRI-extracted boundary of the ROI. When high quality estimates of the anatomical boundary are available and MRI and ECT images are perfectly registered, the performance of this method was shown [9] to be very close to that attainable using ideal side information, i.e. noiseless anatomical boundary estimates. However when the variance of the MRI-extracted boundary estimates becomes significant this penalty function method performs poorly. This poor performance is due in large part to the fact that the Gibbs weights do not account for estimation errors in the anatomical boundary estimates. In [10] a technique called "variance corrected weighting" was proposed which relied on a local Taylor series expansion of the Gibbs weights about the estimated spline parameters. This scheme applies only to weight assignments that are smooth functions of the spline parameters for which Taylor series can be applied. Furthermore, even for the smooth weight assignments presented in [10] the method is accurate only when the uncertainty in the extracted boundary estimates are less than a single MRI pixel width.

Here we derive a variance corrected weighting which applies globally to non-smooth weights. The method is based on marginalization [11, Sec 3.2.2] which averages the penalized likelihood over an asymptotic normal approximation to the posterior distribution of the side information. Jensen's inequality provides a lower bound on the resultant marginal which gives a simple penalty function requiring averaging of the weight maps. This averaging is performed with respect to the asymptotic distribution of the boundary estimates which is a multivariate Gaussian density centered at the boundary estimate $\{\hat{r}_\phi\}_{\phi \in [0,2\pi]}$ and with covariance matrix equal to the inverse (observed) Fisher information.

2. BACKGROUND

Let Y_E and Y_M be two random measurements corresponding to (Poisson) ECT projections and (Gaussian)

NMR spin density images, respectively. In [10] two different methodologies were proposed for extracting anatomical boundary estimates from noisy MRI images for the purpose of incorporation into ECT. One was based on non-linear maximum likelihood estimation of coefficients θ in a periodic B-spline model for the closed boundary. The other was based on direct estimation of the boundary defined by the radial distance function $r(\phi)$, $\phi \in [0, 2\pi]$ which specifies the (polar) coordinates (r, ϕ) of the boundary relative to an origin defined within the ROI. This latter method performed a polar to rectangular coordinate conversion, Canny or Marr-Hildreth edge detection, followed by smoothing of the extracted edge by a median filter. It was observed through simulations that both boundary extraction methdologies yielded boundary estimates which were approximately unbiased over a wide range of SNR. The Fisher information matrix was derived and it was observed that the methods also came close to the CR lower bound on attainable estimator variance for low to moderate additive noise levels (σ_n less than 15% of edge contrast).

The boundary estimates, which are denoted by common notation $\hat{\theta}$ for spline coefficient parameterization or for radial distance parameterization, are incorporated into ECT image reconstruction via maximizing the penalized likelihood (PL) objective over the ECT intensity distribution λ

$$J(\lambda) = \ln f(Y_E|\lambda, \hat{\theta}) - \beta R(\lambda; \hat{\theta}) \tag{1}$$

where $f(Y_E|\lambda)$ is the ECT likelihood function in which the penalty function $R(\lambda; \hat{\theta})$ depends on the MRI measurements and is specified to enforce smoothness within the estimated boundary of the ROI. The penalty is of the form of a Gibbs potential function

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

where w_{ij} are Gibbs weights which depend on $\hat{\theta}$ through one of many possible functional assignments [10, Ch. 4].

3. A BAYESIAN SETTING

Assume that λ and θ are random parameters (ECT intensity and MRI boundary estimates). Under the assumption that Y_M and Y_E are conditionally independent given λ, and that λ is a function of θ, the joint density density function factors in the following manner:

$$f(\lambda, Y_E, \theta, Y_M) = f(Y_E|\lambda) f(\lambda|\theta) f(\theta|Y_M) f(Y_M). \tag{2}$$

The first three factors on the right hand side of the above equation can be recognized as the likelihood function of λ without side information, the conditional density linking θ to λ, and the posterior density of θ given Y_M.

When θ is known exactly, the posterior mode (MAP) estimate $\hat{\lambda}$ of λ maximizes the posterior loglikelihood $\ln f(\lambda|Y_E, \theta)$ which is equivalent to maximizing the objective

$$J_\theta(\lambda) = \ln f(Y_E|\lambda, \hat{\theta}) + \ln f(\lambda|\theta).$$

Comparing this to (1) we can identify $f(\lambda|\theta)$ as a Gaussian "Gibbs-type" distribution of λ:

$$f(\lambda|\theta) = \alpha \exp\{-\beta R(\lambda; \theta)\} \tag{3}$$

where α is a normalizing constant.

