CENTER FOR ADVANCED COMPUTATION

MODIFIED EXPECTATION MAXIMIZATION ALGORITHM FOR TRANSMISSION TOMOGRAPHY

PATIWAT KAMONPET

VENKATARAMA KRISHNAN

Technical Report CAC T94-01
January 28, 1994

Parts of this report have appeared as Chapter 23,
Probabilistic Methods in Transmission Tomography in the book
Probability and Random Processes published by John Wiley & Sons Inc., 2006

UNIVERSITY OF MASSACHUSETTS LOWELL

LIST OF FIGURES

ABSTRACT

A Bayesian reconstruction algorithm of transmission tomographic images is presented. The reconstruction is based on maximum aposteriori (MAP) estimation using the Expectation Maximization (EM) algorithm. The Gibbs prior with sigmoidal potential function that smoothes out high differences in neighboring pixel values while retaining sharp-edged features of images is employed. The Gibbs prior improves the image quality as well as the convergence properties of the EM algorithm. The Bayesian reconstruction problem itself is solved through the One-Step-Late Expectation-Maximization algorithm (EM-OSL) proposed by Green. Approximations are introduced to make the problem more tractable for parallel computation. Computer simulated phantoms are used in this study. Images reconstructed from this algorithm are compared to those reconstructed from currently available reconstruction algorithms in terms of image quality and convergence rate. The reconstruction algorithm works quite well for low-photon-count and low-contrast cases in which the widely used convolution backprojection algorithm performs rather poorly.

KEY WORDS : Imaging, Transmission tomography, Expectation

INTRODUCTION

In transmission tomography (TT) the acquired projection data contains high statistical variations due to limited dose of radiation introduced in patients. When this is compounded with the presence of high contrast objects inside low contrast ones, many practical problems arise. Conventional convolution back projection and other Fourier-based methods fail to produce acceptable reconstructions and the resulting images are usually plagued by various types of artifacts. Transmission tomography models for reconstruction of images have to take into account the statistical nature of the problem. Statistical variations and other physical phenomena of the image reconstruction problem can be incorporated in a stochastic model. The application of the well-established theory of maximum likelihood estimation to image reconstruction problem was recognized by Rockmore and Macovski [1,2]. An alternative to the maximum likelihood approach is provided by Dempster et. al [3] with a general iterative method known as the expectation maximization (EM) algorithm. The relevance of the EM algorithm to maximum likelihood image reconstruction was noticed by Shepp and Vardi [4], who applied it to a stochastic model of positron emission tomography (PET), and independently by Lange and Carson [5] who applied it to general models of emission and transmission tomography. However, all these methods have to address the problems of computational complexity and storage

The initial optimism of the EM method of reconstruction in emission tomography has been tempered by the grainy, speckled appearance of the reconstructed images. This graininess is simply a manifestation of low correlation between parameter estimates for neighboring pixels. An intrinsically satisfying solution is to incorporate some smoothness criterion in the reconstruction process itself.

In emission tomography, this perspective has been advocated by Geman and McClure [6,7], Hebert and Leahy [8], and Green [9], who build on previous works of Geman and Geman [10], and Besag [11]. In particular Geman and McClure [6,7] introduce Gibbs priors with nearest neighbor interactions. Green [9] devises an approximate EM algorithm incorporating nearest neighbor interactions with the One-Step-Late Expectation-Maximization (EM-OSL) algorithm. Lange [12] has demonstrated the conditions under which EM-OSL algorithm actually converges to the unique maximum posterior point in image space for both emission and transmission cases.

In this report the Gibbs prior with sigmoidal potential function is developed that applies the EM-OSL scheme of Green [9] to the transmission algorithm proposed by Lange [12]. In simplifying the resultant transcendental equation three terms have been used in the series expansion rather than two in Lange. The solution of the resulting quadratic equation yields more accurate results. The sigmoidal potential function smoothes out high differences in neighboring pixel values while retaining sharp-edged features of images.

Section II discusses the model in detail with the underlying assumptions, the derivation of the algorithms and the simplifying approximations involved. Section III discusses the computer simulation of the algorithm with implementations of the various parameters. In Section IV the results from other algorithms, such as convolution backprojection, EM algorithm of Lange and Carson [5], Green [9] are presented. They are compared with regard to convergence properties, image quality in terms of visual appeal, analytical criteria, complexity of algorithms and susceptibility to low radiation flux. In Section V, some discussion of results are presented.

.

THE MODEL AND THE ALGORITHM

Before mentioning in detail about the algorithm, some of the notations are defined in Figure 1.

- i projection subscript
- j ordered pixel-subscript for each projection path i starting from the furthest pixel from the detector i
- I_i set of (m - 1) pixels contributing to the projection i
- J_j set of projections to which pixel j contributes
- Y_i random variable for detected photons for projection i
- X_{ij} random variable for photons from projection i *entering* pixel j
- $\tilde{X}_{ij}$ random variable for photons from projection i *leaving* pixel j, or entering pixel j+1, that is, $\tilde{X}_{ij}=X_{ij+1}$.
- U_j random variable for attenuation coefficient of pixel j
- m_j realization of U_j
- l_{ij} path length of projection i through pixel j

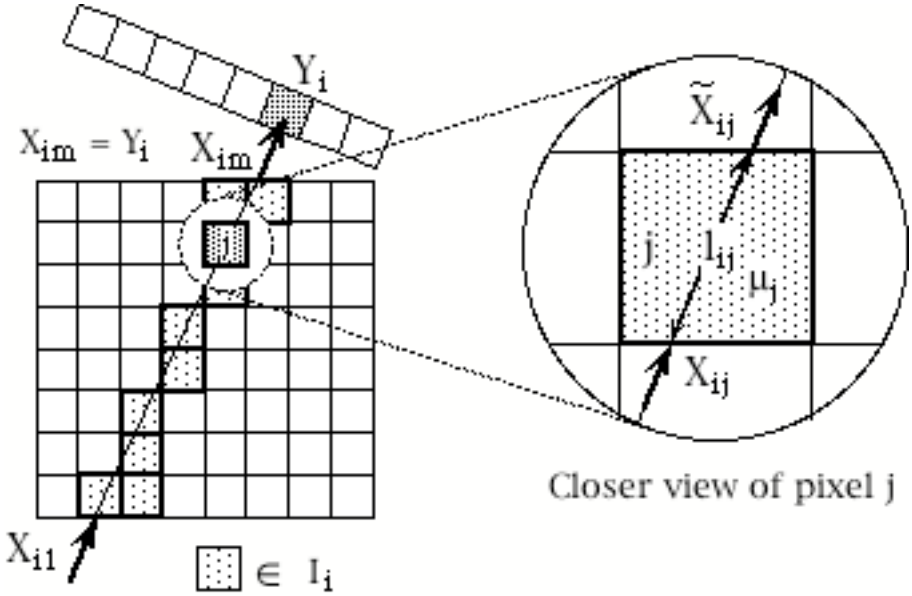


Figure 1 : Model of Transmission Computed Tomographic System

Corresponding to each projection i, the source photons are assumed to travel along a thin path on the way to the detector. This path cuts through a set I_i of (m -1)† pixels ordered by the index j, that is, pixel j=1 is the nearest pixel to the source and pixel j=m-1 is the last pixel along the path. The attenuation of pixel j is a random variable U_j taking the value m_j and the length of the path inside pixel j is l_{ij}. Photons traveling along path i are counted as x_{ij} at the entrance to pixel j and as $\tilde{x}_{ij}=x_{ij+1}$ at the exit with $x_{ij}-\tilde{x}_{ij}$ photons getting attenuated by pixel j. The number of photons x_{im} leaving the last pixel (m-1) is the same as the number of photons y_i getting detected by the detector i, or, $X_{im}= Y_i$. The attenuation coefficients $\{m_j\}$ have to be estimated from the observed photon count $x_{im} = y_i$.

