A CASE STUDY IN STATISTICAL IMAGE PROCESSING: POSITRON-EMISSION TOMOGRAPHY[a]

Donald L. Snyder[b], David G. Politte[b], and Michael I. Miller[b]

ABSTRACT

This paper traces the development of methods for producing estimates of the spatial distribution of a positron-emitting radionuclide within the body.

[a] This work was supported by the National Institute of Health under grants RR01380, from the Division of Research Resources, and HL13851 and by the National Science Foundation under grants MIP-8722463 and ECE-8552518.

[b] Washington University, St. Louis, Missouri 63130.

Introduction

Positron-emission tomography is a technique in which images are formed of the spatial distribution within the body of a positron-emitting radionuclide. Systems for accomplishing this have existed since the early 1970s, but it is only recently that accurate mathematical models have been postulated for the physical phenomena and instrumentation involved. These models are now being used with the estimation theory for spatial Poisson processes to define new methods for processing data to produce more accurate images of radionuclide distributions. The new methods are very demanding computationally, so it is fortuitous that the required computations can be organized for implementation on the massively parallel, multiprocessor computers that are becoming available at the same time.

In this paper, we review the mathematical model for data acquired with a positron-emission tomograph. We then indicate the confidence-weighted method, which is the one in current use, and the maximum-likelihood method, which is under intense current study, for producing images of radioactivity distributions. Approaches are described for addressing the issues of noise, edge enhancement, and high computational load which at first seem to preclude the practical usefulness of the maximum-likelihood method. Our discussion is rather narrow because we focus on developments in our group at Washington University, and we have not attempted a broad review of the many developments by others in the field.

The instrumentation

The Super PETT-I positron-emission tomograph shown in Figure 1 was designed and constructed under the direction of Dr. M. M. Ter-Pogossian in the Division of Radiation Sciences of the W. U. Mallinckrodt Institute of Radiology [1]. It was completed in 1982 and is now located in the Department of Cardiology. The tomograph is used to acquire data that are processed to form an estimate, which is displayed as a two-dimensional image, of the spatial distribution of a positron-emitting radioactive tracer within the body. Such images are used not so much for viewing anatomical structures but, rather, for regional quantification of biophysical and biochemical processes, such as blood-flow rate and glucose-utilization rate, and for mapping regions of the brain that become active when external stimuli, such as a light flash, are applied. The estimation of rate parameters and brain-mapping studies are very demanding because multiple images showing how the radioactivity distribution changes with time are needed.

Shown in Figure 2 is the Super PETT-I tomograph during its construction. Cylindrically shaped scintillation detectors used to sense photons emanating from the radioactivity are arranged in a ring geometry surrounding the patient port. There are four rings, each having 96 closely spaced detectors in the completed instrument. A ring defines a planar section through the radioactivity being imaged. In addition, *cross-plane* data are also acquired permitting a total of seven planar sections through the activity to be imaged concurrently.

A radioactive decay at some location in the activity creates a positron which travels a short distance, about 3 mm in water, where it interacts with an electron in an annihilation that produces two 511 keV photons propagating at the velocity of light in nearly opposite directions along a line. A detection event occurs when two photons are sensed in opposing scintillation detectors within a small time interval. The data acquired for such an event consists in part of the line-of-flight of the annihilation photons, as defined

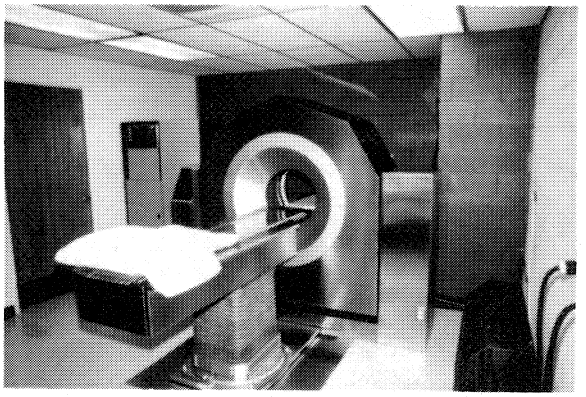

Figure 1. The Super PETT-I Positron-Emission Tomograph

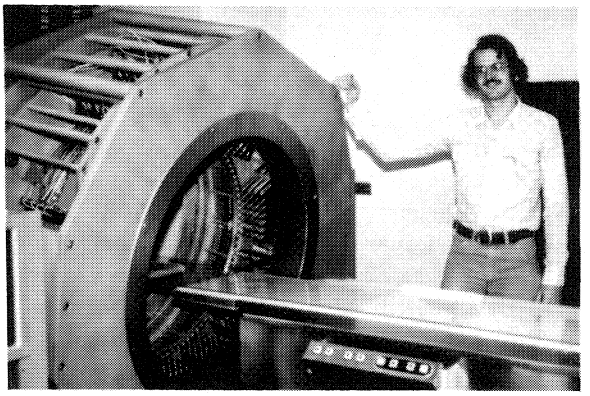

Figure 2. The Super PETT-I Tomograph During Construction. Partially completed detector rings can be seen

