Emission Expectation Maximization Look-Alike Algorithms for X-Ray CT and Other Applications

Gengsheng L. Zeng

Department of Engineering, Weber State University, Ogden, Utah 84408, USA. Department of Radiology and Imaging Sciences, University of Utah, Salt Lake City, Utah 84108, USA

Gengsheng L. Zeng: larryzeng@weber.edu, larry.zeng@hsc.utah.edu

Abstract

Purpose—In emission tomography, the EM (expectation maximization) algorithm is easy to use with only one parameter to adjust — the number of iterations. On the other hand, the EM algorithms for transmission tomography are not so user-friendly and have many problems. This paper develops a new transmission algorithm similar to the emission EM algorithm.

Methods—This paper develops a family of emission-EM-look-alike algorithms by expressing the emission EM algorithm in the additive form and changing the weighting factor. One of the family members can be applied to transmission tomography such as the x-ray CT (computed tomography).

Results—Computer simulations are performed and compared with a similar algorithm by a different group using the transmission CT noise model. Our algorithm has the same convergence rate as theirs, and our algorithm provides better contrast-to-noise ratio for lesion detection.

Conclusions—For any noise variance function, an emission EM look-alike algorithm can be derived. This algorithm preserves many properties of the emission EM algorithm such as multiplicative update, non-negativity, faster convergence rate for the bright objects, and ease of implementation.

Keywords

EM algorithm; Iterative image reconstruction; Transmission tomography; Convergence rate; X-ray CT

I. INTRODUCTION

The EM (expectation maximization) methodology is a general approach to compute maximum likelihood estimates by using iterative techniques.¹ There are many EM algorithms. The most famous EM algorithm in medical imaging community is the one for emission tomography.^{2–5} The emission EM algorithm uses a multiplicative form to update the image; it has a built-in property to enforce the image non-negativity and Poisson noise nature in the data. It is efficient to implement and stable. It has no adjustable parameters

other than the iterations. It is safe to state that it is the most favorite iterative algorithm in nuclear medicine.

An EM algorithm for transmission tomography was developed in Reference 5 by Lange and Carson. Unlike its emission tomography counterpart, this EM algorithm for transmission tomography has many drawbacks. It is complicated to compute and slow in convergence.⁶ There are other versions of the transmission EM algorithms, but there are no fundamental improvements.^{7–10}

In 2001, Nuyts *et al* proposed a method to scale the measurements making the scaled version somewhat like the emission data with Poisson noise in the sense that the variance is approximately the mean value.¹¹ Our proposed method here is very similar, except that their scaling factor was a function of the measured sinogram value while our proposed scaling factor is a function of the forward projection of the image reconstructed at the previous iteration.

Since the original transmission EM algorithm, many user-friendly algorithms have been developed.^{12,13} A faster algorithm may not be in the form of an EM algorithm.¹⁴ An EM algorithm can also be applied to other applications such as simultaneous estimation of the emission activity and the attenuation map in PET (positron emission tomography).¹⁵

This paper develops a family of emission-EM-look-alike algorithms. They are iterative algorithms in the form of multiplicative image update, which intrinsically enforces the image non-negativity. The unique feature of this family is that the scaling factor is formed by the forward projection of the reconstructed image at the previous iteration, which is a unique feature in the “E-step” in an EM algorithm. Each member of the family has its own noise model as explained in detail in the next section.

II. Methods

2.1 The emission EM algorithm

The starting point of our development is the emission EM algorithm

$$x_i^{(k+1)} = \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{P_j}{\sum_n a_{jn} x_n^{(k)}} \right), \quad (1)$$

where $x_i^{(k)}$ is the i^{th} image pixel at the k^{th} iteration, P_j is the j^{th} line-integral (ray-sum) measurement value, and a_{ji} is the contribution of the i^{th} image pixel to the j^{th} measurement. The summation over the second index is the projector and the summation over the first index is the backprojector. Expression (1) is in the form of multiplicative image update, and it can be re-written in the form of additive image update