The random variable X_{i1} of photons leaving the source at projection i is assumed to be Poisson distributed as,

$$P\{X_{i1}\} = \frac{e^{-d_i} d_i^{x_{i1}}}{x_{i1}!} \qquad \ldots(1)$$

where $d_i = E\{X_{i1}\}$ is the mean number of photons leaving the source at projection i.

The probability that a photon entering pixel j along path i will survive the attenuation at pixel j, conditioned on the attenuation coefficient $U_j = m_j$, is assumed to be $p_j = e^{-m_j l_{ij}}$. The probability that $\tilde{x}_{ij}$ photons will emanate from pixel j conditioned on x_{ij} photons entering pixel j conditioned on the attenuation coefficient $U_j = m_j$, can be characterized by a binomial distribution.

$$P\{\tilde{X}_{ij}=\tilde{x}_{ij} \mid X_{ij}=x_{ij} \mid U_j=\mu_j\} = \binom{x_{ij}}{\tilde{x}_{ij}} \left(e^{-\mu_j l_{ij}}\right)^{\tilde{x}_{ij}} \left(1- e^{-\mu_j l_{ij}}\right)^{(x_{ij}-\tilde{x}_{ij})} \qquad \ldots(2)$$

Using eq.(1) the conditional mean value, $g_{ij} = E\{X_{ij} \mid m_1,m_2,\ldots,m_{j-1}\}$ of the photons, x_{ij} entering pixel j can be given by

$$\gamma_{ij} = d_i \left(e^{-\mu_1 l_{i1}}\right)\left(e^{-\mu_2 l_{i2}}\right)\ldots\left(e^{-\mu_{j-1} l_{ij-1}}\right) = d_i\, e^{-\sum_{k=1}^{j-1} \mu_k l_{ik}} \qquad \ldots(3)$$

† Subscript i in m_i has been omitted for simplifying notation

and similarly the conditional mean value, $g_{im} = E\{X_{im} | m_1, m_2, \ldots, m_{m-1}\}$ of the detected photons x_{im} is given by $\gamma_{im} = d_i \, e^{-\sum_{k=1}^{m-1} \mu_k l_{ik}}$. The ratio $\frac{g_{im}}{g_{ij}}$ can be expressed as:

$$\frac{\gamma_{im}}{\gamma_{ij}} = \frac{d_i \, e^{-\sum_{k=1}^{m-1} \mu_k \, l_{ik}}}{d_i \, e^{-\sum_{k=1}^{j-1} \mu_k \, l_{ik}}} = e^{-\sum_{k=j}^{m-1} \mu_k \, l_{ik}} \qquad \ldots(4)$$

The input random variable X_{ij} to the pixel j conditioned on $\{m_1, m_2, \ldots, m_{j-1}\}$ is again Poisson-distributed with mean g_{ij} and the distribution is given by

$$P\{X_{ij} \mid \mathbf{u_{j-1}}\} = P\{X_{ij} \mid \mu_1, \mu_2, \ldots, \mu_{j-1}\} = \frac{e^{-\gamma_{ij}} \gamma_{ij}^{x_{ij}}}{x_{ij}!} \qquad \ldots(5)$$

where the vector $\mathbf{u_{j-1}}$ represents the attenuation coefficients $\{m_1, m_2 \ldots m_{j-1}\}$ up to the jth pixel. The probability of the detected photons X_{im} conditioned on X_{ij} photons entering pixel j can again be given by a binomial distribution

$$P\{X_{im} | X_{ij} | \mathbf{u_{m-1}}\} = \binom{x_{ij}}{x_{im}} \left(e^{-\sum_{k=j}^{m-1} \mu_k l_{ik}}\right)^{x_{im}} \left(1 - e^{-\sum_{k=j}^{m-1} \mu_k l_{ik}}\right)^{x_{ij} - x_{im}}$$

Using the ratio $\frac{g_{im}}{g_{ij}}$ from eq.(4), the above equation can be simplified to yield

$$P\{X_{im} | X_{ij} | \mathbf{u_{m-1}}\} = \binom{x_{ij}}{x_{im}} \left(\frac{\gamma_{im}}{\gamma_{ij}}\right)^{x_{im}} \left(1 - \frac{\gamma_{im}}{\gamma_{ij}}\right)^{x_{ij} - x_{im}} \qquad \ldots(6)$$

To find the probability of X_{ij} conditioned on the detected photons $X_{im} = x_{im} = y_i$ and the attenuation coefficients $\mathbf{u_{m-1}} = \{m_1, m_2 \ldots m_{m-1}\}$, Bayes' theorem as stated by,

$$P\{\mathfrak{X} \mid \mathfrak{Y}\} = \frac{P\{\mathfrak{Y} \mid \mathfrak{X}\} P\{\mathfrak{X}\}}{P\{\mathfrak{Y}\}}$$

is used to express $P\{X_{ij} \mid X_{im} \mid \mathbf{u_{m-1}}\}$ as

$$P\{X_{ij} \mid X_{im} \mid \mathbf{u_{m-1}}\} = \frac{P\{X_{im} \mid X_{ij} \mid \mathbf{u_{m-1}}\} P\{X_{ij} \mid \mathbf{u_{m-1}}\}}{P\{X_{im} \mid \mathbf{u_{m-1}}\}}$$

$$= \left(\frac{x_{ij}!}{x_{im}!(x_{ij}-x_{im})!}\right)\left(\frac{\gamma_{im}}{\gamma_{ij}}\right)^{x_{im}}\left(1-\frac{\gamma_{im}}{\gamma_{ij}}\right)^{x_{ij}-x_{im}} \frac{e^{-\gamma_{ij}}(\gamma_{ij})^{x_{ij}}}{x_{ij}!} \frac{x_{im}!}{e^{-\gamma_{im}}(\gamma_{im})^{x_{im}}}$$

$$= \frac{1}{(x_{ij}-x_{im})!}(\gamma_{ij}-\gamma_{im})^{(x_{ij}-x_{im})} e^{-(\gamma_{ij}-\gamma_{im})} \qquad \ldots(7)$$

The mean of the photons X_{ij} conditioned on the observed detected photons $X_{im} = x_{im} = y_i$ and attenuation coefficients $\mathbf{u_{m-1}}$ can be obtained from by taking the expected value of eq. (7).

$$E\{X_{ij} \mid X_{im}=Y_i=x_{im} \mid \mathbf{u_{m-1}}\} = \sum_{x_{ij}=0}^{\infty} \frac{x_{ij}}{(x_{ij}-x_{im})!}(\gamma_{ij}-\gamma_{im})^{(x_{ij}-x_{im})} e^{-(\gamma_{ij}-\gamma_{im})}$$

$$= \sum_{x_{ij}=0}^{\infty} \frac{x_{ij}+x_{im}}{x_{ij}!}(\gamma_{ij}-\gamma_{im})^{x_{ij}} e^{-(\gamma_{ij}-\gamma_{im})} \qquad \ldots(8)$$

$$= \gamma_{ij}-\gamma_{im}+x_{im} = \gamma_{ij}-\gamma_{im}+y_i \equiv N_{ij}$$

Defining a vector $\mathbf{X_i} = \{X_{im}, X_{im-1}, \ldots, X_{i1}\}^T$, the probability, $P\{\mathbf{X_i} \mid Y_i = X_{im} = x_{im} \mid \mathbf{u_{m-1}}\}$ of $\mathbf{X_i}$ conditioned on the observations $Y_i = X_{im} = x_{im}$ and the attenuation coefficient vector $\mathbf{u_{m-1}}$ can be evaluated by using the Markov property of $\mathbf{X_i}$.