by a line connecting the two detectors involved. Also, high speed timing circuitry is used to measure the differential propagation time of the photons. Thus, data (t, d, θ) acquired for a single detection event are three dimensional, consisting of two parameters (d, θ) that define the line-of-flight by its perpendicular distance d from the system axis and the angle θ made by the line with the abscissa and a third parameter t that is the differential time-of-flight.

The measurement of these parameters is imperfect. Errors in measuring the differential time-of-flight occur because of the finite resolution of the electronic timing-circuitry. These errors have been measured experimentally and are reasonably approximated by a Gaussian density with a full-width-at-half-maximum (FWHM) of about 7.5 cm corresponding to a 500 ps timing resolution [3]. Errors in the direction transverse to the line-of-flight occur because of the finite size of the detectors. These have also been measured experimentally and are reasonably well approximated by a Gaussian density with a FWHM value of about 1 cm. Thus, the measurement errors for a detection event

are described by a two-dimensional Gaussian density with asymmetric concentration ellipsoids having their major axis aligned with the line-of-flight of the photons being detected. Provided the event rate is not too high, measurement errors may be assumed to be independent for separate events [2, Ch. 3].

A large amount of data is acquired to form images, and this amount is even greater in newer instruments. For Super PETT-I, the (t, d, θ) parameters for each event are quantized and stored in a three-dimensional array of dimension $(40, 128, 96)$. Thus, the number of data values to be processed for an image is $40 \times 128 \times 96 = 491,520$. In all, there are then somewhat less than 3.5 million data values for the seven images to be formed. Super PETT-II is the newest instrument, of which there are two models, one for head studies and one for body studies. The number of data values for the head unit is $24 \times 128 \times 192 = 589,824$ per plane, or about 4.1 million values altogether. For the body unit, there are $32 \times 216 \times 320 = 2,211,840$ data values per plane, or about 15.5 million totally. Thus, there are much data to process, and for some neurological studies being performed, the seven images should be available in about five minutes.

Mathematical model

The mathematical model which we use is given by Snyder, Thomas, and Ter-Pogossian [3]; Vardi, Shepp, and Kaufman [4] describe a similar model in discrete form. There is an *emission space* X and a *measurement space* Y where two point processes occur. One, denoted by $\{N(A), A \in \sigma(X)\}$ and called the *emission point process*, describes annihilations, and the other, denoted by $\{M(A), A \in \sigma(Y)\}$ and called the *measurement point process*, describes the data collected from which estimates are formed. A point of the emission point process has a coordinate at the location x of a positron-electron annihilation. It is common to treat the three-dimensional distribution of activity as a collection of two-dimensional distributions defined by the planes of the detector rings, in which case the coordinate of a point of the emission point process is two-dimensional. From the physics of radioactive decay, the emission point process may be taken to be an inhomogeneous Poisson process with an intensity function $\{\lambda(x), x \in X\}$ that is proportional to the concentration of radiotracer.

Points in the emission space are translated to and become points in the measurement space where they represent measured data related to the underlying and unobservable points in the emission space. The measurement space is three dimensional, with each point in it having coordinates $y = (t, d, \theta)$. The translation of a point from the emission space to the measurement space is random with a conditional density $p(y|x)$, called the *point spread function*, that is the two-dimensional asymmetric Gaussian density describing measurement errors oriented at angle θ, centered at x, and scaled by $1/\pi$.[c] Because emission points form a Poisson process and translation errors are independent, points on the measurement space form a Poisson process with an intensity $\mu(y)$ given by

$$\mu(y) = \int_X p(y|x)\lambda(x)m_x(dx). \tag{1}$$

The actual situation is more complicated than this model indicates. Some points leaving the emission space do not become points in the measurement space. One reason

[c] The factor of $1/\pi$ occurs because flight lines of annihilation photons are equidistributed in angle.