$$\begin{aligned}
x_i^{(k+1)} &= \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{p_j}{\sum_n a_{jn} x_n^{(k)}} \right) \quad (2) \\
&= \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{p_j - \sum_n a_{jn} x_n^{(k)} + \sum_n a_{jn} x_n^{(k)}}{\sum_n a_{jn} x_n^{(k)}} \right) \\
&= \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{p_j - \sum_n a_{jn} x_n^{(k)}}{\sum_n a_{jn} x_n^{(k)}} \right) + \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{\sum_n a_{jn} x_n^{(k)}}{\sum_n a_{jn} x_n^{(k)}} \right) \\
&= \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \frac{1}{\sum_n a_{jn} x_n^{(k)}} (p_j - \sum_n a_{jn} x_n^{(k)}) + x_i^{(k)} \\
&= x_i^{(k)} + \lambda_i^{(k)} \sum_j a_{ji} w_j^{(k)} (p_j - \sum_n a_{jn} x_n^{(k)})
\end{aligned}$$

The last line of (2) is in the form of the iterative Landweber algorithm, which is a gradient decent algorithm. In (2),

$$\lambda_i^{(k)} = \frac{x_i^{(k)}}{\sum_j a_{ji}} \quad (3)$$

is the relaxation parameter, which is also known as the step size, and

$$w_j^{(k)} = \frac{1}{\sum_n a_{jn} x_n^{(k)}} \quad (4)$$

is the weighting factor for the j th projection ray. For the Poisson noise model, the noise variance is the mean value of the ray sum. The denominator of the right-hand-side of (4) is the “E-step” estimation of the variance of the sinogram variance for emission tomography by using the forward projection.

We can make two observations from (3) and (4). First, the EM algorithm’s step size $\lambda_i^{(k)}$ is scaled by the image pixel value $x_i^{(k)}$ at the k th iteration. The step is larger for objects with larger image values. Therefore, the bright lesions converge faster than the dark lesions.

Second, the weighting factor $w_j^{(k)}$ is the reciprocal of the estimated mean value of the j th ray by the reconstruction at the k th iteration. The mean value is the same as the variance for Poisson noise.

2.2 Modification of the emission EM algorithm

We propose to introduce a new scaling factor $s_j^{(k)}$ in the backprojector for each projection ray j . Thus the original emission EM algorithm (1) becomes

$$x_i^{(k+1)} = \frac{x_i^{(k)}}{\sum_j a_{ji}s_j^{(k)}} \sum_j a_{ji}s_j^{(k)} \left(\frac{p_j}{\sum_n a_{jn}x_n^{(k)}} \right). \quad (5)$$

As a consequence, the additive version (2) becomes

$$\begin{aligned} x_i^{(k+1)} &= \frac{x_i^{(k)}}{\sum_j a_{ji}s_j^{(k)}} \sum_j a_{ji}s_j^{(k)} \left(\frac{p_j}{\sum_n a_{jn}x_n^{(k)}} \right) \\ &= \frac{x_i^{(k)}}{\sum_j a_{ji}s_j^{(k)}} \sum_j a_{ji}s_j^{(k)} \left(\frac{p_j - \sum_n a_{jn}x_n^{(k)} + \sum_n a_{jn}x_n^{(k)}}{\sum_n a_{jn}x_n^{(k)}} \right) \\ &= \frac{x_i^{(k)}}{\sum_j a_{ji}s_j^{(k)}} \sum_j a_{ji}s_j^{(k)} \left(\frac{p_j - \sum_n a_{jn}x_n^{(k)}}{\sum_n a_{jn}x_n^{(k)}} \right) \\ &\quad + \frac{x_i^{(k)}}{\sum_j a_{ji}s_j^{(k)}} \sum_j a_{ji}s_j^{(k)} \left(\frac{\sum_n a_{jn}x_n^{(k)}}{\sum_n a_{jn}x_n^{(k)}} \right) \\ &= x_i^{(k)} + \frac{x_i^{(k)}}{\sum_j a_{ji}s_j^{(k)}} \sum_j a_{ji}s_j^{(k)} \frac{1}{\sum_n a_{jn}x_n^{(k)}} (p_j - \sum_n a_{jn}x_n^{(k)}) \\ &= x_i^{(k)} + \hat{\lambda}_i^{(k)} \sum_j a_{ji}\hat{w}_j^{(k)} (p_j - \sum_n a_{jn}x_n^{(k)}) \end{aligned} \quad (6)$$