When θ is not known exactly there are several ways to approach the estimation of λ. Many researchers [2, 3, 4, 5] have investigated joint estimation of λ and θ under various models for the prior $f(\theta)$ and $f(\lambda|\theta)$. While this is an important task for applications where both λ and θ are of intrinsic interest, this approach does not generally yield Bayes optimal estimates of λ which are given by the posterior mode estimator of λ obtained by marginalizing $f(\lambda, \theta|Y_E, Y_M)$ by integrating out θ.

For random unknown θ the posterior mode estimate of λ maximizes the posterior $f(\lambda|Y_E, Y_M)$ which is equivalent to maximizing the following joint density function over λ:

$$\begin{aligned} f(\lambda, Y_E, Y_M) &= \int f(\lambda, Y_E, \theta, Y_M) d\theta \\ &= f(Y_E|\lambda) \cdot f(Y_M) \cdot \int f(\lambda|\theta) f(\theta|Y_M) d\theta \\ &= \ln f(Y_E|\lambda) + \ln \int f(\lambda|\theta) f(\theta|Y_M) d\theta \end{aligned}$$

Using the Gibbs conditional distribution (3) for $f(\lambda|\theta)$ this becomes

$$\begin{aligned} \hat{\lambda} &= \\ &\operatorname{argmax}_\lambda \left\{ \ln f(Y_E|\lambda) + \ln \int \exp\{-\beta R(\lambda; \theta)\} f(\theta|Y_M) d\theta \right\} \end{aligned} \tag{4}$$

This is of the form of a penalized likelihood estimator (1) where the penalty is obtained by taking the log of the $f(\theta|Y_M)$-averaged exponentiated Gibbs penalty function. This penalty is convex in λ and for piecewise constant $(0, 1)$ weights w_{ij} can be expressed as:

$$\begin{aligned} &\ln \int \exp\{-\beta R(\lambda; \theta)\} f(\theta|Y_M) d\theta \\ &\quad = \ln \sum_l \exp\{-\beta R(\lambda; \theta_l)\} \int_{\Theta_l} f(\theta|Y_M) d\theta \end{aligned} \tag{5}$$

where $\{\Theta_l\}_l$ is a partition of $\mathbb{R}^p$ into sets over which $R(\lambda; \theta)$ is constant and θ_l is a point within the set Θ_l.

4. SIMPLIFICATION BY JENSENS INEQUALITY

Because of the exponential form of the average in 4) or (5) the computation of the penalty function is numerically unstable for even moderate values of the smoothing constant β. A simplification to (4) can be made by applying Jensen's inequality to lower bound this penalty term

$$\begin{aligned} & \ln \int \exp\{-\beta R(\lambda;\theta)\} f(\theta|Y_M) d\theta \\ \geq & -\beta \int R(\lambda;\theta) f(\theta|Y_M) d\theta \\ = & -\beta E_{\theta|Y_M}[R(\lambda;\theta)] \\ = & -\beta \sum_{ij} \tilde{w}_{ij}(\lambda_i - \lambda_j)^2 \end{aligned}$$

where $E_{\theta|Y_M}$ denotes expectation with respect to the posterior density $f(\theta|Y_M)$ and $\tilde{w}_{ij} = E_{\theta|Y_M}[w_{ij}(\theta)]$ are average weights. Therefore, use of the simpler penalty $-\beta E_{\theta|Y_M}[R(\lambda;\theta)]$ instead of $\ln \int \exp\{-\beta R(\lambda;\theta)\} f(\theta|Y_M) d\theta$ entails an overall reduction in the influence of the penalty factor for any given β. However, it is likely that the underpenalization can be be compensated by increasing the value of β.

The computation of the penalty requires averaging over a posterior density $f(\theta|Y_M)$ which requires knowing the prior distribution $f(\theta)$. When multiple realizations of $f(Y_M, \theta)$ are available empirical Bayes techniques can be used to accurately estimate the posterior density. Empirical Bayes techniques were formalized by Robbins [12] as a methodology for dealing with uncertain parameters. Empirical Bayes results in the substitution of an empirically determined distribution for the unknown prior distribution required for Bayes parameter estimation, e.g. the conditional mean or the posterior mode (maximum a posteriori) estimators. Another technique that eliminates the need for an explicit prior $f(\theta)$ are use of asymptotic approximations to $f(\theta|Y_M)$.

5. ASYMPTOTIC MARGINALIZATION

Let $\hat{\theta} = \text{argmax}_\theta f(Y_M|\theta)$ be the maximum likelihood estimate of θ based on Y_M. We now state the following limit theorem which can be shown using techniques of [13], and [14, Secs 12.4 and 17.7].