for this is that the conversion of a photon into a light flash in the scintillation crystal of the detector is statistical, occurring with a probability less than one. Another source for loss occurs because one or both of the photons created in an annihilation can undergo Compton scatter as they propagate through the tissues and bones of the body, in which case they lose energy and change flight direction. The fraction of annihilations which are deleted is so large that the effect must be taken into account for acceptable images to result; 30% to 50% deletions is not uncommon. There are also some points appearing in the measurement space that are not translated there from the emission space [5]. These are called *randoms*. One way for a random to occur is for two photons to be sensed within a small time interval and be accepted as arising from a common annihilation but originating instead from two distinct annihilations. There are very many randoms in collected data, so this effect must also be taken into consideration for acceptable images; depending on the level of activity, upwards of 30% of the detected events are randoms. The common practice is to *correct* measured data for these two effects. A constant is subtracted to correct for randoms, and the difference is scaled upwards in a nonuniform manner to correct for attenuation, nonuniformity in detector sensitivity, and wobble-cycle dwell time. This correction procedure invalidates the assumed Poisson statistics for the measurement point process. While more fundamental methods for treating these effects have been known for some time [6], it is only recently that we have had the opportunity to begin incorporating them into the image formation process directly in order to learn what benefits there may be in not pre-correcting the data in the usual manner. For the remainder of this paper, the simple translation model without insertions and deletions will be used on the assumption that the data are perfectly pre-corrected. This is thought not to be a major limitation for positron-emission tomography because the effects can be determined with reasonable accuracy through the use of a separate measurement to determine attenuation. As an aside, the situation in single photon emission tomography is quite different in this respect. Here, the effects of attenuation are severe and the frequent practice of assuming attenuation to be a known effect that can be corrected is a deficiency in many mathematical descriptions of this imaging problem.

The loglikelihood functional for the Poisson process on the measurement space, as a functional of the intensity on the emission space, is given by

$$l(\lambda) \doteq -\int_Y \mu(y) m_y(dy) + \int_Y \ln[\mu(y)] M(dy), \tag{2}$$

where $\mu(y)$ as a functional of $\lambda(x)$ is given in (1).

Imaging problem

The imaging problem is simply stated: we are given the above model and the locations of points on Y, and we wish to estimate $\lambda(x)$ for all $x \in X$.

Image formation methods

The *confidence-weighted* and *maximum-likelihood* methods are two approaches which can be used for producing the desired estimate. The confidence weighted method [3,7] is the one presently used on all positron-emission tomographs that acquire time-of-flight information. This method produces an estimate which is a linear functional of the data; since this functional becomes an inverse Radon transform as the point-spread

function tends to an impulsive fence, the confidence-weighted method may be thought of as generalizing Radon inversion methods to accommodate nonideal line integrals. The confidence-weighted method is readily and economically implemented to produce estimates rapidly, and substantial experience has been acquired by neurologists, cardiologists, and other practitioners in interpreting the estimates displayed as images. For these reasons, estimates produced by alternative methods must have a superior performance for them to be adopted, and the degree of improvement must increase with the difficulty in producing them. To describe the confidence-weighted method, let (u, θ) represent a measurement point (t, d, θ) in rectangular coordinates, where

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} t \\ d \end{pmatrix}. \tag{3}$$

Then, the estimate $\hat{\lambda}_{cw}(x)$ produced with the confidence-weighted method is realized in two steps, the first of which forms the *preimage* $c(x)$ according to

$$c(x) = \int_0^{\pi} \int_U p(u|x,\theta) M(du, d\theta), \tag{4}$$

where $M(du, d\theta)$ are the measurements organized by flight-line angle and expressed in rectangular coordinates. Because $p(u|x,\theta)$ is a Gaussian density, this step may be interpreted as placing the mean of the Gaussian error-density at each measurement point with the major axis oriented along the flight line, with the resulting densities for all points then summed. Thus, $c(x)$ is a form of Parzen-Rosenblatt kernel-estimate of the intensity of the point process on the measurement space, averaged over angles. The estimate is formed from the preimage according to the convolution

$$\hat{\lambda}_{cw}(x) = \int_X h(x - x') c(x') m(dx'), \tag{5}$$

where the kernel $h(x)$ satisfies

$$\int_X h(x - x') \left[\int_Y p(y|x') p(y|x'') m_y(dy) \right] m_x(dx') = w(x - x''), \tag{6}$$

in which $w(x)$ is a preselected window function used to stabilize $h(x)$ and control image noise; for example, the Fourier transform of $w(x)$ may be obtained by truncating the tails of a Gaussian function to zero at the Nyquist frequency consistent with the discretization used to compute estimates.