with a new step size (also known as relaxation parameter) and a new weighting factor:

$$\hat{\lambda}_i^{(k)} = \frac{x_i^{(k)}}{\sum_j a_{ji} s_j^{(k)}} \quad (7)$$

and

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}}. \quad (8)$$

Eq. (5) represents a family of emission-EM-look-alike algorithms depending on the definition of the new scaling factor $s_j^{(k)}$, as explained in the following special cases (or examples).

2.3 Special case 1: No weighting

Let's consider a hypothetical imaging system, in which the noise in the sinogram is identically distributed with the same variance. No noise weighting should be used for the image reconstruction algorithm. Thus we require the weighting factor (8) to be

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}} = 1 \quad (9)$$

which results in

$$s_j^{(k)} = \sum_n a_{jn} x_n^{(k)}. \quad (10)$$

Substituting (10) into (5) yields

$$\begin{aligned} x_i^{(k+1)} &= \frac{x_i^{(k)}}{\sum_j a_{ji} \left(\sum_n a_{jn} x_n^{(k)} \right)} \sum_j a_{ji} \left(\sum_n a_{jn} x_n^{(k)} \right) \left(\frac{p_j}{\sum_n a_{jn} x_n^{(k)}} \right) \\ &= x_i^{(k)} \frac{\sum_j a_{ji} p_j}{\sum_j a_{ji} \sum_n a_{jn} x_n^{(k)}}. \end{aligned} \quad (11)$$

In (11), $\sum_j a_{ji} p_j$ is the backprojection of the measured sinogram and $\sum_j a_{ji} \sum_n a_{jn} x_n^{(k)}$ is the backprojection of the forward projection of the current estimate of the image.

2.4 Special case 2: Modified Emission

Let's consider another hypothetical imaging system, in which the noise in the sinogram is not quite Poisson distributed, but the variance is given as

$$\text{var}(p_j) = \bar{p}_j^\gamma, \quad (12)$$

where $\bar{p}_j$ is the mean value of the sinogram p_j and γ is a constant. When $\gamma = 1$, this case is the classic situation of emission tomography using a Poisson noise model.

Using the EM strategy of estimating $\bar{p}_j$ by $\sum_n a_{jn} x_n^{(k)}$, the weighting factor (8) for our modified emission case is given as

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}} = \frac{1}{(\sum_n a_{jn} x_n^{(k)})^\gamma}, \quad (13)$$

which results in

$$s_j^{(k)} = (\sum_n a_{jn} x_n^{(k)})^{1-\gamma}. \quad (14)$$

Substituting (14) into (5) yields

$$\begin{aligned} x_i^{(k+1)} &= \frac{x_i^{(k)}}{\sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)})^{1-\gamma}} \sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)})^{1-\gamma} \left(\frac{p_j}{\sum_n a_{jn} x_n^{(k)}} \right) \quad (15) \\ &= \frac{x_i^{(k)}}{\sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)})^{1-\gamma}} \sum_j a_{ji} \frac{p_j}{(\sum_n a_{jn} x_n^{(k)})^\gamma}. \end{aligned}$$

As expected, when $\gamma = 1$, (15) is the famous emission EM algorithm.

