AN ITERATIVE BAYES ALGORITHM FOR EMISSION TOMOGRAPHY USING A SMOOTHED SINOGRAM

Jun Ma

Department of Statistics, Macquarie University
NSW 2109, Australia

ABSTRACT

In this paper we formulate a new approach to medical image reconstruction from projections in emission tomography. This approach differs from traditional methods such as filtered back projection, maximum likelihood or maximum penalized likelihood. Our method is developed directly from the Bayes formula and the final result is an iterative algorithm, for which the maximum likelihood expectation-maximization of [1] (or [2]) is a special case.

1. INTRODUCTION

Maximum likelihood (ML) and maximum penalized likelihood (MPL) are widely used reconstruction methods in emission tomography. For ML image reconstructions, the iterative expectation maximization (EM) algorithm (see, e.g., [2]) is widely practiced. In the context of image and signal processing it is well known that the EM algorithm is indistinguishable from the Richardson-Lucy (RL) algorithm ([3], [4]). The RL method is a Bayesian based iterative algorithm; its development comes directly from the Bayes conditional probability formula.

This paper develops a new reconstruction method called the iterative Bayes (IB) algorithm. This algorithm differs from RL or EM in that it requires to pre-smooth the camera data. Smoothing the camera data before reconstruction is not a new idea; for example, [5] contains a method which first smooths the camera measurements and then reconstructs using filter back projection (FBP). However, utilizing smoothed measurements in an RL or EM type algorithm is novel.

The observed measurements are denoted by $y_1, \ldots, y_n$, and the unknown emission activities are denoted by $x_1, \ldots, x_p$. Let $\boldsymbol{y}$ be the n-vector for all y_i and $\boldsymbol{x}$ the p-vector for all x_j. The matrix $A = (a_{ij})_{n\times p}$ is called the projection matrix, where a_{ij} represents the conditional probability of a photon being detected in camera bin i given it is emitted from voxel j. It is commonly assumed that the y_i's are independent and each y_i follows a Poisson distribution with mean $\mu_i = \sum_{i=1}^n a_{ij}x_j$.

Our iterative method presented in Section 2 estimates $\boldsymbol{x}$ employing the conditional probability of a photon being emitted from voxel j given it is detected in bin i.

The rest of this paper is organized as follows. Section 2 formulates the IB algorithm. Section 3 discusses how to estimate the μ_i's by smoothing the measurements y_i. Section 4 provides an ordered subset formulation of the algorithm. Simulation studies are given in Section 5 and discussions in Section 6.

2. IMAGE RECONSTRUCTION BY THE ITERATIVE BAYES ALGORITHM

Let $P(j)$ be the probability of a photon being emitted from voxel j and $P(i)$ the probability of a photon being detected in camera bin i. Similarly, we use $P(i\,|\,j)$ to denote the probability of a photon being detected in bin i given that it is released from voxel j and $P(j\,|\,i)$ the probability of a photon being emitted from voxel j given it arrives in bin i.

Probability $P(j)$ is not known; however if x_j is available, then one can estimate $P(j)$ by $\hat{P}(j) \propto x_j$.

For $P(i\,|\,j)$, as a_{ij} contains various information about projection from voxel j to camera bin i, such as system geometry, attenuation, scattering etc, it is commonly accepted to estimate $P(i\,|\,j)$ by $\hat{P}(i\,|\,j) \propto a_{ij}$.

Probability $P(j\,|\,i)$ can be evaluated from $P(j)$ and $P(i\,|\,j)$ using the Bayes formula, namely,

$$P(j\,|\,i) = \frac{P(i\,|\,j)P(j)}{\sum_{t=1}^p P(i\,|\,t)P(t)}. \tag{1}$$

As the $P(j)$'s are unknown, we may compute $P(j\,|\,i)$ iteratively. From the current estimate of $\boldsymbol{x}$ (denoted by $\boldsymbol{x}^k$), first obtain the following update to $P(j\,|\,i)$ from (1):

$$\hat{P}^k(j\,|\,i) = \frac{a_{ij}x_j^k}{\sum_{t=1}^p a_{it}x_t^k} = \frac{a_{ij}x_j^k}{\mu_i^k}, \tag{2}$$

where $\mu_i^k = \sum_{j=1}^p a_{ij}x_j^k$, then from $\hat{P}^k(j\,|\,i)$ we can develop $\boldsymbol{x}^{k+1}$. Note for any emitted photon from voxel j the probability of it being detected is $\sum_{i=1}^n a_{ij}$; thus,

$$\left(\sum_{i=1}^n a_{ij}\right) x_j^{k+1} = \sum_{i=1}^n \hat{P}^k(j\,|\,i)\mu_i, \tag{3}$$

where μ_i is the expected counts (unobservable) for bin i.