The maximum-likelihood method implemented via the expectation-maximization algorithm has been under intense study since Shepp and Vardi [8] published their paper on the subject in 1982. It is straightforward to extend this method for tomographs in which time-of-flight data are collected [9]. A naive application of this idea results in a sequence of estimates defined iteratively by

$$\hat{\lambda}^{(k+1)}(x) = \hat{\lambda}^{(k)}(x) \int_Y \left[\frac{p(y|x)}{\int_X p(y|x') \hat{\lambda}^{(k)}(x') m_x(dx')} \right] M(dy). \tag{7}$$

Estimates produced by this method appear at first to have better resolution and lower noise than those produced from the same data with the confidence-weighted method,

which, in turn, are superior to those obtained from the same data with the filtered back-projection method when the time-of-flight information in the data are ignored [10,11]. However, as iterations proceed and estimates of ever increasing likelihood are obtained, the displayed images of the estimates begin to deteriorate, eventually becoming unacceptable to practitioners as they are much worse than those of the confidence-weighted method.

As experience with the maximum-likelihood method has grown, three difficulties with it have been recognized and frequently articulated in publications and conferences. There is a *noise artifact* and an *edge artifact* which become increasingly severe as iterations proceed, and the computations required to implement the iterations in (7) are so demanding that the estimates may take too long to produce for them to be usable in practice no matter how much better they may be. It is our conclusion that the noise artifact is due to the use of unconstrained maximum-likelihood estimation and can be greatly reduced by introducing suitable constraints. The edge artifact is due to the finite numerical precision in a computer implementation of (7) and can be reduced to acceptable levels by restricting resolution. The issue of long computation time can be addressed effectively through the use of specially designed hardware that exploits both the structure of (7) and contemporary computer architectures.

Noise and edge artifacts

Snyder and Miller [12] and Snyder, Miller, Thomas, and Politte [13] discuss the noise and edge artifacts encountered with the use of the unconstrained maximum-likelihood method. The noise artifact appears, with increasing severity as iterations of the expectation-maximum algorithm proceed, as high peaks and low valleys seemingly randomly distributed throughout the image. The effect is most easily demonstrated in one-dimensional computer simulations, as shown in Figure 3. Points of a Poisson process on a one-dimensional emission space were generated in a Monte Carlo simulation. The intensity, shown as "INT" in the figure, was uniform on the interval $(-2.5, +2.5)$. Measurement errors were introduced according to a Gaussian point spread function having a 1 cm FWHM. The unconstrained maximum-likelihood estimate, generated according to (7) with 5000 iterations and displayed as "INT_EST" in the figure, is very rough with the high peaks and low valleys of the noise artifact. This artifact arises fundamentally because the loglikelihood functional $l(\lambda)$ in (2) is unbounded, so a well behaved maximizer of it does not exist. The noise artifact can be greatly suppressed through the use of Grenander's method of sieves [14]. For this, the estimate of $\lambda(x)$ is constrained to lie in a sieve space S defined in terms of a sieve-kernel $s(x|u)$ according to

$$S = \{\lambda(x) : \lambda(x) = \int_U s(x|u)\psi(u)du\}, \tag{8}$$

where $\psi(u)$ is an intensity function on the sieve space S. With this approach, we typically use a Gaussian convolution-kernel sieve, and in [12,13], we show how the sieve-constrained maximum-likelihood estimate can be produced numerically via the expectation-maximization algorithm. As an aside, we have also made preliminary investigations of the use of penalty methods, based on Good's roughness measure, for controlling the noise artifact [12,15]; we are continuing to investigate this alternative approach.