$$P\{\mathbf{X_i} \mid Y_i \mid \mathbf{u_{m-1}}\} = P\{\mathbf{X_i} \mid Y_i=X_{im}=x_{im} \mid \mathbf{u_{m-1}}\}$$

$$= P\{X_{im}, X_{im-1}, \cdots, X_{i1} \mid Y_i=X_{im}=x_{im} \mid \mathbf{u_{m-1}}\}$$

$$= P\{X_{im} \mid X_{im-1} \mid \mathbf{u_{m-1}}\} P\{X_{im-1} \mid X_{im-2} \mid \mathbf{u_{m-2}}\} \cdots P\{X_{i2} \mid X_{i1} \mid \mathbf{u_1}\} P\{X_{i1}\}$$

$$= \left[\prod_{k=1}^{m-1} \binom{x_{ik}}{x_{ik+1}} \left(e^{-\mu_k l_{ik}}\right)^{x_{ik+1}} \left(1-e^{-\mu_k l_{ik}}\right)^{(x_{ik}-x_{ik+1})}\right] \frac{e^{-d_i} d_i^{x_{i1}}}{x_{i1}!}$$

$\ldots(9)$

Defining vectors $\mathbf{X} = \{\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_i}, \ldots\}^T$, $\mathbf{Y} = \{Y_1, Y_2, \ldots, Y_i, \ldots\}^T$, $\mathbf{U} = \{$the set of attenuation coefficients $\mathbf{u_{m-1}}$ for all projections i$\}$ and the fact that $x_{ik+1} = \tilde{x}_{ik}$, and summing eq.(9) over all projections i the following conditional probability can be written.

$$P\left\{\mathbf{X}|\mathbf{Y}|\mathbf{U}\right\} = \prod_i \left[\prod_k \binom{x_{ik}}{\tilde{x}_{ik}} \left(e^{-\mu_k l_{ik}}\right)^{\tilde{x}_{ik}} \left(1-e^{-\mu_k l_{ik}}\right)^{(x_{ik}-\tilde{x}_{ik})}\right] \frac{e^{-d_i} d_i^{x_{i1}}}{x_{i1}!} \quad \ldots(10)$$

and $\ln P\left\{\mathbf{X}|\mathbf{Y}|\mathbf{U}\right\}$ can be written as,

$$\ln P\left\{\mathbf{X}|\mathbf{Y}|\mathbf{U}\right\} = \sum_i \sum_k \left[\tilde{x}_{ik} \ln\left(e^{-\mu_k l_k}\right) + (x_{ik}-\tilde{x}_{ik}) \ln\left(1-e^{-\mu_k l_k}\right)\right] + K \quad \ldots(11)$$

where the constant K represents the logarithm of terms independent of m_k.

The estimate of the attenuation matrix **U** is obtained by maximizing the aposteriori probability $P\left\{\mathbf{U}|\mathbf{X}|\mathbf{Y}\right\}$. Using Bayes' theorem the likelihood function $P\left\{\mathbf{U}|\mathbf{X}|\mathbf{Y}\right\}$ is given by:

$$P\left\{\mathbf{U}|\mathbf{X}|\mathbf{Y}\right\} = \frac{P\left\{\mathbf{X}|\mathbf{Y}|\mathbf{U}\right\} P\{\mathbf{U}\}}{P\left\{\mathbf{X}|\mathbf{Y}\right\}} \quad \ldots(12)$$

and the corresponding log-likelihood function is:

$$\ln P\left\{\mathbf{U}|\mathbf{X}|\mathbf{Y}\right\} = \ln P\left\{\mathbf{X}|\mathbf{Y}|\mathbf{U}\right\} + \ln P\{\mathbf{U}\} - \ln P\left\{\mathbf{X}|\mathbf{Y}\right\} \quad \ldots(13)$$

The maximum likelihood estimator for **U** is obtained by assuming that the attenuation coefficients $\{m_k\}$ are equiprobable in which case $P\{\mathbf{U}\}$ is constant. Since $P\left\{\mathbf{X}|\mathbf{Y}\right\}$ is independent of **U,** maximizing $P\left\{\mathbf{U}|\mathbf{X}|\mathbf{Y}\right\}$ amounts to maximizing the log-likelihood function $\ln P\left\{\mathbf{X}|\mathbf{Y}|\mathbf{U}\right\}$ given by eq.(11). Since eq.(11) contains the unknown values of random variables X_{ik} and $\tilde{X}_{ik}$, direct maximization of eq.(11) will not give meaningful results. Hence the expected value of eq.(11) is to be maximized with respect to $\{m_k\}$. Since the conditional expectations of X_{ik} and $\tilde{X}_{ik}$ are functions of $\{m_k\}$, an iterative method of maximizing eq.(11) has to be developed.

The projection matrix **Y** is a function **g**(**X**) where **g**(·) is a many to one mapping from the space of **X** to the space of **Y**. Hence the iterative solution to the maximum likelihood estimation problem can proceed in alternating steps. of expectation and maximization. First the expected value $E\left[\ln P\left\{\mathbf{X}|\mathbf{Y}|\mathbf{U}\right\}|\mathbf{U}^{(n)}\right]$ is taken conditioned on the current estimate $\mathbf{U}^{(n)}$. This expected value is maximized with respect to **U** to yield an updated estimate $\mathbf{U}^{(n+1)}$. The updated expected value $E\left[\ln P\left\{\mathbf{X}|\mathbf{Y}|\mathbf{U}\right\}|\mathbf{U}^{(n+1)}\right]$ is maximized and

the process repeated. According to Dempster [3] the objective function $E\left[\ln P\left\{X|Y|U\right\}\right]$ will increase in value at each iteration with $E\left[\ln P\left\{X|Y|U\right\}|U^{(n+1)}\right] \geq E\left[\ln P\left\{X|Y|U\right\}|U^{(n)}\right]$. The equality holds only when the maximum likelihood solution $\mathbf{U}_{max}$ is reached and the expected value is $E\left[\ln P\left\{X|Y|U_{max}\right\}\right]$.

Expectation Step

Defining $\mathbf{U}^{(n)}$ as the estimate of the attenuation matrix **U** at the nth iteration, the expectation of $\ln P\left\{X \mid Y \mid \mathbf{U}\right\}$ at the (n+1)st iteration, on the conditionally expected values of X_{ik} and $\tilde{X}_{ik}$ at the nth iteration, can be written as,

$$E\left[\ln P\left\{X|Y|U\right\}|U^{(n)}\right] = \sum_i \sum_k \tilde{N}_{ik} \ln\left(e^{-\mu_k l_{ik}}\right) + \left(N_{ik} - \tilde{N}_{ik}\right) \ln\left(1 - e^{-\mu_k l_{ik}}\right) + K_1 \quad \ldots(14)$$

where K_1 is yet another constant independent of $\{m_k\}$. N_{ik} and $\tilde{N}_{ik}$ conditioned on **Y** and $\mathbf{U}^{(n)}$ are obtained from eq.(8) as,

$$\begin{aligned} N_{ik} &= E\left\{X_{ik} \mid Y_i \mid \mathbf{U}^{(n)}\right\} = \gamma_{ik} - \gamma_{im} + y_{im} \\ \tilde{N}_{ik} &= N_{ik+1} = E\left\{\tilde{X}_{ik} \mid Y_i \mid \mathbf{U}^{(n)}\right\} = \gamma_{ik+1} - \gamma_{im} + y_{im} \end{aligned} \quad \ldots(15)$$