Equation (3) cannot be implemented yet as the μ_i's are unknown. We estimate μ_i exploiting the smoothness of the μ_i's, as neighboring μ_i values are believed to be similar.

Denote the estimate of μ_i by $\hat{\mu}_i$. Substituting $\hat{\mu}_i$ into (3) gives an algorithm for estimating the emission intensity vector $\boldsymbol{x}$:

$$x_j^{k+1} = \frac{x_j^k}{\sum_{i=1}^n a_{ij}} \sum_{i=1}^n a_{ij} \frac{\hat{\mu}_i}{\mu_i^k}. \tag{4}$$

We call this the Iterative Bayes (IB) algorithm. Note when $\hat{\mu}_i = y_i$, this algorithm coincides with the EM algorithm of Vardi *et al.*[6] for emission tomography. Discussions on how to obtain $\hat{\mu}_i$ are given in Section 3.

Following the convergence proof in [6] for the EM algorithm we can show that, under certain regularity conditions, the IB algorithm converges (with the initial $\boldsymbol{x}^0 > 0$) to one of the global solutions maximizing

$$d(\boldsymbol{x}) = -\sum_{i=1}^n \mu_i + \sum_{i=1}^n \hat{\mu}_i \log \mu_i. \tag{5}$$

for $\boldsymbol{x} \geq 0$. Note $d(\boldsymbol{x})$ is a concave function, thus there exist global maxima (may not be unique).

Let $\boldsymbol{\mu}$ and $\hat{\boldsymbol{\mu}}$ be the vector for all μ_i and $\hat{\mu}_i$. The Kullback-Leibler (KL) distance between $\hat{\boldsymbol{\mu}}$ and $\boldsymbol{\mu}$ is

$$KL(\hat{\boldsymbol{\mu}}, \boldsymbol{\mu}) = \sum_{i=1}^n \left(\hat{\mu}_i \log \frac{\hat{\mu}_i}{\mu_i} - \hat{\mu}_i + \mu_i \right). \tag{6}$$

Maximizing $d(\boldsymbol{x})$ is equivalent to minimizing $KL(\hat{\boldsymbol{\mu}}, \boldsymbol{\mu})$. Thus the IB algorithm finds a positive minimum $KL(\hat{\boldsymbol{\mu}}, \boldsymbol{\mu})$ (minKL) estimate.

3. NONPARAMETRIC SMOOTHING OF THE MEASUREMENTS

Estimation of $\boldsymbol{\mu}$ is equivalent to the problem of nonparametric regression, with an identity link function and Poisson noise in the context of generalized linear models (GLM); see [7]. Gyorfi *et al.*[8] covers almost all known nonparametric regression methods.

For emission tomography, as the likelihood function is readily available, it is natural to use the penalized likelihood method to estimate $\boldsymbol{\mu}$.

As in [5], we estimate $\boldsymbol{\mu}$ on an angle-by-angle basis. That is, only those μ_i on the same projection angle are estimated simultaneously; for different projection angles estimation procedures are independent. Suppose there are M projection angles. Let S_t be the index set for projection angle t and let $\boldsymbol{y}_t$ and $\boldsymbol{\mu}_t$ be, respectively, the vector for measurements and expected measurements of projection angle t. Given $\boldsymbol{y}_t$, our aim is to find non-parametrically a smoothed estimate of $\boldsymbol{\mu}_t$ using MPL.

From set $\boldsymbol{y}_t$, the penalized log-likelihood function is:

$$g(\boldsymbol{\mu}_t) = -\sum_{i \in S_t} \mu_i + \sum_{i \in S_t} y_i \log \mu_i - \frac{1}{2} \lambda\, b(\boldsymbol{\mu}_t), \tag{7}$$

where $\lambda > 0$ is the smoothing parameter and $b(\boldsymbol{\mu}_t)$ is the penalty function. We call λ the projection smoothing parameter.

The function $b(\boldsymbol{\mu}_t)$ is chosen to penalize local differences between neighboring μ_i. In this paper we let $b(\boldsymbol{\mu}_t)$ be a penalty function known as the "roughness penalty" defined below. In this context the maximum of $g(\boldsymbol{\mu}_t)$ is a natural cubic spline [7].

Let $f_t(\xi)$ be a smooth function connecting all μ_i ($i \in S_t$) along angle t. If we assume each angle comprises the same number D of camera bins, the roughness penalty remains unchanged for different projection angles as the supporting interval for $f_t(\xi)$ is always $[1, D]$.