The noise artifact is prominent for a low number of measured points. As the number increases, however, the edge artifact becomes more dominant. Shown in Figure 4 is the

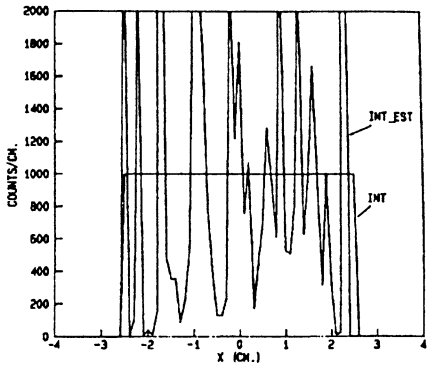

Figure 3. Demonstration of the Noise Artifact (from [13])

estimate produced by 5000 iterations of (7) for mean-value data, which corresponds to an infinite number. This was obtained by replacing $M(dy)$ in (7) by $\mu(y)dy$ for the one-dimensional simulation described for Figure 3. The Gibbs' like edge overshoot is the edge artifact, which in images is seen as a distracting high ridge that follows sharp transitions in the underlying activity distribution. The degree of overshoot is exquisitely sensitive to implementation choices. The result labeled "INT_EST.6SIG" and showing an overshoot of about 18% in the figure was obtained when the Gaussian kernel $p(y|x)$ in (7) was truncated in the numerical implementation to zero for values of $|y-x|$ greater than six standard deviations of the density. The result labeled "INT_EST.3SIG" shows a dramatic increase to about 35% overshoot when the truncation to zero is at three standard deviations. The edge artifact is fundamental because the inverse problem defined by (1) is ill posed. In [13], we have described a way to eliminate the edge artifact at the expense of reduced resolution by estimating not the intensity $\lambda(x)$ on the emission space but, rather, a blurred version of the intensity $d(x)$ defined in terms of a resolution kernel $r(w|x)$ according to

$$d(w) = \int_X r(w)|x)\lambda(x)dx\,. \tag{9}$$

In [13], we show that the maximum-likelihood estimate of $d(w)$ can be produced numerically with the expectation-maximization algorithm.

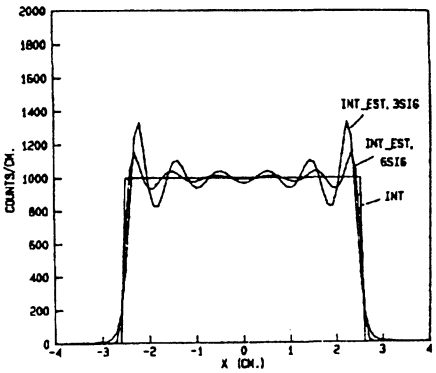

Figure 4. Demonstration of the Edge Artifact (from [13])

Both the noise and edge artifacts can be greatly reduced simultaneously by combining the use of a sieve and resolution kernel [13]. This is demonstrated in Figure 5 showing the result of a one-dimensional simulation for the same data used in Figure 3. A Gaussian convolution-kernel with a 1 cm FWHM and a Gaussian resolution kernel with a 0.8 cm FWHM were used, and 5000 iterations of the expectation-maximization algorithm were performed.

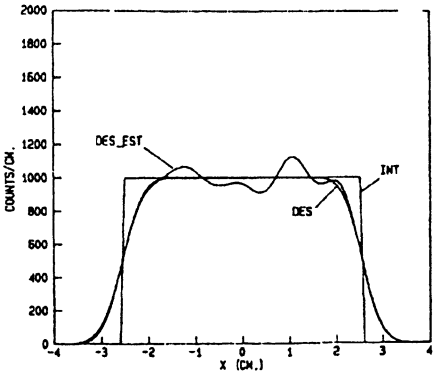

Figure 5. Demonstration of the Use of Sieves and Resolution Kernels to Suppress Noise and Edge Artifacts (from [13])

There are other benefits which accrue with the use of sieve and resolution kernels, and these are discussed in [13]. For example, without a resolution kernel, it is found that very many iterations are needed to reach a stable estimate if one is reached at all, but a stable estimate is achieved with significantly fewer iterations when the width of the resolution kernel exceeds the width of the sieve kernel. Also, the sensitivity seen in Figure 4 to implementation parameters is greatly reduced when a resolution kernel is used.

Performance comparisons of the CW and constrained ML methods

We have performed extensive studies in which the confidence-weighted and constrained maximum-likelihood methods are compared. These studies are quite tedious because more than a day of dedicated computation time is needed to produce an image using the 200 iterations we have found necessary to reach a stable estimate with the constrained maximum-likelihood method. We will describe comparisons in which a phantom motivated by neurological applications was used; studies have also been performed for pie and chest phantoms.