Maximization Step

Since N_{ik} and $\tilde{N}_{ik}$ are known $E\left[\ln P\left\{X|Y|U\right\}|U^{(n)}\right]$ can now be maximized by differentiating eq. (14) with respect to m_k and setting it equal to 0

$$\frac{\partial E\left[\ln P\left\{X|Y|U\right\}|U^{n}\right]}{\partial \mu_k} = \sum_{i \in J_k} \left[-\tilde{N}_{ik} l_{ik} + \left(N_{ik} - \tilde{N}_{ik}\right) \frac{l_{ik}}{\left(e^{\mu_k l_{ik}} - 1\right)}\right] = 0 \quad \ldots(16)$$

Eq. (16) is a transcendental equation that can be solved by approximating the exponential term by using the first three terms in the series expansion,

$$\frac{1}{e^{\mu_k l_{ik}} - 1} = \frac{1}{\mu_k l_{ik}} - \frac{1}{2} + \frac{1}{12} \mu_k l_{ik} + o\left[\left(\mu_k l_{ik}\right)^3\right] \quad \ldots(17)$$

EM Algorithm

Simplification of eq. (16) using eq. (17) yields

$$\underbrace{\left\{\frac{1}{12}\sum_{i\in J_k}\left(N_{ik}-\tilde{N}_{ik}\right)l_{ik}^2\right\}}_{A}\mu_k^2-\underbrace{\left\{\frac{1}{2}\sum_{i\in J_k}\left(N_{ik}-\tilde{N}_{ik}\right)l_{ik}\right\}}_{B}\mu_k+\underbrace{\left\{\sum_{i\in J_k}\left(N_{ik}-\tilde{N}_{ik}\right)\right\}}_{C}=0$$

or, $$A\mu_k^2 - B\mu_k + C = 0 \text{ for all k} \qquad \ldots(18)$$

solving for m_k in the above quadratic yields the (n+1)st estimate $\mu_k^{(n+1)}$. $\mu_k^{(n+1)} = \left.\frac{B \pm \sqrt{B^2 - 4AC}}{2A}\right|_{\mu}$

where the coefficients A, B and C are dependent on $\mu_k^{(n)}$. The smaller root has to be chosen as the solution in this algorithm so that for $\mu_k l_{ik} < 6$ this solution will always fall inside the solution bounds established by Lange [5] for all values of k and i.

Eq.(19) is the desired Expectation-Maximization (EM) algorithm

EM-OSL Algorithm

Eq.(19) represents the maximum likelihood solution with equiprobable attenuation coefficients $\mu_k^{(n+1)}$. This solution can be considerably refined by using prior information $P\{\mathbf{U}\}$ about the attenuation coefficients **U,** yielding the maximum aposteriori solution. In the expected value of eq.(13) shown below,

$$E\left[\ln P\left\{\mathbf{U}\middle|\mathbf{X}\middle|\mathbf{Y}\right\}\right] = E\left[\ln P\left\{\mathbf{X}\middle|\mathbf{Y}\middle|\mathbf{U}\right\}\right] + E\left[\ln P\left\{\mathbf{U}\right\}\right] - E\left[\ln P\left\{\mathbf{X}\middle|\mathbf{Y}\right\}\right]$$

the maximization step using the same iteration schemes as used in eq.(16) becomes

$$\frac{\partial E\left[\ln P\left\{\mathbf{U}\middle|\mathbf{X}\middle|\mathbf{Y}\right\}\middle|\mathbf{U}^{(n)}\right]}{\partial \mu_k} = \frac{\partial E\left[\ln P\left\{\mathbf{X}\middle|\mathbf{Y}\middle|\mathbf{U}\right\}\middle|\mathbf{U}^{(n)}\right]}{\partial \mu_k} - \frac{\partial \ln P\{\mathbf{U}\}}{\partial \mu_k} = 0$$

$$= \sum_{i\in J_k}\left[-\tilde{N}_{ik}l_{ik} + \left(N_{ik}-\tilde{N}_{ik}\right)\frac{l_{ik}}{\left(1-e^{-\mu_k l_{ik}}\right)}\right] - \frac{\partial \ln P\{\mathbf{U}\}}{\partial \mu_k} = 0 \qquad \ldots(20)$$

Using the series approximation of eq.(17) in eq.(20) leads to the equation,

$$\left\{\frac{1}{12}\sum_{i\in J_k}\left(N_{ik}-\tilde{N}_{ik}\right)l_{ik}^2\right\}\mu_k^2-\left\{\frac{1}{2}\sum_{i\in J_k}\left(N_{ik}-\tilde{N}_{ik}\right)l_{ik}-\frac{\partial \ln P(\mathbf{U})}{\partial \mu_k}\right\}\mu_k+\left\{\sum_{i\in J_k}\left(N_{ik}-\tilde{N}_{ik}\right)\right\}=0 \qquad \ldots(21)$$

Using the definitions for A, B and C as given in eq.(17), eq.(21) can rewritten as,

$$A\mu_k^2-\left(B-\frac{\partial \ln P(\mathbf{U})}{\partial \mu_k}\right)\mu_k+C=0 \quad \text{for all k} \qquad \ldots(22)$$

Solving for m_k in the above quadratic yields the result

$$\mu_k^{(n+1)}=\left.\frac{\left(B-\frac{\partial \ln P(\mathbf{U})}{\partial \mu_k^{(n+1)}}\right)\pm\sqrt{\left(B-\frac{\partial \ln P(\mathbf{U})}{\partial \mu_k^{(n+1)}}\right)^2-4AC}}{2A}\right|_{\mu_k^{(n)}}$$

The above equation can not be solved since it involves the derivative of the log-prior distribution at $\mu_k^{(n+1)}$. Green [9] resolves the problem by using the derivative at $\mu_k^{(n)}$ yielding the EM one-step-late algorithm (EM-OSL). In addition a parameter b is introduced in the derivative of the prior P(**U**) to control the influence of the prior on the algorithm. The estimate $\mu_k^{(n+1)}$ at the (n+1)st iteration for the EM-OSL algorithm is,

$$\mu_k^{(n+1)}=\left.\frac{\left(B-\beta\frac{\partial \ln P(\mathbf{U})}{\partial \mu_k^{(n)}}\right)\pm\sqrt{\left(B-\beta\frac{\partial \ln P(\mathbf{U})}{\partial \mu_k^{(n)}}\right)^2-4AC}}{2A}\right|_{\mu_k^{(n)}} \qquad \ldots(23)$$

Again the smaller root of eq.(23) is used in this iterative algorithm for the same convergence reasons as in eq.(19). The optimum value of b will depend upon both the characteristics of the image and the prior used. If b=0 eq.(23) is the EM algorithm of eq.(19).

Prior Distributions P{**U**}

The prior distribution P{**U**} or P{**u**} in eq.(23) has to be carefully chosen. From consideration of Markovian fields a proper choice is the exponential distribution given by $P\{\mathbf{u}\}=\frac{1}{z}e^{-V(\mathbf{u})}$, where $z\equiv\int_{\mathbf{u}\geq 0}e^{-V(\mathbf{u})}\,d\mathbf{u}<\infty$ is a normalization constraint that V(**u**) has to satisfy The potential function V{**u**} is designed to penalize large differences in

estimated parameters for neighboring pixels. The constraints on the potential function have been discussed by Lange [12}. In this particular application a first order neighborhood interaction of the form

$$V(\mu) \equiv \sum_{\{j,\, k\} \in N} w_{jk}\, v\{\mu_j - \mu_k\}, \qquad \ldots(24)$$

is considered where the weighting functions w_{jk} for a first order Markovian field are defined as

$$w_{jk} = \begin{cases} 1 & : \text{orthogonal neighbor} \\ \frac{1}{\sqrt{2}} & : \text{diagonal neighbor} \\ 0 & : \text{otherwise} \end{cases}$$

In this paper two potential functions v(r) are considered. The first is a sigmoidal function with a positive parameter x given by,

$$v(r) = \frac{2}{1 + e^{-\xi r^2}} - 1 \qquad \ldots(25)$$

and the second is a logarithm of hyperbolic cosine with a parameter x given by

$$v(r) = \ln\cosh(\xi r) = \ln \frac{e^{\xi r} + e^{-\xi r}}{2} \qquad \ldots(26)$$

The parameter x = 0 in both of these potential functions amounts to a straight EM algorithm without any prior.