Theorem 1 *Assume that $f(Y_M|\theta)$ is a smooth function of θ in the sense of satisfying the regularity conditions [13, p. 131] and that $f(\theta)$ is a smooth function in the neighborhood of $\hat{\theta}$. Then $\hat{\theta}$ is an asymptotically consistent estimator of $\theta \in I\!R^p$ and*

$$f(\theta|Y_M) = \frac{|F_{\hat{\theta}}|^{\frac{1}{2}}}{(\sqrt{2\pi})^p} \cdot \exp\left\{-\tfrac{1}{2}(\theta-\hat{\theta})^T F_{\hat{\theta}}(\theta-\hat{\theta})\right\}\left(1+O(\Delta^2)\right)$$

where $F_{\hat{\theta}}$ is the observed Fisher information matrix

$$F_{\hat{\theta}} = -\nabla^2_{\hat{\theta}} \ln f(Y_M|\hat{\theta})$$

and $O(\Delta^2)$ decays to zero faster than $\Delta^2 = ||\theta - \hat{\theta}||^2$.

This theorem gives an asymptotic Gaussian form for the posterior $f(Y_M|\theta)$ which depends on the estimate $\hat{\theta}$ and the observed Fisher information at $\hat{\theta}$ but is independent of the explicit form of the prior density $f(\theta)$. This form is identical to the profile posterior approximation proposed in [15] when the joint distribution is specialized to the factorization (2). In some cases it is reasonable to use the (expected) Fisher information in place of F. In particular $F(\hat{\theta})$ can be approximated by the Fisher information evaluated at $\hat{\theta}$ [16]:

$$I_{\hat{\theta}} = -\int f(Y_M|\theta)\nabla^2_\theta \ln f(Y_M|\theta) dY_M|_{\theta=\hat{\theta}}.$$

6. BACK TO PENALIZED LIKELIHOOD

Application of the above results gives the following form for the variance compensated penalized likelihood estimator

$$\hat{\lambda} = \text{argmax}_\lambda \left\{\ln f(Y_E|\lambda) - \beta \sum_{ij} \tilde{w}_{ij}(\hat{\theta})(\lambda_i - \lambda_j)^2\right\} \quad (6)$$

where

$$\tilde{w}_{ij}(\hat{\theta}) = |F_{\hat{\theta}}|^{\frac{1}{2}} \int w_{ij}(\theta)\phi(F_{\hat{\theta}}^{-\frac{1}{2}}[\theta-\hat{\theta}])d\theta$$

Note that the weight averaging corresponds to considering θ as a p-variate Gaussian random vector with mean $\hat{\theta}$ and covariance matrix $F_{\hat{\theta}}^{-1}$ and $\phi(x_1,\ldots,x_p)$ is the standard p-dimensional Gaussian density function with zero mean and identity covariance. This average is over the the posterior uncertainty region of θ.

For the binary weight scheme the computation of $\tilde{w}_{ij}$ can be performed directly in the pixel domain by averaging the weight map over all perturbations of θ which produce changes in the boundary by at least a single pixel. Since this average may be difficult, it may be better to use the reparametrization $w_{ij}(\hat{\theta}) = w_{ij}(\hat{r}, \nu)$, where $\hat{r}, \nu$ index the locations of the boundary in polar coordinates, and do the averaging in the pixel domain

$$\begin{aligned} \tilde{w}_{ij}(\hat{\theta}) = & \quad (7) \\ & \frac{|F_{\hat{r}}|^{\frac{1}{2}}}{(\sqrt{2\pi})^p} \cdot \sum_r w_{ij}(\hat{r},\nu)\exp\left\{-\tfrac{1}{2}(r-\hat{r})^T F_{\hat{r}}(r-\hat{r})\right\} \end{aligned}$$

where the sum is over the variation of radii r_ϕ which give rise to different weight maps and $F_{\hat{r}}^+$ denotes the pseudo-inverse $F_{\hat{r}}$:

$$F_{\hat{r}} = BF_{\hat{\theta}}B^T.$$

As a practical approximation, the sum is truncated by extracting p principle components $\xi_1,\ldots,\xi_p$ (eigenvalues ordered in increasing rank $\lambda_1 \leq \lambda_2 \leq \ldots$) from the eigendecomposition of $F_{\hat{r}}$ and summing over the variations in $(r-\hat{r})^T \xi_i$ which: i) produce changes in the weight maps; ii) have magnitude changes over range: $|(r-\hat{r})^T \xi_i|^2 \leq \frac{3}{\lambda_i}$, $i = 1, \ldots, p$. This will cover a range of changes in r which will account for the majority of the mass in the averaging distribution (plus or minus 3 standard deviations from the mean).