The roughness penalty is $b(f_t) = \int_1^D (f_t''(\xi))^2\, d\xi$. According to [7, Chapter 2], given $D \geq 2$, if $f_t(\xi)$ is the cubic spline interpolating μ_i ($i \in S_t$) with knots $\{1, 2, \ldots, D\}$, then

$$\int_1^D (f_t''(\xi))^2\, d\xi = \boldsymbol{\mu}_t^T K \boldsymbol{\mu}_t \stackrel{\text{def}}{=} b(\boldsymbol{\mu}_t), \tag{8}$$

where K is a $D \times D$ penalty matrix. Construction of the matrix K is very simple; details are given in [7].

For each angle t our aim is to estimate $\boldsymbol{\mu}_t$ maximizing $g(\boldsymbol{\mu}_t)$ of (7), with $b(\boldsymbol{\mu}_t)$ given by (8), subject to the non-negativity constraints $\mu_i \geq 0$ ($i \in S_t$). This is a special case of the general non-negatively constrained inverse problem described in [9]; hence the multiplicative iterative inversion method (MIIL) of [9] can be implemented directly.

4. ORDERED SUBSET FORMULATIONS

The ordered subset EM (OSEM) [10] algorithm has attracted many research activities recently. We develop ordered subset IB (OSIB) algorithms by adopting the idea of OSEM.

Let $\{S_u, u = 1, \ldots, L\}$ be a partition of the measurements index set $\mathcal{N} = \{1, \ldots, n\}$, namely, $\bigcup_{u=1}^L S_u = \mathcal{N}$ and $S_u \bigcap S_v = \emptyset$ for $u \neq v$. Let $\boldsymbol{x}^{k,\,u}$ denote the update of $\boldsymbol{x}$ at sub-iteration u of iteration k. At iteration k of the OSIB algorithm, $\boldsymbol{x}$ is updated through L sub-iterations using the corresponding $\hat{\boldsymbol{\mu}}$ and A subsets. More specifically, starting with $\boldsymbol{x}^{k,\,0} = \boldsymbol{x}^{k-1}$,

$$x_j^{k,\,u} = \frac{1}{\sum_{i \in S_u} a_{ij}} \sum_{i \in S_u} \tilde{P}^{k,\,u}(j \,|\, i)\, \hat{\mu}_i, \tag{9}$$

where $\tilde{P}^{k,\,u}(j \,|\, i)$ is computed for $i \in S_u$ only, and is given by

$$\tilde{P}^{k,\,u}(j \,|\, i) = \frac{a_{ij} x_j^{k,\,u-1}}{\sum_{t=1}^p a_{it} x_t^{k,\,u-1}} = \frac{a_{ij} x_j^{k,\,u-1}}{\mu_i^{k,\,u-1}}. \tag{10}$$

Clearly, OSIB is unsatisfactory. In each sub-iteration, it estimates the entire $\boldsymbol{x}$ by only a single subset and, at the same time, it attempts to match $\hat{\mu}_i$ with μ_i for this subset.

A more reliable approach is offered by the following modification. It updates $\boldsymbol{x}$ using all the data rather than only the current subset:

$$x_j^{k,\,u} = \frac{1}{\sum_{i=1}^{n} a_{ij}} \sum_{i=1}^{n} \tilde{P}^{k,\,u}(j \mid i)\, \hat{\mu}_i, \tag{11}$$

where

$$\tilde{P}^{k,\,u}(j \mid i) = \begin{cases} a_{ij} x_j^{k,\,u-1} / \mu_i^{k,\,u-1} & \text{for } i \in S_u \\ \tilde{P}^{k,\,u-1}(j \mid i) & \text{for } i \notin S_u \end{cases} \tag{12}$$

with $\tilde{P}^{k,\,0}(j \mid i) = a_{ij} x_j^{k-1} / \mu_i^{k-1}$. Note that the $\tilde{P}(j \mid i)$'s for the current subset are updated and the others are unchanged.

This algorithm appears similar to the complete data OSEM (COSEM) algorithm of [11], and hence is named the COSIB algorithm.

Convergence properties of OSEM and COSEM-ML are studied in [10] and [11] respectively; these results apply directly to the OSIB and COSIB algorithms. In particular, we can show that the COSIB algorithm converges to one of the global maxima of $d(\boldsymbol{x})$.