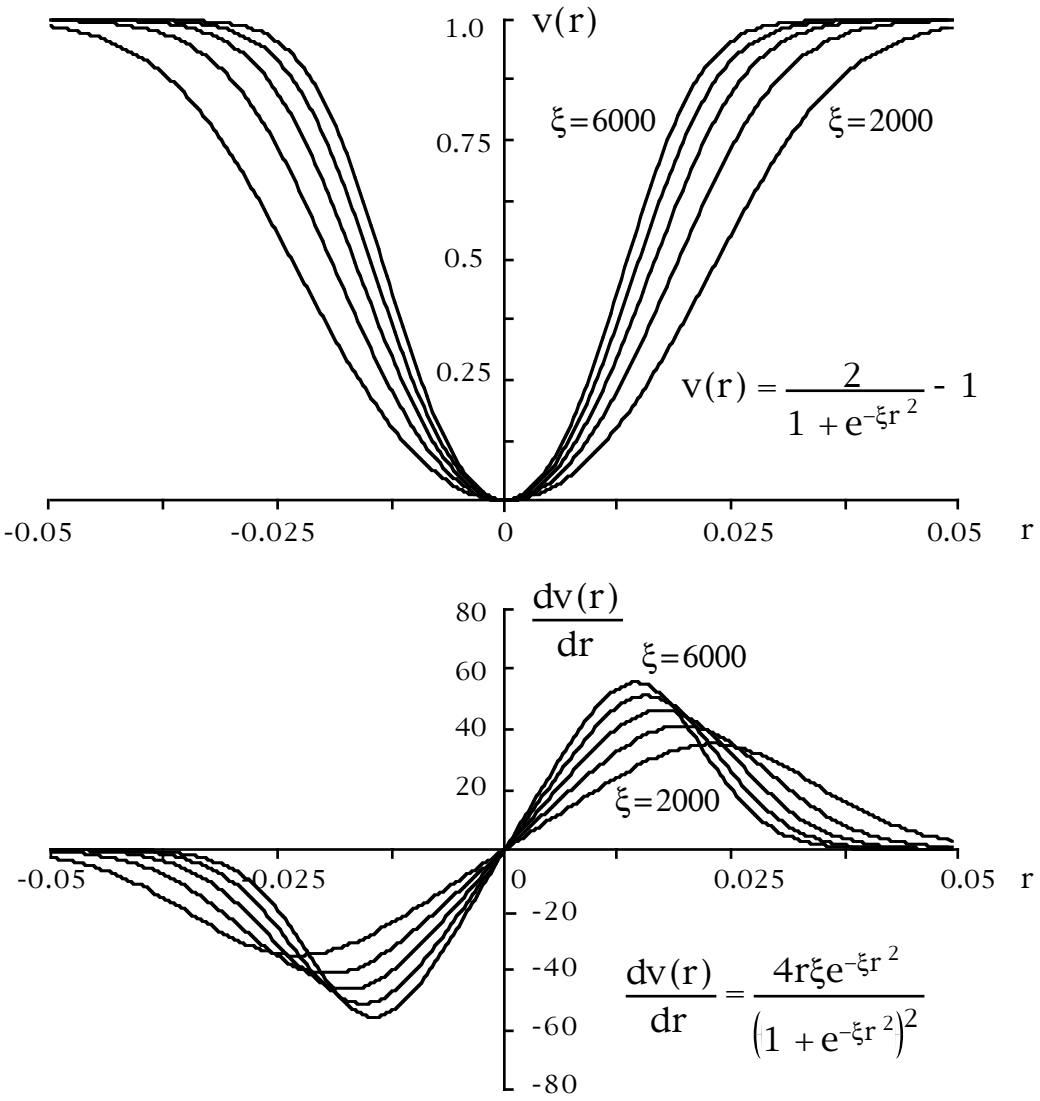


Figure 2 : Sigmoidal Functions and Their First Derivatives

Figure 2 shows the sigmoidal functions and their derivatives used in this study. The choice of the sigmoidal function with an adjustable parameter x (varying between 2000 and 6000) makes it possible to smooth out speckles in reconstructed images while retaining sharp edges. This fact can be demonstrated analytically from the series expansion of eq.(23).

Defining $B_1 = B - \beta \frac{\partial \ln P(U)}{\partial \mu_k^{(n)}}$ eq.(23) can be expressed as

$$\mu_k^{n+1} = \frac{B_1}{2A} - \frac{\sqrt{B_1^2 - 4AC}}{2A}$$

$$= \frac{C}{B_1}\left[1 + \frac{AC}{B_1^2} + \frac{2A^2C^2}{B_1^4} + \frac{5A^3C^3}{B_1^6} + \ldots\right]$$

In this investigation $\frac{AC}{B_1^2} < 1$ and the first order approximation for μ_k^{n+1} can be given by

$$\mu_k^{n+1} \cong \frac{C}{B_1} = \frac{\sum_{i \in J_k} \left(N_{ik} - \tilde{N}_{ik}\right)}{\frac{1}{2}\sum_{i \in J_k} \left(N_{ik} + \tilde{N}_{ik}\right) l_{ik} - \beta \left.\frac{\partial \ln P(U)}{\partial \mu_k^{(n)}}\right|_{\mu_k^n}}$$

$$= \frac{\sum_{i \in J_k} \left(N_{ik} - \tilde{N}_{ik}\right)}{\frac{1}{2}\sum_{i \in J_k} \left(N_{ik} + \tilde{N}_{ik}\right) l_{ik} + \beta \sum_{j \in N} w_{ij} \left.\frac{\partial v\left(\mu_k^{(n)} - \mu_j^{(n)}\right)}{\partial \mu_k^{(n)}}\right|_{\mu_k^n}} \quad \text{for all } k \qquad \ldots(27)$$

N ≡ Neighboring pixels

The denominator of eq.(27) involves the *sum* of two terms, one of which is the derivative of the potential function. This sum directly determines the penalty imposed on successive iterations in this algorithm. The relatively higher intensity of pixel k than those of its nearest neighbors is reflected by a positive value of the derivative and hence a larger value of the denominator. As a consequence, the penalty for the next iteration of μ_k^{n+1} is higher resulting in a lower estimate for the intensity of pixel k. On the other hand, a lower intensity of pixel k than those of its nearest neighbors results in a higher estimate of the intensity of pixel k. The derivative of a sigmoidal function exhibits a band (Figure 2) of high (0-0.05) and low (-0.05-0) values dependent on the parameter x. Because of this characteristic, the penalty imposed is effective only when intensity differences fall into this band. As a consequence the unusually high differences in intensity such as sharp-edged features of the image are preserved while the speckled appearance of the image due to Poisson noise is smoothed. As shown in Figure 2 the degree of edge-separation needed can be adjusted with the parameter x for any particular image reconstruction.

The derivative of the lncosh function on the other hand monotonically increases. As a consequence, there is no discrimination between edges and artifacts resulting in smoothed edges.

COMPUTER SIMULATION

A computer simulation was performed to investigate the quality of reconstructed images using the sigmoid potential function and to compare it with other reconstructed images. The simulation block diagram is shown in Figure 3.