7. APPLICATION TO SPECT/MRI

Here we consider the situation treated in [9] where MRI-derived anatomical side information is applied to SPECT through the Gibbs weights $w_{i,j}$. We compare reconstructed images using side information averaged "smoothed weights" $\tilde{w}_{ij}$ and the "unsmoothed weights" $\tilde{w}_{ij}$ extracted directly from the MRI image. A modified version of the SAGE3 algorithm [17] was used to maximize the penalized likelihood (6). The MRI anatomical boundary estimates were obtained by edge detection and median filtering algorithm discussed in [9]. We have shown that the mean-square error of this edge extraction algorithm virtually achieves the CR bound over the practical range of MRI signal-to-noise ratios. This allows us to use an analytical form of the Fisher information, derived in [18], in the smoothing formula (7).

The ROI boundary was a quadratic spline with $K = 16$ equispaced knots which is a least squares fit to an ellipse with major axis 5 pixels and minor axis 3 pixels located at $(-10, 0)$ pixels relative to image center. The contrast for both the MRI and ECT mean images were identical: for MRI the mean intensity inside the ROI was 6 in a zero mean background while for ECT the mean intensity inside the ROI was 9 in a background of mean 3. In both ECT and MRI images each boundary pixel was assigned mean intensity equal to the interior mean intensity times the fraction of the boundary pixel included inside the ROI. The MRI spatial blurring width was equal to $\sigma_s = 0.75$ (approximately 19% of average ROI radius) and a spatially independent Gaussian noise was added having zero mean and standard deviation $\sigma_n = 0.18$ (appoximately 3% of contrast). The ECT data was sampled by a parallel beam tomograph corresponding to PET projections over 64 radial bins, and 60 equispaced projection angles over 180^o. Poisson noise was added to the ECT data by generating 10^6 Poisson realizations of the mean intensity and adding 15% random coincidences.

Figure 1 and 2 shows the true SPECT phantom and the SAGE3 reconstruction with no MRI boundary information. Note that without the benefit of MRI side information the smoothing penalty is spatially invariant and therefore the hot spot (small white ellipse) is smoothed into the surrounding anatomy. Figure 3 shows the SAGE3 presented in [9] which uses weights derived from the MRI boundary estimate. While the hot spot boundary is sharp in this unsmoothed image it is of irregular shape and underestimates the uptake of the true hot spot. Figure 4 shows the SAGE3 reconstruction using the smoothed weights and Figure 5 shows the result of incorporating a leakage-prevention boundary into $\tilde{w}_{ij}$. The leakage-prevention hard limits $\tilde{w}_{ij}$ to 0 at the 98 percentile of the radial boundary uncertainty distribution (determined from Fisher information and the posterior boundary density). This preserves overall intensity level within the hot spot region and lowers the bias of the uptake estimate.

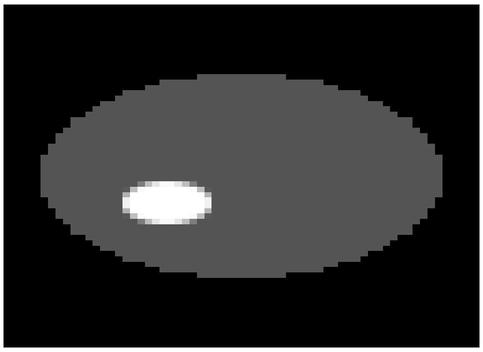

Figure 1: True noiseless phantom for SPECT and MRI data.

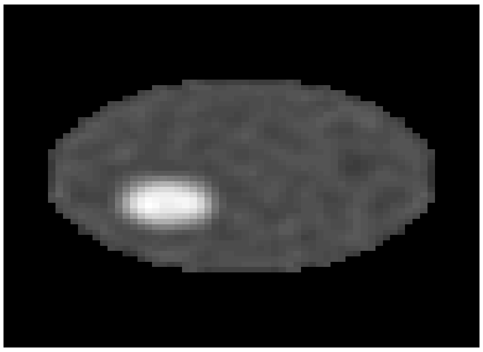

Figure 2: SPECT reconstruction without any side information for $\beta = 1$ for SPECT and MRI data.

8. REFERENCES

[1] S. Titus, A. O. Hero, and J. A. Fessler, "Improved penalized likelihood reconstruction of anatomically correlated emission data," in *IEEE Int. Conf. on Image Processing*, volume 2, Laussane, Sep. 1996.