Shown in Figure 6 is the *Hoffman brain phantom.* This is a plastic phantom made by the Data Spectrum Company. It consists of voids in a known geometry into which a radioactive liquid can be placed for performance studies on emission tomographs. The phantom was filled with food coloring so that a computer model of it could be obtained with a CCD camera. The image on the left in Figure 6 is the resulting computer model. The image on the right was obtained by blurring the digitized image with a circularly symmetric Gaussian kernel with a 7.5 mm FWHM. This is the desired image $d(w)$ when we use blurring to reduce the edge artifact.

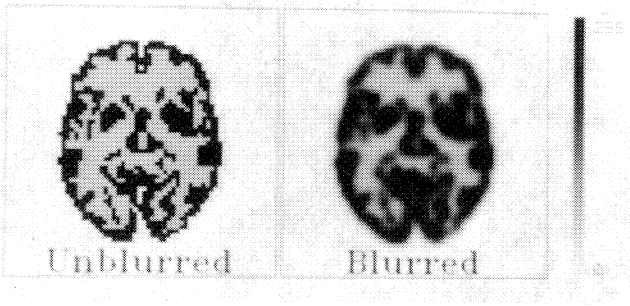

Figure 6. The Hoffman brain phantom (left),
and blurred Hoffman brain phantom (right)

Figure 7 shows the results of two studies we performed. In the left column are images obtained from computer simulated data for a positron emitting radiotracer in the Hoffman brain phantom, and the right column are the corresponding results for real data acquired with ^{15}O-labeled water and the Super PETT-I tomograph. An average of about 100,000 measured events was used. The similarity of the two images produced with the confidence-weighted method gives an indication that the computer simulation is a reasonable reflection of reality. The images shown for the constrained maximum-likelihood method were obtained with 200 iterations of the expectation-maximization algorithm. A circularly symmetric Gaussian density with a 7.5 mm FWHM was used as a convolution kernel for both the sieve and resolution constraints. We have obtained similar images for many other FWHM choices, including a large number in which the two kernels differ; neurologists who have participated in our comparison studies have tended to prefer those in which the two kernels are nearly the same. Inspection of the images produced with the constrained maximum-likelihood method compare favorably with those produced with the confidence-weighted method and to the perfect image on the right in Figure 6.

Figure 8 shows the results of another computer-simulation study we made with the Hoffman brain phantom. For this, twenty-five independent realizations of data from the Super PETT-I tomograph were obtained, taking three full weeks on two DEC VAX/8530 computers. For each realization, an image was produced using 200 iterations of the expectation-maximization algorithm and the same choices for sieve and resolution kernels as in Figure 7. The pooled results for the confidence-weighted and constrained maximum-likelihood methods are shown in the first and second rows of the figure, respectively. The mean images were obtained as the pixel-by-pixel arithmetic average over the twenty-five realizations. The mean-value image for the maximum-likelihood method compares favorably to the ideal result on right in Figure 6, and the mean for the confidence-weighted method contains greater noise. The images on the right are of the pixel-by-pixel numerical variance for the twenty-five realizations. By spatially averaging the gray and white matter areas of the phantom, the variance for the maximum-likelihood method is found to be about five times lower than that for the confidence-weighted method. This is significant because it opens the possibility of be-

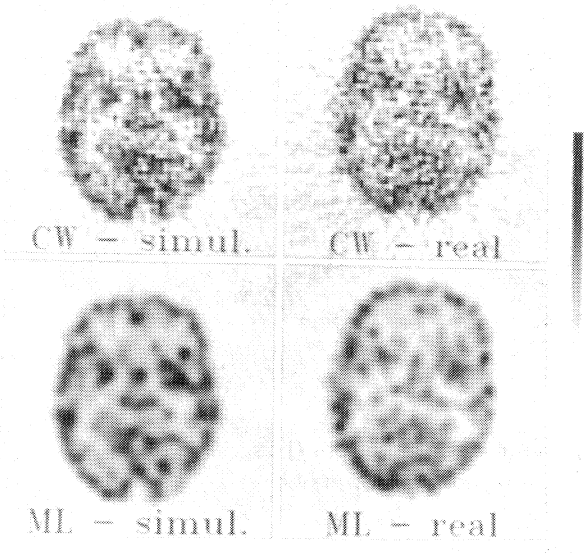

Figure 7. Images Produced with the *CW* and Constrained ML Methods for Computer Simulated Data (left) and Real Data (right)

ing able to perform brain mapping studies in individual patients, which is not presently feasible because images for several (typically five to seven) patients must be averaged to discern regional activity variations when the confidence-weighted method is used.