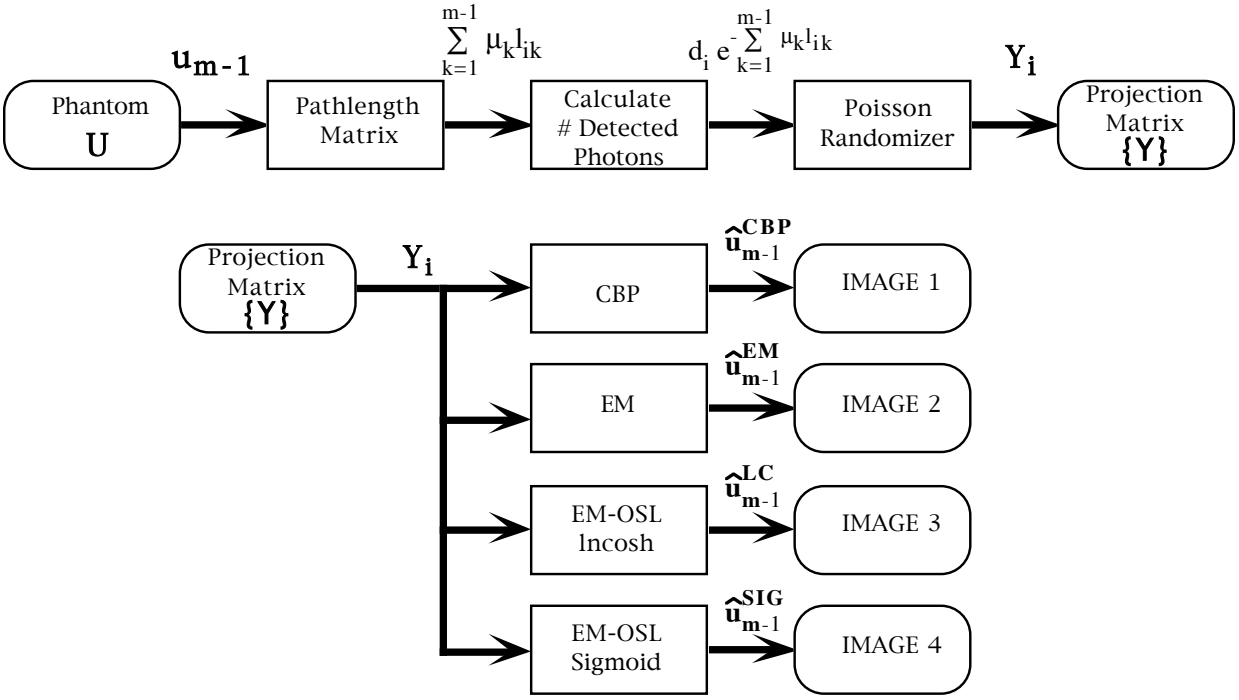


Figure 3: Simulation diagram

A phantom was created with parameters as shown in Figure 4. The numerical values show the uniform attenuation coefficients in that region.

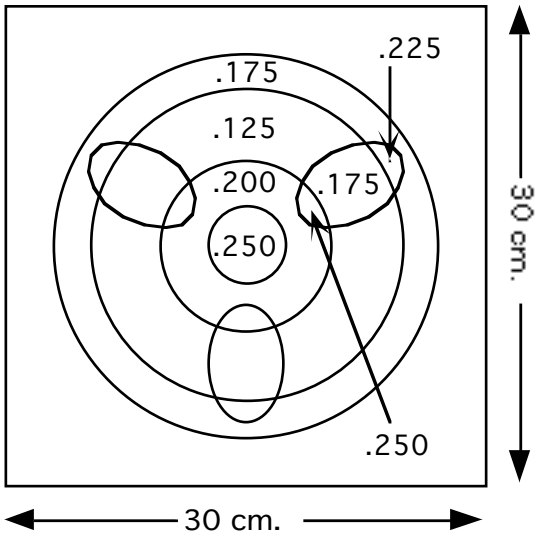


Figure 4: Phantom Parameters

This particular phantom was chosen so that it has a fairly complicated detail including a high contrast outer shielding ring. In the usual reconstructions, especially from projections with low number of photon counts, this shielding ring obscures the delicate features present inside it. The intention in this study was to assess how well the sigmoidal potential function handled the presence of the shielding ring. The photon source was assumed to emit a relatively low number of $d_i = 10^4$ expected number of photons per projection path i. The phantom was divided into cells of 64¥64 pixels.

A major problem was in the implementation of the coefficient matrix $[l_{ij}]$ that has a size of $64^2 ¥ 64^2$. $[l_{ij}]$ is a collection of lengths of the path along the projection path i inside pixel j. This is a highly sparse matrix, as each path i passes through (Figure 1) only a small fraction of the total number of pixels in the image and each pixel has only a small fraction of the total number of paths passing through it. Furthermore, computation for each projection path i requires access to all corresponding l_{ij}. Even though only a fraction of elements of this matrix is nonzero, calculation of all of them on a fly for each iteration would still require an excessive amount of computation. To reduce this computational burden the path lengths are calculated only once and stored in a form that is easily accessible. This stored data is defined as the ‘Pathlength Matrix.’

A semi-sparse matrix scheme is devised to handle the problem of computational burden. Two sets of path coefficients are stored. The first set is accessible in the 'forward' direction, that is, given the path i all coefficients associated with it are stored. This is defined as the ‘Forward Pathlength Matrix.’ The second set is accessible in the

'backward' direction, that is, given pixel j all coefficients associated with it are stored. This is defined as the 'Backward Pathlength Matrix.'

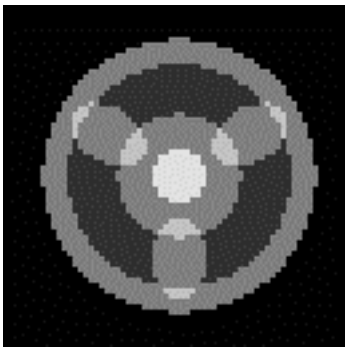
Forward Pathlength Matrix is used to generate noise-free projection data. This matrix in conjunction with a Poisson random number generator routine is used to generate Poisson-randomized data. Poisson random number is generated from a given intensity parameter and acquisition time. This random number is equal to a total number of events which have exponentially distributed interarrival times with mean equal to the given intensity parameter d occurring during the acquisition time. Exponential random variables are generated from a set of random variables uniformly distributed over [0,1] by probability transformation.

The data from the Poisson randomizer is the projection matrix **Y.** From this projection matrix **Y** four different types of reconstruction images are generated. The first of the reconstructed images used the conventional deterministic convolution backprojection algorithm. In the other three images using the EM algorithm, a combination of Forward and Backward Pathlength matrices are used in reconstructing the image. N_{ik} and $\tilde{N}_{ik}$ defined in eq.(15) are computed by using the Forward Pathlength matrix. The computed N_{ik} and $\tilde{N}_{ik}$ are then in turn utilized to determine the updated estimated attenuation coefficient μ_k^{n+1} in eq.(19) and (23) using the Backward Pathlength matrix. Three images are reconstructed using 1) conventional EM algorithm proposed by Lange [12], 2) the Modified EM with lncosh potential function proposed by Green [9], and 3) the Modified EM with sigmoidal potential function. Performances of these algorithms and their reconstructions are then compared. The criteria used for comparison are:

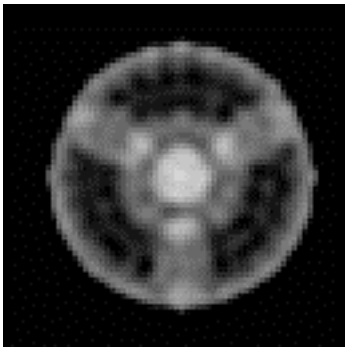
- Convergence and convergence rate
- Image quality (primarily visual tempered with analytical criteria)
- Complexity (actual execution time)

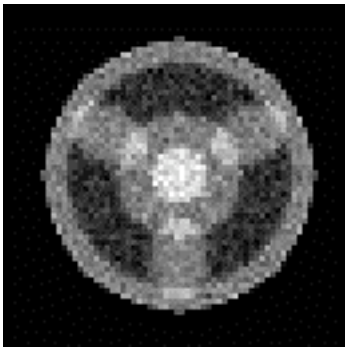
Noise-free projection data was generated over 64 equally spaced projection angles between 0 and 180° using parallel beam scanner geometry. Each projection consisted of 64 equally spaced rays over the field of view of 30 cm. in length. Poisson-randomized projection data was generated by using the previous noise-free data as mean value. The Poisson random number generator algorithm as described previously is used for this purpose.