2.5 Special case 3: Transmission

In transmission tomography, the sinogram variance is proportional to the exponential function of the sinogram's mean value, that is,¹⁶

$$\text{var}(p_j) \propto \exp(\bar{p}_j). \quad (16)$$

The weighting factor can be chosen as the reciprocal of the variance. Using the EM strategy of estimating $\bar{p}_j$ by $\sum_n a_{jn} x_n^{(k)}$, the weighting factor (8) for the transmission tomography case is given as

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}} = \exp\left(-\sum_n a_{jn} x_n^{(k)}\right), \quad (17)$$

which results in

$$s_j^{(k)} = \left(\sum_n a_{jn} x_n^{(k)}\right) \exp\left(-\sum_n a_{jn} x_n^{(k)}\right). \quad (18)$$

Substituting (18) into (5) yields

$$x_i^{(k+1)} = x_i^{(k)} \frac{\sum_j a_{ji} p_j \exp\left(-\sum_n a_{jn} x_n^{(k)}\right)}{\sum_j a_{ji} \left(\sum_n a_{jn} x_n^{(k)}\right) \exp\left(-\sum_n a_{jn} x_n^{(k)}\right)}. \quad (19)$$

Eq. (19) is the main result of this paper. It is an emission EM look-alike algorithm for the transmission tomography. It is in the form of multiplicative image update, has an intrinsic non-negativity constraint, and weights the sinogram with (17). Most importantly, (19) is user friendly and easy to implement. The initial values of Algorithm (19) must be positive.

2.6 Special case 4: Modified Transmission

Again, this is another hypothetical imaging system, in which the noise in the sinogram has a variance given by

$$\text{var}(p_j) \propto \exp(\gamma \bar{p}_j), \quad (20)$$

where $\bar{p}_j$ is the mean value of the sinogram p_j and γ is a constant. When $\gamma = 1$, this case is the classic situation of transmission tomography using a Poisson noise model for its pre-log data.

According to the noise model (20), the weighting factor can be chosen as

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}} = \exp(-\gamma \sum_n a_{jn} x_n^{(k)}), \quad (21)$$

which results in

$$s_j^{(k)} = (\sum_n a_{jn} x_n^{(k)}) \exp(-\gamma \sum_n a_{jn} x_n^{(k)}). \quad (21)$$

Substituting (21) into (5) yields

$$x_i^{(k+1)} = x_i^{(k)} \frac{\sum_j a_{ji} p_j \exp(-\gamma \sum_n a_{jn} x_n^{(k)})}{\sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)}) \exp(-\gamma \sum_n a_{jn} x_n^{(k)})}. \quad (22)$$

When $\gamma = 1$, this degenerates to the usual transmission tomography case (19). When $\gamma = 0$, this is the no-weighting case as (11).

From the examples above, a large family of emission EM look-alike image reconstruction algorithms can be developed as long a noise variance function is provided.

2.7 Computer Simulations

In this paper, computer simulations are used to compare the performance of our proposed algorithm with a similar algorithm developed by Nuyts *et al* by using transmission CT noise model. Gaussian noise was added to the sinogram data. The computer simulations in this paper use a two-dimensional (2D) uniform elliptical phantom containing 3 small hot lesions and 2 small cold lesions, as shown in Figure 1. The large ellipse (with a linear attenuation coefficient of 0.001/unit) had a length of the horizontal semi-axis of 204.8 and a length of the vertical semi-axis of 128. The 3 small circular hot lesions (with a linear attenuation coefficient of 0.004/unit) had a radius of 10.24, 5.12, and 2.56, respectively. The 2 small circular cold lesions (with a linear attenuation coefficient of 0.0004/unit) had a radius of 5.12 and 2.56, respectively. The unit is the pixel size. The detector bin size was the same as the image pixel size. The large ellipse is also referred to as the background.

The performance is characterized by the lesion contrast recovery coefficient (CRC), which is defined as¹⁸

$$CRC = \frac{\left(\frac{M_{les}}{M_{back}}\right)_{rec} - 1}{\left(\frac{M_{les}}{M_{back}}\right)_{phan} - 1} \quad (23)$$

where M_{les} and M_{back} represent the mean of the lesion and the mean of the background, and the subscripts '*rec*' and '*phan*' denote the reconstruction and the true phantom, respectively. In computer simulations, we calculated the CRC for each of the five lesions for every reconstruction using the noiseless projections. The lesion value M_{les} was the value at the center of the lesion. The background value M_{back} was the value at the center of the large disc. For the hot lesions $(M_{les}/M_{back})_{phan} - 1 = (0.002/0.001) - 1 = 1$. For the cold lesions $(M_{les}/M_{back})_{phan} - 1 = (0.0004/0.002) - 1 = -0.8$.