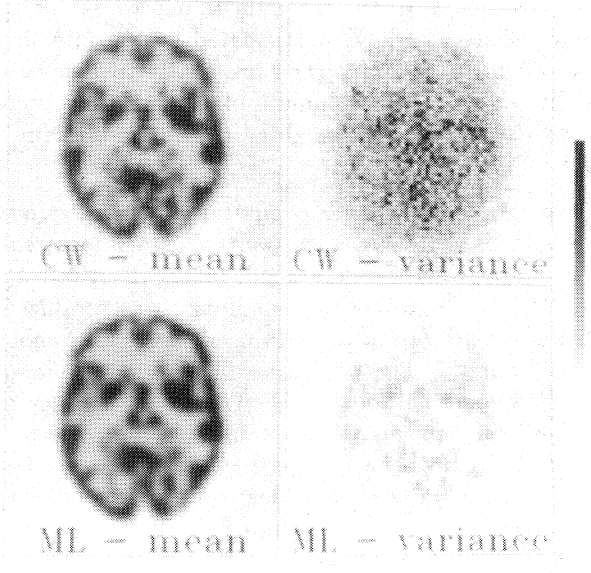

Figure 8. Average Images Produced with the *CW* and Constrained ML Methods for Twenty-Five Independent Computer Simulations

It is our conclusion based upon studies of this type that the noise and edge artifacts present with unconstrained maximum-likelihood estimation can be controlled and that the images which then result compare favorably to those of the confidence-weighted method. We now address the issue of long computation time for the maximum-likelihood method implemented via the expectation-maximization algorithm.

Implementation of the ML method

In our view, the long computation time required to produce images with the maximum-likelihood method is not a fundamental limitation. Table 1 contains a comparison made in our laboratory of the computation time required to perform 200 iterations of the expectation-maximization algorithm under a practical selection of parameters when the Super PETT-I tomograph is used for neurological studies [16,18].

Table 1. EM algorithm performance on Selected Architectures
(Super PETT-I parameters, 200 x 200 image array, 200 iterations)

Processor	Classification	Total Time/Iteration
Mercury ZIP	pipeline array processor	11 min
Masscomp 55020	image processing workstation	5 min
Cray XMP24	supercomputer	3 min
NCUBE	hypercube MIMD parallel processor	17 sec
InMOS Transputer	Mesh MIMD parallel processor	9 sec
multiple TMS320C30 DSPs	custom MIMD parallel processor	16 sec
Image Manipulation Modules	custom image processor	3 sec
NCR GAPP	systolic array processor	0.5 sec

The processors ranged from the Cray XMP supercomputer to NCUBE parallel processor to Image Manipulation Modules, which are being specially designed in the Institute for Biomedical Computing at Washington University [17] and the NCR GAPP, which is a mesh connected systolic array of simple one-bit processors. We have implemented a 2048 element GAPP processor in our Electronic Systems and Signals Research Laboratory [18]. The timing estimates for the Cray XMP, Masscomp, and NCR GAPP were obtained experimentally; the other estimates were obtained by performing operation counts. The Masscomp 55020 is a conventional 4 m.i.p. processor we use for performing the EM algorithm. The time per iteration, which includes both computation and communication time, is indicated in the last column of the table. It ranges

from 2.8 minutes to 0.5 seconds per iteration.

Neurologists in our School of Medicine are presently performing about 60 patient studies a week in their brain-mapping research, with each study producing seven images, for a total of about 420 images per week. Dividing into the number of seconds in a week indicates that a dedicated processor running full time would need to perform an iteration in seven seconds. This can be achieved with the specially designed Image Manipulation Modules and the NCR GAPP. We have also predicted that a specially designed processor based on the use of residue arithmetic can perform an iteration in about 10 ms, which would more than achieve the present requirements by our neurologists [19].

Thus, we conclude that the long computation time of the maximum-likelihood method is not a fundamental limitation precluding the use of the method in a clinical setting.

Acknowledgements

We appreciate the encouragement and support which we have received in this work from Dr. Lewis J. Thomas (Director of the WU Biomedical Computer Laboratory), Dr. Michael M. Ter-Pogossian (Director of the WU Division of Radiation Sciences, Mallinckrodt Institute of Radiology), Dr. Marcus E. Raichle (WU Department of Neurology), and Dr. Mark A. Mintun (WU Department of Radiology).

References

1. Ter-Pogossian, M. M., Ficke, D. C., Yamamoto, M., & Hood, J. T. Sr. (1982). Super PETT I: A Positron Emission Tomograph Utilizing Photon Time-of-Flight Information. *IEEE Transactions on Medical Imaging* **MI-1**, 179-187.