5. SIMULATION STUDIES

In this section we explore the performance of IB, OSIB and COSIB for the positive minKL estimate, in 2D SPECT, by applying them to a simulated image. We will also compare them with the EM algorithm of Shepp and Vardi [1] for positive ML estimate and the multiplicative iterative inversion algorithm (MIIL) of Ma [9] for positive MPL estimate.

The simulation used an elliptical phantom of size 64×64 pixels. There were 64 attenuated parallel beam projections uniformly spaced over 360^0, and each projection contained 64 measurements. Attenuation coefficients were 0.15 /cm (water) within the body, except for within the two lungs, where coefficients were 0.375 /cm.

The projection matrix A (with dimension $64^2 \times 64^2$) was pre-determined by the geometry of pixels, adjusted according to the attenuation coefficients. Poisson noises were added to $\boldsymbol{\mu} = A\boldsymbol{x}$ to form the observed measurements vector $\boldsymbol{y}$. The total measurements in projections was 400,605.

We used the Root Mean Square (RMS), defined as:

$$\text{RMS} = \sqrt{\sum_{j=1}^{p} (\hat{x}_j - x_j)^2 / p}, \tag{13}$$

to evaluate the quality of estimates; for each algorithm, we computed the RMS in each iteration against the phantom image. Note that RMS alone is not a perfect measure for assessing image differences; a better measure, which is not available in a closed from, is given by replacing $\hat{x}_j$ by $E(\hat{x}_j)$ in RMS.

We used a quadratic penalty $J(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T R \boldsymbol{x}$ in the MPL estimate with R given by $r_{jt} = -0.25$ for $j \neq t$ and t in the first-order neighborhood of j, $r_{jj} = 1$ for all j and $r_{jt} = 0$ for all other j and t. The "optimal" smoothing parameter, which was found by minimizing the RMS, was 1.053×10^{-5}.

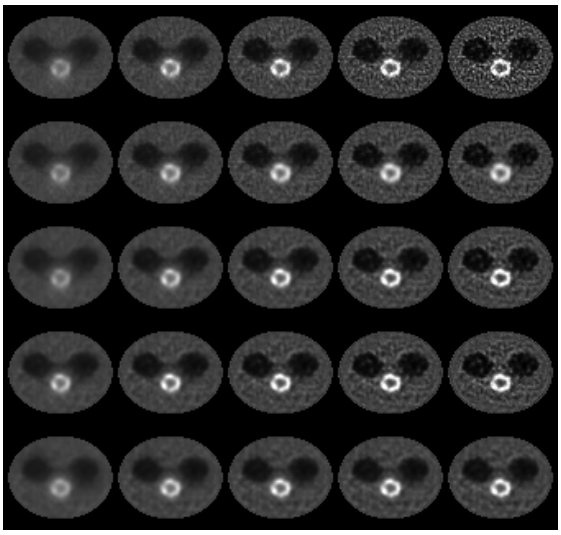


Fig. 1. Row 1: EM; row 2: MIIL; row 3: IB; row 4: COSIB-8; row 5: OSIB-8. Columns are for iterations 8, 16, 32, 64, 128, except OSIB where columns are for iterations 1, 2, 3, 4, 5

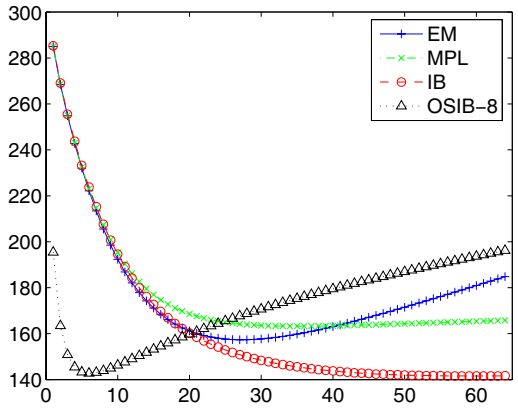
The IB reconstructions were developed from $\hat{\boldsymbol{\mu}}$ with $\lambda = 10^{-3}$. Note that this λ is only a rough choice and it may not be optimal.

We implemented IB, COSIB with 8 and 64 (denoted by COSIB-8 and COSIB-64) subsets and OSIB with 8 (denoted by OSIB-8) subsets. We wish to exam their convergence performance.