The CBP method was implemented on an Apollo Domain series 3500 system. All EM-based algorithms were implemented on the Alliant FX-2800 series system which has both concurrent and vector processing capability. Each iteration takes about 18 seconds. This may be slow but is not deemed a serious drawback because the speed of iteration can be considerably enhanced with a multiprocessor machine. The gray-scale images presented herein are windowed between 0.150 cm^{-1} and 0.275 cm^{-1} for better visual contrast.

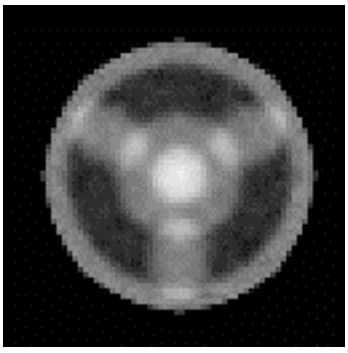


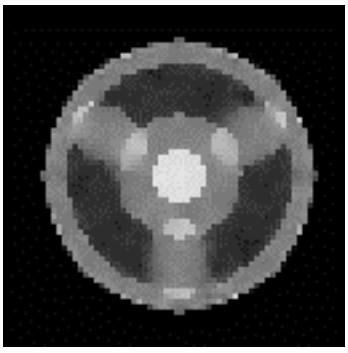
(a) Phantom





(b) CBP **(c) EM**





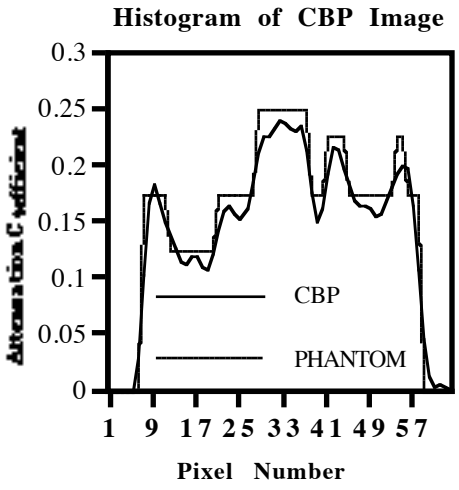
(d) LnCosh **(e) Sigmoid**

Figure 5: Phantom and Reconstructed Images

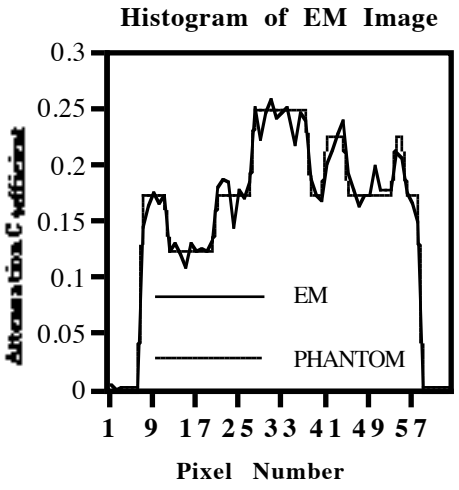
RESULTS AND CONCLUSIONS

Figure 5 shows (a) the phantom image, (b) reconstruction with the convolution back projection (CBP), (c) reconstruction with the conventional EM (d) reconstruction with EM-OSL using lncosh potential function, and (e) reconstruction with EM-OSL using sigmoidal potential function. Cross-sectional histograms of reconstructed images sectioned vertically through the center with cross-sectional histogram of the phantom superimposed are shown in Figure 6. In Figure 7, the progression of RMS errors between the reconstructed images and the phantom as the iteration proceeds are drawn for CBP, EM, lncosh and sigmoid realizations.

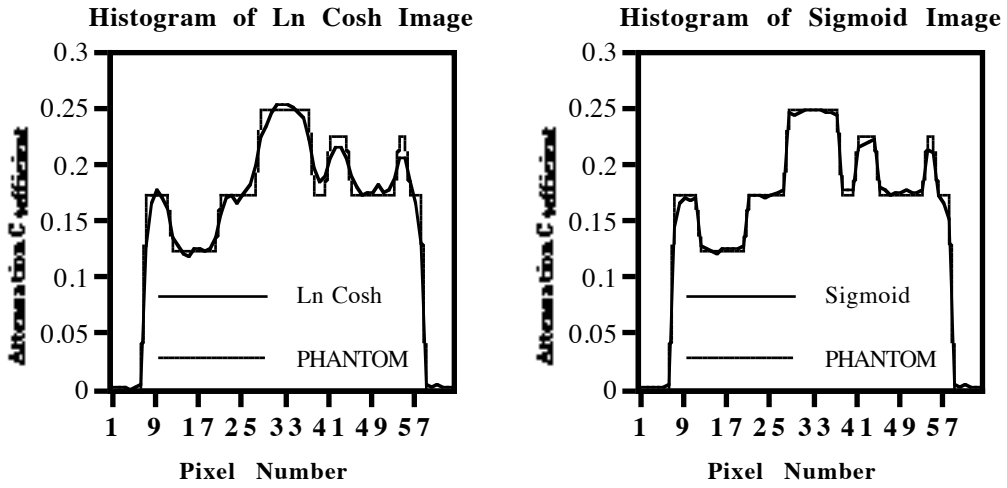
The conventional convolution back projection (CBP) reconstruction (Figure 5.b) achievable in a single step is the quickest in execution time because of its deterministic nature. However, it is the poorest reconstruction among all the four. This poor performance is caused by factors such as intrinsic statistical variation of data, low photon counts, low angular and spatial sampling rates and shielding ring effects. The cross sectional histogram (Figure 6.a) shows well rounded edges and the RMS error (Figure 7) is quite high.



(a) CBP Image



(b) EM Image



(c) EM-OSL Lncosh Image **(d) EM-OSL Sigmoidal Image**

Figure 6: Cross Sectional Histograms

By using a statistical model, better results are obtained for the EM algorithms (Figures 5.c,d,e). However, the straight EM algorithm (Figure 5.c) has a rather grainy and speckled appearance. This graininess is simply a manifestation of low correlation between parameter estimates for neighboring pixels. The cross sectional histogram (Figure 6.b) shows badly jagged edges and the RMS error (Figure 7) remains fairly constant after about 40 iterations. There is also the problem that this algorithm tends to diverge (Rajeevan [13]) after a certain number of iterations and stopping rules are to be devised for the optimum number of iterations.

To smooth the speckles in the EM procedure, the algorithm was modified using Gibbs' prior with a lncosh potential function to penalize large differences between adjacent pixels. The lncosh reconstruction (Figure 5.d) shows a smoother result but the cross-sectional histogram (Figure 6.c) indicates rounded edges. The RMS error curve (Figure 7) is also better than before but tends to saturate at a little higher level. The difficulty stems from the fact that the lncosh potential function tends to oversmooth the image resulting in the loss of sharp-edged features.