The image noise was evaluated using the normalized standard deviation of the noise in the central 70×70 square region of the large disc reconstructed with the noisy data as

$$\frac{\sigma}{M_{back}} = \frac{1}{M_{back}} \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x}_i)^2} \quad (24)$$

where σ is the standard deviation, N is the number of image elements used in the calculation, x_i is the value of the i th pixel of the image reconstructed from noisy data, and $\bar{x}_i$ is the expected mean value of the i th pixel. The purpose of the normalization is to eliminate the influence on the noise measurement of non-uniform values of the image within the regions that are supposed to be uniform.

A parallel-beam imaging geometry is assumed. The image array was 512×512 , the number of views was 400 over 180° , and the number of detection channels was 512. To avoid inverse-problem crime, the projector used to generate projection sinograms is different from the projector used for image reconstruction. In sinogram data generation, the image was first up-sampled 5 times and the detector was also up-sampled 5 times.¹⁷ The noiseless sinogram data was then down-sampled 5 times and noise was introduced. Negative sinogram values were set to zero. Transmission CT noise model was used in sinogram data with extremely low counts, in which the sinogram variance was proportional to the exponential function of the sinogram value.

Computer simulations were performed and compared with a similar algorithm by a different group using the transmission CT noise model.

III. Results

Three algorithms have been implemented for the transmission CT noise model with extremely low counts, because for the high count situations the performance is similar for similar algorithms. Two of the three algorithms are our proposed algorithms, and the third algorithm was developed by Nuyts *et al.*¹¹ The only difference among these three algorithms is the calculation of the scaling factor $s_j^{(k)}$ in (8). In one of our algorithms, the scaling factor is calculated using the forward projection of the current image, that is,

$$s_j^{(k)} = \left(\sum_n a_{jn} x_n^{(k)} \right) \exp \left(- \sum_n a_{jn} x_n^{(k)} \right). \quad (25)$$

In the algorithm by Nuyts *et al*, the scaling factor is calculated using the measured sinogram, that is,

$$s_j^{(k)} = \max (p_j, \varepsilon) \exp \left(- p_j \right) \quad (26)$$

where the purpose of using a small threshold $\varepsilon = 0.01$ is to prevent the scaling factor from being zero.

There is a third algorithm that is also presented in this result section, to show that none of the above algorithm is optimal in the sense of lesion contrast-to-noise ratio, because the third algorithm outperforms the two algorithms discussed above. This third algorithm is obtained by trial-and-error and is a hybrid combination of the above two methods. For the third method (referred to as “Mix” method), the scaling factor is giving as

$$s_j^{(k)} = \max (p_j, \varepsilon) \exp \left(- \sum_n a_{jn} x_n^{(k)} \right). \quad (27)$$

It is important to note that (27) cannot be changed to

$$s_j^{(k)} = p_j \exp \left(- \sum_n a_{jn} x_n^{(k)} \right), \quad (28)$$

because when $s_j^{(k)} = 0$, the measurement p_j is never used in the image reconstruction.

These three algorithms gave almost the same images as shown in Figure 2, all at iteration number 50. However, the differences in performance can be better revealed by the CRC curves for each lesion, as shown in Figures 3, 4, 5, 6, and 7, respectively. Those curves are produced by varying the number of iterations. The “good points” in the curves are close to

the upper left corner, where the contrast is high and the noise is low. The “bad points” in the figures are close to the lower right corner, where the contrast is low and the noise is high. We observe that our proposed emission-EM lookalike method outperforms the Nuyts Method, and the mixed method is the best of the three, in terms of lesion contrast-to-noise ratio.