[2] 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, , December 1993.

[3] K. M. Hanson, "Bayesian reconstruction based on flexible priors," *J. Opt. Soc. Am.*, vol. 10, no. 5, pp. 997–1004, May 1993.

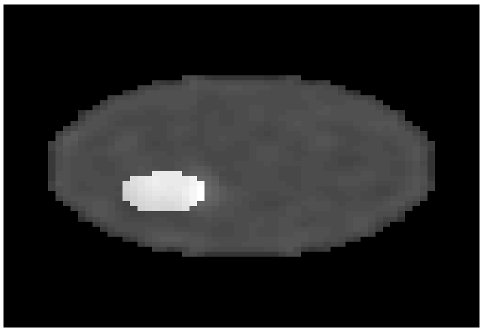

Figure 3: SPECT reconstruction using MRI side information and unsmoothed boundary estimates.

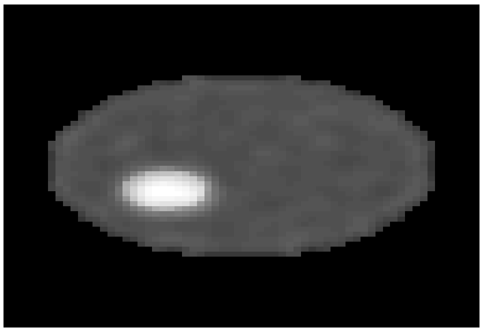

Figure 4: SPECT reconstruction using MRI side information and smoothed boundary estimates.

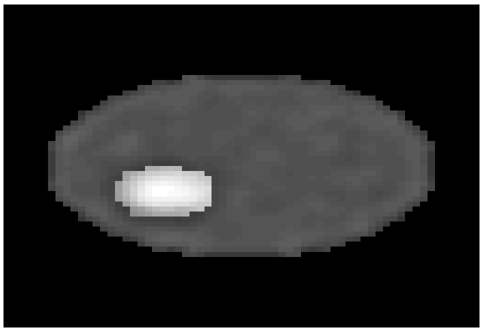

Figure 5: SPECT reconstruction using MRI side information and smoothed boundary estimates with leakeage prevention boundary.

[4] T. Hebert and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Trans. on Medical Imaging*, vol. 8, no. 2, pp. 194–203, 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 *Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. D. Paola, editors, Kluwer, 1995.

[6] Y. Zhang, J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "Joint estimation for incorporating MRI anatomic images into SPECT reconstruction," in *Proc. of IEEE Nuclear Science Symposium*, volume 3, pp. 1256–1260, 1994.

[7] Y. Zhang, J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "Incorporating MRI region information into SPECT reconstruction using joint estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, volume 4, pp. 2307–2310, 1995.

[8] S. Titus, A. O. Hero, and J. A. Fessler, "Nmr object boundaries: B-spline modeling and estimator performance," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, volume 4, pp. 2423–2426, Detroit, MI, May 1995.

[9] S. Titus, A. O. Hero, and J. A. Fessler, "Penalized likelihood emission image reconstructon with uncertain boundary information," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, volume 3, Munich, April 1997.

[10] S. R. Titus, *Improved penalized likelihood reconstruction of anatomically correlated emission computed tomography data*, PhD thesis, The University of Michigan, Ann Arbor, December 1996.

[11] M. A. Tanner, *Tools for Statistical Inference; Methods for the exploration of posterior distributions and likelihood functions*, Springer-Verlag, New York, 1993.

[12] H. Robbins, "The empirical Bayes approach to statistical decision problems," *Ann. Math. Statist.*, vol. 35, pp. 1–68, 1964.

[13] P. J. Huber, *Robust Statistics*, Wiley, New York, 1981.

[14] L. LeCam, *Asymptotic Methods in Statistical Decision Theory*, Springer-Verlag, New York, 1986.

[15] L. Tierney and J. B. Kadane, "Accurate approximations for posterior moments and marginal densities," *jasa*, vol. 81, pp. 82–86, 1986.

[16] 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.

[17] J. A. Fessler and A. O. Hero, "Penalized maximum-likelihood image reconstruction using space alternating generalized EM algorithm," *IEEE Trans. on Image Processing*, vol. IP-4, no. 10, pp. 1417–1429, Oct. 1995.

[18] S. R. Titus, *Improved penalized likelihood reconstruction of anatomically correlated emission computed tomography data*, PhD thesis, The University of Michigan, Ann Arbor, December 1996.