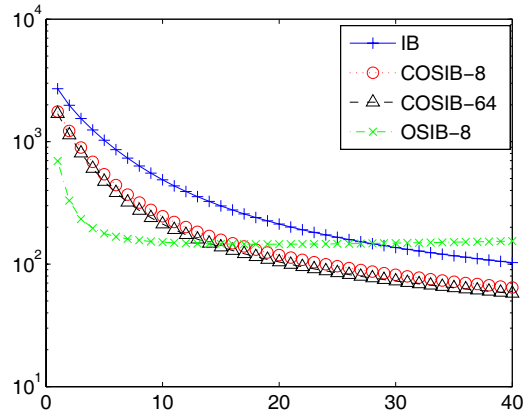
Figure 1 exhibits various reconstructions. The EM results show that the ML estimates at high iterations are less attractive. The MPL solution, even with the optimal smoothing parameter, is also less satisfactory as it loses boundary details. In contrast, IB and its OS algorithms offer the best estimate; they not only smooth the reconstructions, but also preserve boundaries better than MPL.

Figure 2 (a) shows the RMS values of EM, MIIL, IB and OSIB-8. OSIB-8 produces "acceptable" images quickly, only 5 to 6 iterations in this example. However, OSIB-8 deteriorates rapidly, making the issue of optimal stopping rule crucial. The IB method provides satisfactory results as its RMS values degenerate much more slowly than EM and OSIB-8.

Figure 2 (b) displays $d(\boldsymbol{x}^{(k)}) - d(\boldsymbol{x}^{(\infty)})$ against the iteration number k, where $\boldsymbol{x}^{(\infty)}$ represents the true positive



(a) RMS



(b) Plot of $d(\boldsymbol{x}^{(k)}) - d(\boldsymbol{x}^{(\infty)})$

Fig. 2. (a) Plots of RMS values. Algorithms are EM, MIIL for MPL, IB and OSIB-8. (b) Plots of $d(\boldsymbol{x}^{(k)}) - d(\boldsymbol{x}^{(\infty)})$ (on log scale) against iteration number k. Algorithms are IB, COSIB-8, COSIB-64 and OSIB-8.

minKL solution and is approximated by 5000 iterations of IB. For COSIB, the subset number has negligible influence on convergence speed; COSIB-64 appears only about 3 to 4 iterations ahead of COSIB-8. However, COSIB (both 8 and 64 subsets) converges faster than IB; COSIB doubles the convergence speed of IB. OSIB-8 has initial fast convergence, but then is moved away from the minKL solution.

6. DISCUSSION

In this article we have developed a new image reconstruction method for emission tomography using simply the Bayes formula. When the measurements are unsmoothed then it becomes the EM algorithm.

The simulation studies show that, even with non-optimal projection smoothing, this method achieves lower RMS than the optimal MPL solution. Also, it enhances boundary preservation when compared with MPL.

Although we have developed the method in the context of histogram mode observations, it can be easily extended to the list mode. Hence it can be treated as a unifying approach for both type of measurements.

7. REFERENCES

[1] L. A. Shepp and Y. Vardi, "Maximum likelihood estimation for emission tomography," *IEEE Trans. Med. Imaging*, vol. MI-1, pp. 113–121, 1982.

[2] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comp. Assisted Tomography*, vol. 8, pp. 306–316, 1984.

[3] W. H. Richardson, "Bayesian based iterative method for image restoration," *J. Opt. Soc. Am.*, vol. 62, pp. 55–59, 1972.

[4] L. B. Lucy, "An iterative technique for the rectification of obeserved distributions," *Astron. J.*, vol. 79, pp. 745–754, 1974.

[5] P. J. La Riviere and X. Pan, "Nonparametric regression sinogram smoothing using a roughness-penalized Poisson likelihood objective function," *IEEE Trans. Med. Imaging*, vol. 19, pp. 773–786, 2000.

[6] Y. Vardi, L. A. Shepp, and A. Kaufman, "A statistical model for positron emission tomography (with discussion)," *J. Amer. Statist. Assoc.*, vol. 80, pp. 8–37, 1985.

[7] P. J. Green and B. W. Silverman, *Nonparametric Regression and Generalized Linear Models - A roughness penealty approach*, Chapman and Hall, London, 1994.

[8] L. Gyorfi, M. Kohler, A. Krzyzak, and H. Walk, *A distribution-Free Theory of Nonparametric Regression*, Springer, New York, 2002.

[9] J. Ma, "Multiplicative algorithms for maximum penalized likelihood inversion with nonnegative constraints and generalized error distributions (to appear)," *Communications in Statistics - Theory and Methods*, vol. 35, Issue 5, 2006.

[10] H. M. Hudson and R. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imaging*, vol. MI-13, pp. 601–609, 1994.

[11] I. T. Hsiao, A. Rangarajan, and G. Gindi, "A provably convergent OS-EM like reconstruction algorithm for emission tomography," *in Conf. Rec. SPIE Med. Imaging*, vol. 4684, pp. 10–19, 2002.