Convergence rate is an essential property of an iterative algorithm. Figure 8 shows the learning curves of the three algorithms. A learning curve is the common logarithm of the Euclidean norm of the discrepancy that is defined as the difference between the measurements and the forward projection values as a function of the iteration number. The shape of the learning curve represents the convergence rate of an iterative algorithm. It is observed from Figure 8 that all these three algorithms have approximately the same convergence rate. This observation is expected, because these three scaling factors have essentially the same magnitude. The differences are the small fluctuations, which do not affect the convergence rate as measured in the sinogram domain. Due to noise, the sinogram domain discrepancy does not reduce to zero. The converged discrepancy value is different for different scaling strategies.

Our algorithm has the same convergence rate as that of the other group, and our algorithm provides better contrast-to-noise ratio for lesion detection.

The emission-EM lookalike algorithms use small and conservative update step sizes. The emission-EM lookalike algorithms use the multiplicative image update scheme:

$x_i^{(k+1)} = x_i^{(k)}(\text{factor})$; therefore they can be sped up by introducing a power function

$x_i^{(k+1)} = x_i^{(k)}(\text{factor})^\gamma$ with a proper $\gamma > 1$. This speedup strategy may destroy a good

property of the emission EM algorithm which conserves the total forward projection counts at every iteration. One remedy is to re-normalize the image according to the total count at every iteration.^{19, 20}

IV. Discussion and conclusions

The emission EM algorithm has wide applications in nuclear medicine due to its good noise control property, intrinsic non-negativity constraint, and ease of implementation. The emission EM algorithm is a nonlinear algorithm and its convergence rate is non-stationary. If written in the additive update form, it is easy to observe that the step size of the algorithm depends on the image intensity. It converges faster for bright objects and slower for dark objects. Its noise weighting factor is estimated by the previous iteration of the image solution, while many other algorithms (such as the similar algorithm developed by Nuyts *et al*¹¹) use the sinogram data to calculate the weighting factors.

All these properties mentioned above are preserved by our newly suggested emission-EM look-alike algorithms, which are developed in an *ad hoc* manner. For any noise variance function, an emission-EM look-alike algorithm can be easily devised by following the steps in Section II.

One important application is the algorithm for transmission tomography. According to our computer simulations, the proposed method's convergence rate is comparable to that of the algorithm by Nuyts *et al* and the proposed method's lesion contrast-to-noise ratio is better. The images reconstructed by Nuyts method is noisier might be due to the noise propagation from the scaling factor into the reconstructed image. In fact, Nuyts et al already realized this problem and suggested a remedy of applying a lowpass filter to smooth the scaling fact. However, our proposed method does not need this lowpass filter.

We do not claim that the emission-EM look-alike algorithms are the optimal algorithms. The “mix” algorithm implemented in the Results section uses a scaling factor which is a hybrid combination of the forward projection of the image from the previous iteration and the sinogram measurement actually gives better results. Therefore, it is still an open research problem to find the optimal noise weighing strategy.

Acknowledgments

Research reported in this publication was supported by the National Institute Of Biomedical Imaging And Bioengineering of the National Institutes of Health under Award Number R15EB024283. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from complete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B.* 39(1):1–38.1977;
2. Richardson WH. Bayesian-based iterative method of image restoration. *JOSA.* 62(1):55–59.1972; DOI: 10.1364/JOSA.62.000055
3. Lucy LB. An iterative technique for the rectification of observed distributions. *Astronomical Journal.* 79(6):745–754.1974; DOI: 10.1086/111605
4. Shepp LA, Vardi Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imag.* 1:113–122.1982; DOI: 10.1109/TMI.1982.4307558
5. Lange K, Carson R. EM reconstruction algorithms for emission and transmission tomography. *J Comp Ass Tomogr.* 8:302–316.1984;
6. Lange K, Fessler JA. Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Trans Imag Proc.* 4:1430–1438.1995;
7. Kent, JT, Wright, C. Some suggestions for transmission tomography based on the EM algorithm. In: Barone, MP, Frigessi, A, editors *Stochastic Models, Statistical Methods, and Algorithms in Image Analysis*. New York: Springer; 1992. 219–232. 74 of *Lecture Notes in Statistics*
8. Browne JA, Holmes TJ. Developments with maximum likelihood x-ray computed tomography. *IEEE Trans Med Imag.* 12:40–52.1992;
9. Ollinger JM, Johns G. The use of maximum a-posteriori and maximum likelihood transmission images for attenuation correction PET. *Proc IEEE Nucl Sci Symp Med Imag Conf.* 2:1185–1187.1992;
10. Ollinger JM. Maximum likelihood reconstruction of transmission images in emission computed tomography via the EM algorithm. *IEEE Trans Med Imag.* 13:89–101.1994;
11. Nuyts J, Michel C, Dupont P. Maximum-likelihood expectation-maximization reconstruction of sinograms with arbitrary noise distribution using NEC-transformations. *IEEE transactions on medical imaging.* 20(5):365–375.2001; [PubMed: 11403196]
12. Lange K, Fessler JA. Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Transactions on Image Processing.* 4(10):1430–1438.1995; [PubMed: 18291974]