2. Holmes, T. J. (1985). *Data Collection and Preprocessing Considerations in Time-of-Flight Positron-Emission Tomography.* D. Sc. Thesis, Department of Electrical Engineering, Washington University.

3. Snyder, D. L., Thomas, L. J. Jr., & Ter-Pogossian, M. M. (1981). A Mathematical Model for Positron-Emission Tomography Systems Having Time-of-Flight Measurements. *IEEE Transactions on Nuclear Science* **NS-28**, 3575-3583.

4. Vardi, Y., Shepp, L. A., & Kaufman, L. (1985). A Statistical Model for Positron Emission Tomography. *Journal of American Statistical Association* **80**, 8-20.

5. Holmes, T. J., Ficke, D. C., & Snyder, D. L. (1984). Modeling of Accidental Coincidences in both Conventional and Time-of-Flight Positron-Emission Tomography. *IEEE Transactions on Nuclear Science* **NS-31**.

6. Snyder, D. L. (1984). Utilizing Side Information in Emission Tomography. *IEEE Transactions on Nuclear Science* **NS-31**, 533-537.

7. Allemand, R., Compagnolo, R., Garderet, P., Gariod, R., Gresset, C., Janin, C., Lavel, M., Odru, R., Tournier, E., & Vacher, J. (1980). A New Time-of-Flight Method for Positron Computed Tomography. *Proceedings of the NSR/CNRS U.S./ France Seminar on Biomedical Image Processing*, Grenoble.

8. Shepp, L. A., & Vardi, Y. (1982). Maximum Likelihood Reconstruction for Emission Tomography. *IEEE Transactions on Medical Imaging* **MI-1**, 113-121.

9. Snyder, D. L., & Politte, D. G. (1983). Image Reconstruction from List-Mode Data in an Emission Tomography System Having Time-of-Flight Measurements *IEEE Transactions on Nuclear Science* **NS-30**, 1843-1849.

10. Politte, D. G., & Snyder, D. L. (1984). Results of a Comparative Study of a Reconstruction Procedure for Producing Improved Estimates of Radioactivity Distributions in Time-of-Flight Emission Tomography. *IEEE Transactions on Nuclear Science* **NS-31**, 614-619.

11. Shepp, L. A., Vardi, Y., Ra, J. B., Hilal, S. E., & Cho, Z. H. (1984). Maximum-Likelihood with Real Data. *IEEE Transactions on Nuclear Science* **NS-31**, 910-913.

12. Snyder, D. L., & Miller, M. I. (1985). The Use of Sieves to Stabilize Images Produced with the EM Algorithm for Emission Tomography. *IEEE Transactions on Nuclear Science* **NS-32**, 3864-3872.

13. Snyder, D. L., Miller, M. I., Thomas, L. J. Jr., & Politte, D. G. (1987). Noise and Edge Artifacts in Maximum-Likelihood Reconstructions for Emission Tomography. *IEEE Transactions on Medical Imaging* **MI-6**, 228-238

14. Grenander, U. (1981). *Abstract Inference.* John Wiley and Sons, New York.

15. Roysam, B., Shrauner, J. A., & Miller, M. I. (1988). Bayesian Imaging Using Good's Roughness Measure–Implementation on a Massively Parallel Processor. *Proceedings of the IEEE International Conference on Acoustic, Speech, and Signal Processing,* New York.

16. Smith, K. R. (1987). Mapping the EM Algorithm for Time-of-Flight Positron Emission Tomography onto Parallel Architectures. *Biomedical Computer Laboratory Monograph* **484**, Institute for Biomedical Computing, Washington University.

17. Rosenberger, F. U. (1988). Image Manipulation Modules. *Computer Systems Laboratory Memo.* **328**, Institute for Biomedical Computing, Washington University.

18. McCarthy, A. W., Barrett, R. C., & Miller, M. I. (1988). Systolic Implementation of the EM Algorithm for Emission Tomography on a Mesh Connected Processor. *Proceedings of the Twenty-Second Annual Conference on Information Sciences and Systems,* Princeton University, Princeton, NJ.

19. Klotz, H., & Snyder, D. L. (1988). A Hardware Architecture Using Finite-Field Arithmetic for Computing Maximum-Likelihood Estimates in Emission Tomography. *IEEE Transactions on Medical Imaging.*