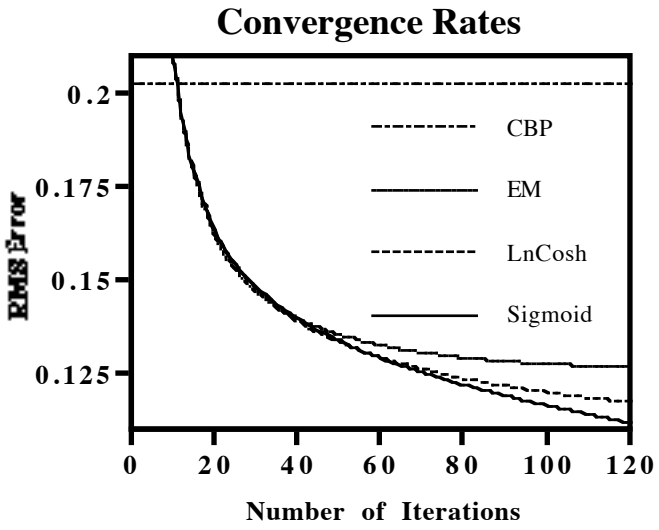


Figure 7: Convergence Rates

To overcome the oversmoothing difficulty a sigmoidal potential function that preserves sharp edges while retaining the smoothing effect is utilized. Convergence curves for increasing iterations were plotted using log-likelihood function and RMS error for different values of b and x. b was varied between 0 and 0.0004 and x was varied between 0 and 7000. From these curves the optimum vaue of b was 0.0002 and the optimum value of x was 5000. With these values, the reconstructed image (Figure 5.e) is superior in quality to the rest. The image is visually very close to the phantom. The cross-sectional histogram (Figure 6.d) also indicates a better edge performance. The RMS error curve monotonically decreases as the iteration proceeds and at the120th iteration is almost negligible. Thus, the modified EM algorithm with the sigmoidal potential function clearly performs better with regard to accuracy and smoothness of reconstruction.

DISCUSSION OF RESULTS

In comparing reconstructions in medical imaging visual accuracy is the most important. The analytical criteria like the RMS error is secondary. The RMS error curves (Figure 7) indicate a big difference between the CBP reconstruction and the EM reconstructions. But visual inspection of the reconstructed images do not show such a big discrepancy. On the other hand the error curves for EM and lncosh reconstructions are not that far apart; but visually the differences are quite discernible. The sigmoidal reconstruction is more appealing both visually and analytically.

The error curves also show that during the first 30 iterations all the EM algorithms converge almost linearly at the same rate. But after that, their convergence become sublinear and their graphs begin to fan-out separately. Among all the four convergence rates presented the sigmoidal algorithm again performs better.

The speed of each iteration is dependent on the number of processors in the computer and can be considerably enhanced with more processors. The structure of this algorithm is more conducive to a multiprocessor machine.

The convergence rates and the quality of reconstruction indicate the decided superiority of this modified EM algorithm using sigmoidal potential functions. Further studies are in progress for SPECT and missing data reconstructions.

REFERENCES

[1] Rockmore, A.J. and A. Macovski, "A Maximum Likelihood Approach to Emission Image Reconstruction from Projections," *IEEE Trans. Nucl. Sci.*, vol. NS-23, pp. 1428-1432, 1976

[2] Rockmore, A.J. and A. Macovski, "A Maximum Likelihood Approach to Transmission Image Reconstruction from Projections," *IEEE Trans. Nucl. Sci.*, vol. NS-24, pp. 1929-1935, 1977

[3] Dempster, A.P., N.M. Laird, and D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. Roy. Statist. Soc.*, B, vol. 39, pp. 1-38, 1977

[4] Shepp, L.A. and Y. Vardi, "Maximum Likelihood Reconstruction for Emission Tomography," *IEEE Trans. Med. Imaging*, vol. MI-1, pp. 113-121, 1982

[5] Lange, K. and R. Carson, "EM Reconstruction Algorithms for Emission and Transmission Tomography," *J. Comput. Assist. Tomog.*, vol. 8, No. 2, pp. 306-316, 1984

[6] Geman, S. and D.E. McClure, "Bayesian Image Analysis: An Application to Single Photon Emission Tomography," *Proc. Amer. Statist. Assoc. Stat. Comp.*, pp. 12-18, 1985

[7] Geman, S. and D.E. McClure, "Statistical methods for Tomographic Image Reconstruction," *Bull. Int. Statist. Inst.*, vol. LII-4, pp. 5-21, 1987

[8] Hebert, T. and R. Leahy, "A Generalized EM Algorithm for 3-D Bayesian Reconstruction from Poisson Data Using Gibbs Priors," *IEEE Trans. Med. Imaging*, vol. 8, No. 2, pp. 194-202, 1989

[9] Green, P.J., "Bayesian Reconstructions from Emission Tomography Data Using a Modified EM Algorithm," IEEE Trans. Med. Imaging, vol. 9, No. 1, pp. 84-93, 1990

[10] Geman, S. and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721-741, 1984

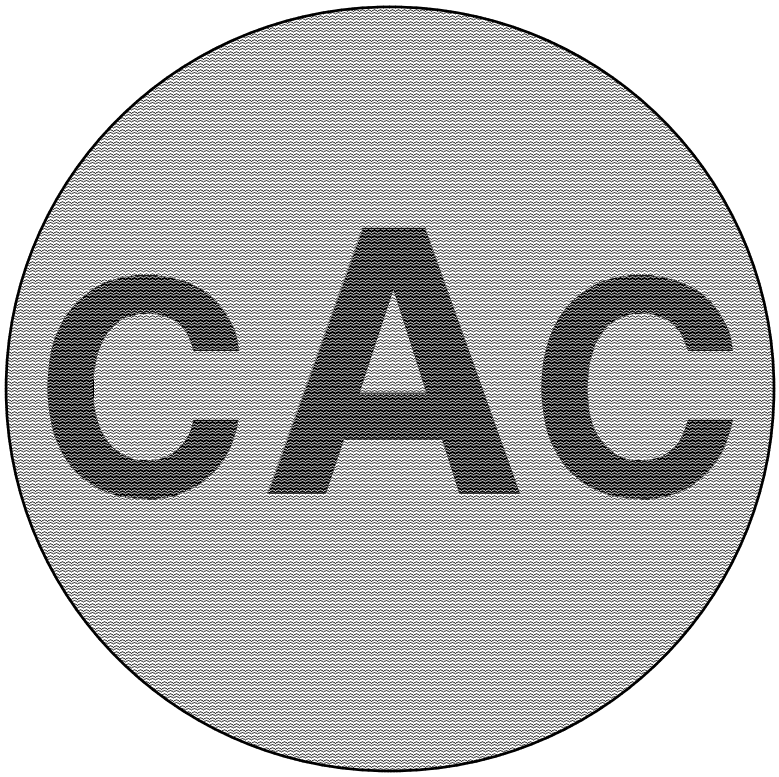
[11] Besag, J., "Spatial Interaction and the Statistical Analysis of Lattice System," J. Roy. Statist. Soc., B, vol. 36, pp. 192-236, 1974

[12] Lange, K., "Convergence of EM Image Reconstruction Algorithms with Gibbs Smoothing," *IEEE Trans. Med. Imaging*, vol. 9, No. 4, pp. 439-446, 1990

[13] Rajeevan, N., "A Stochastic Estimation Approach to Emission Tomography," Ph.D. Thesis, EE Dept., Indian Institute of Science, Bangalore, India, 1991

MODIFIED EM ALGORITHM FOR TRANSMISSION TOMOGRAPHY

PATIWAT KAMONPET
VENKATARAMA KRISHNAN



Technical Report CAC T94-01

Parts of this report has appeared as a Chapter in the book
Probability and Random Processes published by John Wiley & Sons Inc., 2006

CENTER FOR ADVANCED COMPUTATION
UNIVERSITY OF MASSACHUSETTS LOWELL