13. Fessler JA, Fiacaro EP, Clinthorne NH, Lange K. Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction. *IEEE transactions on medical imaging*. 16(2):166–175.1997; [PubMed: 9101326]
14. Yu Z, Thibault JB, Bouman CA, Sauer KD, Hsieh J. Fast model-based X-ray CT reconstruction using spatially nonhomogeneous ICD optimization. *IEEE Transactions on image processing*. 20(1):161–175.2011; [PubMed: 20643609]
15. Mihlin A, Levin CS. An expectation maximization method for joint estimation of emission activity distribution and photon attenuation map in PET. *IEEE transactions on medical imaging*. 36(1): 214–224.2017; [PubMed: 27576244]
16. Zeng GL, Wang W. Noise weighting with an exponent for transmission CT. *Biomedical Physics & Engineering Express*. 2:045004.2016;
17. Kaipio J, Somersalo E. Statistical inverse problems: Discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics*. 198(2):493–504.2007;
18. Liow JS, Strother SC. Practical tradeoffs between noise, quantitation, and number of iterations for maximum likelihood-based reconstructions. *IEEE Trans Med Imaging*. 10:563–571.1991; [PubMed: 18222862]
19. Hwang DS, Zeng GL. Convergence study of an accelerated ML-EM algorithm using bigger step size. *Phys Med Biol*. 51:237–252.2006; [PubMed: 16394336]
20. Tanaka E, Nohara N, Tomitani T, Yamamoto M. Utilization of non-negativity constraints in reconstruction of emission tomograms. in *Proc 9th Conference of Information Processing in Medical Imaging*. :379–393.1985

Figure 1.
The computer generated lesion phantom with 3 hot lesions and 2 cold lesions.

Figure 2.

Images reconstructed from sinogram with transmission CT noise model. Left image: 25 iterations of the proposed algorithm. Right image: 25 iterations of the algorithm by Nuyts *et al.* Middle image: 25 iterations of our “mix” method expressed by (27). Each image is displayed from its own minimum to its own maximum.

Figure 3. Contrast recovery coefficient (CRC) vs. normalized noise standard deviation curves for three algorithms. The contrast is measured at the largest hot lesion. The Mix method gives the best performance and the Nuyts gives the poorest performance.

Figure 4. Contrast recovery coefficient (CRC) vs. normalized noise standard deviation curves for three algorithms. The contrast is measured at the mid-size hot lesion. The Mix method gives the best performance and the Nuyts gives the poorest performance.

Figure 5. Contrast recovery coefficient (CRC) vs. normalized noise standard deviation curves for three algorithms. The contrast is measured at the small hot lesion. The Mix method gives the best performance and the Nuyts gives the poorest performance.

Figure 6. Contrast recovery coefficient (CRC) vs. normalized noise standard deviation curves for three algorithms. The contrast is measured at the small cold lesion. The Mix method gives the best performance and the Nuyts gives the poorest performance.

Figure 7. Contrast recovery coefficient (CRC) vs. normalized noise standard deviation curves for three algorithms. The contrast is measured at the mid-size cold lesion. The Mix method gives the best performance and the Nuyts gives the poorest performance.

Figure 8. Learning curves for three algorithms. The three algorithms have approximately the same convergence rate.